

JOSÉ HENRIQUE FORTE STORNIOLI

Caracterização do satelitoma de *Prochilodus lineatus* (Teleostei,
Characiformes) e a diferenciação de cromossomos B

Bauru

2021

JOSÉ HENRIQUE FORTE STORNIOLI

Caracterização do satelitoma de *Prochilodus lineatus* (Teleostei,
Characiformes) e a diferenciação de cromossomos B

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. Fábio Porto Foresti

Bauru

2021

Dados Internacionais de Catalogação na Publicação (CIP)
Ana Cláudia Inocente Garcia - CRB 8/6887

S885c Stornioli, José Henrique Forte
Caracterização do satelitoma de *Prochilodus lineatus* (Teleostei, Characiformes) e a diferenciação de cromossomos B / José Henrique Forte Stornioli. Assis, 2021.
51 p. : il.

Dissertação de Mestrado - Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Assis
Orientador: Dr. Ricardo Utsunomia
Coorientador: Dr. Fábio Porto Foresti

1. Citogenômica. 2. DNA Satélite. 3. Satelitoma.
I. Título.

CDD 572.8

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE JOSÉ HENRIQUE FORTE STORNIOLI, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM BIOCÊNCIAS, DA FACULDADE DE CIÊNCIAS E LETRAS - CÂMPUS DE ASSIS.

Aos 17 dias do mês de março do ano de 2021, às 09:00 horas, por meio de Videoconferência, realizou-se a defesa de DISSERTAÇÃO DE MESTRADO de JOSÉ HENRIQUE FORTE STORNIOLI, intitulada **Caracterização do satelitoma de *Prochilodus lineatus* (Teleostei, Characiformes) e a diferenciação de cromossomos B**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. RICARDO UTSUNOMIA (Participação Virtual) do(a) UFRJ/Rio de Janeiro, Prof. Dr. ROBERTO FERREIRA ARTONI (Participação Virtual) do(a) Biologia Estrutural Molecular e Genética / UEPG, Prof. Dr. VLADIMIR PAVAN MARGARIDO (Participação Virtual) do(a) UNIOESTE. 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

Agradecimentos

Meus sinceros agradecimentos a todos os animais utilizados para o desenvolvimento deste trabalho, estes sempre manejados respeitando o código de ética na experimentação animal, com todo o respeito que eles merecem. Muito obrigado!

Agradeço ao meu orientador Profº Dr. Ricardo Utsunomia pela orientação durante todo este trabalho, pelos inúmeros ensinamentos que tive durante o período do mestrado com a convivência, as conversas e as discussões, irei levar comigo todos os ensinamentos ao longo da minha jornada.

Ao meu co-orientador Profº Dr. Fábio Foresti pelos inúmeros ensinamentos ao longo da graduação e ao longo da minha vida acadêmica, pelos cafés e conversas que foram deveras importantes.

Aos meus colegas de laboratório Caio Goes, Zeni, Léo, Milan, Pedro, Rafael, Fernanda, Lilian, Carol, Caio, Diego e Yasmin, sem vocês este trabalho não teria sido realizado com êxito, obrigado por todas as conversas, ensinamentos, discussões e cafés, sou imensamente grato por todo o aprendizado e auxílio na realização deste trabalho.

Ao Profº Dr. Sergio Pereira pelos inúmeros ensinamentos ao longo de minha vida acadêmica e por me auxiliar a chegar no final do meu curso de mestrado, sou muito grato por tudo que realizamos ao longo destes anos.

Meus agradecimentos à minha mãe Rosimeire e as minhas irmãs Carol e Victória por acreditarem em mim e permitirem que eu pudesse seguir na carreira acadêmica e que finalizasse este trabalho, sem o apoio de vocês teria sido impossível.

Aos servidores da Unesp – Assis por toda a assistência dada ao longo do meu curso de mestrado e também a todos os servidores da Unesp – Bauru por todo o auxílio ao longo da minha formação, em especial àqueles que estão trabalhando presencialmente em um momento tão difícil quanto este em que vivemos com a pandemia.

A todas as agências de fomento que, desde o início da minha jornada acadêmica, me possibilitaram continuar longe de casa nos meus estudos. A Pró-Reitoria de Extensão da UNESP, a CAPES, a FAPESP e ao CNPq pelo suporte financeiro tanto para mim quanto para os laboratórios que já atuei.

STORNIOLI, José Henrique Forte. **Caracterização do satelitoma de *Prochilodus lineatus* (Teleostei, Characiformes) e a diferenciação de cromossomos B**. 2021. 51 f. Dissertação (Mestrado Acadêmico em Biociências) – Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Assis, 2021.

Resumo

O genoma dos organismos eucarióticos apresenta, além dos genes de cópia única, uma série de elementos repetitivos que podem ser classificados em duas grandes categorias: os elementos transponíveis e as repetições *in tandem*. As repetições *in tandem* são organizadas em uma orientação *head-to-tail* e podem formar longos arrays cromossômicos. Dentro desta categoria, podemos encontrar os DNA satélites (satDNA) que, de uma maneira geral, se localizam em regiões de heterocromatina. Além desta característica, o genoma dos organismos eucarióticos possui um conjunto padrão de cromossomos, denominados A, e pode possuir cromossomos supranumerários, denominados B. Uma particularidade dos cromossomos B é que geralmente são altamente heterocromáticos e podem possuir uma elevada abundância de satDNAs em sua composição. Entre os Characiformes, *Prochilodus lineatus* foi a primeira espécie na qual houve descrição de cromossomo B e desde então é uma espécie que vem sendo utilizada para estudos referentes à dinâmica evolutiva desses elementos. Graças a utilização do sequenciamento de nova geração (NGS) aliado a ferramentas bioinformáticas é possível analisar o conjunto de satDNA de um organismo sequenciado. Dessa forma, o objetivo do trabalho foi caracterizar o satelitoma de *P. lineatus* e mapear os satDNAs através da técnica de FISH em duas populações. Ambas as populações estudadas apresentaram cariótipos com $2n=54$ cromossomos; além disso, os indivíduos de uma população apresentam microcromossomos B, enquanto os indivíduos da outra não, caracterizando-se pela primeira ocorrência de uma população de *P. lineatus* sem cromossomos B. A caracterização do satelitoma desta espécie revelou a existência de 51 famílias de satDNAs. Observamos que PliSat01 e PliSat02 estão colocalizados nos centrômeros mas não apresentam similaridade entre si. Por fim é possível que os cromossomos B de *P. lineatus* tenham se originado a partir do cromossomo 4 do complemento padrão A, pois este apresenta sinal de FISH de quase todos os satDNA presentes em cromossomos B exceto o PliSat14, exclusivo dos cromossomos B.

Palavras-chave: Citogenômica. DNA Satélite. Satelitoma.

STORNIOLI, Jose Henrique Forte. **Characterization of the *Prochilodus lineatus* satellitome (Teleostei, Characiformes) and the differentiation of B chromosomes.** 2021 51 p. Dissertation (Masters in Bioscience) Sao Paulo State University (UNESP), School of Science, Humanities and Languages, Assis, 2021

Abstract

The genome of eukaryotic organisms has, in addition to single copy genes, a series of repetitive elements that can be classified into two broad categories: the transposable elements and in tandem repetitions. In tandem repetitions are arranged in a head-to-tail orientation and can form long chromosomal arrays. Within this category, we can find the satellite DNA (satDNA) that are usually located in heterochromatin regions. In addition to this characteristic, the genome of eukaryotic organisms has a standard set of chromosomes, named A chromosomes, and may have supernumerary chromosomes, named B chromosomes. A particularity of B chromosomes is that they are generally highly heterochromatic and they may have a high abundance of satDNAs in their composition. Among Characiformes, *Prochilodus lineatus* was the first species, in which there was a description of the B chromosomes, and since then it has been used for studies related to the evolutionary dynamics of these elements. Due the use of new generation sequencing (NGS) combined with bioinformatics tools it is possible to analyze the satDNA set of a sequenced organism. Thus, the objective of the study was to characterize the *P. lineatus* satellite and cytolocalizing the satDNAs using the FISH technique in two populations. Both populations studied showed karyotypes with $2n = 54$ chromosomes; in addition, individuals from one population had B microchromosomes, while the other did not present, characterized the first occurrence of a population of *P. lineatus* without B chromosomes. The characterization of this species' satellitome revealed the existence of 51 families of satDNAs. We observed that PliSat01 and PliSat02 are colocalized in the centromeres but are not similar to each other. Finally, it is possible that the B chromosomes of *P. lineatus* originated from chromosome 4 of the standard complement A, as this shows a FISH signal from almost all satDNAs present on B chromosomes except PliSat14, exclusive to B chromosomes.

Keywords: Cytogenomics. Satellite DNA. Satellitome.

Sumário

1. Introdução Geral.....	8
2. Objetivos	13
2.1 Objetivos gerais	13
2.2 Objetivos específicos	13
3 Material e métodos	14
3.1 Amostragem dos indivíduos de <i>P. lineatus</i>	14
3.2 Análises citogenéticas.....	14
3.2.1 Estimulação de mitoses	15
3.2.2 Obtenção de cromossomos metafásicos <i>in vivo</i>	15
3.3 Obtenção de cromossomos metafásicos <i>in vitro</i>	16
3.4 Coloração convencional com Giemsa	17
3.5 Extração de DNA genômico.....	17
3.6 Sequenciamento de nova geração (NGS).....	17
3.7 Análises bioinformáticas em servidor remoto	18
3.7.1 Identificação de DNAs satélites com Tarean – RepeatExplorer – satMiner.....	18
3.8 Desenho de <i>primers</i> e preparação das sondas:.....	19
3.9 Hibridação <i>in situ</i> fluorescente – FISH	20
4 Capítulo I	22
4.1 Introdução.....	22
4.2 Material e Métodos.....	24
4.2.1 Amostragem dos indivíduos	24
4.2.2 extração de DNA e sequenciamento do genoma	25
4.2.3 Caracterização do Satelitoma.....	25
5. Resultados	27
5.1 <i>P. lineatus</i> - Cariótipos e a primeira descrição de uma população sem cromossomos B	27
5.2 O satelitoma de <i>P. lineatus</i>	27
5.3 Distribuição cromossômica das 37 famílias de DNA satélite em <i>P. lineatus</i>	32
6 Discussão	37
6.1 Macroestrutura cariotípica em <i>Prochilodus lineatus</i>	37
6.2 O satelitoma do curimatá e a diferenciação dos cromossomos B.....	38
7. Referências.....	40
8. Material suplementar.....	47

1. Introdução Geral

1.1 Considerações sobre DNAs repetitivos - DNAs satélites

O genoma dos organismos eucarióticos apresenta, além dos já amplamente conhecidos "genes de cópia única", uma série de elementos repetitivos que contemplam a maior parte dos genomas das espécies (BISCOTTI; OLMO; HESLOP-HARRISON, 2015). Tais sequências de DNA repetitivo podem ser divididas em duas categorias básicas, baseadas em suas características particulares de organização genômica: os elementos dispersos e as repetições *in tandem* (YUAN et al., 2018). Enquanto os elementos dispersos são sequências móveis, como os transposons e retrotransposons (CHUONG; ELDE; FESCHOTTE, 2017), as repetições *in tandem* são unidades repetidas em uma orientação *head-to-tail*, que podem formar longos *arrays* cromossômicos, que se estendem de algumas dezenas até centenas de milhares de repetições (CHARLESWORTH; SNIEGOWSKI; STEPHAN, 1994; PALOMEQUE; LORITE, 2008). Dentre estas sequências, destacam-se aquelas que são codificantes, como as famílias multigênicas (*e.g.* genes de histona, genes Us e genes ribossômicos) e as não-codificantes, como os DNAs satélites (satDNA) (LÓPEZ-FLORES; GARRIDO-RAMOS, 2012).

Os satDNAs constituem parte das repetições *in tandem* que podem estar presentes no genoma de um organismo, podendo ser classificados, de acordo com a quantidade de pares de bases, em microsatélites, minissatélites e satélites; entretanto, ainda não há uma definição precisa na literatura sobre essa classificação, principalmente no que diz respeito aos micro e minissatélites (RICHARD; KERREST; DUJON, 2008; RUIZ-RUANO et al., 2016). Estas sequências são um dos principais componentes da heterocromatina, região do genoma que se mantém condensada na maior parte do tempo e com baixa presença de genes (PLOHL; MEŠTROVIĆ; MRAVINAC, 2012) e são encontradas, de uma maneira geral, mas não restritas, nas regiões pericentroméricas e subteloméricas dos cromossomos. (LÓPEZ-FLORES; GARRIDO-RAMOS, 2012). Um genoma eucariótico pode apresentar uma diversificada coleção de famílias de satDNAs, sendo este conjunto denominado "satelitoma" (GARRIDO-RAMOS, 2017; RUIZ-RUANO et al., 2016).

Por não apresentarem uma função definida nos organismos em geral, os satDNAs eram considerados, junto com outros componentes do genoma, como parte do

“DNA lixo” (do inglês: *junk DNA*; GARRIDO-RAMOS, 2015). Embora muitas sequências satélites obedeam a esta regra (ausência de papel funcional bem definido), algumas delas podem desempenhar um papel relevante do ponto de vista estrutural e funcional dos genomas como RNAs não-codificantes (PLOHL; MEŠTROVIĆ; MRAVINAC, 2012, 2014; LÓPEZ-FLORES; GARRIDO-RAMOS, 2012; HALBACH et al., 2020).

Os satDNAs podem ser classificados em famílias que diferem entre si em sua composição geral, bem como em relação ao tamanho da unidade de repetição, abundância genômica e organização cromossômica (RUIZ-RUANO et al., 2016). Por serem sequências de evolução rápida e estando, *a priori*, livres de seleção, geralmente elas são espécie-específicas, porém podem ser encontradas em organismos próximos ou até mesmo em uma mesma ordem (GARRIDO-RAMOS, 2015; CHAVES et al., 2017; DOS SANTOS et al., 2021). Em estudo pioneiro, Fry & Salser (1977) propuseram a hipótese *library* como uma ideia de evolução a longo prazo de satDNAs, que prediz que grupos de organismos próximos compartilham várias famílias de satDNAs e que as diferenças observadas nesta fração genômica seriam majoritariamente quantitativas.

Apesar da sua abundância nos genomas eucarióticos, os satDNAs foram pouco explorados até o início dos anos 2010. Em peixes neotropicais, por exemplo, destaca-se que menos de 10 satDNAs haviam sido caracterizados (revisão em VICARI et al., 2010). Esta ausência de dados pode ser explicada pela indisponibilidade de ferramentas genéticas simples e não laboriosas até aquele momento. Por exemplo, para que um satDNA fosse caracterizado a partir de uma espécie, experimentos de isolamento utilizando-se de restrição enzimática, seguidos de clonagem, sequenciamento e *southern blot* eram necessários (VICARI et al., 2010). Além disso, destaca-se que este processo não garantia a caracterização de um satDNA da espécie em estudo, uma vez que a seleção correta (e aleatória) da enzima de restrição é mandatória.

Com a expansão das técnicas de sequenciamento massivo, bem como de ferramentas bioinformáticas específicas, a análise de coleções completas de satDNAs em espécies não-modelo se tornou possível desde o início dos anos 2010 (NOVÁK et al., 2013; NOVÁK; NEUMANN; MACAS, 2010; RUIZ-RUANO et al., 2016; UTSUNOMIA et al., 2016). Assim, a partir de uma biblioteca de sequenciamento com baixa cobertura genômica é possível caracterizar coleções inteiras de DNAs satélites de qualquer espécie-alvo. Usando esta metodologia, o catálogo do gafanhoto *Locusta*

migratoria foi caracterizado, sendo encontrados 62 famílias de satDNAs (RUIZ-RUANO et al., 2016), enquanto os satelitomas dos peixes neotropicais *Psalidodon paranae*, *Characidium gomesi* e *Megaleporinus macrocephalus* também foram prospectados, sendo encontrados, respectivamente, 45 (SILVA et al., 2017), 59 (SERRANO-FREITAS et al., 2020) e 164 famílias de satDNAs (UTSUNOMIA et al., 2019).

Além de permitir a caracterização dos catálogos dos satDNAs de qualquer espécie de interesse, as informações geradas nos trabalhos citados anteriormente têm permitido algumas inferências sobre a persistência evolutiva de determinados satDNAs em alguns clados (UTSUNOMIA et al., 2017; RODRIGUES et al., 2019; DOS SANTOS et al., 2021), bem como compreender o papel deste tipo de sequência na evolução e diferenciação de cromossomos sexuais e supranumerários (FERRETTI et al., 2020; RUIZ-RUANO et al., 2016; SERRANO-FREITAS et al., 2020; SILVA et al., 2017; UTSUNOMIA et al., 2016, 2019).

1.2 Cromossomos B

Cromossomos supranumerários, ou cromossomos B, são elementos genéticos extras presentes nos genomas de algumas espécies eucarióticas (CAMACHO; SHARBEL; BEUKEBOOM, 2000). Em princípio, estes elementos não apresentam qualquer função definida e têm sido caracterizados como parasitas intracelulares, capazes até mesmo de modular o funcionamento das células em prol da sua própria manutenção nas populações, beneficiando as suas próprias taxas de transmissão, processo coletivamente conhecido como *drive* (CAMACHO; SHARBEL; BEUKEBOOM, 2000; HOUBEN, 2017; VALENTE et al., 2014).

Em linhas gerais, os cromossomos B podem se originar por dois mecanismos básicos: i) do genoma da própria espécie hospedeira (origem intraespecífica), sendo este o processo mais comumente observado; ou ii) do genoma de uma segunda espécie (origem interespecífica; CAMACHO; SHARBEL; BEUKEBOOM, 2000). Uma vez originados e estabelecidos em determinada população, estes cromossomos começam a evoluir de forma análoga aos cromossomos sexuais heteromórficos (Y e W) por não se recombinarem recorrentemente com os cromossomos A (GREEN, 1990). Por esta razão, o acúmulo de sequências repetitivas e pseudogenes nestes elementos é bastante

documentado e estudos anteriores têm confirmado esta hipótese (revisão em CAMACHO; SHARBEL; BEUKEBOOM, 2000; HOUBEN, 2017).

Por apresentarem este acúmulo exacerbado de sequências repetitivas a caracterização da ocorrência de satDNAs, famílias multigênicas e elementos transponíveis nestes elementos constituem uma abordagem primária para o conhecimento do conteúdo completo de cromossomos B, bem como permitem uma melhor compreensão dos processos de origem e evolução destes cromossomos.

1.3 Citogenética de peixes Neotropicais com foco em *Prochilodus*

A região Neotropical (que compreende a América Central e a América do Sul) possui 20% das espécies de peixes descritas atualmente, contemplando mais de 6000 espécies e com estimativas do número total ser acima de 8000, desta forma sendo a região com maior biodiversidade de peixes no mundo (MALABARBA; MALABARBA, 2019). Dentre os peixes que se distribuem nesta região, a ordem Characiformes se destaca como uma das mais ricas em número de espécies, sendo representada atualmente por 23 famílias e mais de mil espécies válidas (ARCILA et al., 2017)

A família Prochilodontidae está inserida na ordem Characiformes e é composta atualmente por três gêneros (*Ichthyolephas*, *Prochilodus* e *Semaprochilodus*) e 21 espécies válidas, amplamente distribuídas pela região Neotropical (CASTRO; VARI, 2004; FRICKE 2021). Dentre estas espécies, o curimbatá *Prochilodus lineatus* representa um peixe migrador de grandes distâncias, capaz de percorrer milhares de quilômetros em época reprodutiva (SVERLIJ et al., 1993).

Estudos citogenéticos anteriores nesta espécie revelaram um número diploide conservado de $2n=54$ cromossomos, além da ampla ocorrência de um sistema de microcromossomos B (ARTONI et al., 2006a; DE JESUS et al., 2003; OLIVEIRA et al., 1997; PAULS; BERTOLLO, 1983, 1990), o qual estudos posteriores também revelaram estar presente em outras espécies do gênero (MELO et al., 2017; VOLTOLIN et al., 2013). Apesar destes microcromossomos supranumerários não serem facilmente identificados e classificados, evidências apontam para a existência de pelo menos três variantes (ARTONI et al., 2006a; PENITENTE et al., 2013), sendo que as taxas de

transmissão de cada uma destas variantes diferem entre si (PENITENTE et al., 2013). Embora o conteúdo gênico destes elementos não seja conhecido, é sabido que os cromossomos B desta espécie são compostos em abundância por um DNA satélite nomeado SATH1 (DE JESUS et al., 2003). Além dos cromossomos B, este DNA satélite está presente em regiões centroméricas de alguns pares cromossômicos, sugerindo a origem intraespecífica deste sistema de cromossomos B a partir destas regiões cromossômicas (ARTONI et al., 2006a; DE JESUS et al., 2003; MELO et al., 2017). No entanto, deve-se destacar que este é o único satDNA encontrado nestes elementos supranumerários de um total de dois satDNAs caracterizados até o momento (DE JESUS et al., 2003). Nesse contexto, a caracterização de elementos repetidos adicionais presentes nos cromossomos B desta espécie permitiria uma melhor compreensão dos eventos envolvidos na origem destes cromossomos, bem como sobre o processo de diferenciação destes elementos.

Diante do exposto, no presente trabalho nós caracterizamos o satelitoma de *P. lineatus* e mapeamos cromossomicamente 37 famílias de satDNAs, sendo que seis estão presentes nos cromossomos B desta espécie, expandindo o conhecimento sobre o conteúdo destes elementos supranumerários.

2. Objetivos

O presente projeto se insere no programa geral de estudos citogenéticos de peixes Neotropicais que vêm sendo desenvolvidos no Laboratório de Genética de Peixes (Faculdade de Ciências, UNESP/Bauru), visando um melhor conhecimento da estrutura e organização cromossômica de peixes. Assim, os objetivos do presente trabalho foram:

2.1 Objetivos gerais

- a) Caracterizar o satelitoma da espécie *Prochilodus lineatus*, o curimatá;
- b) Verificar se as variantes de cromossomos supranumerários apresentam algum grau de diferenciação citogenética.

2.2 Objetivos específicos

- a) Citolocalizar as regiões cromossômicas onde os satDNAs isolados se encontram em indivíduos de duas populações provenientes de localidades distintas.

3 Material e métodos

3.1 Amostragem dos indivíduos de *P. lineatus*.

As análises citogenéticas e moleculares nos indivíduos de *P. lineatus* foram realizadas em animais coletados em duas localidades: nos tanques de cultivo de uma piscicultura localizada em Sales de Oliveira, São Paulo, em que foram coletados indivíduos F1 da população do Rio Sapucaí, bacia do rio Sapucaí, destinados ao repovoamento desta mesma população. A outra coleta foi realizada em uma população natural do Rio Apa, na divisa entre o Estado do Mato Grosso do Sul e o Paraguai, bacia do Rio Paraguai, onde foram coletados indivíduos adultos. Para a população do Rio Sapucaí foram utilizados 7 indivíduos que originaram as amostras e, para os estudos, foi escolhido a amostra que apresentava as três variantes de cromossomos B. Para a população do rio Apa foi coletado sangue de 20 indivíduos, que originaram as amostras analisadas. 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 1204-CEUA/2019.

Os exemplares provenientes da população do rio Sapucaí foram anestesiados e dissecados, sendo retirados fragmentos da porção anterior do rim para estudos citogenéticos e fragmentos de tecido muscular para posterior extração de DNA e análises moleculares. Os exemplares utilizados na amostragem foram fixados em formol 10%, conservados em álcool 70% e depositados na coleção de peixes do Laboratório de Genética de Peixes, Faculdade de ciências, UNESP - Bauru-SP, sob os números de *vouchers* LG13393 até LG13407. Os indivíduos coletados na população do rio Apa foram devolvidos à natureza após a coleta de amostras de sangue da veia caudal para a obtenção de cromossomos mitóticos a partir de cultura de linfócitos. Fragmentos de nadadeiras foram retirados e fixados em álcool absoluto, e depositados no banco de tecidos do Laboratório de Genética de Peixes, sob os *vouchers* LGP13501 até LGP13520.

3.2 Análises citogenéticas

Para a análise citogenética dos exemplares provenientes do rio Sapucaí, utilizamos o método de estimulação de mitoses (OLIVEIRA et al., 1988) e preparação direta de células renais *in vivo* (FORESTI; TOLEDO; TOLEDO FILHO, 1981) em

todos os exemplares analisados, conforme detalhado abaixo. Para os indivíduos coletados no rio Apa, utilizamos o método de obtenção de cromossomos metafásicos *in vitro* (FENOCCHIO & BERTOLLO, 1988).

3.2.1 Estimulação de mitoses

Para a obtenção das preparações cromossômicas, adotamos a técnica de estimulação de divisão celular, que consiste na inserção de uma solução de fermento biológico na musculatura dorsal do animal, como descrito em Oliveira et al. (1988) para peixes. O procedimento utilizado consiste em:

- 1) preparar uma solução de fermento biológico (Fleischmann) na seguinte proporção: 0,5g de fermento, 0,5g de açúcar e 7mL de água destilada;
- 2) incubar a solução em uma estufa (37°C) por cerca de 30 min;
- 3) injetar a solução na região dorsolateral do peixe, na proporção de 1mL para cada 100g de peso do animal. A dose administrada pode ser realizada em duas doses, dividida e aplicada em um período de 48h;
- 4) manter o animal em aquário bem aerado a 28°C.

3.2.2 Obtenção de cromossomos metafásicos *in vivo*

Para a obtenção de cromossomos metafásicos *in vivo*, seguimos o protocolo descrito por Foresti et al. (1981), utilizado com algumas adaptações. Essa técnica consiste na inibição da polimerização dos microtúbulos através da ação da colchicina, seguida consequentemente de uma hipotonização das células em suspensão, seguindo-se a fixação celular através de uma mistura de metanol/ácido acético (3:1). O procedimento consiste em:

1. injetar, na região intra-abdominal, solução aquosa de colchicina (0,025%) na proporção de aproximadamente 1mL para cada 100g de peso do animal;
2. deixar o peixe em aquário bem aerado, por um período aproximado de 50 min;
3. induzir a morte do animal e retirar fragmentos de rim anterior. Transferir o material para uma pequena placa de vidro, contendo 7mL de uma solução hipotônica de KCl (0,075M);

4. dissociar o material com o auxílio de pinças de dissecação, complementando esse processo com o auxílio de uma pipeta seringa de 3mL, até que se obtenha uma solução aquosa 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, colocando 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 cerca de 6mL de fixador e novamente agitar a mistura; levar à centrífuga (900 ± 100 rpm) por 10 min;
8. retirar o sobrenadante, e logo após ressuspender o material precipitado em 7mL de fixador; centrifugar por 7 min a 900 ± 100 rpm;
9. repetir o item 8 por três vezes para uma completa fixação das células em suspensão;
10. pingar o material em lâminas e deixar secar.

3.3 Obtenção de cromossomos metafásicos *in vitro*.

Para obtenção de cromossomos mitóticos *in vitro*, utilizamos o protocolo estabelecido em Fenocchio & Bertollo (1988). Esta técnica consiste em isolar os linfócitos de uma amostra sanguínea, e produzir uma cultura celular utilizando meio para cariótipos (RPMI - Cultilab) por 72h em estufa a 37°C. A partir daí, é feita a inibição da polimerização dos microtúbulos utilizando colchicina, seguido de hipotonização das células e subsequente fixação do material. O procedimento consiste em:

1. anestesiocar o animal com solução de benzocaína (100mg/L);
2. limpar a região caudal com álcool iodado e puncionar a veia caudal, retirando de 3 a 5 ml de sangue;
3. aguardar até a separação do soro, linfócitos e hemácias e retirar os linfócitos para serem colocados na cultura. Descartar o soro e as hemácias;
4. colocar os linfócitos em 5,2 ml de meio de cultura para cariótipo (RPMI - Cultilab);
5. incubar a cultura a cerca de 28°C em estufa por 72 h, protegendo da luz;

6. colocar 0,1ml de solução aquosa de colchicina a 0,025% por 15 min;
7. para colher o material, centrifugar por 7 minutos a 1000 rpm e descartar o sobrenadante;
8. acrescentar 7ml de solução hipotônica (KCL 0,075M) e agitar levemente;
9. deixar o tubo no interior de uma estufa a 28°C por 30 minutos;
10. retirar da estufa, colocar 6 gotas de fixador gelado (metanol e ácido acético na proporção de 3:1, respectivamente). Agitar levemente a mistura com uma pipeta de Pasteur e deixar repousar por 5 minutos a temperatura ambiente;
11. adicionar 6 ml de fixador gelado e novamente agitar a mistura. Levar a centrifuga (1000 rpm) por 10 minutos;
12. retirar o sobrenadante e ressuspender o precipitado em 6 ml de fixador. Centrifugar por 7 minutos a 1000 rpm;
13. repetir o item 12 por duas ou três vezes;
14. pingar o material em lâminas e deixar secar.

3.4 Coloração convencional com Giemsa

As preparações cromossômicas fixadas nas lâminas foram coradas por dez minutos com solução de Giemsa a 5% em tampão fosfato (pH=6,8) para a verificação do número diploide dos indivíduos de *P. lineatus* e a verificação da presença de cromossomos B.

3.5 Extração de DNA genômico

O DNA genômico foi obtido a partir de amostras de nadadeiras preservadas em etanol utilizando-se os Kits Wizard Genomic DNA Purification (PROMEGA) ou DNeasy Blood & Tissue (QIAGEN), conforme instruções recomendadas, com um passo para a remoção de RNA com RNase A (Invitrogen). Os produtos gerados foram visualizados em eletroforese de gel de agarose a 1%.

3.6 Sequenciamento de nova geração (NGS)

Para o presente estudo, uma biblioteca genômica de *Prochilodus lineatus* proveniente do rio Mogi-Guaçu foi utilizada. Esta biblioteca está depositada no banco de dados do NCBI e foi gerada através de sequenciamento massivo de nova geração (NGS), realizado na plataforma BGISEQ-500 da BGI (BGI Shenzhen Corporation, Shenzhen, China) (*paired-end* 2x100pb) (SRA accession number SRR11676686).

3.7 Análises bioinformáticas em servidor remoto

As análises de bioinformática foram realizadas remotamente através de um servidor de alta capacidade localizado no Laboratório de Biologia e Genética de Peixes do Departamento de Biologia Estrutural e Funcional da Universidade Estadual Paulista, *campus* de Botucatu, sob os cuidados do Prof. Dr. Claudio de Oliveira.

3.7.1 Identificação de DNAs satélites com Tarean – RepeatExplorer – satMiner

A caracterização dos satelitomas foi realizada utilizando as ferramentas TAREAN (NOVÁK et al., 2017) e satMiner (RUIZ-RUANO et al., 2016) em conjunto, seguindo os seguintes passos: i) preparar os *reads* para aplicação no programa TAREAN (NOVÁK et al., 2017), já com os cortes nos adaptadores e seleção de qualidade; ii) executar o *software* TAREAN com a opção padrão; iii) selecionar os clusters *high* e *low confidence*, identificados como DNAs satélites; iv) montar uma base de dados customizada com os satélites já isolados e utilizar um *script* (*deconseq_run.py*), para exclusão de *reads* semelhantes aos satélites já identificados da biblioteca inicial; v) selecionar novos *reads* para uma nova rodada de TAREAN; vi) realizar as etapas (i a v) até que nenhum DNA satélite novo seja encontrado; vii) utilizar as bibliotecas descontaminadas para preparar *reads* para aplicação no programa RepeatExplorer1; viii) executar o *software* Repeatexplorer1 com a opção padrão; ix) selecionar manualmente, com auxílio do *software* Geneious 8.4, os clusters que apresentarem a forma de gráficos circulares (formato padrão de DNAs satélites); x) repetir as etapas iv e v; xi) repetir as etapas (vii a x) até não ser mais encontrado sequências repetidas *in tandem*.

Após a identificação de todos os satélites, nós realizamos uma busca de homologia entre as sequências identificadas e as agrupamos em variante, família e superfamília se a identidade entre os monômeros foi maior do que 95%, 80% e 50%, respectivamente (RUIZ-RUANO et al., 2016). Os valores de abundância e divergência de cada variante foram calculados com o *software* RepeatMasker v4.0.5 (SMIT et al., 2017), com as opções *cross_match* e uma seleção de 10,000,000 de *reads*.

3.8 Desenho de *primers* e preparação das sondas:

Com os satélites caracterizados e suas sequências bem estabelecidas, *primers* foram desenhados para o preparo de sondas por PCR. Para os satélites menores (com até 600 pb), *primers* divergentes foram desenhados, enquanto para os satélites maiores do que 600 pb, *primers* convergentes foram desenhados, de forma a gerar um *amplicon* de aproximadamente 600 pb, ideal para hibridação *in situ* fluorescente (FISH). A qualidade dos *primers* foi verificada com o *Multiple Primer Analyzer* (ThermoFisher Scientific), de maneira a evitar a formação de grampos e auto-pareamento.

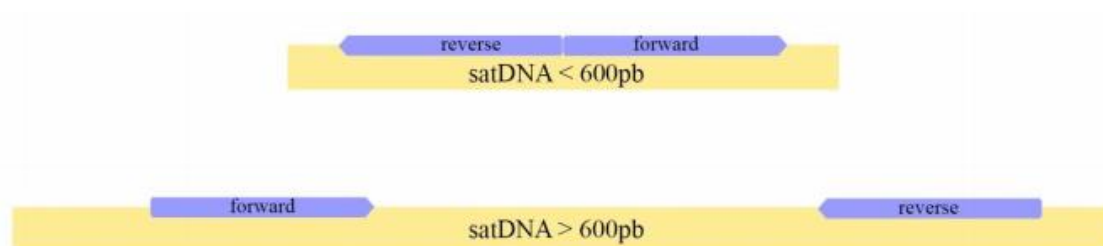


Figura 1. Exemplo dos *primers* divergente e convergente.

Reações de Amplificação em Cadeia da Polimerase (PCR) foram realizadas para cada um dos DNAs satélites, a partir do DNA genômico de *P. lineatus* em diluições de 1:10 (100ng/μL) até 1:1000 (1ng/μL) com os seguintes parâmetros: 8.75 μL de água autoclavada, 1.25 μL de Buffer, 0.75 μL de MgCl₂ (25 mM), 0.25 μL de dNTP (2 mM), 0.5 μL de Primer A + Primer B (10 mM), 0.1 μL de Taq polimerase (5 U/μL) e 1 μL de DNA, em um total da reação de 12 μL. A produção das sondas segue o mesmo protocolo, com algumas pequenas alterações, como a troca dos dNTPs regulares (0.25ul) por 0.35ul de dNTP + ½ dTTP, além da incorporação de digoxigenina-11-dUTP (Roche Applied Science).

Devido ao alto número de DNAs satélites identificados, as reações em termociclador variaram de acordo com a especificação de cada satélite e elas estão sumarizadas na figura abaixo.

95° C → 6 minutos	} 30x
95° C → 20 segundos	
50°-56° C → 10-50 segundos	
72° C → 10-50 segundos	
72° C → 14 minutos	

3.9 Hibridação *in situ* fluorescente – FISH

As sondas utilizadas na FISH de todos os DNAs satélites deste trabalho foram marcadas utilizando digoxigenina-11-dUTP, tendo como base a amplificação das sequências satélites no genoma de *P. lineatus*. Para a realização da técnica de FISH utilizamos os procedimentos descritos por Pinkel et al. (1986), utilizando algumas modificações já estabelecidas em nosso laboratório:

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 cada e secar as lâminas.

Mix de Hibridação

1. Preparar o *mix* de hibridação (15 ul de formamida 70%, 6 ul de sulfato dextrano 10%, 3 ul 20xSSC, entre 1 ou 2 ul de sonda por lâmina, completando de água até o total de 30ul);

Pré-Hibridação

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

Hibridação

1. Montar as lâminas com o 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 ul de NFDM/0,5 ul 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 ul de DAPI e cobrir com lamínula, mantendo estas em um local escuro até sua visualização.

ReFISH

Este protocolo tem como objetivo hibridizar novas sondas em lâminas previamente analisadas, permitindo a localização de sequências distintas em uma mesma metáfase.

1. Lavar em 2x SSC por 5 minutos a temperatura ambiente;
2. A partir daqui seguimos o protocolo da FISH a partir da Pré-Hibridação, a diferença é que a lâmina é desnaturada em formamida 70% em 2xSSC por 30 segundos a 70°C.

4. Capítulo I

Caracterização do satelitoma de *Prochilodus lineatus* (Teleostei, Characiformes) e a diferenciação de cromossomos B

José Henrique Forte Stornioli, Caio Augusto Gomes Goes, Rodrigo Milan Calegari, Rodrigo Zeni dos Santos, Leonardo Moura Giglio, Fausto Foresti, Claudio Oliveira, Manolo Penitente, Fábio Porto-Foresti, Ricardo Utsunomia

4.1 Introdução

A porção das sequências repetitivas dos genomas eucarióticos pode ser classificada em elementos dispersos, como os elementos transponíveis (TEs), ou em sequências organizadas *in tandem*, que abrangem os grupos das famílias multigênicas e os DNAs satélites (satDNAs; CHARLESWORTH; SNIEGOWSKI; STEPHAN, 1994; JURKA et al., 2005). Com exceção das famílias multigênicas, os satDNA constituem uma coleção diversa e dinâmica de sequências que se diferem em sua organização, composição e localização, embora não desempenhem um papel funcional geral e definido (BISCOTTI; OLMO; HESLOP-HARRISON, 2015; CHARLESWORTH; SNIEGOWSKI; STEPHAN, 1994).

Os DNAs satélites são unidades monoméricas com tamanhos variáveis e arrançadas *in tandem*, ou seja, com monômeros organizados em uma orientação cabeça-a-cauda de forma sequencial, podendo se estender de dezenas até milhões de repetições, sendo localizados majoritariamente, porém não exclusivamente, na heterocromatina centromérica e pericentromérica (GARRIDO-RAMOS, 2017; MIGA et al., 2020; PLOHL; MEŠTROVIĆ; MRAVINAC, 2012). Cada espécie carrega seu próprio catálogo de satDNAs, resultando em uma ampla diversidade de famílias destas sequências, que geralmente surgem independentemente de regiões genômicas distintas, inicialmente por uma duplicação *de novo*, sendo seguida, ocasionalmente, por clusterização (RUIZ-RUANO et al. 2016). Como os principais componentes da heterocromatina, os satDNAs são propensos a se acumularem em regiões de baixa, ou nenhuma, recombinação, o que inclui os cromossomos sexuais e os cromossomos B, devido a processos análogos à catraca de Müller (CAMACHO; SHARBEL;

BEUKEBOOM, 2000; CHARLESWORTH; CHARLESWORTH; MARAIS, 2005; GARRIDO-RAMOS, 2017; GREEN, 1990; PLOHL; MEŠTROVIĆ; MRVINAC, 2012).

Os cromossomos B são cromossomos supranumerários que usualmente não obedecem às leis Mendelianas (CAMACHO; SHARBEL; BEUKEBOOM, 2000). Eles geralmente se originam a partir de quebras cromossômicas que ocorrem nos cromossomos A de suas próprias espécies, e podem se acumular em uma população natural por mecanismos coletivamente conhecidos como *drive*, que ocorrem quando as taxas de transmissão dos cromossomos são superiores a 0,5 (CAMACHO; SHARBEL; BEUKEBOOM, 2000; HOUBEN, 2017). Como já ressaltado, os DNAs repetitivos desempenham um papel fundamental na diferenciação dos cromossomos B (devido à sua natureza não recombinante), além de também poderem fornecer algumas indicações sobre a origem desses elementos (ARTONI et al., 2006b; DA SILVA et al., 2014; LÓPEZ-LEÓN et al., 1994; MESTRINER et al., 2000; SERRANO-FREITAS et al., 2020; TERUEL et al., 2010; UTSUNOMIA et al., 2016).

Pertencendo a Ordem Characiformes, *Prochilodus lineatus* é um peixe migratório que possui um número diplóide de 54 cromossomos. Notavelmente, todas as populações analisadas desta espécie até o momento apresentam sistemas de cromossomos B composto por, pelo menos, três variantes de microcromossomos B (ARTONI et al., 2006a; MAISTRO; OLIVEIRA; FORESTI, 2000; PENITENTE et al., 2013, 2016; VOLTOLIN et al., 2010a). Da mesma forma, outras espécies deste mesmo gênero também apresentam microcromossomos B semelhantes, tendo sido sugerido que esses elementos se originaram no último ancestral comum do gênero *Prochilodus* (MELO et al., 2017). Em um estudo pioneiro, Jesus et al. (2003) isolaram dois satDNAs pericentroméricos do genoma de *P. lineatus* (SATH1 e SATH2), dos quais um (SATH1) também se localizava em abundância nos cromossomos B, sugerindo uma origem intraespecífica para esses elementos (ARTONI et al., 2006a; DE JESUS et al., 2003). Embora este seja um dos sistemas de cromossomos B mais bem estudados em peixes, sua composição molecular é restrita a um único DNA satélite (DE JESUS et al., 2003).

Nos últimos anos, a caracterização massiva de sequências de DNA repetitivo se tornou viável devido à expansão de tecnologias de sequenciamento de nova geração, bem como o desenvolvimento de *pipelines* para o estudo dessas porções genômicas

(HARRIS; CECHOVA; MAKOVA, 2019; LOWER et al., 2018; NOVÁK et al., 2013; RUIZ-RUANO et al., 2016). No presente trabalho, o genoma de um espécime de *P. lineatus* foi sequenciado e o seu catálogo de DNAs satélites foi caracterizado, combinando ferramentas computacionais e citogenéticas, com foco no acúmulo de satDNAs nos cromossomos B desta espécie.

4.2 Material e Métodos

4.2.1 Amostragem dos indivíduos

As análises citogenéticas e moleculares nos indivíduos de *P. lineatus* foram realizadas em animais coletados em duas localidades: nos tanques de cultivo de uma piscicultura localizada em Sales de Oliveira, São Paulo, em que foram coletados indivíduos F1 da população do Rio Sapucaí, bacia do rio Sapucaí, destinados ao repovoamento desta mesma população. A outra coleta foi realizada em uma população natural do Rio Apa, na divisa entre o Estado do Mato Grosso do Sul e o Paraguai, bacia do Rio Paraguai, onde foram coletados indivíduos adultos. Para a população do Rio Sapucaí foram utilizados 7 indivíduos que originaram as amostras e, para os estudos, foi escolhido a amostra que apresentava as três variantes de cromossomos B. Para a população do rio Apa foi coletado sangue de 20 indivíduos, que originaram as amostras analisadas. 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 1204-CEUA/2019.

Os exemplares utilizados na amostragem foram fixados em formol 10%, conservados em álcool 70% e depositados na coleção de peixes do Laboratório de Genética de Peixes, Faculdade de ciências, UNESP- Bauru-SP, sob os números de *voucher* LG13393 até LG13407. Os indivíduos coletados na população do rio Apa foram devolvidos à natureza após a coleta de amostras de sangue da veia caudal para a obtenção de cromossomos mitóticos a partir de cultura de linfócitos. Fragmentos de nadadeiras foram retirados e fixados em álcool absoluto, e depositados no banco de tecidos do Laboratório de Genética de Peixes, sob os vouchers LGP13501 até LGP13520.

4.2.2 Extração de DNA e sequenciamento do genoma

Foi utilizado um indivíduo de *P. lineatus* para a realização do sequenciamento de nova geração (NGS), na plataforma BGISEQ-500 da BGI (BGI Shenzhen Corporation, Shenzhen, China) (*paired-end* 2x100pb). Para isso, extraímos o DNA genômico do músculo utilizando o Wizard Genomic DNA Purification Kit (Promega), seguindo as instruções fornecidas pelo fabricante, com uma etapa extra de remoção do RNA, utilizando uma RNase A (Invitrogen). As amostras obtidas do processo anterior foram corridas em gel de agarose 1%, para verificação da integridade do DNA. No total, 3.64 Gb foram sequenciados, resultando em uma cobertura de sequenciamento de aproximadamente 2.14x, considerando que o genoma de *P. lineatus* tem tamanho estimado em 1.7 Gb (CARVALHO; OLIVEIRA; FORESTI, 1998).

4.2.3 Caracterização do Satelitoma

Para realizar a caracterização do satelitoma de *P. lineatus*, utilizamos a *pipeline* TAREAN (NOVÁK et al., 2017), seguida por uma etapa de clusterização pelo RepatExplorer1 (NOVÁK et al., 2013), tendo como base o protocolo satMiner descrito por Ruiz-Ruano et al. (2016). Para a clusterização dos satélites pelo TAREAN, as opções padrão para uma identificação dos satDNAs com alta e baixa confiança foram utilizadas. Para realizar a caracterização, primeiramente nós realizamos uma filtragem de índice de qualidade phred33 nas bibliotecas de gDNA utilizando o Trimmomatic (BOLGER; LOHSE; USADEL, 2014). Após a filtragem das bibliotecas, selecionamos aleatoriamente 2 x 500.000 *reads* para a clusterização utilizando o TAREAN (NOVÁK et al., 2017). Os satDNAs resultantes deste processo foram removidos utilizando o programa DeconSeq (SCHMIEDER; EDWARDS, 2011) e esse processo foi repetido até que não fossem encontrados mais satélites.

Posteriormente, realizamos o mesmo processo de clusterização utilizando o RepeatExplorer1, em busca de clusterizações com o formato circular, típico de satDNAs, passando para um processo de análise dos contigs destes clusters com a utilização do *software* Geneious 8.4, para a identificação de sequências repetidas em *tandem*, utilizando a opção de *dotplot*, oferecida pelo Geneious. Os monômeros aqui encontrados, passaram também pelo processo de extração da biblioteca original, já filtrada dos satDNAs encontrados no TAREAN, com a utilização do DeconSeq,

realizando então o *workflow* até que nenhum satélite fosse também encontrado pelo RepeatExplorer1.

Após a fase de caracterização dos satDNAs, realizamos um processo para a filtragem e remoção de outros elementos repetidos *in tandem* (i.e. famílias multigênicas). Posteriormente, fizemos uma busca de similaridade dentro do próprio catálogo de satDNAs de *P. lineatus*, tendo como base a classificação de variantes estabelecida em Ruiz-Ruano et al. (2016).

Com o catálogo de satDNAs do *P. lineatus* caracterizado, passamos para a etapa de estudo de abundância e divergência de cada um dos satélites encontrados, utilizando o *software* RepeatMasker (SMIT et al., 2017), utilizando a ferramenta de busca *cross_match*, usando todos os *reads* da biblioteca (36,469,178 *paired-end reads*). Para estimar a abundância de cada satélite do nosso catálogo, calculamos a proporção de *reads* mapeados de cada satDNA pelo total de nucleotídeos utilizados e classificamos as famílias de satDNA em ordem decrescente de abundância no genoma. Também estimamos a divergência média entre as famílias de satDNAs considerando a distância entre as sequências com base no modelo de parâmetros Kimura-2 usando o *script* calcDivergenceFromAlign.pl disponível no pacote RepeatMasker (SMIT et al., 2017).

Posteriormente à caracterização do satelitoma de *P. lineatus*, realizamos uma busca de similaridade entre os satelitomas de *P. lineatus*, *Psalidodon paranae*, *Characidium gomesi* e *Megaleporinus macrocephalus* (SERRANO-FREITAS et al., 2020; SILVA et al., 2017; UTSUNOMIA et al., 2019) utilizando o *script* *rm_homology* (<https://github.com/fjruizruano/ngs-protocols>) e o algoritmo "Muscle Alignment" de Geneious 8.4 (KEARSE et al., 2012).

Finalmente, uma pesquisa utilizando BLAST para todas as sequências de satélite foi conduzida contra a coleção de nucleotídeos do NCBI (<http://www.ncbi.nlm.nih.gov>). Filtramos os resultados excluindo alinhamentos menores que o tamanho dos monômeros do satélite e, posteriormente, realizamos uma análise entre os satDNAs e os resultados do BLAST, usando diferentes algoritmos de alinhamento Geneious 8.4 e a ferramenta de anotação deste, para verificar se haviam semelhanças com algum significado biológico.

5. Resultados

5.1 *Prochilodus lineatus* - Cariótipos e a primeira descrição de uma população sem cromossomos B

Todos os indivíduos de *P. lineatus* analisados apresentaram um conjunto de cromossomos do complemento A representado por $2n=54$ cromossomos, composto de metacêntricos e submetacêntricos. Adicionalmente, microcromossomos B mitoticamente estáveis foram identificados nos indivíduos da população do rio Sapucaí. A presença dos cromossomos B variou entre 1 e 3 cromossomos, com as morfologias metacêntrico (Bm), submetacêntrico (Bsm), e acrocêntrico (Ba), representados na **Figura 2**. Os cromossomos B não foram identificados nas amostras provenientes da população do rio Apa. Todas as análises realizadas nas amostras de ambas as localidades estão sumarizadas, respectivamente, nas **tabelas S1** e **S2** encontradas no material suplementar.

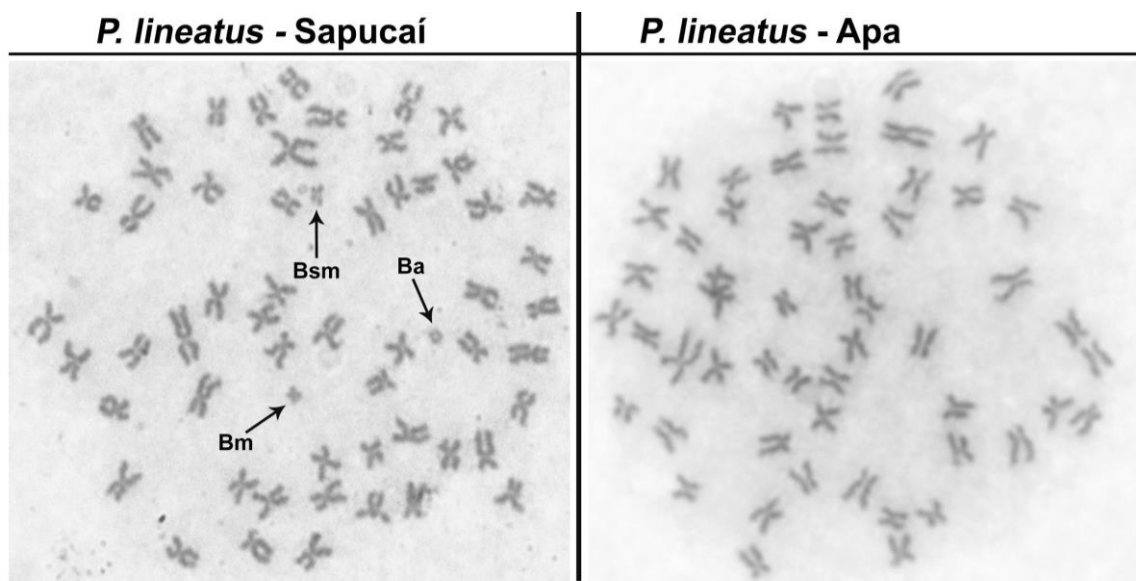


Figura 2. Metáfases de *P. lineatus* em ambas as populações estudadas. As setas indicam os cromossomos B com as diferentes morfologias observadas nessa população.

5.2 O satelitoma de *P. lineatus*

Após três iterações de TAREAN e uma iteração usando o protocolo satMiner, nós encontramos 51 famílias de satDNAs para *P. lineatus*, com o comprimento de

unidades repetidas (RUL - “repeat unit lengths”) variando de 6 a 3008 pb, com uma mediana de 59 pb (**Tabela 1**). O conteúdo A + T das famílias de satDNAs variou de 40% a 76,2%, com 55,24% de média e 39 dos satélites possuindo um conteúdo A + T maior do que 50%.

Tabela 1: Principais características presentes nos 51 satélites identificados em *P. lineatus*. SF: Superfamília; RUL: *Repeat unit length* (comprimento da unidade de repetição)

SF	satDNA	RUL	A+T (%)	Abundância (%)	Divergência (%)
1	PliSat01-225	225	64,5	0,0133150436	7,25
	PliSat02-184	184	54,9	0,009881882723	10,89
	PliSat03-900	900	56,5	0,004828665044	1,84
	PliSat04-441	441	57,9	0,003190691328	6,43
	PliSat05-178	178	64,6	0,003096656579	8,99
	PliSat06-44	44	47,8	0,002020573921	13,85
	PliSat07-3008	3008	63,4	0,001919532708	3,31
	PliSat08-54	54	62,9	0,001846954571	4,04
1	PliSat09-495	495	58	0,001809277275	5,78
	PliSat10-25	25	40	0,001214838048	11,61
	PliSat11-709	709	50,4	0,001119510563	9,9
	PliSat12-42	42	52,4	0,0009840660516	13,58
	PliSat13-1928	1928	62,4	0,000842032003	7,86
	PliSat14-44	44	50	0,0005822412559	7,52
	PliSat15-75	75	61,3	0,0004322211485	3,33
	PliSat16-16	67	58,2	0,0003912575161	3,79
	PliSat17-21	21	76,2	0,0003087390673	15,05
	PliSat18-40	40	57,5	0,0002978053138	10,47
	PliSat19-30	30	56,7	0,0002641511415	12,96
	PliSat20-54	54	57,4	0,0002557563815	9,63
	PliSat21-59	59	49,2	0,0002382236309	4,69
	PliSat22-37	37	45,9	0,000212001625	4,59

	PliSat23-39	39	53,8	0,0001860081683	10,87
	PliSat24-31	31	51,6	0,0001846132918	17,13
	PliSat25-34	34	47,1	0,0001610760188	5,42
	PliSat26-52	52	59,6	0,0001568538232	8,83
	PliSat27-1683	1683	52	0,0001367892087	2,25
	PliSat28-32	32	56,8	0,0001360984884	14,45
	PliSat29-60	60	45	0,0001251175719	3,31
	PliSat30-31	31	54,8	0,000122501253	16,69
	PliSat31707	707	53,6	0,0001187574066	4,27
	PliSat32-30	30	66,3	0,0001184738795	12,19
2	PliSat33-6	6	57,1	0,0001040484104	25,51
	PliSat34-39	39	53,8	0,0001012587945	7,97
	PliSat35-1128	1128	61	9,47E-05	10,43
	PliSat36-68	68	63,2	9,04E-05	5,55
	PliSat37-554	554	50	8,82E-05	5,47
	PliSat38-32	32	53,1	8,49E-05	6,29
	PliSat39-915	915	54,3	7,85E-05	1,35
	PliSat40-27	27	48,8	7,71E-05	14
2	PliSat41-162	162	45,8	7,49E-05	1,75
	PliSat42-24	24	55,6	6,70E-05	7,75
	PliSat43-32	32	53,1	6,28E-05	9,61
	PliSat44-1134	1134	51,6	6,22E-05	3,43
	PliSat45-340	340	50,6	6,19E-05	4,83
	PliSat46-29	29	65,5	6,01E-05	13,97
	PliSat47-574	574	53,8	5,28E-05	6,89
	PliSat48-387	387	48,8	4,80E-05	5
	PliSat49-67	67	49,3	4,52E-05	4,3
	PliSat50-48	48	58,3	4,28E-05	9,07
	PliSat51-911	911	55	3,90E-05	5
TOTAL				0,05183406656	

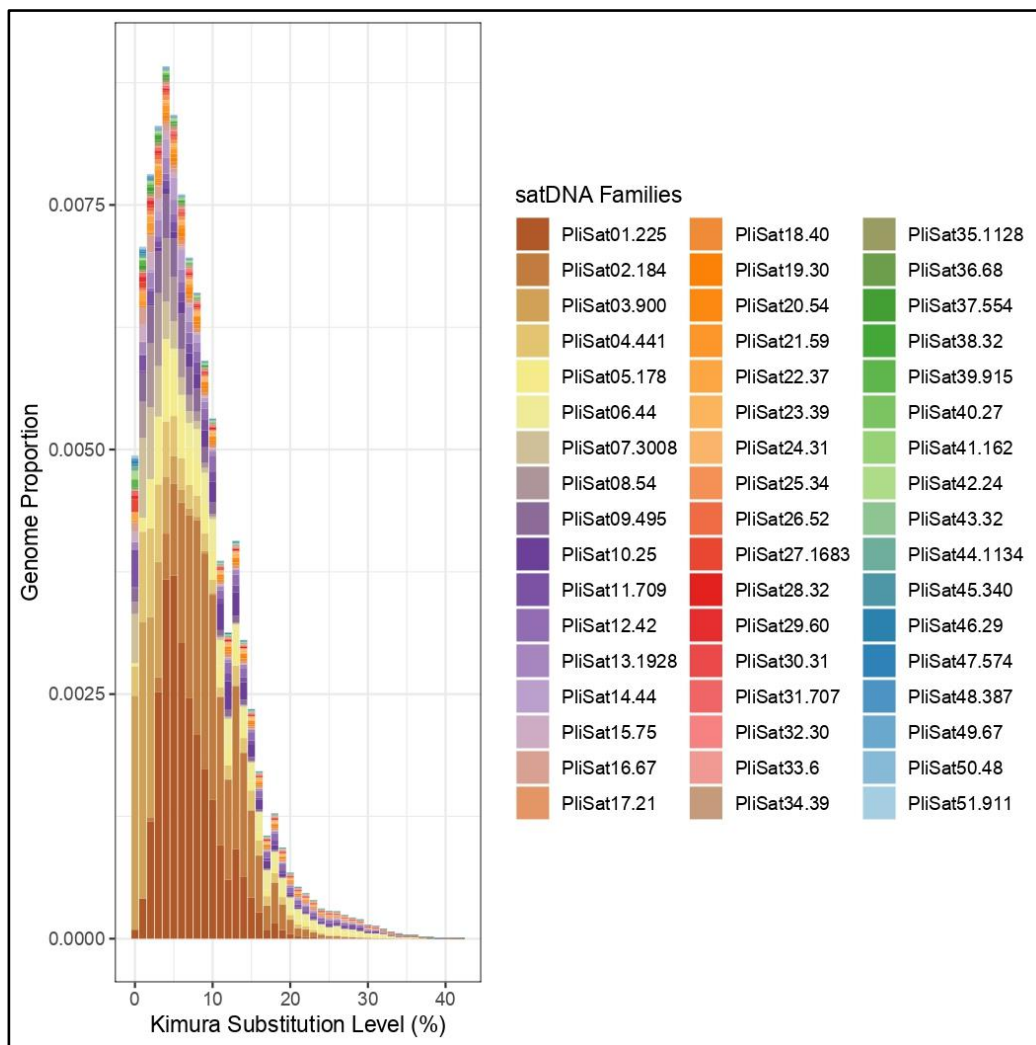


Figura 3. Gráfico evidenciando os perfis de abundância e divergência dos satélites extraídos do genoma de *P. lineatus*.

A busca por homologia das famílias de satDNA de *P. lineatus* detectou a existência de duas superfamílias (SF1 - PliSat01 e PliSat09; SF2 - PliSat33 e PliSat42). O resultado da busca por homologia realizada contra os satelitomas de *M. macrocephalus*, *P. paranae* e *C. gomesi* é mostrado na **Tabela 2**.

Tabela 2: Satélites de *P. lineatus* similares aos satélites de *M. macrocephalus*, *P. paranae* e *C. gomesi*

Famílias de satDNA de <i>P. lineatus</i>	satDNAs Homólogos	Similaridade(%)
PliSat12	MmaSat07	88,10%
PliSat15	MmaSat18	78,67%
PliSat18	MmaSat16	87,50%
PliSat21	MmaSat149	62,71%
PliSat26	MmaSat85	92,31%
	ApaSat29	92,31%
	CgomSat02	88,62%
PliSat30	MmaSat51	96,88%
PliSat36	MmaSat79	88,24%
PliSat40	MmaSat27	77,78%
PliSat43	MmaSat26	65,63%

PliSat12 MmaSat07	AGTGGCCAGTGAGTGGAGACAAGGTGGGTGTTTCTAATAA AGTAGCCAGTGAGTGGAGACAAGGTGGGTGTTTCTAATAA
PliSat15 MmaSat18	GAGTTATACACCTCACCTAATGAGAACTGTAGAACAGACTCAGTCTATACACAGTACCTACATTAGAGCTTA GAGTTATACACCTCACCTAATGAGAACTGTAA-----CTCAGTTATATAACA-TACCTACATTAGAGCTTA
PliSat18 MmaSat16	AAAGGAATGAAGGGATGAAGGAACAGTGTTGGTGATGGAG AAAGGAATGAAGGAATGAAGGAACAGTGTTGGTGAT-GAG
PliSat21 MmaSat16	AGGTAGGGGGCGATACGACCCACACGCTCATCTGACAGACTGTAATACACTACAGGA AGCTCGGATGCG-----ACGGACGATCTACAGACTGGAATACACTACAGGA
PliSat26 MmaSat16 ApaSat29 CgomSat02	TGATATCTCCTTTCTAGGTATAGGAGTGTTTATTGAGGGTGTATTATGGAC TAATATCTCCTATTCTAGTGATAGGAGTGTTTATTGAGGGTGTATTATGGAC TAATATCTCCTATTCTAGTGATAGGAGTGTTTATTGAGGGTGTATTATGGAC TAATATCTCCTATTCTAGTGATAGGAGTGTTTATTGAGGGTGTATTATGGAC
PliSat30 MmaSat51	GATTCAGACAGAAATTTGTGGGCTACAGTCG G-GTTCAGACAGAAATTTGTGGGCTACAGTCG
PliSat36 MmaSat79	CTGAGTAGACTGATAAATCAGGTGATGACAGTTATTAATCTAGTGTTACTGAGCTCTGCTAAAG CTGAGTAGACTGATAAATCAAGTGATGACAGTTATTAATCTAGTGTTACTGAGCTCTGCTAAAG
PliSat40 MmaSat27	CTAAACCCCTAACCCCTACATTACA CTAAACCC--TAACCCCTACATTGAC
PliSat43 MmaSat26	CTCTACAGCAGAGTGATCACTGTAATGGTC -----AGAGTGATCACTGTAATGGTC

Figura 4. Alinhamento dos satélites de *P. lineatus* com os satélites similares. As cores indicam que há alteração na respectiva base em relação ao satDNA similar.

Utilizando BLAST, obtivemos 72 *hits* com o banco de dados, muitos sendo variantes de uma mesma sequência entre diferentes amostras, além de algumas sequências não caracterizadas. Além disso, mostramos que o PliSat03 e o PliSat04 correspondem a, respectivamente, SATH1 e SATH2, dois satélites de *P. lineatus*, identificados anteriormente pelo método clássico de isolamento de satDNAs por enzimas de restrição (DE JESUS et al., 2003).

5.3 Distribuição cromossômica das 37 famílias de DNA satélite em *P. lineatus*.

Nós desenhamos *primers* para amplificação em PCR de 40 famílias de satDNAs em *P. lineatus*. Destes, 37 tiveram sucesso na amplificação do PCR, representando uma grande porcentagem do satelitoma total da espécie. Dentre os *primers* que resultaram em amplificação positiva foram construídas sondas para serem utilizadas na FISH e, de todas as 37, apenas oito mostraram sinais conspícuos nos cromossomos após FISH. Todas as 28 famílias de satDNAs restantes foram consideradas não-clusterizadas em nível citológico, como apresentado na **Tabela 3**.

Tabela 3: Padrão de clusterização das famílias de DNAs Satélites mapeadas em *P. lineatus* (C = clusterizado; NC = não clusterizado; 0 = ausência; 1 = presença)

Satélite	Presença de Cluster	Presença no B
PliSat01-225	C	1
PliSat02-184	C	1
PliSat03-900	C	1
PliSat04-441	C	0
PliSat05-178	C	1
PliSat08-54	C	0
PliSat12-42	NC	0
PliSat13-1928	NC	0
PliSat14-44	C	1
PliSat15-75	C	1
PliSat16-67	NC	0
PliSat18-40	NC	0
PliSat19-30	NC	0
PliSat20-54	NC	0

PliSat21-59	NC	0
PliSat22-37	NC	0
PliSat23-39	NC	0
PliSat25-34	NC	0
PliSat26-52	NC	0
PliSat27-1683	NC	0
PliSat28-32	NC	0
PliSat29-60	NC	0
PliSat31-707	NC	0
PliSat32-30	NC	0
PliSat34-39	NC	0
PliSat35-1128	NC	0
PliSat36-68	NC	0
PliSat37-554	NC	0
PliSat39-915	NC	0
PliSat41-162	NC	0
PliSat44-1134	NC	0
PliSat45-340	NC	0
PliSat47-574	NC	0
PliSat48-387	NC	0
PliSat49-67	NC	0
PliSat50-48	NC	0
PliSat51-911	NC	0

satDNAs clusterizados foram observados nas regiões centroméricas e teloméricas, além daqueles que foram observados nos cromossomos B. Em geral, os seguintes padrões de clusterização foram observados: PliSat01 apresentando grande acúmulo na região centromérica de todos os cromossomos, inclusive os supranumerários; PliSat02 em todos os centrômeros e com clusters pequenos nos cromossomos B; PliSat03 clusterizado em 08 centrômeros, incluindo nos cromossomos B; PliSat04 em 7 centrômeros dos cromossomos A; PliSat05 com clusters nos três cromossomos B e no par 4 de cromossomos A; PliSat08 clusterizado em todos os telômeros, exceto dos cromossomos B; PliSat14 exclusivos dos telômeros dos cromossomos B; e PliSat15, disperso em cromossomos A e B. A figura com a técnica

de reFISH para os PliSat03, PliSat04 e PliSat05 na mesma metáfase encontra-se no material suplementar (**Figura S1**) juntamente com um idiograma que sumariza todas as marcações encontradas na FISH (**Figura S2**) e um cariótipo contendo essas três marcações, na mesma metáfase, está presente na **Figura 6**. Destacamos a ocorrência conjunta de PliSat03 e PliSat04 nos pares cromossômicos 1, 5, 18 e 21. (**Figura S1** e **Figura 6**).

Os indivíduos do rio Apa apresentaram um mesmo padrão de clusterização nas famílias de satDNAs estudadas, com exceção daquela que estão clusterizadas exclusivamente nos cromossomos B (PliSat14). Os indivíduos desta população não apresentaram nenhum tipo de cromossomo B, com isso, não apresentam um resultado positivo de clusterização dos satDNAs PliSat14 e PliSat15, exclusivos dos cromossomos B dos indivíduos provenientes do Sapucaí, enquanto o PliSat05 apresentou clusterização somente em um par de cromossomos nos indivíduos do rio Apa (**Figura 5**, **Figura 6**, **Figura S1**).

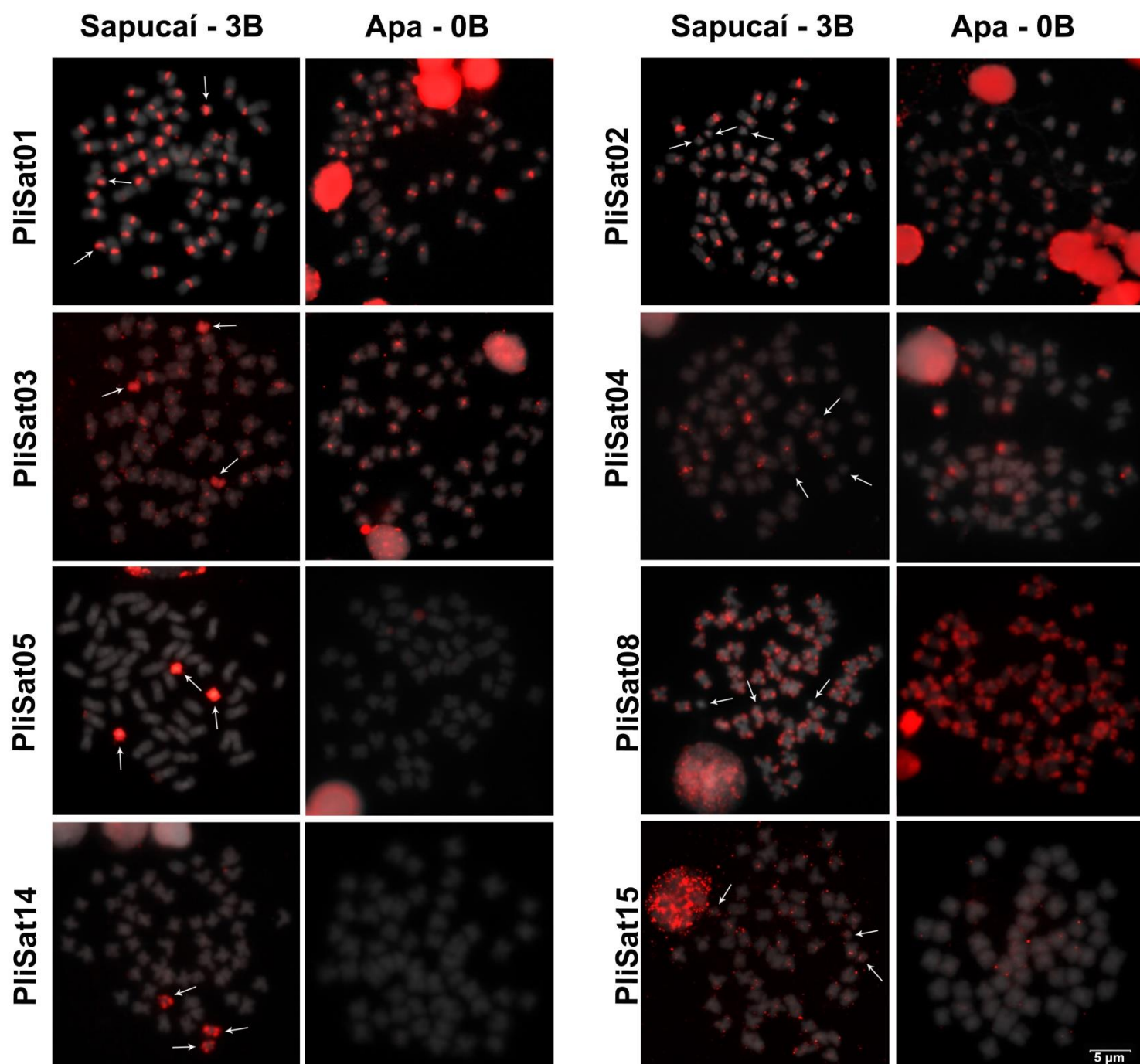


Figura 5. Metáfases de *P. lineatus* com as hibridações realizadas nas quais os satDNA estão visíveis na FISH em indivíduos proveniente das duas localidades. As setas indicam os cromossomos B.

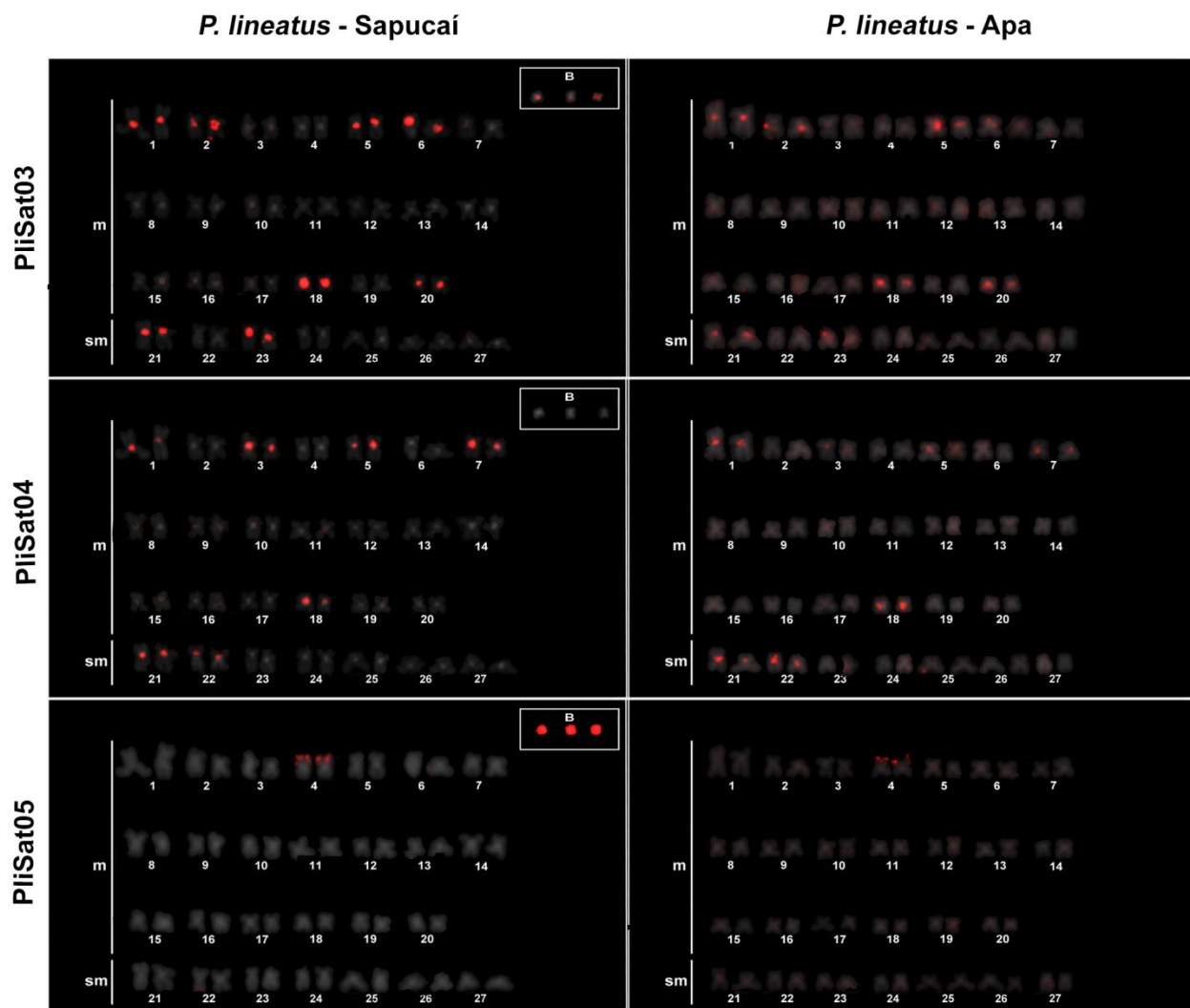


Figura 6. Cariótipos de *P. lineatus* obtidos através da técnica de reFISH com os PliSat03, PliSat04 e PliSat05 com destaque para a ocorrência conjunta de PliSat03 e PliSat04 nos pares cromossômicos 1, 5, 18 e 21.

6 Discussão

6.1 Macroestrutura cariotípica em *Prochilodus lineatus*

Os resultados obtidos aqui revelaram que o cariótipo dos indivíduos de ambas as localidades amostradas é conservado, com $2n=54$ cromossomos, corroborando dados relatados para esta espécie (ARTONI et al., 2006a; DE JESUS et al., 2003; MELO et al., 2017; PAULS; BERTOLLO, 1983, 1990; VOLTOLIN et al., 2010a). Adicionalmente, os indivíduos provenientes do rio Sapucaí apresentaram microcromossomos B mitoticamente estáveis, com a ocorrência de três variantes (acrocêntrico, submetacêntrico e metacêntrico), confirmando dados obtidos nesta espécie para uma bacia hidrográfica adjacente (PENITENTE et al., 2013). A ocorrência de três cromossomos B na população do Rio Sapucaí está de acordo com a média da população constatada em revisão bibliográfica realizada por Penitente (2018) e pelo trabalho realizado por Voltolin et al. (2010b), em que uma frequência média de 2.614 cromossomos B foi constatada em um total de 275 indivíduos. Por outro lado, a análise citogenética de espécimes coletados no rio Apa ($n = 20$ indivíduos) revelou que os indivíduos desta população não apresentam cromossomos B, caracterizando-se como a primeira população de *P. lineatus* analisada até o momento em que cromossomos supranumerários não estão presentes.

A ocorrência exclusiva de cromossomos B apenas em determinadas populações de uma mesma espécie é comum e já foi relatada múltiplas vezes em peixes (MOREIRA-FILHO; GALETTI; BERTOLLO, 2004; PANSONATO-ALVES et al., 2011; SCACCHETTI et al., 2015; UTSUNOMIA et al., 2016). Em geral, esta diferença populacional está diretamente ligada ao isolamento geográfico e à baixa vagilidade destas populações. Por estas razões, o presente caso é notável, uma vez que *P. lineatus* é considerado um migrador de grandes distâncias e vários estudos utilizando diferentes marcadores moleculares evidenciaram a ausência de estruturação genética na bacia do rio Paraná (CARVALHO-COSTA; HATANAKA; GALETTI, 2008; FERREIRA et al., 2017; GARCEZ; CALCAGNOTTO; DE ALMEIDA-TOLEDO, 2011; MELO et al., 2013; ORTÍ et al., 2008; REVALDAVES; RENESTO; MACHADO, 1997), bacia em que os indivíduos analisados aqui foram coletados. Neste contexto, nossos resultados indicam um inédito grau de diferenciação populacional nesta bacia.

Uma hipótese que poderia explicar a ausência de diversidade genética desta espécie em uma grande bacia, como a do rio Paraná, estaria relacionada ao tipo de marcador (e.g. sequências mitocondriais) utilizado na maioria dos estudos populacionais desta espécie. De fato, Melo e colaboradores (2018) revelaram que linhagens mitocondriais no gênero *Prochilodus* são pouco divergentes e sugerem que episódios recentes de diversificação seriam responsáveis por esta situação. Por esta razão, nossos resultados abrem uma janela de oportunidade para avaliar a diversidade genética desta espécie utilizando outros tipos de marcadores.

6.2 O satelitoma do curimatá e a diferenciação dos cromossomos B

No presente trabalho, nós caracterizamos 51 famílias de DNAs satélites de *P. lineatus*, sendo a quarta espécie de peixe a ter o seu próprio satelitoma caracterizado. Destaca-se que as características gerais do satelitoma, como tamanho médio de monômeros e conteúdo A+T, foram bastante parecidas com aquelas já descritos (SERRANO-FREITAS et al., 2020; SILVA et al., 2017; UTSUNOMIA et al., 2019). Além disso, a robustez de nossas análises é reforçada pela confirmação da presença dos satDNAs anteriormente caracterizados para *P. lineatus* e nomeados como SATH1 e SATH2 (DE JESUS et al., 2003), aqui referidos como PliSat03 e PliSat04, respectivamente; bem como pela presença de um satDNA conservado em Characiformes, denominado CharSat01 (DOS SANTOS et al., 2021), sendo aqui referido como PliSat26.

Em relação à conservação e persistência de satDNAs entre espécies distantes, é notória a existência de nove satélites compartilhados com *Megaleporinus macrocephalus* (Tabela 2), cuja diferenciação ocorreu há mais de 50 milhões de anos (KOLMANN et al., 2020). Estes dados estão de acordo com a hipótese *library* de evolução de satDNAs, uma vez que as outras duas espécies com satelitomas caracterizados (*Psalidodon paranae* e *Characidium gomesi*) são ainda mais distantes de *P. lineatus*, por volta de 80 – 140 milhões de anos (ARCILA et al., 2017). É possível perceber as diferenças quantitativas previstas pela hipótese *library*, uma vez que dentre os satDNA similares entre *P. lineatus* e *M. macrocephalus*, encontramos satélites que possuem abundância similar em seus respectivos genomas (como PliSat18 e MmaSat16) enquanto outros a diferença é maior (PliSat21 e MmaSat149; UTSUNOMIA et al., 2019).

Considerando um total de 37 sondas de satDNAs produzidas a partir do genoma de *P. lineatus*, apenas oito foram visualizadas por FISH, indicando que vários satDNAs não estão clusterizados formando longos *arrays* de sequências perceptíveis através da FISH (DOS SANTOS et al., 2021). É interessante notar que duas famílias (PliSat01 e PliSat02) estão colocalizadas nos centrômeros de todos os cromossomos do complemento A, mas não apresentam qualquer similaridade entre si, possivelmente indicando que dois satélites com origem independente se tornaram os satDNAs centroméricos desta espécie. Destaca-se também que apenas o PliSat01 é encontrado em grande abundância nos cromossomos B e nos centrômeros dos cromossomos A, o que pode indicar que este satDNA desempenha alguma função centromérica nesta espécie.

Em trabalho pioneiro, foi verificado que os cromossomos B de *P. lineatus* e de outras espécies do gênero apresentam clusters de PliSat03 (SATH1), o que poderia indicar uma origem intraespecífica a partir da região pericentromérica de cromossomos portadores de clusters deste satDNA (ARTONI et al., 2006a; DE JESUS et al., 2003; MELO et al., 2017). Nós expandimos o conhecimento sobre o conteúdo molecular destes cromossomos B para seis satDNAs, indicando um grande acúmulo de sequências satélites nesses elementos. Além disso, o par N° 4 do complemento A é o único que apresenta a maioria dos sinais de FISH de satDNAs presentes nos cromossomos B, exceto o PliSat14, que é exclusivo do cromossomo B. Assim, nós sugerimos que os cromossomos B de *Prochilodus* possam ter se originado como um fragmento do cromossomo N° 4.

No presente trabalho, nós caracterizamos pela primeira vez uma população de *P. lineatus* que não apresenta cromossomos B, dando indícios de que possa haver um certo grau de diferenciação populacional desta espécie na bacia do rio Paraná, bem como caracterizamos o satelitoma desta espécie. Além disso os cromossomos B estão enriquecidos com elementos repetitivos e a localização cromossômica dos satDNAs dão indícios da origem intraespecífica destes cromossomos supranumerários possivelmente do par 4 do cromossomo A.

7. Referências

- ARCILA, D. et al. Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. **Nature Ecology and Evolution**, v. 1, n. 2, p. 1–10, 2017.
- ARTONI, R. F. et al. Banding pattern of A and B chromosomes of *Prochilodus lineatus* (Characiformes, Prochilodontidae), with comments on B chromosomes evolution. **Genetica**, v. 127, n. 1–3, p. 277–284, 2006a.
- ARTONI, R. F. 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, 2006b.
- BISCOTTI, M. A.; OLMO, E.; HESLOP-HARRISON, P. Repetitive DNA in eukaryotic genomes. **Chromosome Research**, v. 23, n. 3, p. 415–420, 2015.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: A flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.
- CAMACHO, J. P. M.; SHARBEL, T. F.; BEUKEBOOM, L. W. B-chromosome evolution. **Philosophical Transactions of the Royal Society B**, v. 355, n. 1394, p. 163–178, 2000.
- CARVALHO, M. L.; OLIVEIRA, C.; FORESTI, F. Nuclear DNA content of thirty species of neotropical fishes. **Genetics and Molecular Biology**, v. 21, n. 1, p. 47–54, mar. 1998.
- CARVALHO-COSTA, L. F.; HATANAKA, T.; GALETTI, P. M. Evidence of lack of population substructuring in the Brazilian freshwater fish *Prochilodus costatus*. **Genetics and Molecular Biology**, v. 31, n. 1 SUPPL. 1, p. 377–380, 2008.
- CASTRO, R. M. C.; VARI, R. P. **Detritivores of the South American Fish Family Prochilodontidae (Teleostei: Ostariophysi: Characiformes): a Phylogenetic and Revisionary Study**. [s.l: s.n.].
- CHARLESWORTH, B.; SNIEGOWSKI, P.; STEPHAN, W. The evolutionary dynamics of repetitive DNA in eukaryotes. **Nature**, v. 371, p. 215–220, 1994.
- CHARLESWORTH, D.; CHARLESWORTH, B.; MARAIS, G. Steps in the evolution of heteromorphic sex chromosomes. **Heredity**, v. 95, n. 2, p. 118–128, 2005.

CHAVES, R. et al. FA-SAT is an Old Satellite DNA Frozen in Several Bilateria Genomes. **Genome Biology and Evolution**, v. 9, n. 11, p. 3073–3087, 2017.

CHUONG, E. B.; ELDE, N. C.; FESCHOTTE, C. Regulatory activities of transposable elements: from conflicts to benefits. **Nature Reviews Genetics**, v. 18, n. 2, p. 71–86, 2017.

DA SILVA, M. et al. New approach data in electric fish (Teleostei: *Gymnotus*): Sex chromosome evolution and repetitive DNA. **Zebrafish**, v. 11, n. 6, p. 528–535, 2014.

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

DOS SANTOS, R. Z. et al. A long-term conserved satellite DNA that remains unexpanded in several genomes of Characiformes fish is actively transcribed. **Genome Biology and Evolution**, v. 13, n. 2, 2021.

FERREIRA, D. G. et al. Genetic structure and diversity of migratory freshwater fish in a fragmented Neotropical river system. **Reviews in Fish Biology and Fisheries**, v. 27, n. 1, p. 209–231, 2017.

FERRETTI, A. B. S. M. et al. High dynamism for neo-sex chromosomes: satellite DNAs reveal complex evolution in a grasshopper. **Heredity**, v. 125, n. 3, p. 124–137, 2020.

FORESTI, F.; TOLEDO, L. F. A.; TOLEDO FILHO, S. DE A. Polymorphic nature of nucleolus organizer regions in fishes. **Cytogenetics and Cell Genetics**, v. 31, p. 137–144, 1981.

FRICKE, R.. **ESCHMEYER'S CATALOG OF FISHES: GENERA, SPECIES, REFERENCES.** 2021. Disponível em: <https://www.calacademy.org/scientists/projects/eschmeyers-catalog-of-fishes>. Acesso em: 20 fev. 2021.

FRY, K.; SALSER, W. Nucleotide Sequences of HS- α Satellite DNA from Kangaroo rat *Dipodomys ordii* and Characterization of Similar Sequences in other Rodents. **Cell**, v. 12, n. 4, p. 1069–1084, 1977.

GARCEZ, R.; CALCAGNOTTO, D.; DE ALMEIDA-TOLEDO, L. F. Population

structure of the migratory fish *Prochilodus lineatus* (Characiformes) from rio Grande basin (Brazil), an area fragmented by dams. **Aquatic Conservation: Marine and Freshwater Ecosystems**, v. 21, n. 3, p. 268–275, 2011.

GARRIDO-RAMOS, M. A. Satellite DNA in plants: More than just Rubbish. **Cytogenetic and Genome Research**, v. 146, n. 2, p. 153–170, 2015.

GARRIDO-RAMOS, M. A. Satellite DNA: An evolving topic. **Genes**, v. 8, n. 9, 2017.

GREEN, D. M. Muller's ratchet and the evolution of supernumerary chromosomes. **Genome**, v. 33, n. 6, p. 818–824, 1990.

HALBACH, R. et al. A satellite repeat-derived piRNA controls embryonic development of *Aedes*. **Nature**, v. 580, n. 7802, p. 274–277, 2020.

HARRIS, R. S.; CECHOVA, M.; MAKOVA, K. D. Noise-cancelling repeat finder: Uncovering tandem repeats in error-prone long-read sequencing data. **Bioinformatics**, v. 35, n. 22, p. 4809–4811, 2019.

HOUBEN, A. B chromosomes – a matter of chromosome drive. **Frontiers in Plant Science**, v. 8, n. February, p. 1–6, 2017.

JURKA, J. et al. Repbase Update, a database of eukaryotic repetitive elements. **Cytogenetic and Genome Research**, v. 110, n. 1–4, p. 462–467, 2005.

KEARSE, M. et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. **Bioinformatics**, v. 28, n. 12, p. 1647–1649, 2012.

KOLMANN, M. A. et al. Phylogenomics of Piranhas and Pacus (Serrasalminae) Unveils How Dietary Convergence and Parallelism Obscure Traditional Morphological Taxonomy. **Systematic Biology**, v. 0, n. 0, p. 1–14, 2020.

LÓPEZ-FLORES, I.; GARRIDO-RAMOS, M. A. The repetitive DNA content of eukaryotic genomes. **Genome Dynamics**, v. 7, p. 1–28, 2012.

LÓPEZ-LEÓN, M. D. et al. Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. **Chromosome Research**, v. 2, n. 2, p. 87–92, 1994.

LOWER, S. S. et al. Satellite DNA evolution: old ideas, new approaches. **Current**

Opinion in Genetics and Development, v. 49, n. i, p. 70–78, 2018.

MAISTRO, E. L.; OLIVEIRA, C.; FORESTI, F. Cytogenetic analysis of A- and B-chromosomes of *Prochilodus lineatus* (Teleostei, Prochilodontidae) using different restriction enzyme banding and staining methods. **Genetica**, v. 108, n. 2, p. 119–125, 2000.

MALABARBA, L. R.; MALABARBA, M. C. **Phylogeny and classification of Neotropical fish**. Elsevier Inc., 2019.

MELO, B. F. et al. The roles of marginal lagoons in the maintenance of genetic diversity in the Brazilian migratory fishes *Prochilodus argenteus* and *P. costatus*. **Neotropical Ichthyology**, v. 11, n. 3, p. 625–636, 2013.

MELO, B. F. et al. Little divergence among mitochondrial lineages of *Prochilodus* (Teleostei, Characiformes). **Frontiers in Genetics**, v. 9, n. APR, p. 1–9, 2018.

MELO, S. et al. B chromosome dynamics in *Prochilodus costatus* (Teleostei, Characiformes) and comparisons with supernumerary chromosome system in other *Prochilodus* species. **Comparative Cytogenetics**, v. 11, n. 2, p. 393–403, 2017.

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

MIGA, K. H. et al. Telomere-to-telomere assembly of a complete human X chromosome. **Nature**, v. 585, n. 7823, p. 79–84, 2020.

MOREIRA-FILHO, O.; GALETTI, P. M.; BERTOLLO, L. A. C. B chromosomes in the fish *Astyanax scabripinnis* (Characidae, Tetragonopterinae): An overview in natural populations. **Cytogenetic and Genome Research**, v. 106, n. 2–4, p. 230–234, 2004.

NOVÁK, P. et al. RepeatExplorer: A Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. **Bioinformatics**, v. 29, n. 6, p. 792–793, 2013.

NOVÁK, P. et al. TAREAN: A computational tool for identification and characterization of satellite DNA from unassembled short reads. **Nucleic Acids Research**, v. 45, n. 12, 2017.

- NOVÁK, P.; NEUMANN, P.; MACAS, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. **BMC Bioinformatics**, v. 11, p. 1–12, 2010.
- OLIVEIRA, C. et al. Cromosome formulae of Neotropical Freshwater Fishes. **Revista Brasileira de Genética**, v. 11, n. 3, p. 577–624, 1988.
- OLIVEIRA, C. et al. Increased B chromosome frequency and absence of drive in the fish *Prochilodus lineatus*. **Heredity**, v. 79, n. 5, p. 473–476, 1997.
- ORTÍ, G. et al. Phylogeny of the Serrasalminae (Characiformes) based on mitochondrial DNA sequences. **Genetics and Molecular Biology**, v. 31, n. 1 SUPPL. 1, p. 343–351, 2008.
- PALOMEQUE, T.; LORITE, P. Satellite DNA in insects: A review. **Heredity**, v. 100, n. 6, p. 564–573, 2008.
- PANSONATO-ALVES, J. C. et al. Chromosomal diversification in populations of *Characidium cf. gomesi* (Teleostei, Crenuchidae). **Journal of Fish Biology**, v. 78, n. 1, p. 183–194, 2011.
- PAULS, E.; BERTOLLO, L. A. C. Evidence for a system of supernumerary chromosomes in *Prochilodus scrofa* Steindachner, 1881 (Pisces, prochilodontidae). **Caryologia**, v. 36, n. 4, p. 307–314, 1983.
- PAULS, E.; BERTOLLO, L. A. C. Distribution of a supernumerary chromosome system and aspects of karyotypic evolution in the genus *Prochilodus* (Pisces, Prochilodontidae). **Genetica**, v. 81, n. 2, p. 117–123, 1990.
- PENITENTE, M. et al. Transmission rate variation among three B chromosome variants in the fish *Prochilodus lineatus* (Characiformes, Prochilodontidae). **Anais da Academia Brasileira de Ciências**, v. 85, n. 4, p. 1371–1377, 2013.
- PENITENTE, M. et al. B chromosome variants in *Prochilodus lineatus* (Characiformes, Prochilodontidae) analyzed by microdissection and chromosome painting techniques. **Caryologia**, v. 69, n. 2, p. 181–186, 2016.
- PENITENTE, M. **CROMOSSOMOS B: OCORRÊNCIA NOS PEIXES NEOTROPICAIS E SUA DISTRIBUIÇÃO EM *Prochilodus lineatus***. Universidade Estadual Paulista “Julio de Mesquita Filho”, 2018.

- PINKEL, D.; STRAUME, T.; GRAY, J. W. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. **Proceedings of the National Academy of Sciences of the United States of America**, v. 83, n. 9, p. 2934–2938, 1986.
- PLOHL, M.; MEŠTROVIĆ, N.; MRAVINAC, B. Satellite DNA evolution. **Genome Dynamics**, v. 7, p. 126–152, 2012.
- PLOHL, M.; MEŠTROVIĆ, N.; MRAVINAC, B. Centromere identity from the DNA point of view. **Chromosoma**, v. 123, n. 4, p. 313–325, 2014.
- REVALDAVES, E.; RENESTO, E.; MACHADO, M. F. P. S. Genetic variability of *Prochilodus lineatus* (Characiformes, Prochilodontidae) in the upper Parana river. **Brazilian Journal of Genetics**, v. 20, n. 3, p. 381–388, set. 1997.
- RICHARD, G.-F.; KERREST, A.; DUJON, B. Comparative Genomics and Molecular Dynamics of DNA Repeats in Eukaryotes. **Microbiology and Molecular Biology Reviews**, v. 72, n. 4, p. 686–727, 2008.
- RODRIGUES, P. H. D. M. et al. Chromosomal and Genomic Dynamics of Satellite DNAs in Characidae (Characiformes, Teleostei) Species. **Zebrafish**, v. 16, n. 4, p. 408–414, 2019.
- RUIZ-RUANO, F. J. et al. High-throughput analysis of the satellitome illuminates satellite DNA evolution. **Scientific Reports**, v. 6, n. June, p. 1–14, 2016.
- SCACCHETTI, P. C. et al. Chromosomal mapping of repetitive DNAs in *Characidium* (Teleostei, Characiformes): Genomic organization and diversification of ZW sex chromosomes. **Cytogenetic and Genome Research**, v. 146, n. 2, p. 136–143, 2015.
- SCHMIEDER, R.; EDWARDS, R. Quality control and preprocessing of metagenomic datasets. **Bioinformatics**, v. 27, n. 6, p. 863–864, 2011.
- SERRANO-FREITAS, É. A. et al. Satellite DNA content of B chromosomes in the characid fish *Characidium gomesi* supports their origin from sex chromosomes. **Molecular Genetics and Genomics**, v. 295, n. 1, p. 195–207, 2020.
- SILVA, D. M. Z. DE A. et al. High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*. **Scientific Reports**, v. 7, n. 1, p. 1–12, 2017.

- SMIT, A. F. A. et al., 1996–2010. RepeatMasker Open-3.0. 2017.
- TERUEL, M. et al. B chromosome ancestry revealed by histone genes in the migratory locust. **Chromosoma**, v. 119, n. 2, p. 217–225, 2010.
- UTSUNOMIA, R. et al. Uncovering the Ancestry of B Chromosomes in *Moenkhausia sanctaefilomenae* (Teleostei, Characidae). **PLoS ONE**, v. 11, n. 3, p. 1–20, 2016.
- UTSUNOMIA, R. et al. A glimpse into the satellite DNA library in Characidae fish (Teleostei, Characiformes). **Frontiers in Genetics**, v. 8, n. AUG, p. 1–11, 2017.
- UTSUNOMIA, R. et al. Satellitome landscape analysis of *Megaleporinus macrocephalus* (Teleostei, Anostomidae) reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome. **Scientific Reports**, v. 9, n. 1, p. 1–10, 2019.
- VALENTE, G. T. et al. Origin and Evolution of B Chromosomes in the Cichlid Fish *Astatotilapia latifasciata* Based on Integrated Genomic Analyses. **Molecular Biology and Evolution**, v. 31, n. 8, p. 2061–2072, 2014.
- VICARI, M. R. et al. Satellite DNA and chromosomes in Neotropical fishes: Methods, applications and perspectives. **Journal of Fish Biology**, v. 76, n. 5, p. 1094–1116, 2010.
- VOLTOLIN, T. A. et al. Origin and molecular organization of supernumerary chromosomes of *Prochilodus lineatus* (Characiformes, Prochilodontidae) obtained by DNA probes. **Genetica**, v. 138, n. 11, p. 1133–1139, 2010a.
- VOLTOLIN, T. A. et al. B-chromosome frequency stability in *Prochilodus lineatus* (Characiformes, Prochilodontidae). **Genetica**, v. 138, n. 3, p. 281–284, 2010b.
- VOLTOLIN, T. A. et al. Karyotypic conservatism in five species of *Prochilodus* (Characiformes, Prochilodontidae) disclosed by cytogenetic markers. **Genetics and Molecular Biology**, v. 36, n. 3, p. 347–352, 2013.
- YUAN, Z. 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.

8. Material suplementar

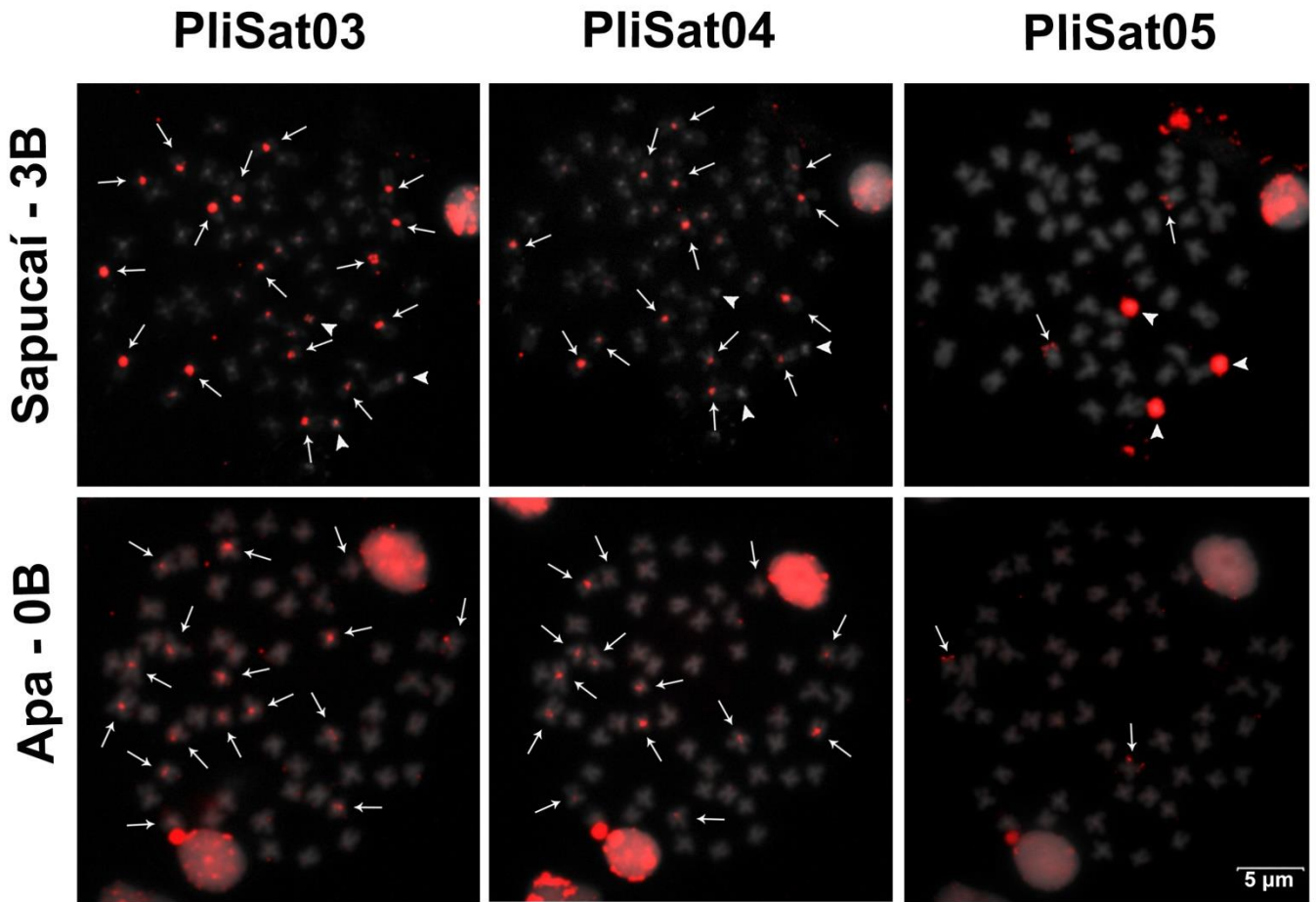


Figura S1. Metáfases de *P. lineatus* após FISH com as sondas PliSat03, PliSat04 e PliSat05 indicando, na mesma metáfase, as respectivas marcações na FISH para os satélites em ambas as amostras. As setas indicam as marcações nos cromossomos e a cabeça da seta os cromossomos B. (Obs: Estas metáfases foram utilizadas para a montagem dos cariótipos que estão apresentados na figura 6).

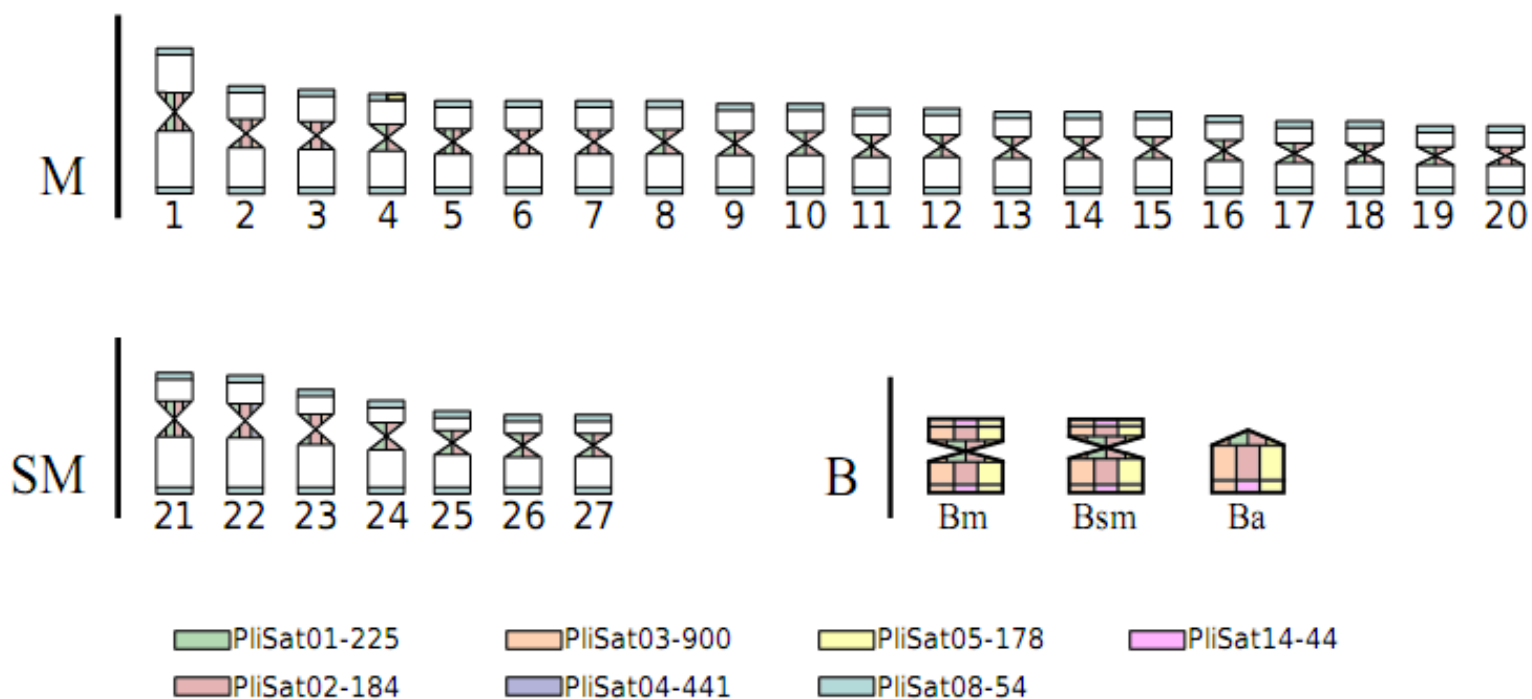


Figura S2. Idiograma representando um cari6tipo de *P. lineatus* e as localiza76es dos satDNA observadas na FISH.

Tabela S1: Análise da ocorrência de cromossomos B nas amostras do Rio Sapucaí e suas respectivas morfologias, quando possível foi analisada cinco metáfases de cada amostra. Bm: B metacêntrico, Bsm: B submetacêntrico, Ba: B acrocêntrico, B: não foi possível especificar a morfologia do cromossomo, apenas sua presença na metáfase.

Amostra	Bm	Bsm	Ba	B	Total
13394	1		1	1	3
				3	3
				2	2
				2	2
13395	1			2	3
	1	1	1		3
	3	1	1		5
	3	1	1		5
				4	4
13397	1	1	1	3	3
	1	1	1	3	3
	1	1	1	3	3
	1	1	1	3	3
	1	1	1	3	3
13402				2	2
	2				2
				1	1
				3	3
13403	2				2
				2	2
				2	2
13404		2			2
	1		1		2
	1				1
	1	1			2
	1	1			2
13405	1				1
	1				1
				1	1
	1				1
	1				1

Tabela S1: Análise da ocorrência de cromossomos B nas amostras do Rio Sapucaí e suas respectivas morfologias, quando possível foi analisado cinco metáfases de cada amostra. Bm: B metacêntrico, Bsm: B submetacêntrico, Ba: B acrocêntrico, B: não foi possível especificar a morfologia do cromossomo, apenas sua presença na metáfase. Continuação

13406	2	1			3
				3	3
	1			2	3
				3	3
				3	3
13407	1	1		1	3
	2		1		3
	2			1	3
	1		1	1	3
	1	1	1		3

Tabela S2: Análise das amostras provenientes do Rio Apa. Foram analisadas 5 metáfases de cada amostra. 1B: Presença de cromossomo B. 0B: Ausência de cromossomo B.

Amostra	1B	0B
13501		X
13502		X
13503		X
13504		X
13505		X
13506		X
13507		X
13508		X
13509		X
13510		X
13511		X
13512		X
13513		X
13514		X
13515		X
13516		X
13517		X
13518		X
13519		X
13520		X