Papel dos pseudogenes na evolução dos cromossomos supranumerários em *Psalidodon* (Teleostei, Characiformes)

Lucas Fortino Lasmar

Botucatu, SP

2022

Universidade Estadual Paulista "Júlio de Mesquita Filho"

Instituto de Biociências de Botucatu

Papel dos pseudogenes na evolução dos cromossomos supranumerários em *Psalidodon* (Teleostei, Characiformes)

Aluno: Lucas Fortino Lasmar

Orientador: Prof. Dr. Fausto Foresti

Co-orientador: Dr. Duílio M. Z. A. Silva

Dissertação apresentada ao Instituto de Biociências, Campus de Botucatu, UNESP, para qualificação de Mestrado no Programa de Pós-Graduação em Ciências Biológicas (Zoologia).

Botucatu, SP

2022

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Lasmar, Lucas Fortino.
Papel dos pseudogenes na evolução dos cromossomos supranumerários em Psalidodon (Teleostei, Characiformes) / Lucas Fortino Lasmar. - Botucatu, 2022

Dissertação (mestrado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu
Orientador: Fausto Foresti
Coorientador: Duílio Mazzoni Zerbinato de Andrade Silva
Capes: 20400004

1. Astyanax (Peixe). 2. Cromossomos. 3. Sequenciamento de nucleotídeo. 4 DNA. 5. Characiformes. 6. Pseudogenes.

Palavras-chave: Astyanax; Cromossomo B; Psalidodon; Retrogene; Retrotransposão.

Agradecimentos

Agradeço primeiramente ao meu orientador Professor Fausto Foresti, por aceitar me orientar e servir de fonte de inspiração.

Ao meu Co-Orientador e amigo Duílio Mazzoni Zerbinato de Andrade Silva, por me incentivar, direcionar sempre ao caminho e me apoiar nos momentos difíceis.

Ao Professor Cláudio Oliveira, por todo o apoio fornecido para o desenvolvimento deste trabalho.

Aos meus pais, Ana Lúcia Fortino Lasmar e Luiz Antônio Menezes Lasmar, minhas irmãs Sarah Fortino Lasmar Cunha e Samira Fortino Lasmar e cunhados Helder Teixeira Cunha e Gustavo Henrique Silva Ferreira por todo o apoio e incentivo, todo o carinho e suporte todos esses anos, sem vocês eu não teria conseguido.

Ao meu amigo e companheiro de laboratório Mateus Rossetto Vidal, pela colaboração neste trabalho, por sempre estar disposto a me ajudar e por toda amizade.

Aos meus colegas de republica, por toda amizade e parceria.

A todos os colegas do Laboratório de Biologia e Genética de Peixes, por todo o acolhimento e pelos momentos de descontração.

Ao Professor Rafael Henrique Nóbrega e a colega Maira da Silva Rodrigues membros do Laboratório de Biologia Molecular e Reprodutiva da UNESP de Botucatu, pelo suporte.

À Professora Maeli Dal Pai e a colega Bruna Tereza T. Zanella membros do Laboratório de Biologia do Músculo Estriado Esquelético da UNESP Botucatu, pelo suporte.

Ao Professor Erick José Ramo da Silva e ao colega Alexandre Dorth de Andrade membros do Laboratório de Farmacologia da Reprodução e Andrologia da UNESP de Botucatu, pelo suporte

Ao Professor Fábio Porto Foresti e a todos os colegas do Laboratório de Genética de Peixes da UNESP Bauru, pelo apoio no desenvolvimento deste projeto.

Ao Instituto de Biociências de Botucatu.

Ao programa de Pós-graduação em Zoologia.

Aos funcionários do Departamento de Biologia Estrutural e Funcional e da seção técnica de Pós-Graduação que, pelo seu trabalho, permitiram que este estudo fosse realizado.

A CAPES, FAPESP, e ao CNPq pelo financiamento deste projeto.

Esta pesquisa foi financiada com recursos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)
Processo: 2020/01775-0





"A sabedoria é um paradoxo.

O homem que mais sabe é aquele que mais reconhece a vastidão da sua ignorância."

Friedrich Nietzsche

Resumo

Cromossomos B são elementos genômicos adicionais ao complemento de cromossomos padrão presentes em cerca de 15% dos organismos eucariotos e são dispensáveis para sua sobrevivência. Estes elementos são compostos principalmente por sequências de DNA repetitivo, mas podem também possuir genes codificadores de proteínas e pseudogenes. Os pseudogenes podem desempenhar diversos papéis regulatórios no organismo, o que poderia levar a um interessante efeito funcional quando associados aos cromossomos B. Desta forma, o presente estudo visou caracterizar os pseudogenes *sbno2-B* e *simc1-B* encontrados previamente no cromossomo B de *Psalidodon paranae*, visando identificar os processos evolutivos envolvidos na sua formação, além de verificar se estes pseudogenes estão presentes nos cromossomos B de outras espécies de peixes do gênero *Psalidodon* e comparar suas estruturas genômicas, bem como sua expressão. Para isso foram coletadas amostras de indivíduos das espécies *P. paranae*, *P. bockmanni* e *P. fasciatus* (duas populações), às quais foram aplicadas e utilizadas técnicas de sequenciamento de *reads* curtos de DNA e RNA e *reads* longos de DNA. As análises bioinformáticas foram validadas por PCR e qPCR. Adicionalmente, foram padronizados genes de referência para futuras análises de expressão gênica por RT-qPCR. Com isso, foi possível verificar que os genes *sbno2* e *simc1* apresentam junções éxon-éxon em bibliotecas genômicas 1B das quatro espécies analisadas, o que pode indicar que estas sequências sofreram processo de retrotransposição para um cromossomo B ancestral em *Psalidodon*. A análise comparativa das variantes de cromossomos B de *Psalidodon* mostrou que ambos os pseudogenes possuem estrutura altamente conservada, com algumas exceções em que sofreram degeneração. A análise de RNAseq mostrou que ambos os genes *sbno2* e *simc1* apresentam subexpressão em gônadas de indivíduos de *P. paranae* com cromossomos B, o que poderia indicar que alterações na expressão da cópia canônica seriam decorrentes de alterações determinadas pela presença dos pseudogenes. Ainda, a presença de SNPs-B específicos em regiões de baixa cobertura indica que cópias com diferentes estruturas existam no cromossomo B, o que sugere que provavelmente os genes *sbno2-B* e *simc1-B* se inseriram no cromossomo B durante sua origem, e posteriormente sofreram processo de retrotransposição para o próprio cromossomo B.

Abstract

B chromosomes are genomic additional elements to the complement of standard chromosomes present in about 15% of eukaryotic organisms and are dispensable for their survival. These elements are mainly composed of repetitive DNA, but may also have protein-coding genes and pseudogenes. Pseudogenes can play several regulatory roles in the organism, which could lead to an interesting functional effect when associated with B chromosomes. Thus, the present study aimed to characterize the *sbno2-B* and *simc1-B* pseudogenes previously found in the B chromosome of *Psalidodon paranae*, in order to identify the evolutionary processes involved in their formation, in addition to verifying whether these pseudogenes are present in the B chromosomes of other fish species of the genus *Psalidodon* and comparing their genomic structures, as well as their expression. For this, individuals of the species *P. paranae*, *P. bockmanni* and *P. fasciatus* (two populations) were collected. Sequencing techniques of short DNA and RNA reads and long DNA reads were used. Bioinformatic analyzes were validated by PCR and qPCR. Additionally, reference genes were standardized for future gene expression analysis by RT-qPCR. Thus, it was possible to verify that the *sbno2* and *simc1* genes present exon-exon junctions in 1B genomic libraries in the four analyzed species, which indicates that these sequences underwent a retrotransposition process to an ancestral B chromosome in *Psalidodon*. Comparative analysis of the B chromosome variants of *Psalidodon* showed that both pseudogenes have a highly conserved structure, with some exceptions where they suffered degeneration. RNAseq analysis showed that both genes *sbno2* e *simc1* are underexpressed in the gonads of individuals with B chromosomes, which could indicate that alterations in the expression of the canonical copy would be due to alterations determined by the presence of pseudogenes. The presence of B-specific SNPs in low coverage regions indicates that copies with different structures exists in the B chromosome, which suggests that probably *sbno2-B* and *simc1-B* were acquired by the B chromosome in its origin, and after these copies went to retrotransposition processes to the B chromosome.

Sumário

Introdução 10
Pseudogenes e sua presença em cromossomos B 10
Objetivos 12
Referências 13
Capítulo 1 – Análise estrutural dos pseudogenes *sbno2-B* e *simc1-B* em *Psalidodon paranae* (Characiformes, Characidae). 16
1.1 Introdução 16
1.2 Material e Métodos 18
1.2.1 Obtenção dos cromossomos metafásicos mitóticos 18
1.2.2 Coloração convencional com Giemsa 18
1.2.3 Detecção da heterocromatina constitutiva (Banda-C) 18
1.2.4 Extração do DNA genômico 19
1.2.5 Bibliotecas de DNA e RNA de *reads* curtos 19
1.2.6 Bibliotecas de DNA de *reads* longos 19
1.2.7 Análise de junções éxon-éxon em *reads* curtos 19
1.2.8 Análise de cobertura em RNAseq 20
1.2.9 Analise da vizinhança das regiões genicas – *sbno2* e *simc1* 20
1.3 Resultados 21
1.3.1 Análise cariotípica de *P. paranae* - Ribeirão Cascatinha 21
1.3.2 Identificação de junções éxon-éxon por análise de *reads* curtos 21
1.3.3. Análise de cobertura dos transcritos dos genes *simc1* e *sbno2* 24
1.3.4 Análise do sequenciamento de *reads* longos e análise da vizinhança das regiões gênicas – *sbno2* e *simc1* 25
1.3.5 PCR com *primers* de junção éxon-éxon para os genes *sbno2* e *simc1* 26
1.4 Discussão 27
1.5 Conclusões 29
1.6 Referências 30
Capítulo 2 – Análise da origem e estrutura dos pseudogenes *sbno2-B* e *simc1-B* em espécies do gênero *Psalidodon* (Characiformes, Characidae). 36
2.1 Introdução 36
2.2 Materiais e Métodos 37
2.2.1 Obtenção dos cromossomos metafásicos mitóticos 38
2.2.2 Coloração convencional com Giemsa 38
2.2.3 Detecção da heterocromatina constitutiva (banda-C) 38
2.2.4 Extração do DNA genômico 38
2.2.5 Sequenciamento de DNA e RNA por *reads* curtos 38
2.2.6 Análise de cobertura em DNAseq 39
2.2.7 Análise de junções éxon-éxon em *reads* curtos 39
2.3 Resultados 41
2.3.1 Análise cariotípica de *P. fasciatus* - Córrego Alambari 41
2.3.2 Análise cariotípica de *P. fasciatus* - Córrego das Araras 41
2.3.3 Análise cariotípica de *P. bockmanni* - Cachoeira Segredo 41
2.3.4 Identificação de junções éxon-éxon por análise de *reads* curtos 41
2.3.5 Análise de cobertura dos genes *simc1* e *sbno2* em *P. fasciatus* e *P. bockmanni* 42
2.3.6 Busca de SNPs B-específicos em espécies de *Psalidodon* 44
2.3.7 Validação por qPCR dos genes *sbno2* e *simc1* na variante *BfMa* de *P. fasciatus* 45
2.4 Discussão 47
2.5 Conclusões 49
2.6 Referências 50
3 Considerações finais 52
Material suplementar 53

Introdução

Cromossomos B, também chamados de cromossomos supranumerários, são elementos adicionais encontrados em diversas espécies de eucariotos que foram descritos pela primeira vez em percevejos *Metapodius* a cerca de um século atrás (Wilson, 1907). Estes cromossomos caracterizam-se por não possuírem par homólogo e por disseminarem comportando-se como elementos parasitas do genoma. De maneira geral, esses elementos são dispensáveis para a sobrevivência, uma vez que não são encontrados em todos os indivíduos de uma população. Esses cromossomos possuem diversos morfotipos, sendo encontrados desde microcromossomos, macrocromossomos, a isocromossomos B, aqueles que apresentam uma similaridade quanto ao conteúdo de ambos os braços (Houben et al., 2001; Rubtsov et al., 2004; Silva et al., 2014). A respeito da origem desses elementos, é sabido que podem surgir a partir de cromossomos do complemento padrão, também denominados de cromossomos A (Camacho, 2005) ou serem adquiridos através de processos de hibridização entre espécies próximas (Tosta et al., 2015). Após sua origem, os cromossomos B que apresentam maiores taxas de transmissão, através de um mecanismo chamado de *drive*, persistem nas populações por um longo período de tempo e se tornam verdadeiros cromossomos B. Esse mecanismo é o fator responsável pelo sucesso evolutivo dos cromossomos B, e pode ocorrer em diferentes maneiras, como drive meiótico, aumento de número por indivíduo e aumento de indivíduos portadores de B na população (Jones, 2018).

Os primeiros estudos realizados com cromossomo B que visaram descobrir a composição desses elementos, revelaram que estes apresentavam um alto grau de condensação da cromatina, o que fez surgir a ideia de que esses elementos eram elementos inertes (Rhoades e McClintock, 1935). Entretanto, com o avanço dos estudos de cromossomos B, foi possível verificar que esses cromossomos apresentam uma alta diversidade de sequências, como elementos repetitivos, DNAs ribossomais, DNAs satélite e elementos transponíveis (Camacho et al., 2000; Camacho, 2005) bem como genes codificadores de proteínas e pseudogenes (Navarro-Domínguéz et al., 2017a; Carmello et al., 2017; Clark e Kocher, 2019).

Pseudogenes e sua presença em cromossomos B

Pseudogenes são sequências de DNAs oriundas de genes funcionais que passaram por importantes mutações como deleções, inserções e mutações pontuais. Pseudogenes podem ser classificados quanto aos seus mecanismos de origem (Pink et al., 2011), sendo divididos em

pseudogenes não processados e processados (Garewal et al., 2021). Pseudogenes não processados são divididos em pseudogenes unitários, quando a sequência original sofre processos de mutação, causando a perda da função gênica; e duplicados, quando a sequência de um gene passa por um processo de duplicação e esta cópia sofre os efeitos de mutações (Garewal et al., 2021). Pseudogenes processados, por sua vez, surgem através de processos de retrotransposição de uma sequência de DNA. Este processo se inicia com a transcrição de um segmento de mRNA a partir de determinadas sequências do genoma e, em seguida, estes mRNA são processados para remoção dos introns pelo processo de *splicing*. Finalmente, o mRNA processado é transcrito reversamente gerando sequências de cDNA que se inserem em novos sítios no genoma. Dessa forma, pseudogenes retroinseridos não possuem introns, sendo caracterizados pela presença de junções éxon-éxon no DNA (Yusuf Tutar, 2012).

Pseudogenes podem desempenhar um grande papel evolutivo e regulatório, uma vez que essas sequências são passíveis de sofrerem alterações e ganharem novas funções, enquanto a sequência do gene original é mantida conservada (Yusuf Tutar, 2012). Com relação aos mecanismos regulatórios, os pseudogenes podem formar pequenos siRNAs ou competirem por fatores de transcrição (Pink et al., 2011), o que pode levar a alteração na expressão das cópias canônicas. É interessante destacar que a presença de pseudogenes vem sendo notada em diversas espécies com cromossomos B, assim como em centeio (Banaei-Moghaddam et al., 2013), em que 15% dos pseudogenes no cromossomo B possuem transcrição ativa, no gafanhoto *Eyprepocnemis plorans*, em que a presença de pseudogenes foi encontrada no cromossomo B24, (Navarro Domínguez et al., 2017), e por fim a presença de pseudogene retroinserido, evidenciado pela ausência de íntrons, em cromossomo B do peixe ciclídeo *Astatotilapia latifasciata* (Carmello et al., 2017).

Os peixes do gênero *Psalidodon* representam um ótimo modelo de estudos para cromossomos B, devido à alta diversidade de cromossomos B presentes em 8 espécies das 33 descritas atualmente (Fricke et al., 2020; Moreira-Filho et al., 2004; Oliveira et al., 2009). Silva et al. (2021) descreveram a presença de genes codificadores de proteínas nos cromossomos B de quatro espécies em *Psalidodon*, sendo que o compartilhamento de genes revelou uma origem em comum para o cromossomo B dessas espécies. Alguns destes genes possuem regiões codificadoras incompletas, assim indicando a presença de pseudogenes nos cromossomos B. Sendo assim o presente projeto visou caracterizar sequências pseudogenicas encontradas em quatro variantes de cromossomos B presentes em três espécies, bem como analisar o histórico evolutivo desses pseudogenes dentro do gênero.

Objetivos

O presente projeto se insere no programa geral de estudos citogenéticos de peixes Neotropicais que vem sendo desenvolvido no Laboratório de Biologia e Genética de Peixes do Departamento de Biologia Estrutural e Funcional, Instituto de Biociências da UNESP, Botucatu-SP, visando um melhor conhecimento do genoma e da estrutura cromossômica dos peixes e, em particular, das células eucariontes. Como já demonstrado, as espécies do gênero *Psalidodon* constituem um interessante modelo para o estudo dos cromossomos B, uma vez que estes elementos genômicos ocorrem em diferentes espécies, apresentando expressiva variação morfológica e estrutural, bem como uma estrutura molecular específica, constituída principalmente por regiões repetidas, genes funcionais e pseudogenes. Dessa forma, o presente estudo tem como objetivos gerais a identificação dessas regiões e caracterizar possíveis atividades funcionais determinadas pela sua presença.

De modo específico, pretende-se:

- caracterizar os pseudogenes *sbno2* e *simc1* encontrados no cromossomo B de *P. paranae,* visando identificar os processos evolutivos envolvidos na sua formação;
- investigar se esses pseudogenes são expressos a partir dos cromossomos B que os portam em *P. paranae*;
- verificar se os pseudogenes *sbno2* e *simc1* estão presentes nos cromossomos B de outras espécies de *Psalidodon*, bem como comparar a estrutura genômica dessas sequências entre as espécies.

Materiais e Métodos

Os materiais e os métodos utilizados, juntamente com os resultados obtidos e sua discussão, estão organizados na forma de artigos em dois capítulos e deverão ser publicados após as correções e aprimoramento do texto.

Referências

BANAEI-MOGHADDAM, A.M.; MEIER, K.; KARIMI-ASHTIYANI, R.; Houben, A. American Society of Plant Biologists. Formation and expression of pseudogenes on the B chromosome of rye. The Plant Cell v. 25, n. 7, p. 2536–2544, 2013.

CAMACHO, J. P. M. The Evolution of the Genome: CHAPTER 4 - B Chromosomes. Burlington: Academic Press, 2005.

CAMACHO, J.P., SHARBEL, T.F., BEUKEBOOM, L.W., B-chromosome evolution. Philos Trans R Soc Lond B Biol Sci v.355, n.1394, p.163-178, 2000.

CARMELLO, B.O., et al. The hnRNP Q-like gene is retroinserted into the B chromosomes of the cichlid fish *Astatotilapia latifasciata*. Chromosome Res. v. 25, n. 3-4, p. 277-290, 2017.

CLARK, F.E., KOCHER, T.D. Changing sex for selfish gain: B chromosomes of Lake Malawi cichlid fish. Sci Rep v.9, n.20213, p. 1-10, 2019.

FRICKE R., ESCHMEYER W.N., VAN DER LAAN, R. (2021). Eschmeyer's Catalog of fishes: genera, species, references. Disponível em: https://researcharchive.calacademy. org/research/ichthyology/catalog/fishcatmain.aspAcesso em: 06 Jan 2022

FRICKE R., ESCHMEYER W.N., VAN DER LAAN, R. (2021). Eschmeyer's Catalog of fishes: genera, species, references. Disponível em: https://researcharchive. calacademy.org/research/ichthyology/catalog/fishcatmain.asp Acesso em: 06 Jan 2022.

GAREWAL, N.; GOYAL, N.; PATHANIA, S.; KAUR, J.; SINGH, K. Gauging the trends of pseudogenes in plants. Crit Rev Biotechnol. v.41, n.7, p.1114-1129, 2021.

HOUBEN, A et al. The genomic complexity of micro B chromosomes of *Brachycome dichromosomatica*. Chromosoma v.110, n.7, p 451–459, 2001.

JONES, R N. Transmission and Drive Involving Parasitic B Chromosomes. Genes v.9, n.388, p.1-9, 2018.

MOREIRA-FILHO, O.; GALETTI, P.M. JR.; BERTOLLO, L.A. B chromosomes in the fish *Astyanax scabripinnis* (Characidae, Tetragonopterinae): an overview in natural populations. Cytogenet Genome Res. V.106, n.2-4, p.230-234, 2004.

NAVARRO-DOMÍNGUEZ, B. et al. Protein-coding genes in B chromosomes of the grasshopper *Eyprepocnemis plorans*. Scientific Reports v.7, n. 45200 p.1–12, 2017a

OLIVEIRA, C.; FORESTI, F.; HILSDORF, A.W.S. Genetics of Neotropical fish: From chromosomes to populations. Fish Physiol Biochem v.35, p.81–100, 2009.

PINK, R.C.; WICKS, K.; CALEY, D.P.; PUNCH, E.K.; JACOBS, L.; CARTER, D.R. Pseudogenes: Pseudo-functional or key regulators in health and disease? RNA. v.17, n.5, p.792-798, 2011.

RHOADES, M. M., MCCLINTOCK, B. The Cytogenetics of maize. Botanical Review, v.1, n.8, p. 292–325, 1935.

RUBTSOV, N B et al. Comparative analysis of micro and macro B chromosomes in the Korean field mouse *Apodemus peninsulae* (Rodentia, Murinae) performed by chromosome microdissection and FISH. Cytogenetic and genome research v.106, n.2-4 p289-294, 2004.

SILVA, D. M., et al. Delimiting the origin of a B chromosome by FISH mapping, chromosome painting and DNA sequence analysis in *Astyanax paranae* (Teleostei, Characiformes). PloS one, v.9, n.4, 2014.

TOSTA, V.C. et al., Possible Introgression of B Chromosomes between Bee Species (Genus *Partamona*). Cytogenetic and Genome Research, v.144, n.3, p.220–226, 2015.

WILSON, E. B. The supernumerary chromosomes of Hemiptera. Science, v.12, n.5, p.303-313, 1907.

YUSUF TUTAR. Pseudogenes. International Journal of Genomics. v.2012, p.1-4, 2012.

Resultados

As informações obtidas nos estudos citogenéticos e moleculares efetuados sobre a presença e estrutura genômica dos pseudogenes *sbno2* e *simc1* nos cromossomos B de espécies do gênero *Psalidodon*, resultaram na proposta de artigos científicos que são apresentados a seguir, na forma de capítulos:

Capítulo I – Análise estrutural dos pseudogenes *sbno2-B* e *simc1-B* em *Psalidodon paranae* (Characiformes, Characidae).

Capítulo II – Análise da origem e estrutura dos pseudogenes *sbno2-B* e *simc1-B* em espécies do gênero *Psalidodon* (Characiformes, Characidae).

Capítulo 1 - Análise estrutural dos pseudogenes *sbno2-B* e *simc1-B* em *Psalidodon paranae* (Characiformes, Characidae).

Capítulo 1 – Análise estrutural dos pseudogenes *sbno2-B* e *simc1-B* em *Psalidodon paranae* (Characiformes, Characidae).

1.1 Introdução

Cromossomos B são elementos adicionais ao complemento de cromossomos padrão. Cerca de 15% dos genomas eucariotos abrigam cromossomos B. Esses cromossomos são dispensáveis para sobrevivência, visto que não são todos os indivíduos que os possuem, mas se perpetuam ao se comportarem como parasitas genômicos (Östergren, 1945; Nur, 1962). Os cromossomos B geralmente carregam informações genéticas redundantes já presentes nos cromossomos padrão (A) (Martis et al., 2012; Valente et al., 2014; Makunin et al., 2014; Navarro-Dominguez et al., 2017; Ruiz- Ruano et al., 2019). No entanto, os cromossomos B também podem abrigar sequências exclusivas, assim como o cromossomo supranumerário paternal sex ratio (PSR) da vespa *Nasonia vitripennis* que possui sequências muito mais parecidas com espécies do gênero *Trichomalopsis* do que com os cromossomos A da própria espécie (McAllister e Werren, 1997; Dalla Benetta et al., 2020). Em alguns casos, os cromossomos B podem se tornar elementos genômicos benéficos para o portador (Pokorná e Reifová, 2021). Estudos com centeio (*Secale cereale*) mostraram que as plantas portadoras de cromossomos B apresentam uma maior resistência ao dano físico causado por estresse térmico (Pereira et al., 2017). Em cebolinha *Allium schoenoprasume*, as sementes portadoras desses cromossomos B apresentam maior taxa de germinação em ambientes secos de em detrimento às não portadoras (Holmes e Bougourd, 1989; Plowman e Bougourd, 1994). Um tipo especial de cromossomos B, os cromossomos restritos à linhagem germinal em aves canoras (GRC), podem até mesmo ser indispensáveis (Kinsella et al., 2019).

Apesar de serem compostos principalmente por sequências de DNA repetitivo (Ziegler et al., 2003; Carmello et al., 2017; Li et al., 2017; Ramos et al., 2017; Coan et al., 2018; Hanlon et al., 2018; Milani et al., 2018) os cromossomos B também possuem genes codificadores de proteínas que podem ser ativamente transcritos (Banaei-Moghaddam et al., 2013; Miao et al., 1991; Ruiz-Estévez et al., 2012; 2013; 2014; Trifonov et al., 2013; Valente et al., 2014; Navarro-Domínguez et al., 2017a, b; Dalla Benetta et al., 2019). Além de expressarem seus próprios transcritos, os cromossomos B podem influenciar na expressão de transcritos do genoma A (Ma et al., 2021). Como exemplo disso, em *Scilla autumnalis* a presença do cromossomo B altera a expressão do gene para a enzima esterase do complemento A (Rejón et

al., 1980; Oliver et al., 1982). Os efeitos dos transcritos B-específicos podem também serem vistos em nível traducional, como ocorre em centeio, em que os cromossomos B codificam uma proteína funcional do tipo Argonauta com atividade *slicer* (Ma et al., 2017).

É interessante destacar que diversos pseudogenes vêm sendo encontrados em cromossomos B. Apesar de terem sofrido alterações na sua sequência original, os pseudogenes podem desempenhar diversas funções como tradução de proteínas, produção de RNAs não codificantes e regulação de interações da cromatina (Cheetam et al., 2020). Ainda, os pseudogenes podem formar pequenos siRNAs ou competirem por fatores de transcrição (Pink et al., 2011). São exemplos da presença desses elementos em cromossomo B os pseudogenes HYI, SLIT, TOP2A, MYCB2, e CKAP2, B-CAP-G localizados na região distal do cromossomo B24 do gafanhoto *Eyprepocnemis plorans*. Dentre eles, os genes CKAP2, CAP-G e MYCB2 são relacionados a atividades do ciclo celular e são expressos a partir do cromossomo B (Navarro-Domínguez et al., 2017). Em *A. latifasciata* foi demostrado que o pseudogene derivado do gene *hnRNP* sofreu processo de retroinserção no cromossomo B, o que foi comprovado pela ausência de introns e presença de elementos transponíveis com domínio da transcriptase reversa flanqueando suas sequências neste cromossomo (Carmello et al., 2017). Além destes, seguimentos pseudogênicos também foram descritos em cromossomos B em centeio, *Secale cereale* (Banaei-Moghaddam et al., 2013).

Um bom modelo para estudos de cromossomos B são as espécies de peixes pertencentes ao gênero *Psalidodon* (Characiformes), que antigamente pertenciam ao gênero *Astyanax*. O gênero *Psalidodon,* recentemente reposicionado por Terán et al. (2020), é um grupo amplamente distribuído pela América do Sul e composto por cerca de 33 espécies (Fricke et al., 2020). Até o presente, cromossomos B foram registrados em oito espécies do gênero, sendo predominantes os cromossomos B metacêntricos grandes (Moreira-Filho et al., 2004; Oliveira et al., 2009). Essa coincidência levou Moreira-Filho et al. (2001) a sugerirem uma origem comum dos cromossomos B em várias espécies de *Psalidodon*. Essa hipótese ganha consistência a partir dos resultados recentes sobre o conteúdo repetitivo de DNA e de genes codificadores de proteínas nesses componentes genômicos (Silva et al., 2017; Silva et al., 2021).

Com o intuito de testar a hipótese da origem comum dos cromossomos B em *Psalidodon* de forma mais consistente, Silva et al. (2021) analisaram o conteúdo de genes codificadores de proteínas (GCP) do cromossomo B de *P. paranae* e *P. scabripinnis.* Nesta análise foram encontrados 20 GCPs nos cromossomos supranumerários das espécies, alguns dos quais com

suas regiões codificadoras incompletas. Devido a isto, foi postulado que estes genes podem ter sofrido processos evolutivos que levou à formação de pseudogenes. De fato, a análise prévia destes genes demonstrou que algumas sequências provenientes do cromossomo B não possuem introns (Silva et al., 2021), indicando que poderiam se tratar de pseudogenes retroinseridos nesses cromossomos. Nesse contexto, o presente estudo teve por objetivo caracterizar os pseudogenes *sbno2-B* e *simc1-B* encontrados no cromossomo B de *P. paranae*, visando identificar os processos evolutivos envolvidos na sua formação, bem como verificar se esses pseudogenes estão sendo expressos a partir dos cromossomos B que os portam em *P. paranae*.

1.2 Material e Métodos

Para as análises foram coletados exemplares de *Psalidodon paranae* da população do Ribeirão Cascatinha (22°53′30″S 48°28′36″W). Além disso foi feito uso de tecidos e preparações cromossômicas *de P. paranae*, conforme Silva et al. (2021) disponíveis no Laboratório de Biologia e Genética de Peixes de Botucatu (LBGP). Os procedimentos de amostragem, manutenção e análise dos animais foram realizados de acordo com as regras internacionais em experimentação animal seguidas pela Universidade Estadual Paulista (CEEAA/IBB/UNESP, Número de Protocolo 1227).

1.2.1 Obtenção dos cromossomos metafásicos mitóticos

Para obtenção dos cromossomos metafásicos mitóticos foi utilizado o protocolo proposto por Foresti et al. (1981), com algumas adaptações. Essa metodologia envolve basicamente a inibição da polimerização dos microtúbulos durante a divisão celular pela ação da colchicina, a hipotonização das células em suspensão e a fixação celular pela mistura de metanol/ácido acético.

1.2.2 Coloração convencional com Giemsa

As suspensões celulares foram gotejadas em lâminas limpas para obtenção das preparações cromossômicas e posteriormente coradas com solução de Giemsa a 5% em tampão fosfato (pH 6,7) por 5 minutos.

1.2.3 Detecção da heterocromatina constitutiva (Banda-C)

Para obtenção do padrão de distribuição da heterocromatina constitutiva nos cromossomos e detecção do cromossomo B nas preparações, foi aplicada a técnica de bandamento-C descrita originalmente por Sumner (1972), com algumas adaptações.

1.2.4 Extração do DNA genômico

O DNA das amostras foi extraído a partir de fragmentos de tecido muscular dos exemplares com a utilização do Kit NucleoSpin® Tissue Columns, Collection Tubes (2ml) e tratado com RNase A (20 mg/ml - Invitrogen) para remoção do RNA residual das amostras.

1.2.5 Bibliotecas de DNA e RNA de *reads* curtos

As bibliotecas de DNA e RNA de *P. paranae* utilizadas foram obtidas por Silva et al. (2021). As informações referentes aos dados de sequenciamento estão sumarizadas na Tabela 2.

Tabela 2. Descrição das bibliotecas utilizadas obtidas em projetos anteriores e publicadas em Silva et al. (2021).

Espécie e tamanho do genoma	Origem	Tecido	2n	Fêmea	Macho	Plataforma	Bases (Gb)	Cobertura
P. paranae (1.8 Gb)	DNA	Fígado	50	3		DNBseq 2x150bp	18	10x
			50+BpM	3		DNBseq 2x150bp	18	10x
	RNA	Ovário	50	3		Hiseq 2500 2x100bp	12	
			50+BpM	3		Hiseq 2500 2x100bp	12	

1.2.6 Bibliotecas de DNA de *reads* longos

Os *reads* longos utilizados foram produzidos por PacBio – Sequel obtidos por Santos et al. (2021), sob o número de acesso SRA: SRR11678219 (3,04 Gb).

1.2.7 Análise de junções éxon-éxon em *reads* curtos

Para verificar a presença de junções éxon-éxon utilizando *reads* curtos, foi realizado o mapeamento dos genes *sbno2* e *simc1* em bibliotecas de gDNA utilizando o programa SSAHA2 (Ning et al., 2001), com pontuação mínima de alinhamento de 40 e 80% de identidade mínima. Posteriormente, a sequência de cada gene foi utilizada como referência para realizar o

mapeamento dos *reads* selecionados no software Geneious® 7.1.3. Assim, uma análise manual permitiu verificar a presença/ausência de *reads* entre junções éxon-éxon.

1.2.8 Análise de cobertura em RNAseq

Para análise de cobertura em RNAseq nas bibliotecas de *P. paranae* foram utilizadas bibliotecas previamente processadas por Silva et al. (2021). Para realizar o mapeamento de transcritos dos genes *sbno2* e *simc1* foi utilizado o programa SSAHA2 (Ning et al., 2001), com pontuação mínima de alinhamento de 40 e 80% de identidade mínima. Além disso, foram utilizados *scripts* customizados para contar o número de *reads* mapeados como medida de abundância e para gerar os gráficos (https://github.com/fjruizruano/ngs-protocols/blob/master/mapping_blat_gs.py; https://github.com/fjruizruano/ngs-protocols/blob/master/coverage_graphics.py

1.2.9 Analise da vizinhança das regiões genicas – *sbno2* e *simc1*

Para análise da vizinhança dos pseudogenes *sbno2* e *simc1*, foram utilizados *reads* PacBio gerados para indivíduos 1B de *P. paranae*, número de acesso SRA: SRR11678219. Para isto, foi utilizada a ferramenta BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Madden, 2003), sendo que a busca por elementos transponíveis foi realizada contra o banco de dados Repbase (Jurka et al., 2005) do Genetic Information Research Institute (Giri) (http://www.girinst.org/repbase/).

Confecção de *primers* e reações de PCR para junção éxon-éxon dos genes *sbno2* e *simc1*

Como uma técnica adicional para comprovar as junções éxon-éxon dos pseudogenes retroinseridos, além das análises de *reads* curtos e *reads* longos, foram realizadas reações de PCR com *primers* desenhados com auxílio do *software* Geneious® 7.1.3, especificamente nas junções éxon-éxon para as regiões que mostraram alta cobertura nos indivíduos 1B de *P paranae*. Para o pseudogene *sbno2-B* foram desenhados *primers* nas junções dos éxons 4-5, 9-10 e 12-13 para amplificação em indivíduos 1B, sendo que os *primers* da junção 9-10 também foram utilizados em indivíduos 0B, uma vez tal junção também se mostrou presente em alguns *reads* destes indivíduos. Para o pseudogene *simc1-B*, foram desenhados *primers* das junções 1-2 e 2-3 para amplificação em indivíduos 1B. Apenas estas regiões foram selecionadas para se comprovar a validade das análises baseadas em bioinformática, ainda que outras junções éxon-éxon tenham sido encontradas nas regiões presentes nos cromossomos B. As reações com o

pseudogene *sbno2-B* foram realizadas em cinco indivíduos 0B e cinco indivíduos 1B, sendo utilizados os mesmos espécimes para cada junção.

As PCRs com primers do pseudogene *sbno2-B* foram realizadas usando os mesmos parâmetros para todas as junções. Os parâmetros utilizados nas reações foram de 10 ng de DNA da amostra em um volume final de 12,8 μL, sendo que cada reação continha 10X PCR buffer, 200 mM dNTPs, 0,25 para cada primer 2,5 μM e 1U Taq DNA polimerase. As condições de termociclagem seguiram o protocolo estabelecido, com tratamentos que envolveram 31 ciclos a 94°C por 2 min, a 94°C por 30 s, a 60°C 45 s, a 68°C por 30 s e a 68°C por 10 min.

Já as reações para o pseudogene *simc1-B* também foram realizadas com cinco indivíduos 0B e cinco indivíduos 1B, sendo utilizados os mesmos para cada junção. A reação de PCR da junção 1-2 foi realizada usando 10 ng de DNA da amostra em um volume final de 12,8 μL. Cada reação continha 10X PCR buffer, 200 mM dNTPs, 0,25 para cada primer a 2,5 μM e 1U de Taq DNA polimerase. As reações da PCR foram realizadas com 35 ciclos a 94°C por 2 min, a 94°C por 30 s, a 60°C por 45 s, a 68°C por 30 s, e a 68°C por 10 min. Para a junção 2-3 foram utilizados os mesmos parâmetros anteriores, com alteração apenas da temperatura de anelamento dos primers para 58°C.

1.3 Resultados

1.3.1 Análise cariotípica de *P. paranae* - Ribeirão Cascatinha

Nos espécimes de *P. paranae* da população do ribeirão Cascatinha, 26 entre 35 fêmeas e cinco entre 28 machos analisados apresentaram cromossomos B metacêntricos grandes heterocromáticos (*BpM*) em 100% das células.

1.3.2 Identificação de junções éxon-éxon por análise de *reads* curtos

A análise dos *reads* curtos de *P. paranae* mostrou *reads* entre todas junções éxon-éxon do gene *sbno2* na região alta cobertura dos indivíduos 1B (Figura 1). A mesma análise feita para indivíduos 0B mostrou que a junção dos éxons 8-9 e 9-10 apresenta alguns *reads* (Figura 2). Não houveram *reads* nas demais junções.

Para o gene *simc1* a mesma análise mostrou *reads* em todas junções éxon-éxon do gene na região alta cobertura das bibliotecas 1B de *P. paranae* (Figura 3). Não houve *reads* nas junções nos indivíduos 0B (Figura 4).

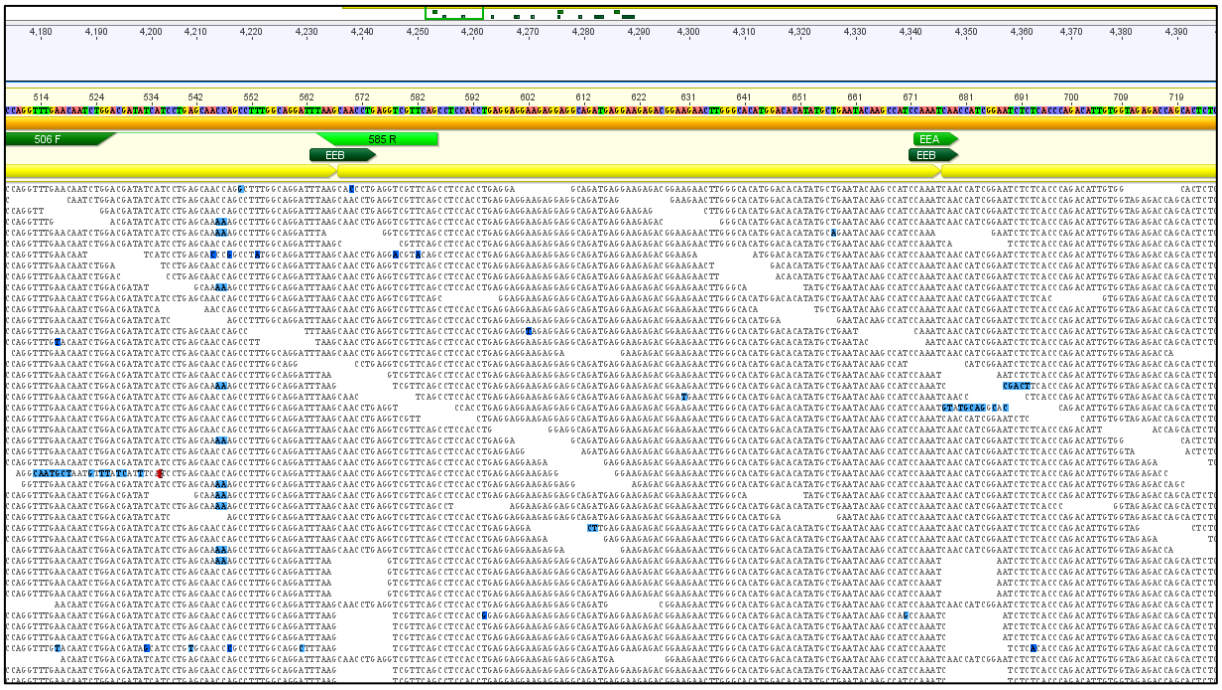


Figura 1. Análise de *reads* curtos de *Psalidodon paranae* de indivíduos 1B do gene *sbno2*. Notar leitura entre todas junções éxon-éxon (junção das anotações em amarelo) do gene *sbno2* na região alta cobertura anotada em laranja. Aqui exemplificadas em apenas duas junções.

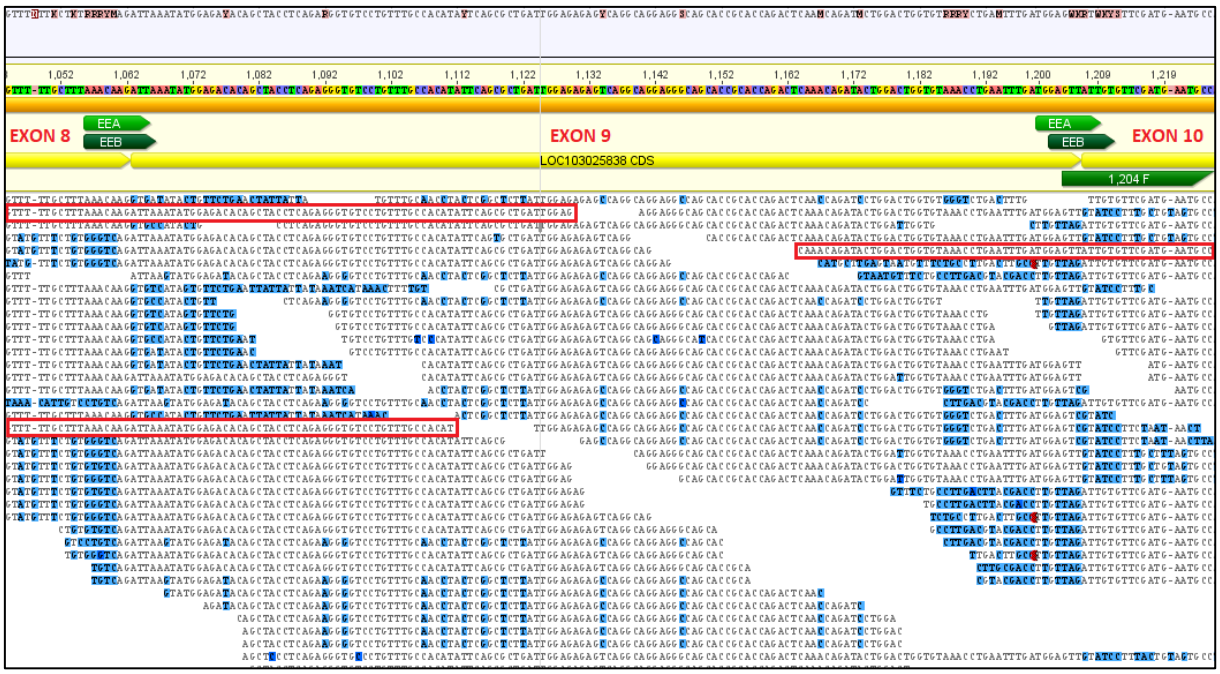


Figura 2. A análise de *reads* curtos de *Psalidodon paranae* de indivíduos 0B do gene *sbno2*.

Notar alguns *reads* entre as junções dos éxons 8-9 e 9-10 (junção das anotações em amarelo) destacados nas caixas vermelhas.

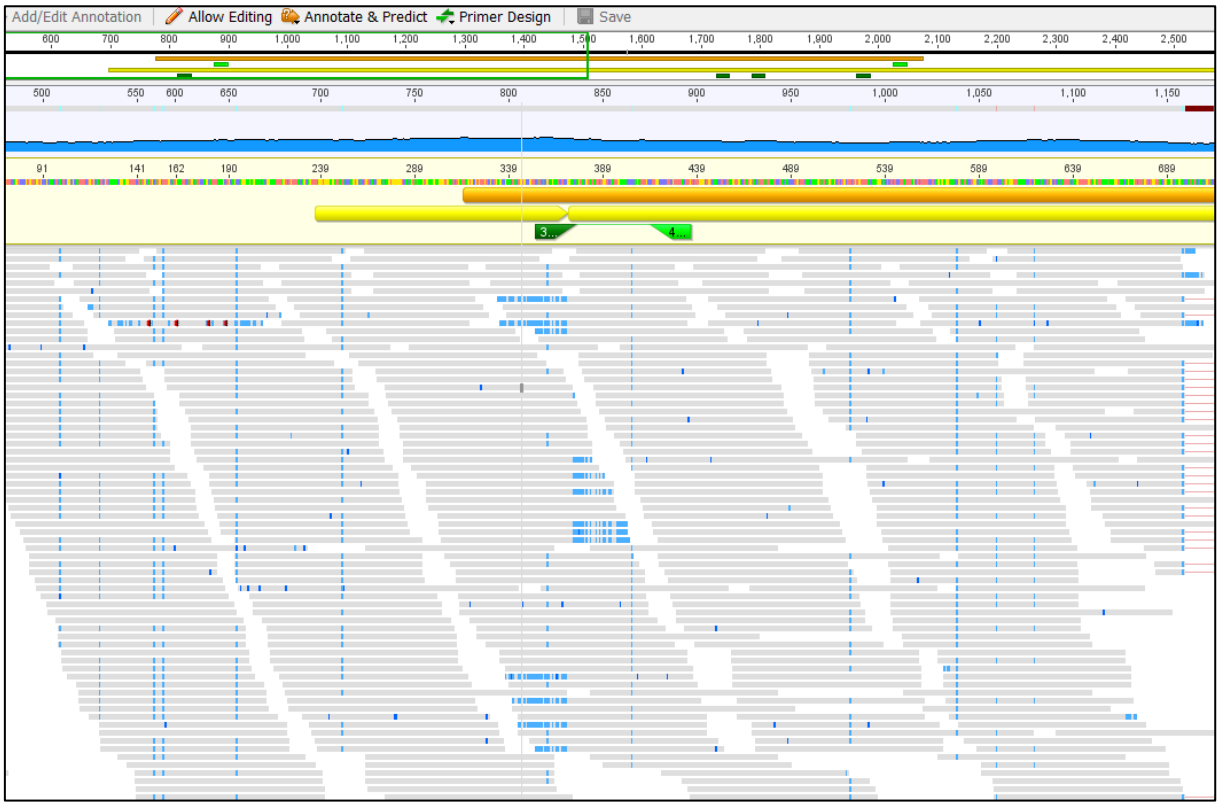


Figura 3. A análise de *reads* curtos de *Psalidodon paranae* de indivíduos 1B do gene *simc1*. Notar leitura entre todas junções éxon-éxon (junção das anotações em amarelo) do gene *simc1* na região alta cobertura anotada em laranja. Aqui exemplificadas em apenas uma junção.

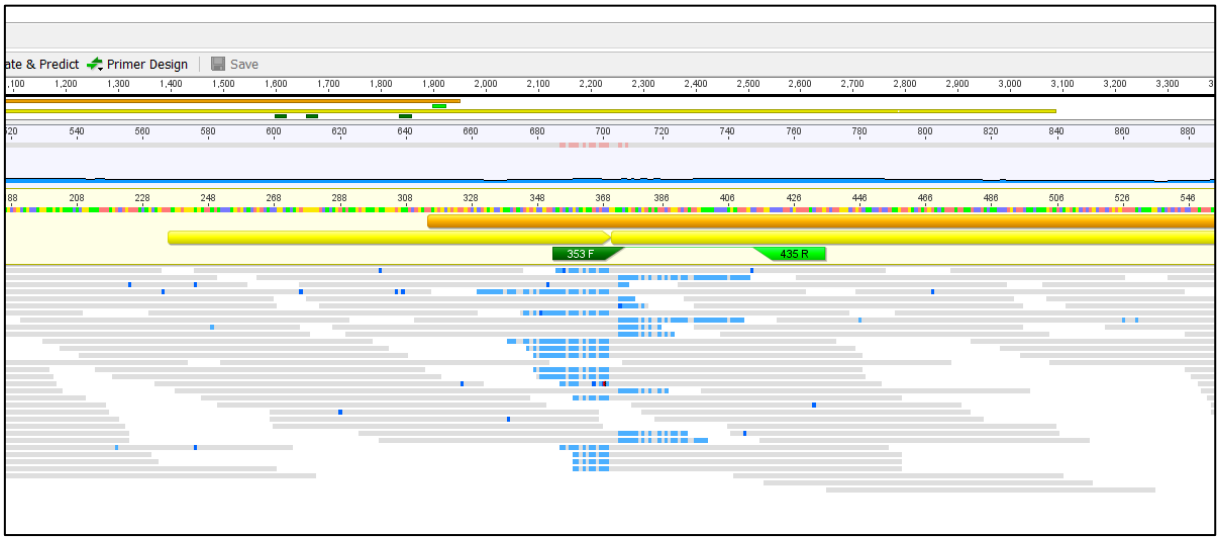


Figura 4. A análise de *reads* curtos de *Psalidodon paranae* de indivíduos 0B do gene *simc1*. Notar que não há *reads* entre a junção éxon-éxon (junção das anotações em amarelo) na região alta cobertura anotada em laranja.

1.3.3. Análise de cobertura dos transcritos dos genes *simc1* e *sbno2*

A análise de cobertura dos transcritos dos genes *sbno2* e *simc1* em espécimes 0B e 1B de *P. paranae* mostrou uma subexpressão em gônadas de fêmeas portadoras de cromossomo B (Figura 5). No entanto, em tecido muscular não houve diferença significativa entre os indivíduos 0B e 1B (Figura 6).

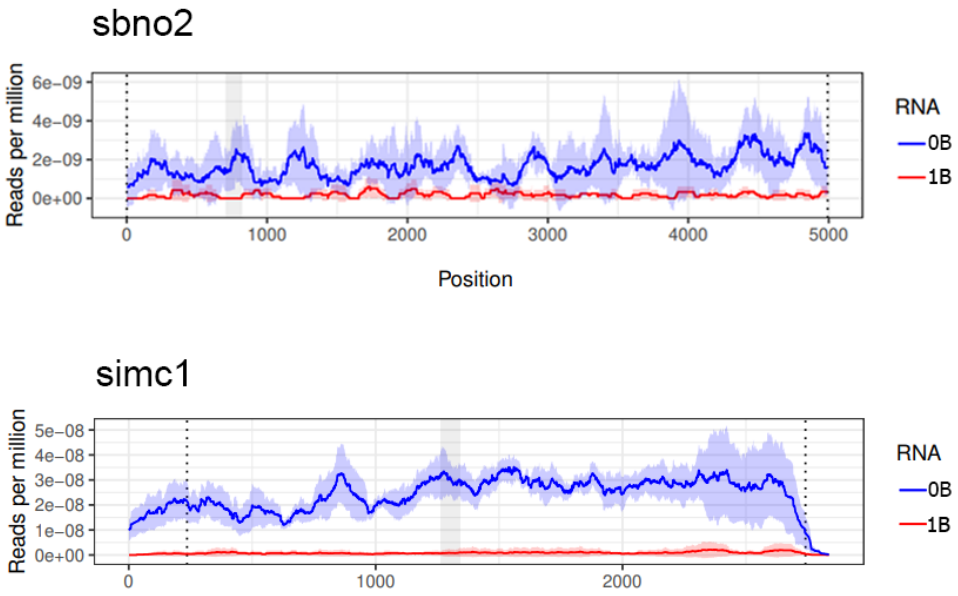


Figura 5. Gráfico de cobertura dos transcritos de RNA dos genes *sbno2* e *simc1* (0B e 1B) em gônada a partir das bibliotecas de indivíduos de *P. paranae*. Notar subexpressão de ambos os genes em bibliotecas 1B. A CDS encontra-se delimitada pelas linhas pontilhadas verticais.

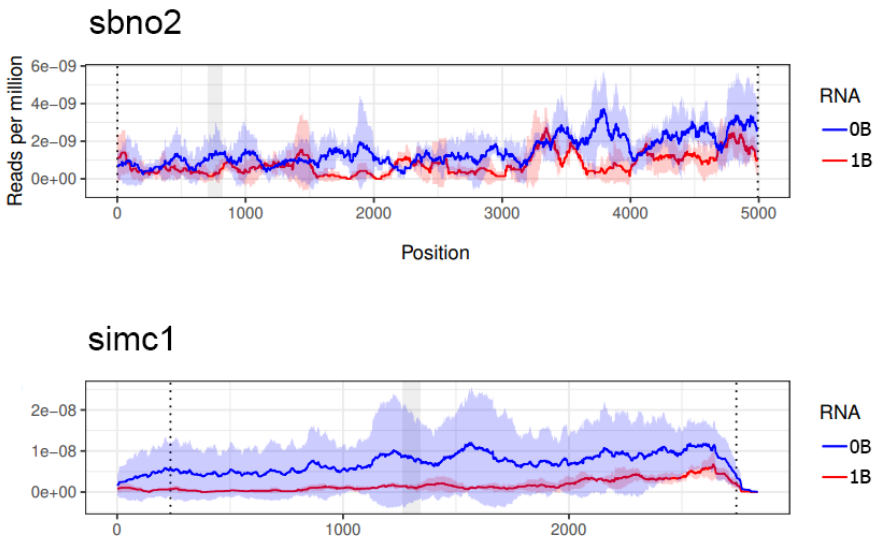


Figura 6. Gráfico de cobertura dos transcritos de RNA dos genes *sbno2* e *simc1* (0B e 1B) em músculo a partir das bibliotecas de indivíduos de *P. paranae*. Notar subexpressão no gene *simc1* em bibliotecas 1B. A CDS encontra-se delimitada pelas linhas pontilhadas verticais.

1.3.4 Análise do sequenciamento de *reads* longos e análise da vizinhança das regiões gênicas – *sbno2* e *simc1*

A análise dos *reads* PacBio gerados para indivíduos de *P. paranae* 1B confirmou a presença das junções éxon-éxon dos genes *sbno2* e *simc1* em indivíduos 1B. Além disso, a análise de vizinhança dos *reads* com leitura éxon-éxon desses genes através da plataforma BLAST revelou a presença de diversos outros genes (Tabela 6), sendo que alguns destes se repetem em mais de um *read*. Além disso, a plataforma REPBASE revelou a presença de diversos elementos transponíveis e apresentados na Tabela 7, como o transposon *Gypsy* que está relacionado em ambos os genes. (Tabelas S1 e S2).

Tabela 6. Analise de *reads* PacBio do gene *sbno2* para indivíduos de *Psalidodon paranae* 1B. Notar a presença de genes na vizinhança, número de *reads* e sua homologia.

Sbno2	**Número de *reads***	**Homologia (%)**
organic cation transporter protein-like mRNA	2	97,96
RNA éxonuclease 1 homolog (variante 1)	1	91,84
RNA éxonuclease 1 homolog (variante 2)	1	91,67
ST8 alphe-N-acetyl-neuraminide alphe-2,8-sialyltransferase1 (st8sia1) mRNA	1	88,41
non-compact myelin associated protein (ncmap) mRNA	4	88,34
nucleolar and spindle associated protein 1 (nusap1) mRNA	1	87,28
engulfment and cell motility protein 3-like transcript variant	1	87,07
stimulated by retinoic acid gene 6 protein-like (variante 1)	2	85,27
NACHT LRR e PYD domains-containing protein 12-like	2	85,02
transcription factor 15-like	1	84,93
ankyrin repeat domain-containing protein 34C-like mRNA	2	83,33
ras-related protein rab-32-like mRNA	1	81,82
stimulated by retinoic acid gene 6 prontein-like (variante 2)	1	81,62
C-type lectin domain family 3 member B (clec3b) mRNA -	4	81,45
sodium channel protein type 4 subunit alpha B-like transcript variant	2	80,86
asialoglycoptrotein receptor 1like mRNA	1	79,08

uncharacterized mRNA	1	84,57
uncharacterized transcript variant	1	93,33

Tabela 7. Analise de *reads* PacBio do gene *simc1* para indivíduos de *Psalidodon paranae* 1B. Notar a presença de genes na vizinhança, número de *reads* e sua homologia.

Simc1	**Número de *reads***	**Homologia (%)**
TATA-box binding protein associated factor 13 (taf13) transcript variant	1	100
von Hippel-Lindau disease tumor suppressor like mRNA	1	100
ubiquitin carboxyl-terminal hydrolase 21-like transcript variant	2	96,47
protein FAM mRNA	1	95,65
malignant brain tumors 1 protein-like	1	94,12
myosinIIIA (myo3a) transcript variant (1 a 6)	1	91,67
solute carrier family 50 member 1 (slc50a1) mRNA	1	90,91
transmembrane protein 109 (tmem109) transcript variant	1	90,70
Cas scaffolding protein family member 4 (cass4) transcript variant	1	88,57
macrophage colony-stimulating factor 1 receptor-like	1	86,04

1.3.5 PCR com *primers* de junção éxon-éxon para os genes *sbno2* e *simc1*

As reações de PCR utilizando *primers* das junções éxon-éxon 4-5, 12-13 do gene *sbno2* para os indivíduos 0B não mostraram amplificação ou apresentaram dímero de primer. Nas reações utilizando os *primers* das junções 4-5, 9-10 e 12-13 para os indivíduos 1B foi observada amplificação com fragmentos de aproximadamente 80 pb, o que corresponde ao tamanho do amplicon esperado. A reação de PCR com o *primer* da junção éxon-éxon 9-10 para os indivíduos 0B, por sua vez, apresentou amplificação, corroborando as analises *in silico* (Figura 6).

Para o gene *simc1*, as reações utilizando *primers* das junções 1-2 e 2-3 geraram amplificação em indivíduos 1B, com amplicons em torno de 80pb, enquanto que em indivíduos 0B não foi observada amplificação específica em nenhuma das junções, com exceção de um

indivíduo na junção 1-2, a qual pode-se tratar de uma contaminação do DNA gnômico com DNA ou amplicons 1B (Figura 7).

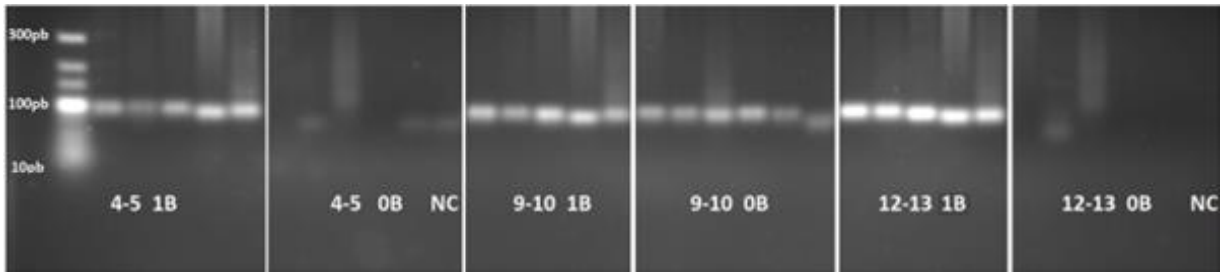


Figura 6. Eletroforese em gel de agarose a 3% evidenciando amplicons gerados através das reações de PCR utilizando os primers das junções éxon-éxon do gene *sbno2* em *P. paranae* 0B e 1B. As reações de controle negativo estão indicadas como NC.

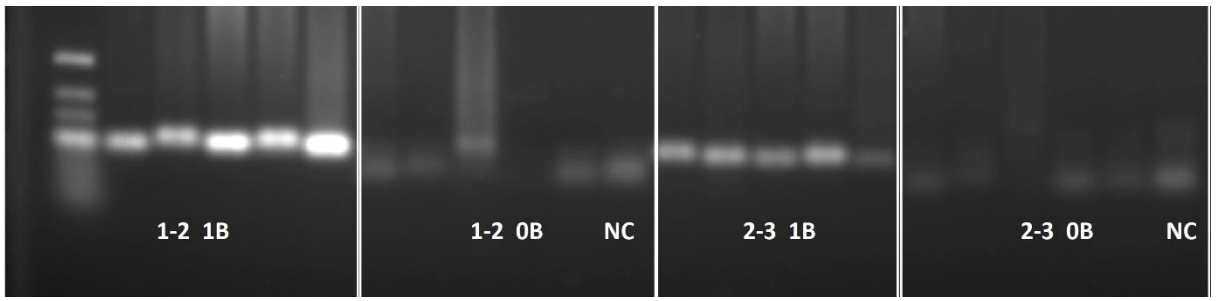


Figura 7. Gel de eletroforese a 3% evidenciando os amplicons gerados através das reações de PCR utilizando os primers das junções éxon-éxon do gene *simc1* em *P. paranae* 0B e 1B. As reações de controle negativo estão indicadas como NC.

1.4 Discussão

Assim como os cromossomos B, os pseudogenes já foram considerados como DNA lixo, sem importância para os organismos portadores. No entanto, esses elementos genômicos podem apresentar papéis fundamentais na regulação da expressão gênica, como por exemplo a estimulação da proliferação de células tronco mesenquimais, conduzidas pela transcrição do pseudogene Oct4P1 associado ao fator de transcrição Oct4 (fator de transcrição associado a plenipotência) (Lin et al., 2007). Outra atividade regulatória dessas sequências gênicas é revelada pela superexpressão de mRNA do gene ABCC6, que se apresenta na atuação de transportadores transmembrânicos, determinada pela expressão do pseudogene ABCC6P (Piehler et al., 2008). Com o crescente avanço e acessibilidade às tecnologias de sequenciamento de alto rendimento, diversos estudos vêm revelando a presença de pseudogenes

nos cromossomos B de diferentes organismos (Banaei-Moghaddam et al., 2013; Navarro-Domínguez et al., 2017a, b). No entanto, até então poucos estudos encontraram pseudogenes retrotranspostos em cromossomos B (Carmello et al., 2017).

A presença de junções éxon-éxon observada nas análises bioinformáticas das regiões de alta cobertura dos genes *sbno2-B* e *simc1-B* nas bibliotecas 1B e sua posterior validação por PCR mostraram que estes fragmentos são pseudogenes processados retroinseridos, do mesmo modo como foi observado ocorrer com o gene hnRNP no cromossomo B de *Astatotilapia latifasciata* (Carmello et al., 2017). Devido ao processo de retrotransposição ocorrer a partir de um mRNA, os genes retrotranspostos perdem muito de sua característica original, não possuindo introns e elementos regulatórios (Pan e Zhang, 2009). Além disso, Carmello et al. (2017) observaram a presença de elementos transponíveis com sequências nucleotídicas da transcriptase reversa na vizinhança do gene hnRNP, o que representa uma forte indicação de que este gene teria sido retrotransposto dos cromossomos do complemento A para o cromossomo B. Similarmente, no presente estudo a análise de *reads* longos mostrou que diversos fragmentos de elementos transponíveis também se encontram na vizinhança dos pseudogenes *sbno2* e *simc1*, o que reforça as evidências de que estes fragmentos foram retrotranspostos.

Foi observado ainda que alguns *reads* foram mapeados nas junções éxon-éxon 8-9 e 9-10 do pseudogene *sbno2*-B em um indivíduo 0B. Uma possível explicação para a presença de junções éxon-éxon no genoma 0B seria que, após a retrotransposição destas sequências para o cromossomo B em um ancestral, alguns fragmentos foram posteriormente transpostos para algum cromossomo do complemento A. Alternativamente, uma hipótese remota a ser também considerada seria que esse processo de retrotransposição no genoma A poderia ter ocorrido sem envolvimento do cromossomo B, sendo estas sequências retrotranspostas a partir da cópia canônica do gene para os componentes do próprio genoma A. No entanto, para melhor entendimento e sustentação de ambas estas hipóteses, outros estudos sobre a presença de SNPs B-específicos (Silva et al., 2021) nos *reads* encontrados são necessários. Ainda, apesar de estes *reads* terem sido encontrados em apenas um indivíduo 0B, as reações de PCR com cinco indivíduos 0B mostraram que a junção 9-10 está presente em todas as amostras, o que contraria a possibilidade de erros no sequenciamento como troca de adaptadores (Costello et al., 2018).

A análise de cobertura de transcritos dos genes *sbno2* e *simc1* realizada no presente estudo demonstrou que ambos os genes apresentam subexpressão em gônadas, quando na

presença de cromossomos B. Silva et al. (2021) verificaram que ambos os genes não possuem transcritos com SNPs B-específicos, indicando que estes genes não estariam sendo expressos a partir dos cromossomos B. No entanto, a metodologia empregada se limita à análise de sequências de mRNA. Desta forma, a expressão de RNAs não codificantes a partir dos cromossomos B não pode ser descartada, uma vez que já foi visto em *Astatotilapia latifasciata* que o cromossomo B pode codificar long non-coding RNAs (Ramos et al., 2017). Ainda, outros tipos de RNAs não codificantes como small non-coding RNAs, small interfering RNAs e exo-siRNAs, entre outros (Patil et al., 2014), poderiam estar regulando a expressão das cópias canônicas presentes nos cromossomos do complemento A em *P. paranae*. Nesse sentido, a análise de *reads* longos mostrou que, na vizinhança de ambos os pseudogenes existem diversos elementos transponíveis no cromossomo B. Estes elementos poderiam apresentar *motifs* de ligação a fatores de transcrição, que poderiam estar competindo com os fatores de transcrição dos genes canônicos e provocando assim, a subexpressão desses genes (Ali et al., 2021).

A implicação de possíveis efeitos funcionais gerados pela subexpressão de ambos os genes nessas espécies ainda é uma relação complexa de ser entendida. O gene *sbno2* é um corregulador transcricional responsável por fatores como regulação de cascata inflamatória (Grill et al., 2015), enquanto que o gene *simc1* é responsável pela inibição da protease CAPN3 (Ono et al., 2013). Assim, para um melhor entendimento das causas e efeitos dessa subexpressão, estudos mais aprofundados são necessários envolvendo principalmente a análise global dos níveis de transcritos para identificar possíveis redes de interação gênica. No entanto, a ausência de expressão diferencial do gene *sbno2* e a alta variabilidade em bibliotecas 1B do gene *simc1* em amostras de tecido muscular parecem indicar que ambos os genes não sofrem influência com a presença destes cromossomos em tecidos somáticos. Neste contexto, pode-se especular que ambos os genes possam estar relacionados a alterações em tecidos reprodutivos.

1.5 Conclusões

Os pseudogenes *sbno2-B* e *simc1-B* são pseudogenes do tipo processados que passaram por um processo de retroinserção, o que é comprovado pelas junções éxon-éxon em regiões de alta cobertura de bibliotecas +B. Ainda, a análise de *reads* longos revelou a presença de diversos elementos transponíveis na vizinhança de ambos os pseudogenes, o que reforça a hipótese de retrotransposição. A presença de junções éxon-éxon no gene *sbno2* de indivíduos 0B parece indicar que processos de retrotransposição também possam ter ocorrido nos cromossomos A. Ambos os genes apresentam menores níveis de expressão associado a presença de cromossomos

B, o que indica que a presença dos pseudogenes *sbno2-B* e *simc1-B* estejam levando a regulação das cópias canônicas.

1.6 Referências

ALI, A.; HAN, K.; LIANG, P. Role of Transposable elements in gene regulation in the human genome. Life (Basel). v.2, n.118, p.1-23, 2021.

BANAEI-MOGHADDAM, A.M.; MEIER, K.; KARIMI-ASHTIYANI, R.; Houben, A. American Society of Plant Biologists. Formation and expression of pseudogenes on the B chromosome of rye. The Plant Cell v. 25, n. 7, p. 2536–2544, 2013.

CARMELLO, B.O., et al. The hnRNP Q-like gene is retroinserted into the B chromosomes of the cichlid fish *Astatotilapia latifasciata*. Chromosome Res. v. 25, n. 3-4, p. 277-290, 2017.

CHEETHAM, S.W., FAULKNER, G.J., DINGER, M.E. Overcoming challenges and dogmas to understand the functions of pseudogenes. Nat Rev Genet v.21, p.191–201, 2020.

COAN, R.L.B.; MARTINS, C. Landscape of transposable elements focusing on the B chromosome of the cichlid fish Astatotilapia latifasciata. Genes. v. 9 n. 269. p. 1-18, 2018.

COSTELLO, M.: FLEHARTY, M.; ABREU, J. et al. Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. BMC Genomics v.19, n.332, p.1-10, 2018.

DALLA BENETTA E.; AKBARI, O.S.; FERREE, P.M. Sequence Expression of Supernumerary B chromosomes: Function or fluff? Genes v. 10, n. 123, p.1-13, 2020.

FRICKE R., ESCHMEYER W.N., VAN DER LAAN, R. (2021). Eschmeyer's Catalog of fishes: genera, species, references. Disponível em: https://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.aspAcesso em: 06 Jan 2022.

GRILL, M.; SYME, T.E.; NOÇON, A.L.; LU, A.Z.; HANCOCK, D.; ROSE-JOHN, S.; CAMPBELL, I.L. Strawberry notch homolog 2 is a novel inflammatory response factor predominantly but not exclusively expressed by astrocytes in the central nervous system.Glia. v.63, n.10, p.1738-1752 2015.

HANLON, S.L.; MILLER, D.E.; ECHE, S. et al. Origin, Composition, and structure of the supernumerary B chromosome of *Drosophila melanogaster*. Genetics v. 210, n. 4, p. 1197-121, 2018.

HOLMES, D.; BOUGOURD, S. B-chromosome selection in Allium schoenoprasum. I. Natural populations. Heredity, v. 63, p.83–87, 1989.

JURKA J, KAPITONOV VV, PAVLICEK A, KLONOWSKI P, KOHANY O, WALICHIEWICZ J. 2005 Repbase update, a database of eukaryotic repetitive elements. Cytogenet Genome https://www.girinst.org/repbase/ Acesso em: 10/04/2022.

KINSELLA, C.M.; RUIZ-RUANO, F.J.; DION-CÔTÉ, A.M. et al. Programmed DNA elimination of germline development genes in songbirds. Nat Commun, v.10, n.5468, p. 1-10, 2019.

LI, Y.; JING, X.A.; ALDRICH, J.C. et al. Unique sequence organization and small RNA expression of a "selfish" B chromosome. Chromosoma v.126 p.753–768, 2017.

LIN, H.; SHABBIR, A.; MOLNAR, M.; LEE, T. Stem cell regulatory function mediated by expression of a novel mouse Oct4 pseudogene. Biochem Biophys Res Commun. v.355, n.1, p.111-116, 2007.

MA, W. et al. Rye B chromosomes encode a functional Argonaute-like protein with in vitro slicer activities similar to its A chromosome paralog. New Phytologist, v.213, n. 916-928, 2017.

MA, W.; LIU, Z.; BEIER, S.; HOUBEN, A.; CARPENTIER, S. Identification of rye B chromosome-associated peptides by mass spectrometry. New Phytologist, v.230, n.6 2179-2185, 2021.

MADDEN, T. The BLAST Sequence Analysis Tool. 2002 Oct 9 [Updated 2003 Aug 13]. In: McEntyre J, Ostell J, editors. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2002-. Chapter 16. Available from: http://www.ncbi.nlm.nih.gov/books/NBK21097/ Acesso em: 06/04/2022.

MAKUNIN, A.I., et al. Genes on B chromosomes of vertebrates. Mol Cytogenet v.7, n.99, p.1-10, 2014.

MARTIS, M.M. et al. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. Proc Natl Acad Sci U S A v. 109, n. 33 p. 13343–13346, 2012.

MCALLISTER, B., WERREN, J. Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. Chromosoma v. 106, p. 243–253, 1997.

MIAO, V.P. et al. A fungal gene for antibiotic resistance on a dispensable ("B") chromosome. Science v.54, n. 5039, p. 1773-1776, 1991.

MILANI, D.; BARDELLA, V.B.; FERRETTI, A.B.S.M. et al. Satellite DNAs unveil clues about the ancestry and composition of B chromosomes in three grasshopper species. Genes v.9, n.523, p 1-15, 2018.

MOREIRA-FILHO, O.; FENOCCHIO, A. S.; PASTORI, M. C.; BERTOLLO, L. A. C. Occurrence of a metacentric macrochromosome B in different species of the genus *Astyanax* (Pisces, Characidae, Tetragonopterinae). Cytologia v.66, n. 1, p. 59–64, 2001.

MOREIRA-FILHO, O.; GALETTI, P.M. JR.; BERTOLLO, L.A. B chromosomes in the fish *Astyanax scabripinnis* (Characidae, Tetragonopterinae): an overview in natural populations. Cytogenet Genome Res. V.106, n.2-4, p.230-234, 2004.

NAVARRO-DOMÍNGUEZ, B. et al. Protein-coding genes in B chromosomes of the grasshopper *Eyprepocnemis plorans*. Scientific Reports v.7, n. 45200 p.1–12, 2017a.

NAVARRO-DOMÍNGUEZ, B.; RUIZ-RUANO, F.J.; CAMACHO, J.P.M.; CABRERO, J.; LÓPEZ-LEÓN, M.D. Transcription of a B chromosome CAP-G pseudogene does not influence normal Condensin complex genes in a grasshopper. Sci Rep v.7 n. 45200, p. 1-12, 2017b.

NING, Z. SSAHA: a fast search method for large DNA databases. Genome Res. v. 11, n. 10, p.1725-1729, 2001.

NUR, U. A supernumerary chromosome with an accumulation mechanism in the lecanoid genetic system. Chromosoma v.13, p. 249-271, 1962.

OLIVEIRA, C.; FORESTI, F.; HILSDORF, A.W.S. Genetics of Neotropical fish: from chromosomes to populations. Fish Physiol Biochem v.35, p.81–100, 2009.

OLIVER, J.L.; POSSE, F.; MARTINEZ-ZAPATER, J.M. et al. B-Chromosomes and E-1 isozyme activity in mosaic bulbs of *Scilla autumnalis* (Liliaceae). Chromosoma v.85, p.399–403, 1982.

ONO, Y.; IEMURA, S.; NOVAK, S.M.; DOI, N.; KITAMURA, F.; NATSUME, T.; GREGORIO, C.C.; SORIMACHI, H. PLEIAD/SIMC1/C5orf25, a novel autolysis regulator for a skeletal-muscle-specific calpain, CAPN3, scaffolds a CAPN3 substrate, CTBP1. J Mol Biol. v.425, n.16, p.2955-2972, 2013.

ÖSTERGREN, G. Parasitic nature of extra fragment chromosomes. Bot Notiser v.2, p 157-163, 1945.

PAN, D.; ZHANG, L. Burst of young retrogenes and independent retrogene formation in mammals. PLoS One, v.4, n.3, p.1-9 2009.

PATIL, VEENA S et al. Gene regulation by non-coding RNAs. Critical reviews in biochemistry and molecular biology, v.49, n.1, p.16-32, 2014.

PEREIRA, H. S.; DELGADO, M.; VIEGAS, W.; RATO, J. M.; BARÃO, A.; CAPERTA, A. D. Rye (*Secale cereale*) supernumerary (B) chromosomes associated with heat tolerance during early stages of male sporogenesis. Annals of Botany, v. 119, n 3, p. 325–337, 2017.

PIEHLER, A.P.; HELLUM, M.; WENZEL, J.J. et al. The human ABC transporter pseudogene family: Evidence for transcription and gene-pseudogene interference. BMC Genomics v.9, n.165, p.1-13, 2008.

PINK, R.C.; WICKS, K.; CALEY, D.P.; PUNCH, E.K.; JACOBS, L.; CARTER, D.R. Pseudogenes: Pseudo-functional or key regulators in health and disease? RNA. v.17, n.5, p.792-798, 2011.

PLOWMAN, A.; BOUGOURD, S. Selectively advantageous effects of B chromosomes on germination behaviour in *Allium schoenoprasum* L. Heredity, v.72, p.587–593, 1994.

POKORNÁ, J. M., REIFOVÁ, R. Evolution of B chromosomes: From dispensable parasitic chromosomes to essential genomic players. Frontiers in genetics v. 12, p.1-11, 2021.

RAMOS, E., CARDOSO, A.L.; BROWN, J. et al. The repetitive DNA element BncDNA, enriched in the B chromosome of the cichlid fish *Astatotilapia latifasciata*, transcribes a potentially noncoding RNA. Chromosoma v.126, p.313–323, 2017.

REJÓN, R. M., POSSE, F., OLIVER, J.L. The B chromosome system of *Scilla autumnalis* (Liliaceae): effects at the isozyme level. Chromosoma v.79, p.341–348, 1980.

RUIZ-ESTÉVEZ M, LÓPEZ-LEÓN, M. D., CABRERO, J.; CAMACHO, J.P.M. B-chromosome ribosomal DNA is functional in the grasshopper *Eyprepocnemis plorans*. PLOS ONE, v.7, n.5, p.1-7, 2012.

RUIZ-ESTÉVEZ, M.; BADISCO, L.; BROECK, J.V. et al. B chromosomes showing active ribosomal RNA genes contribute insignificant amounts of rRNA in the grasshopper *Eyprepocnemis plorans*. Mol. Genet. Genomics v. 289, n. 6, p. 1209–1216, 2014.

RUIZ-ESTÉVEZ, M.; LÓPEZ-LEÓN, M.; CABRERO, J. et al. Ribosomal DNA is active in different B chromosome variants of the grasshopper *Eyprepocnemis plorans*. Genetica v.141, n. 7-9, p. 337– 345, 2013.

RUIZ-RUANO, J.F; NAVARRO-DOMÍNGUEZ, B.; LÓPEZ-LEÓN, M. D.; CABRERO, J.; CAMACHO, J. P. M. Evolutionary success of a parasitic B chromosome rests on gene content. bioRxiv, p.1-25, 2019.

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, p.1-16, 2021.

SILVA, D.M.; RUIZ-RUANO, F.J.; UTSUNOMIA, R. et al. Long-term persistence of supernumerary B chromosomes in multiple species of *Astyanax fish*. BMC Biol 1 v.9, n.52, p.1-17, 2021.

SILVA, D.M.Z.A.; UTSUNOMIA, R.; RUIZ-RUANO, F.J. et al. High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*. Sci Rep v.7, n. 12726, p. 1-12, 2017.

TERÁN, G. E.; BENITEZ, M. F.; MIRANDE J. M. Opening the Trojan horse: Phylogeny of *Astyanax*, two new genera and resurrection of *Psalidodon* (Teleostei: Characidae), Zoological Journal of the Linnean Society, v. 190, n. p.1217–1234, 2020.

TRIFONOV, V.A. et al. Transcription of a protein-coding gene on B chromosomes of the Siberian roe deer (*Capreolus pygargus*). BMC Biol, v.11, n.90, p.1-11, 2013.

VALENTE, G.T. et al. Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. Mol Biol Evol, v.31, n. 8, p.2061–2072, 2014.

ZIEGLER, C.G.; LAMATSCH, D.K.; STEINLEIN, C. et al. The giant B chromosome of the cyprinid fish *Alburnus alburnus* harbours a retrotransposon-derived repetitive DNA sequence. Chromosome Res. v.11, n.1, p.23-35. 2003.

Capítulo 2 - Análise da origem e estrutura dos pseudogenes *sbno2-B* e *simc1-B* em espécies do gênero *Psalidodon* (Characiformes, Characidae).

Capítulo 2 – Análise da origem e estrutura dos pseudogenes *sbno2-B* e *simc1-B* em espécies do gênero *Psalidodon* (Characiformes, Characidae).

2.1 Introdução

Algumas espécies de organismos eucariotos podem apresentam junto aos cromossomos do complemento padrão, cromossomos com características parasitas denominados cromossomos B. Em geral, esses cromossomos possuem origem intraespecifica, surgindo a partir dos cromossomos A e adquirindo várias sequências de diferentes porções do genoma (Jones e Rees, 1982; Marques et al., 2013). No entanto, também podem ter origem interespecífica, a partir de processos de hibridação entre espécies próximas (McAllister e Werren, 1997). Devido ao modo como surgem, esses cromossomos são compostos principalmente por sequências redundantes do complemento A (Marques et al., 2018), principalmente por sequências repetitivas como DNA ribossomal 5S e 18S, DNA satélite e elementos transponíveis (Camacho, 2005). Entretanto diversos genes codificadores de proteína também são encontrados (Banaei-Moghaddam et al., 2013; Dalla Benetta et al., 2019). A presença de um conjunto apropriado de genes ativos nestes cromossomos parece ser um fator crucial para que estes cromossomos consigam sobreviver por gerações (Ruiz-Ruano et al., 2019).

Se tratando de peixes, a presença de cromossomos B já foi relatado em várias espécies como em *Prochilodus lineatus*, *Characidium gomesi* (Oliveira et al., 1997; Maistro et al., 2004) e em algumas espécies de ciclídeos (Poletto et al., 2010; Valente et al., 2014). No entanto, o gênero de peixes *Psalidodon* fornece uma ótima oportunidade para o estudo de cromossomos B. Esse gênero apresenta 33 espécies, dentre as quais pelo menos oito apresentam cromossomos B (Fricke et al., 2020; Moreira-Filho et al., 2004; Oliveira et al., 2009), dos quais diversas variantes com diferentes morfologias e frequências são encontradas. No entanto, a variante metacêntrica grande é a mais recorrente dentro do gênero, o que levou Salvador e Moreira-Filho (1992) a proporem que esta seria uma variante ancestral, o que foi posteriormente confirmado por Silva et al. (2021) através de análises do conteúdo gênico.

Silva et al. (2021) revelaram que a variante de cromossomo B metacêntrica grande teria sido originada em uma espécie ancestral há cerca de 4 milhões de anos atrás. Além de analisar a estrutura de cerca de 21 genes em *P. paranae* e *P. scabripinnis*, Silva et al. (2021) verificaram através de PCR quantitativa em tempo real a presença de alguns destes genes em outras variantes de cromossomo B em diferentes espécies de *Psalidodon*. Dentre eles, foi verificada a

presença dos genes *sbno2* e *simc1* em *P. fasciatus* e *P. bockmanni*, no entanto a estrutura e composição destes genes permaneceu desconhecida até então. Desta forma, tendo em vista que em *P. paranae sbno2-B* e *simc1-B* são pseudogenes retroinseridos (Capítulo 1), análises similares foram conduzidas em *P. fasciatus* e *P. bockmanni*, a fim de elucidar a origem e história evolutiva destas sequências dentro do gênero.

2.2 Materiais e Métodos

Foram coletados exemplares de duas populações de *P. fasciatus* e uma de *P. bockmanni*, (Tabela 1), com presença de cromossomos B. Todos os procedimentos de amostragem, manutenção e análise dos animais foram realizados de acordo com as regras internacionais em experimentação animal seguidas pela Universidade Estadual Paulista (CEEAA/IBB/UNESP, Número de Protocolo 1227). Devido à baixa taxa de cromossomos B na população, fez-se necessário um grande esforço amostral.

Tabela 1: Localidade, número de indivíduos analisados, número diploide (2n) dos exemplares e tipos de variantes do cromossomo B das espécies de *Psalidodon* analisadas.

Espécie	Localidade	Coordenadas	Indivíduos	2n	Tipo de variante	Nomenclatura
P. fasciatus	Córrego Alambari	22°27′6″S 49°14′25″W	39	46-47	M grande	*BfMa*
P. fasciatus	Córrego das Araras	22º22’43”S 47º25’37”W	35	48	M grande M grande	*BfMb*
P. bockmanni	Cachoeira Segredo	23°00′23"S, 48°24′11"W	94	50	M grande	*BbM*

M = metacentrico.

Os exemplares foram anestesiados antes de serem submetidos aos procedimentos de análise. As análises citogenéticas e cariotipagem dos indivíduos portadores de cromossomos B (+B) e sem cromossomos B (0B) foram realizadas com a retirada de fragmentos de tecido renal. Para extração de DNA genômico total foram utilizados fragmentos de tecido hepático, os quais foram armazenados em etanol 95%. Os exemplares utilizados na amostragem foram identificados e conservados em álcool 70% e serão depositados na coleção de peixes do Laboratório de Biologia e Genética de Peixes do Departamento de Biologia Estrutural e Funcional, UNESP- Botucatu-SP.

2.2.1 Obtenção dos cromossomos metafásicos mitóticos

Os cromossomos metafásicos mitóticos foram obtidos por meio do protocolo adaptado proposto por Foresti et al. (1981), que consiste na inibição da formação dos microtúbulos durante a divisão celular por meio da ação da colchicina, a hipotonização de células por meio de uma solução de KCl 0,075 M e fixação do material com uso da mistura de metanol e ácido acético.

2.2.2 Coloração convencional com Giemsa

Para melhor visualização dos cromossomos metafásicos mitóticos foi realizado o gotejamento do preparado celular em lâminas previamente limpas e secas, posteriormente o material foi corado com solução de Giemsa a 5% em tampão fosfato (pH 6,7) por 5 minutos.

2.2.3 Detecção da heterocromatina constitutiva (banda-C)

Afim de determinar quais indivíduos possuíam cromossomos B e qual a variante do mesmo utilizamos a técnica de bandamento-C, protocolo descrito originalmente por Sumner (1972).

2.2.4 Extração do DNA genômico

Para extração de DNA foi feito o uso de Kit NucleoSpin® Tissue Columns, Collection Tubes (2ml) em tecido muscular que posteriormente foi tratado com RNase A (20 mg/ml - Invitrogen) para remoção do RNA residual.

2.2.5 Sequenciamento de DNA e RNA por *reads* curtos

Foram enviadas para sequenciamento de alto rendimento amostras de DNA de indivíduos 0B e 1B de uma população de *P. bockmanni* e de duas populações de *P. fasciatus* contendo diferentes variantes de cromossomos B. As bibliotecas de DNA e RNA de *P. paranae* utilizadas foram obtidas por Silva et al. (2021). As informações referentes aos dados de sequenciamento estão sumarizadas na Tabela 2.

Tabela 2. Descrição dos materiais sequenciados das espécies de *Psalidodon*.

Espécie e tamanho do genoma	Origem	Tecido	2n	Fêmea	Macho	Plataforma	Bases (Gb)	Cobertura
P. fasciatus (1.7 Gb)	DNA	Fígado	46		3	DNBseq 2x150bp	40	20x
			45+BfMa		3	DNBseq 2x150bp	40	20x
	RNA	Ovário	48+BfMb	1		DNBseq 2x150bp	12	
	DNA	Fígado	48	3		DNBseq 2x150bp	40	20x
			48+BfMb	3		DNBseq 2x150bp	40	20x
P. bockmanni (1.8 Gb)	RNA	Ovário	50+BbM	1		DNBseq 2x150bp	12	
	DNA	Fígado	50	3		DNBseq 2x150bp	40	20x
			50+BbM	3		DNBseq 2x150bp	40	20x

2.2.6 Análise de cobertura em DNAseq

Para a análise de cobertura em DNAseq nas bibliotecas de *P. fasciatus* e *P. bockmanni*, primeiramente foi realizada a checagem de qualidade dos *reads* e remoção de adaptadores com Trimmomatic (Bolger et al., 2014). Posteriormente, foi realizado o mapeamento das sequências dos genes *sbno2* e *simc1* descritos em Silva et al. (2021) com uso do programa SSAHA2 (Ning et al., 2001), com pontuação mínima de alinhamento de 40 e 80% de identidade mínima. Além disso, foram utilizados *scripts* customizados para contar o número de *reads* mapeados como medida de abundância e para gerar os gráficos (https://github.com/fjruizruano/ngs-protocols/blob/master/mapping_blat_gs.py;https://github.com/fjruizruano/ngs-protocols/blob/master/coverage_graphics.py.

2.2.7 Análise de junções éxon-éxon em *reads* curtos

Para verificar a presença de junções éxon-éxon utilizando *reads* curtos, foi realizado o mapeamento dos genes *sbno2* e *simc1* em bibliotecas de gDNA utilizando o programa SSAHA2 (Ning et al., 2001), com pontuação mínima de alinhamento de 40 e 80% de identidade mínima. Posteriormente, a sequência de cada gene foi utilizada como referência para realizar o mapeamento dos *reads* selecionados no software Geneious® 7.1.3. Assim, uma análise manual permitiu verificar a presença/ausência de *reads* entre junções éxon-éxon.

2.2.8 Busca de SNPs B-específicos em espécies de *Psalidodon*

A busca de SNPs B-específicos dos pseudogenes *sbno2-B* e *simc1-B* foi realizada com *script* customizado, utilizando os mapeamentos previamente obtidos com o software SSAHA2 (https://github.com/fjruizruano/whatGene#protocol-2-snp-calling).

2.2.9 Validação da presença dos genes nos cromossomos B de *Psalidodon fasciatus* por qPCR

Foram utilizados *primers* (Tabela 3) desenhados para a variante *BpM* de *P. paranae* por Silva et al. (2021), a fim de quantificar a abundância dos genes *sbno2* e *simc1* no genoma de *P. fasciatus* em indivíduos 0B e +*BfMa*, sendo utilizados três indivíduos de cada grupo, checados previamente por citogenética. Esta análise foi realizada apenas para a variante *BfMa*, uma vez que a presença de ambos os genes em outras variantes de *Psalidodon* já foi analisada por qPCR e apresentada por Silva et al. (2021). Para obtenção da quantificação relativa (RQ) dos genes alvo foi aplicado o método $2^{-\Delta Ct}$ (Bel et al., 2011), sendo para isso utilizado o gene autossômico de cópia única *hprt1* como gene de referência. Foi assumida a presença de uma cópia do gene por cromossomo do complemento A para calcular um limiar mínimo teórico para a presença de uma cópia do gene adicional (paráloga) nos cromossomos B. Desta forma só foram considerados validados os genes que apresentaram *fold change* (média do grupo +B/média grupo 0B) > 1.5, considerando 3 cópias no genoma +B e 2 cópias no genoma 0B.

Os ensaios de qPCR foram realizados com auxílio do termociclador StepOne RealTime PCR Systems (Life Technologies). Os genes alvos e o gene de referência foram analisados simultaneamente em triplicatas técnicas e biológicas independentes. As reações foram realizadas com a obtenção do volume final de 10 μL, tendo por composição 3 ng de gDNA, 5 μL de μL Power SYBR™ Green PCR Master Mix (Thermo Fisher Scientific, USA) e 1 μl de cada *primer* a 5 μM. As condições de termociclagem seguiram o protocolo estabelecido, com com tratamentos a 95 °C por 10 minutos; 45 ciclos a 95 °C por 15 segundos e 60 °C por 1 minuto. A especificidade dos produtos amplificados foi confirmada pela análise da curva de dissociação.

Tabela 3. *Primers* utilizados nos experimentos de qPCR em DNA de *Psalidodon fasciatus.*

Genes	Sequência Foward	Sequência Reverse
hprt1	GGCCAGGGAGATCATGAAGG	TGGAGCGGTCACTATTTCGG
sbno2	AGAGACCAGCACTCTCTCCA	GATGGCCTCAAGCTGTAGGG
simc1	ACAGAAATCCTTGCTGGCGA	GGGCTATTGGGCAGACTCTC

2.3 Resultados

2.3.1 Análise cariotípica de *P. fasciatus* - Córrego Alambari

A análise cariotípica de *P. fasciatus* do rio Alambari mostrou indivíduos com números diploides de 2n=45+B, 2n=46, 2n=46+B e 2n=46+2B. Duas variantes de cromossomo B foram encontradas em oito indivíduos machos e em quatro fêmeas, sendo classificadas conforme Silva et al. (2016) e estando presentes em 100% das células dos indivíduos amostrados. A variante metacêntrica grande *BfMa* mostrou possuir um padrão peculiar de heterocromatina, com bandas presentes na região intersticial dos braços longos. A variante subtelocêntrica de tamanho médio *Bfst* apresentou uma natureza totalmente heterocromática; contudo, não foi possível obter quantidade suficiente de indivíduos para análise.

2.3.2 Análise cariotípica de *P. fasciatus* - Córrego das Araras

Na população de *P. fasciatus* do córrego das Araras foram observados indivíduos com cromossomos B metacêntricos grandes heterocromáticos (*BfMb*) (Silva et al., 2016), que ocorreram em 100% das células de três das 22 fêmeas e um entre 10 machos analisados.

2.3.3 Análise cariotípica de *P. bockmanni* - Cachoeira Segredo

Foram observados cromossomos B metacêntricos grandes heterocromáticos (*BbM*) (Silva et al., 2016) em 100% das células de sete entre 49 fêmeas e em um entre 33 machos de *P. bockmanni* analisados na população da Cachoeira Segredo.

2.3.4 Identificação de junções éxon-éxon por análise de *reads* curtos

A análise dos *reads* curtos de *P. bockmanni*, variante *BbM*, não mostrou *reads* entre todas junções éxon-éxon do gene *sbno2* na região alta cobertura dos indivíduos 1B (Figura S1), assim como para os indivíduos 0B (Figura S2). Para o gene *simc1*, entretanto a mesma análise mostrou alguns *reads* entre as junções éxon-éxon do gene na região alta cobertura das bibliotecas 1B de *P. bockmanni* (Figura S3), enquanto nos indivíduos 0B não houve *reads* nas junções (Figura S4).

Para os indivíduos *P. fasciatus* da população de Araras, variante *BfMb*, a análise de *reads* curtos mostrou *reads* entre as junções éxon-éxon presentes na região de alta cobertura nos indivíduos 1B para o gene *sbno2* (Figura S5). Para os indivíduos 0B a mesma análise não mostrou *reads* entre as junções para o mesmo gene. (Figura S6). Já para o gene *simc1* os

indivíduos 0B não apresentaram *reads* enquanto os indivíduos 1B apresentaram alguns *reads* entre as junções éxon-éxon (Figuras S7 e S8).

Por fim, para os indivíduos *P. fasciatus* da população de Bauru, variante *BfMa*, os indivíduos 1B apresentaram *reads* entre todas as junções éxon-éxon da região de alta cobertura para o gene *sbno2* (Figura S9). Já para os indivíduos 0B, não houve *reads* entre as junções (Figura S10). Ainda, a análise de *reads* curtos dos indivíduos 0B e 1B para o gene *simc1*, não demonstrou *reads* entre as junções éxon-éxon (Figuras S11 e S12).

2.3.5 Análise de cobertura dos genes *simc1* e *sbno2* em *P. fasciatus* e *P. bockmanni*

Através dos dados obtidos pelo sequenciamento do DNA de três indivíduos 0B e de três indivíduos +B das variantes *BfMa*, *BfMb* e *BbM* foram gerados gráficos de cobertura das regiões codificadoras (CDS). As análises de cobertura realizadas no presente estudo, somadas às análises já realizadas por Silva et al. (2021) nas espécies *P. paranae* e *P. scabripinnis*, mostram que os genes *sbno2* e *simc1* apresentam maiores valores de cobertura em quase todas as variantes analisadas, com exceção apenas do gene *sbno2* na variante *BbM* e do gene *simc1* na variante *BfMa* (Figuras 2 e 3). Os gráficos de cobertura mostram um alto grau de conservação em sua estrutura para ambos os genes, com cobertura irregular coincidindo em apenas uma região da CDS (Figuras 2 e 3).

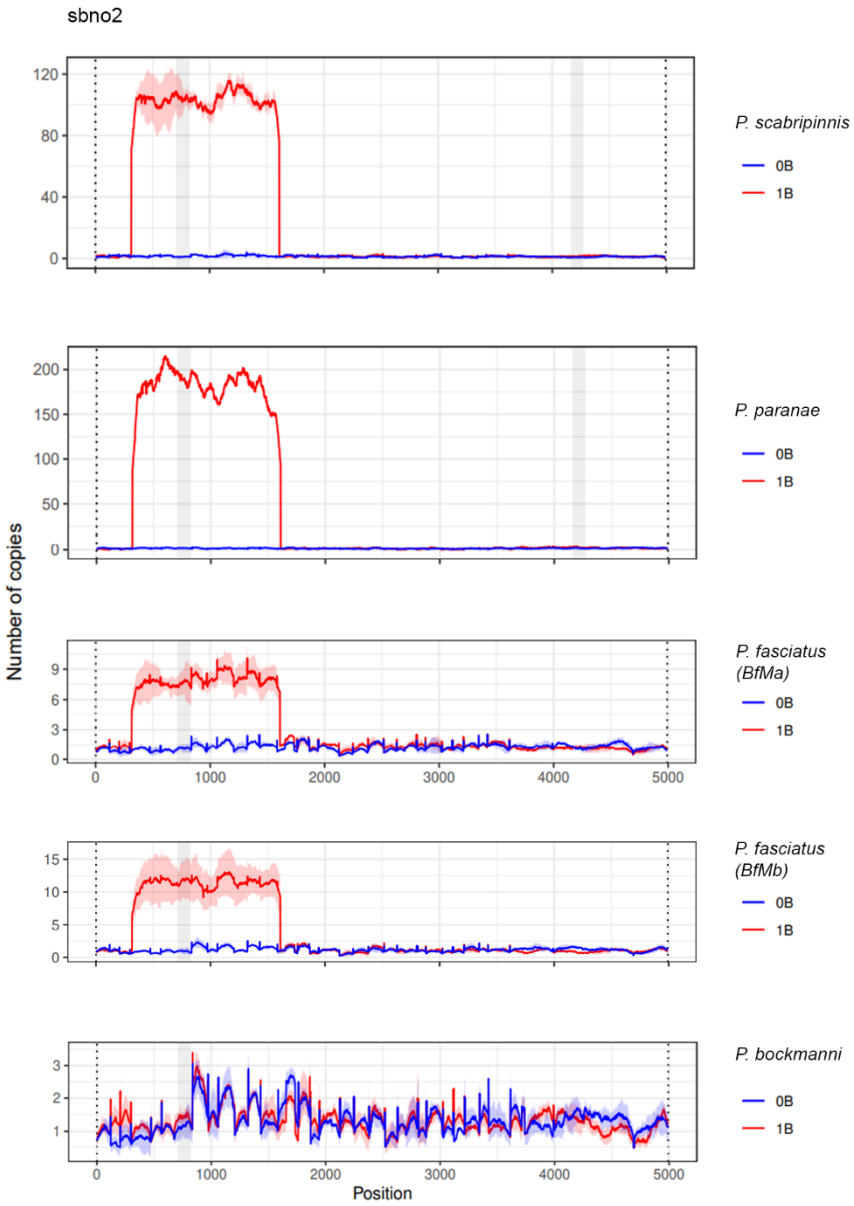


Figura 2. Gráfico de cobertura do gene *sbno2* em bibliotecas de gDNA 0B e +B em cinco variantes de cromossomo B do gênero *Psalidodon*. Os dados de *P. scabripinnis* e *P. paranae* foram obtidos em Silva et al. (2021).

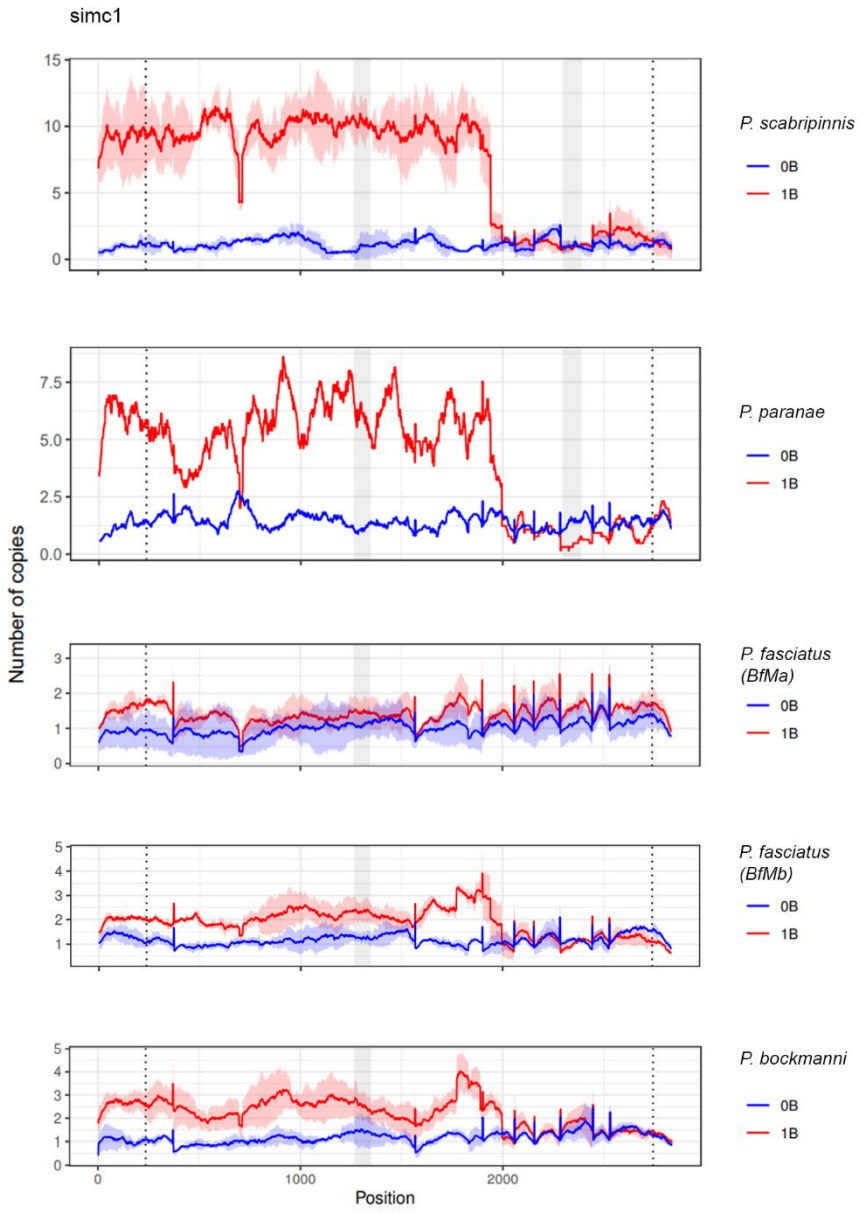


Figura 3. Gráfico de cobertura do gene *simc1* em bibliotecas de gDNA 0B e +B em cinco variantes de cromossomo B do gênero *Psalidodon*. Os dados de *P. scabripinnis* e *P. paranae* foram obtidos em Silva et al. (2021).

2.3.6 Busca de SNPs B-específicos em espécies de *Psalidodon*

A análise de SNPs revelou a existência de SNPs B-específicos com uma variante de referência (Ref), presente apenas em indivíduos 0B, e uma variante alternativa (Alt) exclusiva de indivíduos com B (Tabela 4). A presença de SNPs B-específicos em regiões de baixa cobertura pode indicar a presença de cópias dos genes *sbno2* e *simc1* com diferentes estruturas no cromossomo B.

Tabela 4. Análise de SNPs encontrados apenas em bibliotecas de gDNA com cromossomo B.

Variante de B	Gene	SNPs na região de alta cobertura	SNPs na região de baixa cobertura	Total de SNPs
BfMa	*sbno2*	4	68	72
	simc1	29	12	41
BfMb	*sbno2*	20	64	84
	simc1	134	14	148
BbM	*sbno2*	4	110	114
	simc1	18	18	36

2.3.7 Validação por qPCR dos genes *sbno2* e *simc1* na variante *BfMa* de *P. fasciatus*

Foi realizada a validação dos dados de NGS através da técnica de qPCR na variante *BfMa* para os genes *sbno2* e *simc1* em *P. fasciatus*. Devido à alta variabilidade cariotípica da população, optamos por comparar três indivíduos sem cromossomo B com 2n=46 com três indivíduos 2n=45+B. Análises cariotípicas mostraram que em todos os indivíduos 45+B o par 10 submetacêntrico encontra-se sem seu par homólogo. Dentre as amostras obtidas, esta foi a melhor configuração possível para o experimento, de forma que, caso um gene tivesse quantificação relativa maior nos indivíduos do grupo com B, ela somente poderia ser atribuída a sua presença neste cromossomo, já que este grupo apresenta um cromossomo a menos que o grupo 0B. De acordo com os resultados da qPCR, o teste de plotagem de Gardner-Altman indicou que ambos os pseudogenes estudados se encontram presentes na variante *BfMa*, uma vez que indivíduos portadores de B apresentaram valores de RQ significativamente mais altos do que aqueles 0B (Figuras 4 e 5). No entanto, consideramos validado como presente na variante *BfMa* apenas o gene *sbno2*, uma vez que este gene apresentou valor de *fold change* igual a 4,21 enquanto que o gene *simc1* apresentou valor 1,26 e não ultrapassou o limiar de 1,5 estabelecido, o que corrobora com os dados obtidos na análise de NGS.

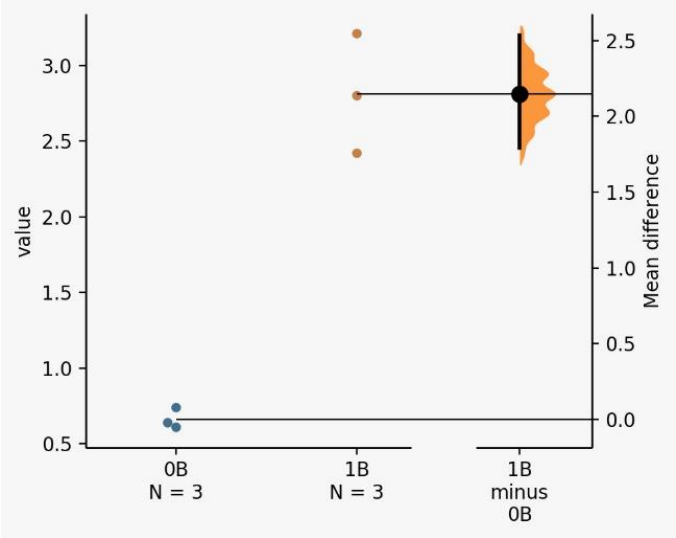


Figura 4. Gráficos de estimativa de Gardner-Altman mostrando os níveis de quantificação relativa do gene *sbno2* em indivíduos 0B e 1B de *Psalidodon fasciatus*, conforme analisado por qPCR. Ambos os grupos são plotados nos eixos esquerdos e a diferença média (tamanho do efeito) é plotada em um eixo flutuante à direita, como uma distribuição de amostragem *bootstrap*. A diferença média é representada como um ponto preto e o intervalo de confiança de 95% é indicado pelas extremidades da barra de erro vertical. NREQ = *Normalized Relative Expression Quantity*.

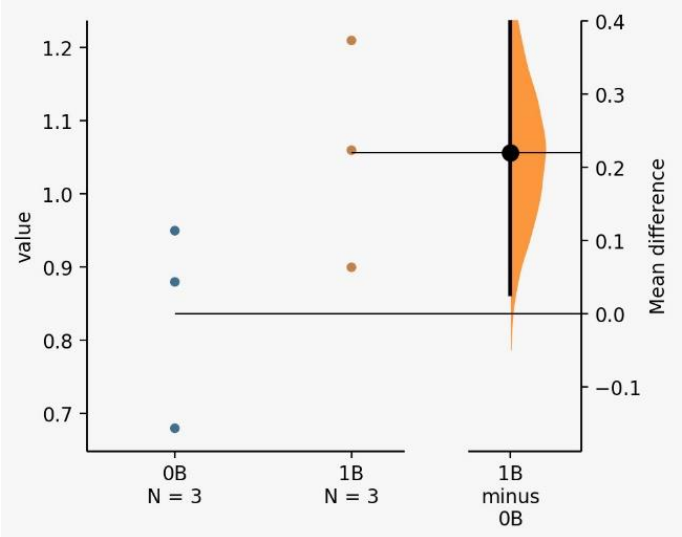


Figura 5. Gráficos de estimativa de Gardner-Altman mostrando os níveis de quantificação relativa do gene *simc1* em indivíduos 0B e 1B de *Psalidodon fasciatus*, conforme analisado por qPCR. Ambos os grupos são plotados nos eixos esquerdos e a diferença média (tamanho do efeito) é plotada em um eixo flutuante à direita, como uma distribuição de amostragem *bootstrap*. A diferença média é representada como um ponto preto e o intervalo de confiança de 95% é indicado pelas extremidades da barra de erro vertical. NREQ = *Normalized Relative Expression Quantity*.

2.4 Discussão

Cromossomos B são elementos genéticos com uma intensa dinâmica evolutiva. Apesar de terem origem a partir dos cromossomos A, diversos cromossomos B apresentam sequências que sofreram alterações ao longo de sua história evolutiva e se tornam B-específicas (Carmello, et al. 2017). De acordo com o modelo proposto por Camacho (2005), para se fixarem na população os cromossomos B precisam apresentar mecanismos de *drive*. Com o passar do tempo, estes mecanismos tendem a ser neutralizados pelo genoma A, o que pode levar ao desaparecimento dos cromossomos B ou ao surgimento de novas variantes capazes de burlar os mecanismos de neutralização. Assim, a presença de diversas variantes que compartilham uma mesma origem dentro do gênero *Psalidodon* (Silva et al., 2021a) revela uma interessante dinâmica coevolutiva entre o genoma A e o genoma B, e pode-se especular que diferentes mecanismos de acúmulo tenham evoluído entre as variantes. Assim, a análise comparativa da estrutura dos pseudogenes *sbno2-B* e *simc1-B* em espécies do gênero *Psalidodon* permitiu um melhor entendimento da origem e evolução destas sequências dentro dos cromossomos B. As análises do presente estudo revelaram que, quando presentes, os pseudogenes *sbno2-B* e *simc1-B* apresentam junções éxon-éxon em todas as variantes analisadas, o que indica que estas sequências poderiam ter sido adquiridas através de um processo de retrotransposição ocorrido em uma variante ancestral de cromossomo B do gênero *Psalidodon* (Figura 6).

A presença de sequências possivelmente pseudogenicas de *sbno2* e *simc1* é descrita por Silva et al. (2021) em quatro espécies de *Psalidodon*. Entretanto, análises bioinformáticas de cobertura foram realizadas apenas para as espécies *P. paranae* e *P. scabripinnis*. No presente trabalho, foi possível observar que de fato trata-se dos pseudogenes *sbno2-B* e *simc1-B* que, quando presentes, apresentam um padrão de cobertura altamente conservado entre as diferentes variantes de cromossomos B analisadas, indicando que provavelmente estariam presentes desde o processo de especiação ocorrido neste grupo. No entanto, os pseudogenes *sbno2-B* e *simc1-B* não estão presentes na variante *Bbm* de *P. bockmanni* e na variante *BfMa* de *P. fasciatus*, respectivamente. Assim, a hipótese mais parcimoniosa para explicar a ausência desses genes nos cromossomos B dessas espécies seria a de que estas sequências teriam sido perdidas nestas variantes após serem adquiridas (Figura 6). Isto pode ter ocorrido por processos de degeneração como a formação de DNAs satélites ou pela inserção de elementos transponíveis, assim como observado ocorrer com cromossomos sexuais (Crepaldi e Parise-Maltempi, 2020; Srikulnath et al., 2022), dentre outros. Ainda, é interessante destacar que Silva et al. (2021) encontraram

sbno2-B em uma diferente população de *P. bockmanni* analisada, no entanto com baixo número de cópias. Isto sugere que a perda do pseudogene *sbno2-B* seja um evento recente nesta espécie, e que não foi completamente fixado.

A busca por SNPs B-específicos revelou a presença de sequências alternativas presentes apenas nas bibliotecas B+ para as três variantes de cromossomo B, assim como observado para o cromossomo B das espécies *P. paranae* e *P. scabripinnis* (Silva et al., 2021). Dentre estas sequências alternativas, diversas foram encontradas em regiões de baixa cobertura dos pseudogenes *sbno2-B* e *simc1-B,* o que indica que cópias com diferentes estruturas, mesmo que em baixo número, existam no cromossomo B. Ainda, cabe destacar que não foram encontradas junções éxon-éxon em regiões de baixa cobertura para ambos os genes, assim no caso de existência de cópias em regiões de baixa cobertura, estas não teriam se originado por processo de retrotransposição. Desta forma, podemos hipotetizar dois cenários para a origem dos pseudogenes *sbno2-B* e *simc1-B* no gênero *Psalidodon*. Em um primeiro cenário, os genes *sbno2* e *simc1* teriam sido adquiridos pelo cromossomo B ancestral de *Psalidodon* em sua origem e posteriormente teriam sofrido processo de retrotransposição dentro do próprio cromossomo B. Em uma hipótese alternativa, os pseudogenes *sbno2-B* e *simc1-B* teriam sido adquiridos ainda em uma variante de cromossomo B ancestral em *Psalidodon*, mas ambos os genes teriam sofrido processo de retrotransposição a partir dos cromossomos As. Assim, acreditamos que futuras análises envolvendo análise de SNPs para construção de árvores filogenéticas, assim como em Ruiz-Ruano et al. (2019), poderão melhor elucidar a história evolutiva destes pseudogenes dentro do gênero *Psalidodon*.

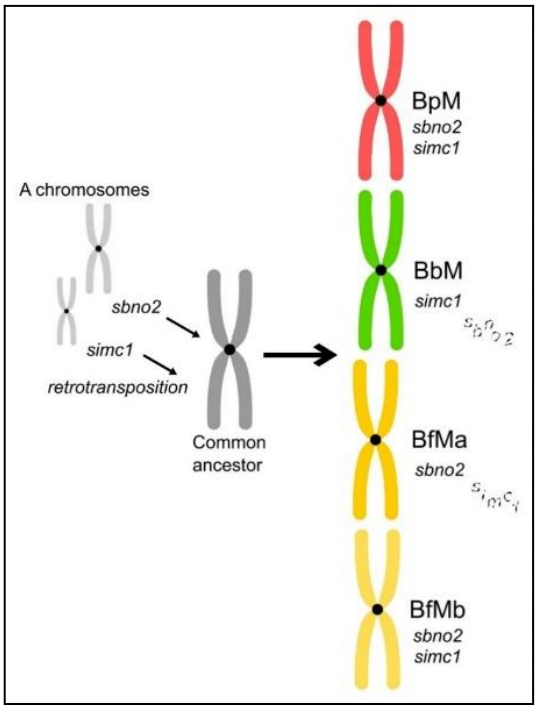


Figura 6. Esquema representando uma das hipóteses do processo de aquisição e perda de pseudogenes nos cromossomos B de *Psalidodon*. Os pseudogenes *sbno2-B* e *simc1-B* teriam sido retrotranspostos em um ancestral em comum das quatro variantes de cromossomo B analisadas. Posteriormente, a variante *BbM* de *P. bockmanni* teria perdido o pseudogene *sbno2*, enquanto que a variante *BfMa* de *P. fasciatus* teria perdido o gene *simc1*.

2.5 Conclusões

Os pseudogenes *sbno2-B* e *simc1-B* se originaram no cromossomo B de *Psalidodon* através de um processo de retrotransposição em uma espécie ancestral do gênero, o que é revelado pelo alto grau de conservação dos padrões de cobertura, e pela presença de junções éxon-éxon em diversas variantes de cromossomo B do gênero. A ausência dos pseudogenes *sbno2-B* em *P. bockmanni* e *simc1-B* em *P. fasciatus* (*BfMa*) sugere que, apesar de uma origem em comum, as diferentes de cromossomo B do gênero seguiram caminhos evolutivos distintos e sofreram diferentes processos de erosão gênica. A presença de SNPs B-específicos de ambos os genes em regiões de baixa cobertura indica que cópias com diferentes estruturas, mesmo que em baixo número, existam nos cromossomos B. Isto indica que, os genes *sbno2-B* e *simc1-B* foram adquiridos em uma variante ancestral de cromossomo B do gênero Psalidodon, e que posteriormente teriam sofrido retrotransposição para o cromossomo B. Alternativamente, ambos os pseudogenes poderiam ter sido adquiridos através da retrotransposição a partir dos cromossomos A para um cromossomo B ancestral.

2.6 Referências

BANAEI-MOGHADDAM, A.M.; MEIER, K.; KARIMI-ASHTIYANI, R.; Houben, A. American Society of Plant Biologists. Formation and expression of pseudogenes on the B chromosome of rye. The Plant Cell v. 25, n. 7, p. 2536–2544, 2013.

BEL, Y.; FERRÉ, J.; ESCRICHE, B. Quantitative real-time PCR with SYBR Green detection to assess gene duplication in insects: Study of gene dosage in Drosophila melanogaster (Diptera) and in *Ostrinia nubilalis* (Lepidoptera). BMC Res. Notes v. 4, n. 84, p.1-8, 2011.

BOLGER, A. M ET AL. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England) v.30, n.15, p.2114-20. 2014.

CAMACHO, J. P. M. The Evolution of the Genome: CHAPTER 4 - B Chromosomes. Burlington: Academic Press, 2005.

CARMELLO, B.O., et al. The hnRNP Q-like gene is retroinserted into the B chromosomes of the cichlid fish *Astatotilapia latifasciata*. Chromosome Res. v. 25, n. 3-4, p. 277-290, 2017.

CREPALDI, C.; PARISE-MALTEMPI, P. Heteromorphic Sex Chromosomes and Their DNA Content in Fish: An Insight through Satellite DNA Accumulation in Megaleporinus elongatus. Cytogenetic and genome research. v.160, n.1, p.1-9, 2020.

DALLA BENETTA, E., AKBARI, O.S., FERREE, P.M., Sequence expression of supernumerary B chromosomes: Function or fluff? Genes (Basel) v.10, n.123, p.1-13, 2019.

FRICKE R., ESCHMEYER W.N., VAN DER LAAN, R. (2021). Eschmeyer's Catalog of fishes: Genera, species, references. Disponível em: https://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.aspAcesso em: 06 Jan 2022.

JONES, R. N., REES, H. B-Chromosomes. Londres: Academic Press, 1982.

MAISTRO, E. L., et al. Cytogenetic analysis of A-, B-chromosomes and ZZ/ZW sex chromosomes of *Characidium gomesi* (Teleostei, Characiformes, Crenuchidae). Cytologia v.69, n.2, p.181–186, 2004.

MARQUES, A. et al. Evolution of plant B chromosome enriched sequences. Genes v.9, n.515. p. 1-15, 2018.

MARQUES, A., et al., B chromosomes of rye are highly conserved and accompanied the development of early agriculture, Annals of Botany, v.112, n.3, p. 527–534, 2013.

MATTHEWS, R., JONES, R. Dynamics of the B-chromosome polymorphism in rye. I. simulated populations. Heredity v.48, p.345–369, 1982.

MCALLISTER, B., WERREN, J. Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. Chromosoma v. 106, p. 243–253, 1997.

NING, Z. SSAHA: a fast search method for large DNA databases. Genome Res. v. 11, n. 10, p.1725-1729, 2001.

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.

POLETTO, A. B.; FERREIRA, I. A.; MARTINS, C. The B chromosomes of the African cichlid fish *Haplochromis obliquidens* harbor 18S rRNA gene copies. BMC Genetics. London: Biomed Central Ltd., v. 11, p. 8, 2010.

SALVADOR, L., MOREIRA-FILHO, O. B chromosomes in *Astyanax scabripinnis* (Pisces, Characidae). Heredity v.69, p.50–56, 1992.

SILVA, D. M. Z. A. et al. Origin of B chromosomes in the genus *Astyanax* (Characiformes, Characidae) and the limits of chromosome painting. Molecular genetics and genomics, MGG v. 291, n.3, 2016.

SILVA, D.M.; RUIZ-RUANO, F.J.; UTSUNOMIA, R. et al. Long-term persistence of supernumerary B chromosomes in multiple species of *Astyanax* fish. BMC Biol 1 v.9, n.52, p.1-17, 2021.

SRIKULNATH, K.; AHMAD, S.F.; SINGCHAT, W.; PANTHUM, T. Do Ty3/Gypsy transposable elements play preferential roles in sex chromosome differentiation? Life. v.12, n.522, p.1-19 2022.

3 Considerações finais

- Pseudogenes retrotranspostos também fazem parte da constituição de cromossomos B de peixes do gênero *Psalidodon*;
- *sbno2-B* e *simc1-B* são pseudogenes retrotranspostos que se originaram em um cromossomo B ancestral no gênero *Psalidodon*;
- Os pseudogenes *sbno2-B* e *simc1-B* podem estar envolvidos com alterações em mecanismos regulatórios, como demonstrado pela subexpressão de ambos os elementos genômicos em *P. paranae*
- O alto grau de conservação de ambos os pseudogenes analisados em diferentes variantes de cromossomos B encontradas nas espécies de peixes do gênero *Psalidodon* sugere que estas sequências já desempenharam ou ainda desempenham papel importante para o sucesso evolutivo destes cromossomos nesse grupo de organismos.

Material suplementar

Tabela S1. Análise de *reads* PacBio do gene *sbno2* para indivíduos de *Psalidodon paranae* 1B. Notar presença de elementos transponíveis na vizinhança, homologia e classe dos elementos.

E.T – *sbno2*	Homologia	Classe
Mariner-3_SSa	1	DNA/Mariner
hAT-N34_XL	0,9667	DNA/hAT
Mariner-3_SSa	0,9667	DNA/Mariner
hAT-N34_XL	0,9667	DNA/hAT
Mariner-3_SSa	0,9655	DNA/Mariner
CR1-1_OL	0,9412	NonLTR/CR1
Gypsy-212C_ZMLTR	0,9333	LTR/Gypsy
LTR6_BT	0,931	ERV/ERV1
Gypsy-8-I_HM	0,9143	LTR/Gypsy
Daphne-28_HRo	0,9143	NonLTR/Daphne
EnSpm-N6B_HM	0,9062	DNA/EnSpm/CACTA
BEL-39_CGi-I	0,8947	LTR/BEL
hAT-10_HM	0,8889	DNA/hAT
RTEX-8_RPr	0,8857	NonLTR/RTEX
Sola1-1_PPa	0,8857	DNA/Sola/Sola1
Sola1- 7_AP	0,8837	DNA/Sola/Sola1
Penelope5_HR	0,881	NonLTR/Penelope/Poseidon
LINE1-11B_ZM	0,8788	NonLTR/L1
EnSpm9_RIr	0,8788	DNA/EnSpm/CACTA
hAT-18_SSa	0,8784	DNA/hAT
EnSpm-5_VV	0,878	DNA/EnSpm/CACTA
DNA36a_AFC	0,875	DNA/hAT
MuDR-21_Cas	0,875	DNA/MuDR
Gypsy-17- I_NVi	0,875	LTR/Gypsy
MuDR21_Cas	0,875	DNA/MuDR
DNA36a_AFC	0,8734	DNA/hAT
hAT-N48B_ALy	0,8718	DNA/hAT
P-1_EuTe	0,8718	DNA/P
HelitronN115_OS	0,8667	DNA/Helitron
DNA-8-1_DR	0,8667	DNA/hAT
ARNOLD1	0,8649	DNA

DNA7_AFC	0,8627	DNA
Gypsy-1_DWil-I	0,8611	LTR/Gypsy
hAT-N24_CCa	0,8588	DNA/hAT
EnSpm-3_ALy	0,8571	DNA/EnSpm/CACTA
DIRS-27_XT	0,8571	LTR/DIRS
Zisupton8_DR	0,8537	DNA/Zisupton
SHALINE16_MT	0,8529	NonLTR/L1
Gypsy-18_RO-I	0,8511	LTR/Gypsy
ERV1- 13_AmRa-I	0,85	ERV/ERV1
Merlin-2_HMa	0,85	DNA/Merlin
DNA36a_AFC	0,8481	DNA/hAT
Merlin-5_HMa	0,8476	DNA/Merlin
2 CR1- 37_HM	0,8462	NonLTR/CR1
Crack-8_SP	0,8444	NonLTR/Crack
HARBINGER1_DR	0,8421	DNA/Harbinger
piggyBac-3_LVa	0,8421	DNA/piggyBac
Harbinger2_AM	0,8421	DNA/Harbinger
Copia-10_PR-I	0,8378	LTR/Copia
DNA6-9_DR	0,8376	DNA/hAT
LTR4A2_Tbel	0,8367	ERV/ERV1
BRODYAGA1	0,8333	DNA/MuDR
hAT-N86C_OS	0,8333	DNA/hAT
Caulimovirus6_ArHy	0,8333	IntegratedVirus/Caulimovirus
Copia143B_MT-I	0,8302	LTR/Copia
Polinton-3_DSuz	0,8293	DNA/Polinton
hAT-18_SSa	0,8269	DNA/hAT
hAT-1_EI	0,825	DNA/hAT
Gypsy47_ALy-I	0,825	LTR/Gypsy
HelitronN103_OS	0,825	DNA/Helitron
hAT-N24_EL	0,8235	DNA/hAT
TE-X-1_DR	0,8226	Interspersed_Repeat
Dada-5S_NoCe	0,8222	DNA/Dada
SINE2-1B_CCa	0,8205	NonLTR/SINE/SINE2
L1-1C_Cpo	0,8205	NonLTR/L1
DNA7_AFC	0,82	DNA
LTR2_AMe	0,8182	ERV/ERV1
DIRS-68_CGi	0,8163	LTR/DIRS

Gypsy18_SiCu-I	0,8148	LTR/Gypsy
PN7_LSal	0,8148	DNA
Dada-tA_LaCr	0,8138	DNA/Dada
Transib-1_DMo	0,8125	DNA/Transib
hAT-N69_CPB	0,8125	DNA/hAT
hAT-2_IlMP	0,8108	DNA/hAT
hAT-N24_CCa	0,8105	DNA/hAT
ERV1-1_SSc-I	0,8085	ERV1-1_SSc-I
SCL288_Ale	0,8077	LTR/Copia
hAT-N8b_XL	0,8077	DNA/hAT
Chapaev-20_HM	0,8077	DNA/EnSpm/CACTA
ISL2EU-7_CVi	0,8049	DNA/ISL2EU
L2B-3_CQ	0,8043	NonLTR/L2B
MuDR-15_GMa	0,8043	DNA/MuDR
L2-3_GA	0,8039	NonLTR/L2
Gypsy-44_SB-I	0,8033	LTR/Gypsy
DNA6-1_SP	0,8028	DNA
hAT-34_DR	0,8	DNA/hAT
TEC2B	0,8	DNA/Mariner
Crack-9_SP	0,8	NonLTR/Crack
ERV1-6_UCy-I	0,8	ERV/ERV1
EnSpm-6_CGi	0,8	DNA/EnSpm/CACTA
MuDR-1_DF	0,8	DNA/MuDR
TE-X-1_DR	0,8	interspersed_Repeat

Tabela S2. Análise de *reads* PacBio do gene *simc1* para indivíduos de *Psalidodon paranae* 1B. Notar presença de elementos trasponiveis na vizinhança, homologia e classe dos elementos.

E.T – simc1	Homologia	Classe
Tc1-9_Xt	0,8.217	DNA/Mariner
Athema-JY_As	0,807	NonLTR/Penelope/Athena
Copia-2_PeEX-i	0,8095	LTR/Copia
CR1_Ele38B_AAe	0,8070	NonLTR/CR1
Dada-tA_DaKy	0,9138	NDA/Dada

DMA-2-8_DR	0,8140	DNA/EnSpm/CACTA
DNA2-2_CQ	0,8197	DNA
DNA6-9_CGi	0,8.889	DNA/IS3EU
DNAV-1f_LVa	0,9048	IntegratedVirus/DNAV
EnSpm-14_DR	0,8125	DNA/EnSpm/CACTA
EnSpm-4 ArHy	0,8788	DNA/EnSpm/CACTA
Gypsy-13_DeGi-I	0,8333	LTR/Gypsy
Gypsy-15_SiMe-I	0,8750	LTR/Gypsy
Gypsy-1B_XyLGra-I	0,8864	LTR/Gypsy
Gypsy-7HyPe	0,8298	LTR/Gypsy
Harbinger-3_VV	0,8026	DNA/Habinger
Helitron-1_DR	0,8000	DNA/Helitron
Helitron-N72_OS	0,8611	DNA/Helitron
I-1_NVi	0,8095	NonLTR/I
L1-105B_XT	0,8000	NonLTR/L1
Mariner-15_EL	0,8165	DNA/Mariner
Mariner-2_DR	0,9286	DNA/Mariner
MuDR-2_DaCa	0,8718	DNA/MuDR
MuDR-20_ALy	0,8039	DNA/MuDR
Penelope-4_HRo	0,8780	Non/LTR/Penelope/Poseidon
REX1-1_GA	0,8400	NonLTR/Rex
Tc1-4_AFC	0,8136	DNA/Mariner
Tc1-9_Xt	0,8182	DNA/Mariner
TZF28B	0,8250	DNA/Mariner
Zator-3_RIr	0,8571	DNA/Zator

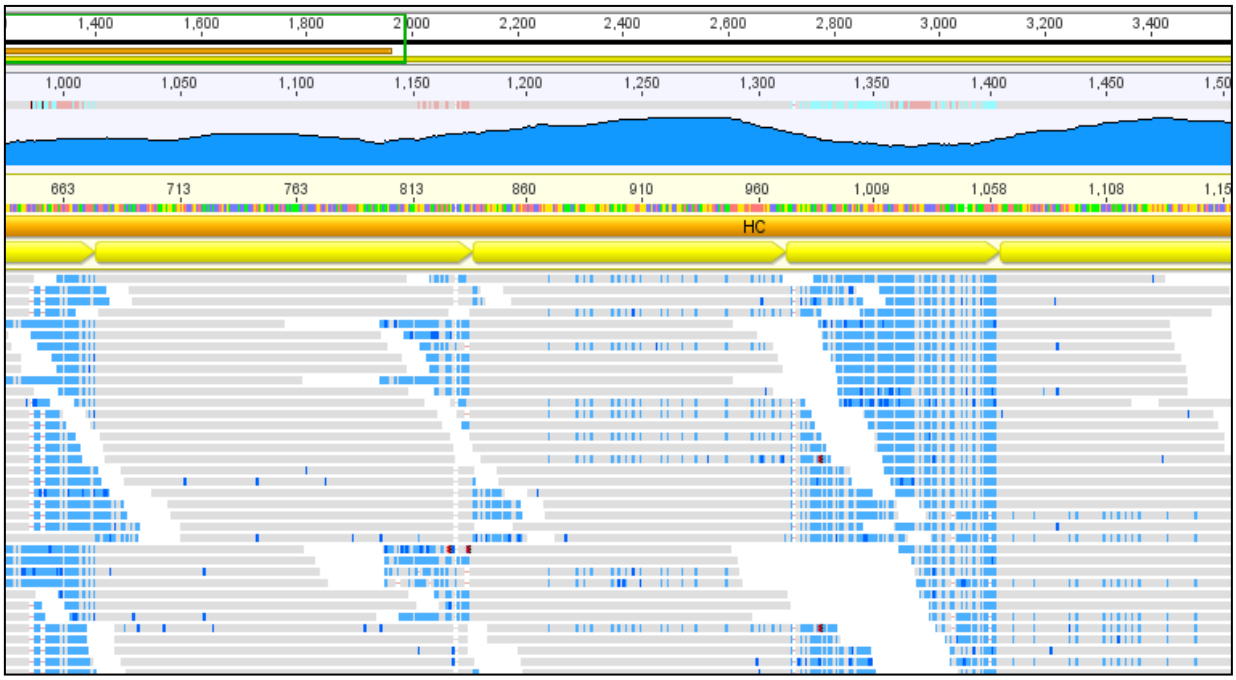


Figura S1. Análise de *reads* curtos de *Psalidodon bockmanni* de indivíduos 1B do gene *sbno2*. Notar ausência de leitura de entre as junções éxon-éxon (junção das anotações em amarelo) do gene *sbno2* na região alta cobertura anotada em laranja.

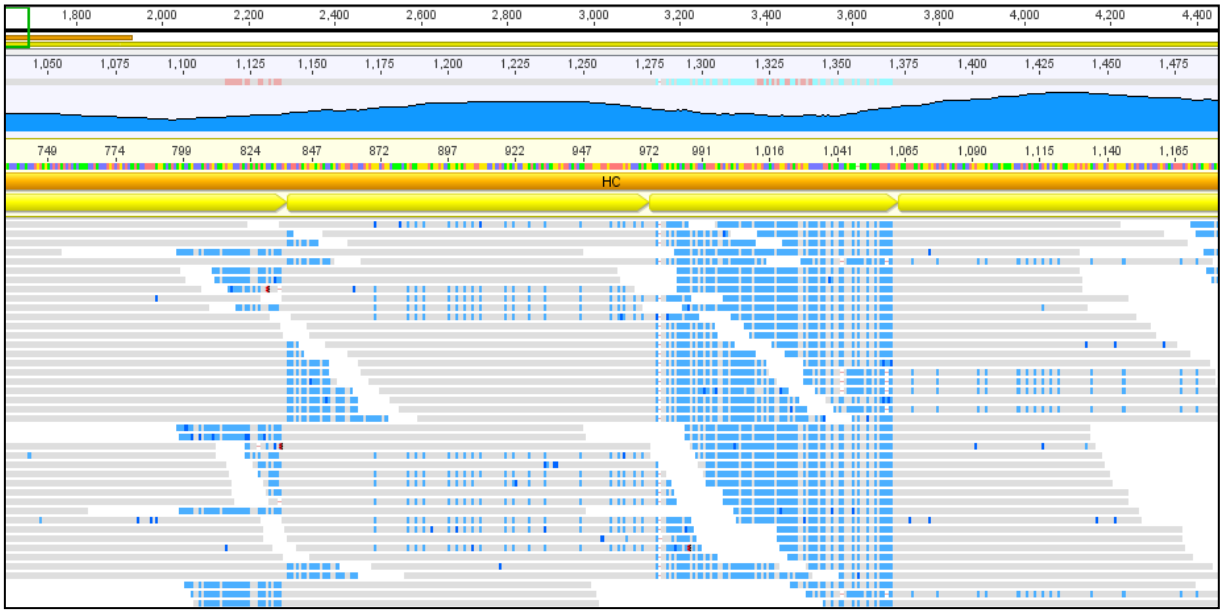


Figura S2. Análise de *reads* curtos de *Psalidodon bockmanni* de indivíduos 0B do gene *sbno2*. Notar ausência leitura entre todas junções éxon-éxon (junção das anotações em amarelo) do gene *sbno2* na região alta cobertura anotada em laranja.

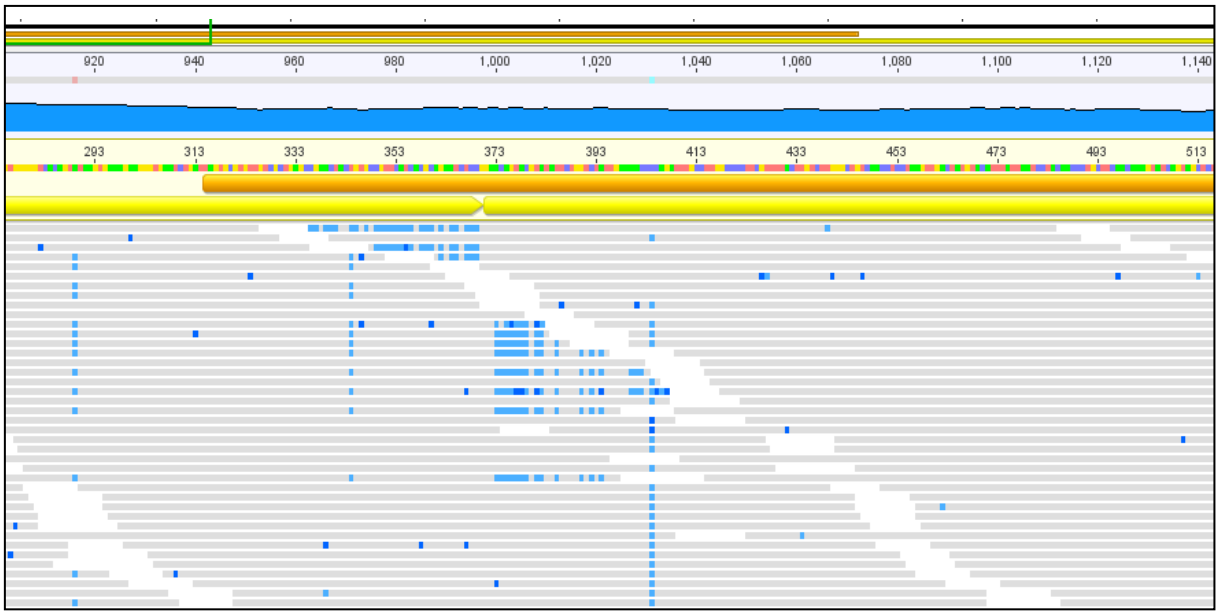


Figura S3. Análise de *reads* curtos de *Psalidodon bockmanni* de indivíduos 1B do gene *simc1*. Notar leitura entre as junções éxon-éxon (junção das anotações em amarelo) do gene *simc1* na região alta cobertura anotada em laranja.

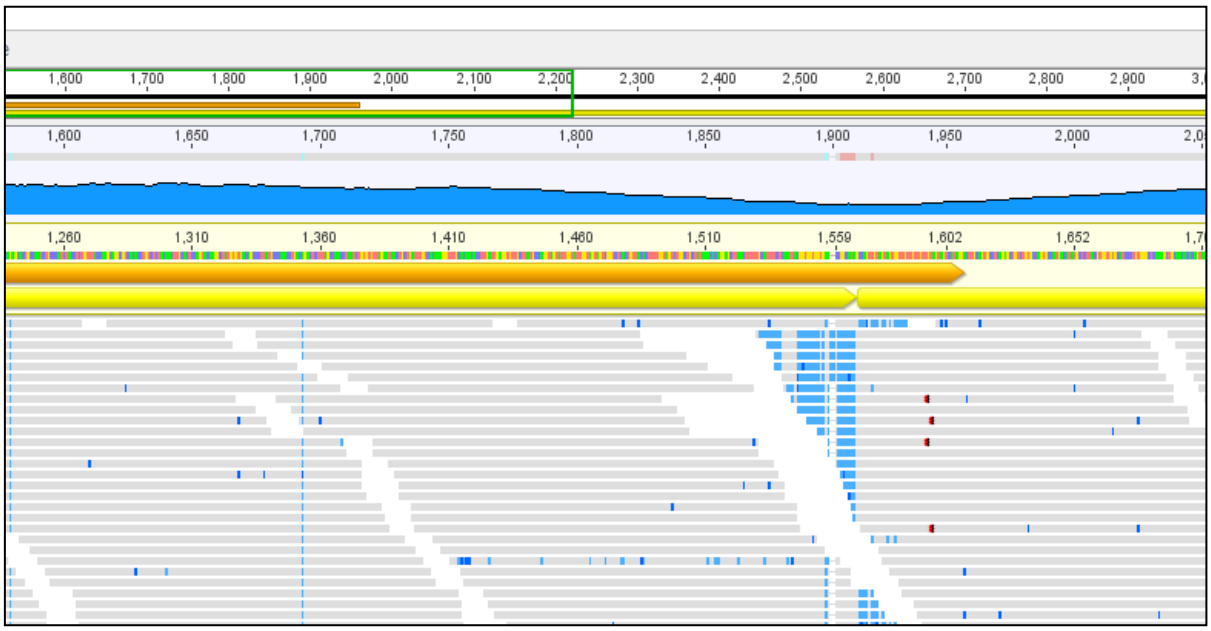


Figura S4. Análise de *reads* curtos de *Psalidodon bockmanni* de indivíduos 0B do gene *simc1*. Notar ausência de leitura entre as junções éxon-éxon (junção das anotações em amarelo) do gene *simc1* na região alta cobertura anotada em laranja.

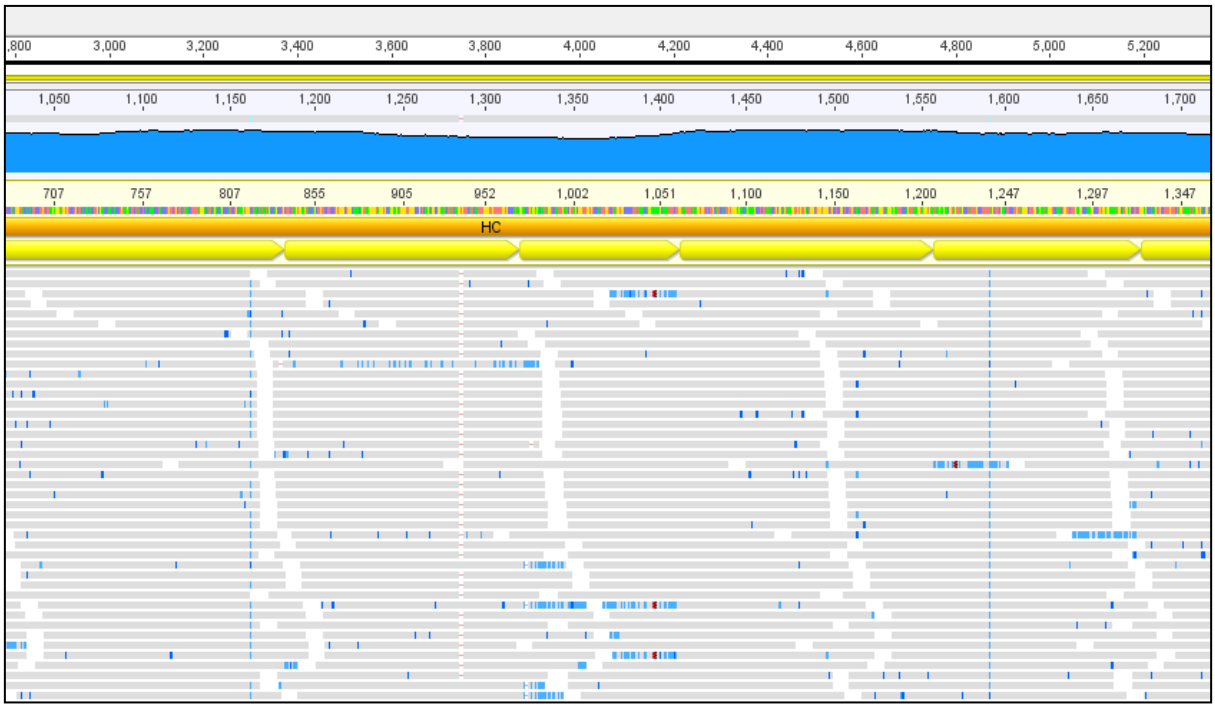


Figura S5. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMb* de indivíduos 1B do gene *sbno2*. Notar leitura entre todas junções éxon-éxon (junção das anotações em amarelo) do gene *sbno2* na região alta cobertura anotada em laranja

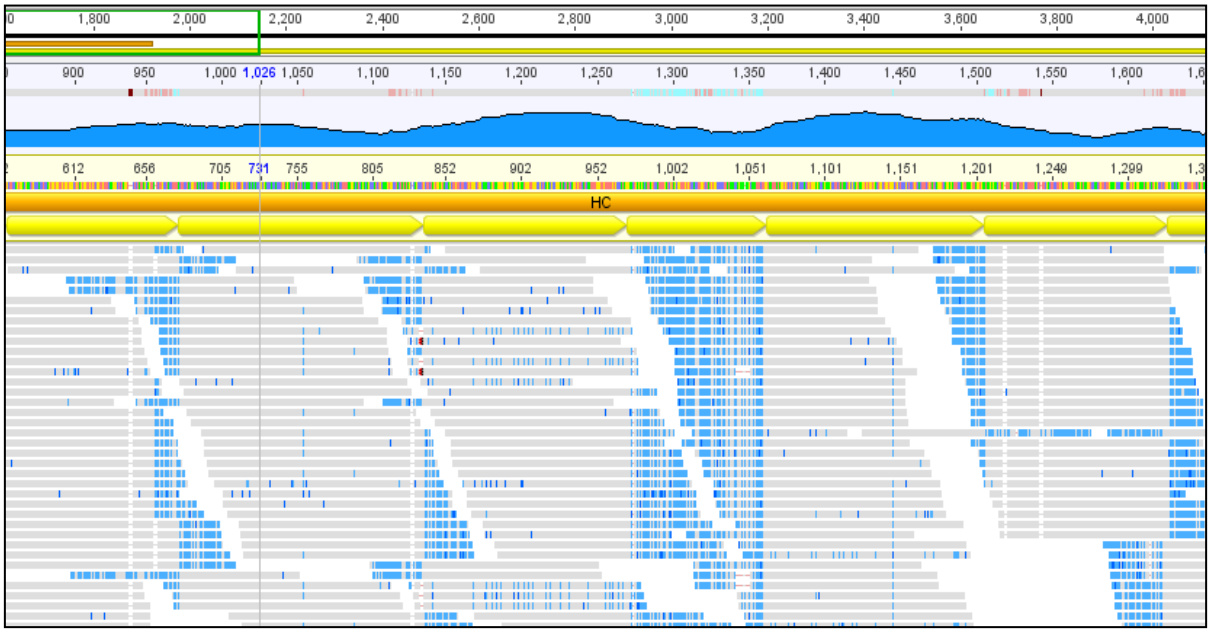


Figura S6. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMb* de indivíduos 0B do gene *sbno2*. Notar ausência de leitura entre todas junções éxon-éxon (junção das anotações em amarelo) do gene *sbno2* na região alta cobertura anotada em laranja.

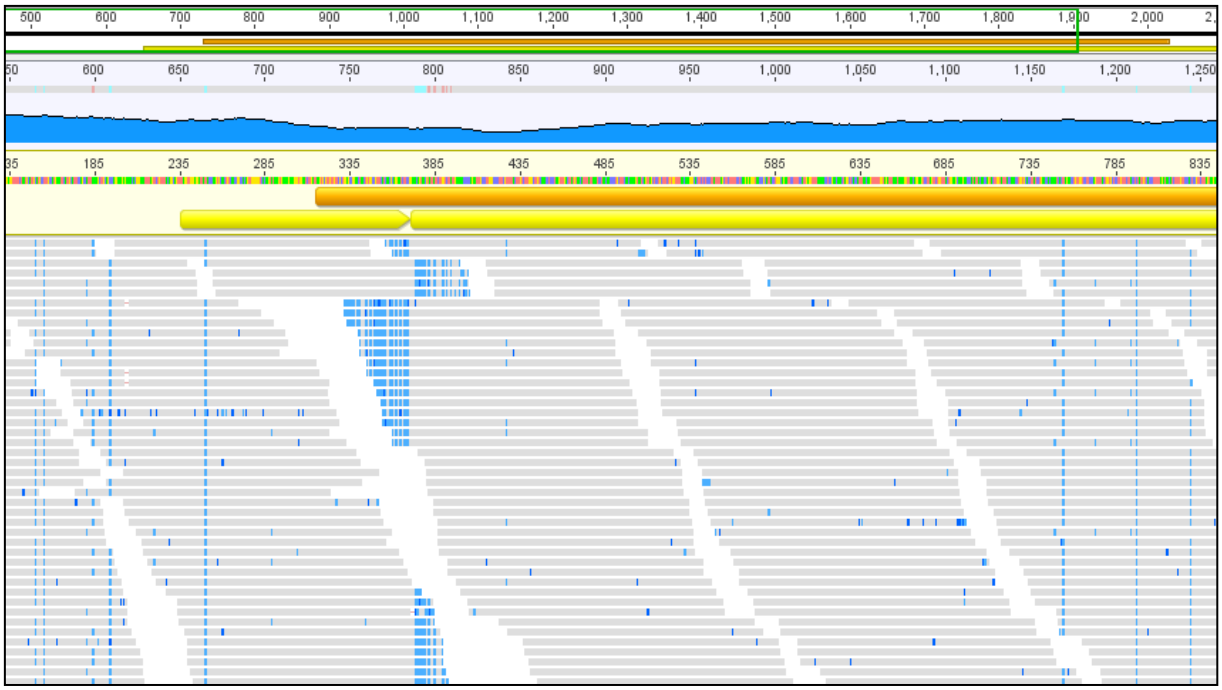


Figura S7. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMb* de indivíduos 0B do gene *simc1*. Notar ausência de leitura entre a junção éxon-éxon (junção das anotações em amarelo) do gene *simc1* na região alta cobertura anotada em laranja.



Figura S8. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMb* de indivíduos 1B do gene *simc1*. Notar leitura entre a junção éxon-éxon (junção das anotações em amarelo) do gene *simc1* na região alta cobertura anotada em laranja.



Figura S9. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMa* de indivíduos 1B do gene *sbno2*. Notar leitura entre todas junções éxon-éxon (junção das anotações em amarelo) do gene *sbno2* na região alta cobertura anotada em laranja.

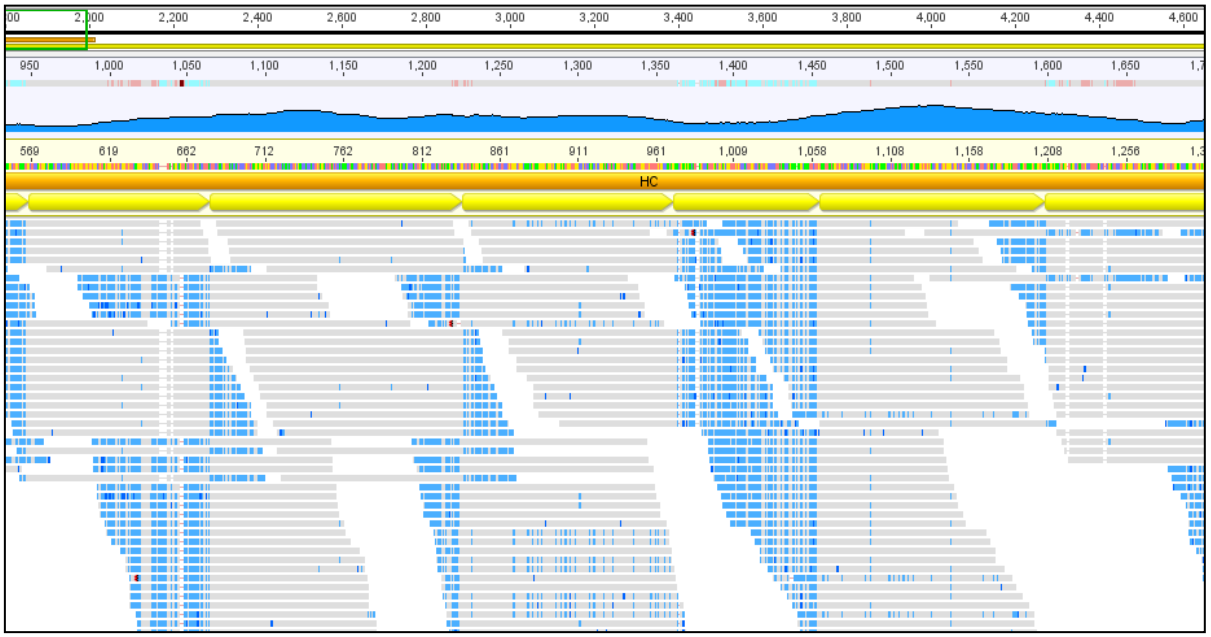


Figura S10. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMa* de indivíduos 0B do gene *sbno2*. Notar ausência de entre as junções éxon-éxon (junção das anotações em amarelo) do gene *sbno2* na região alta cobertura anotada em laranja.

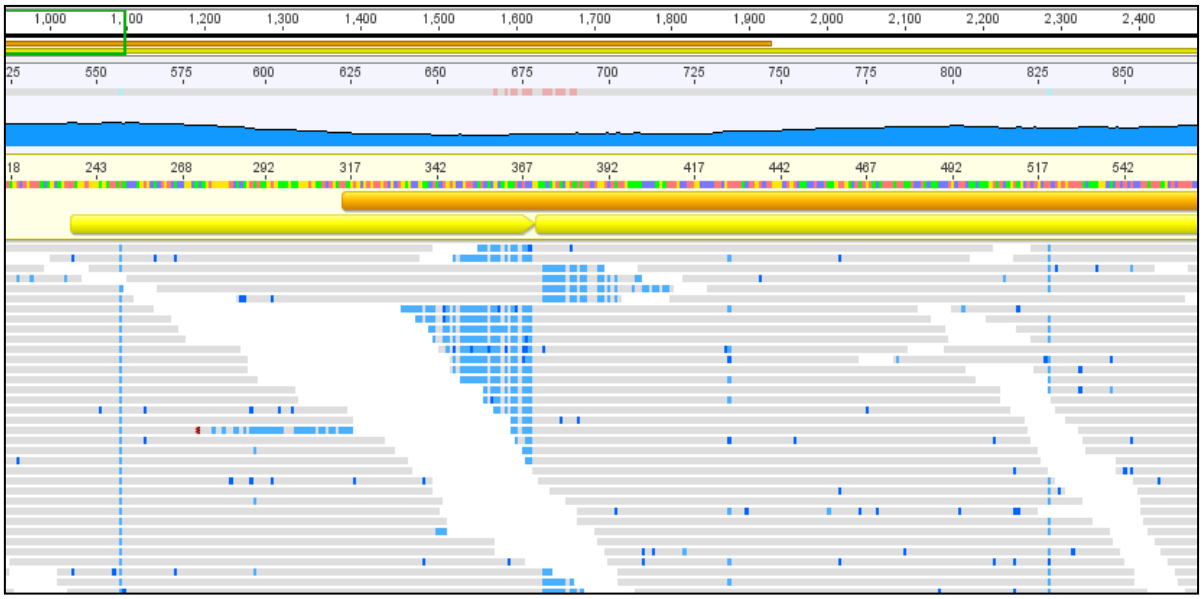


Figura S11. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMa* de indivíduos 0B do gene *simc1*. Notar ausência de leitura entre as junções éxon-éxon (junção das anotações em amarelo) do gene *simc1* na região alta cobertura anotada em laranja.

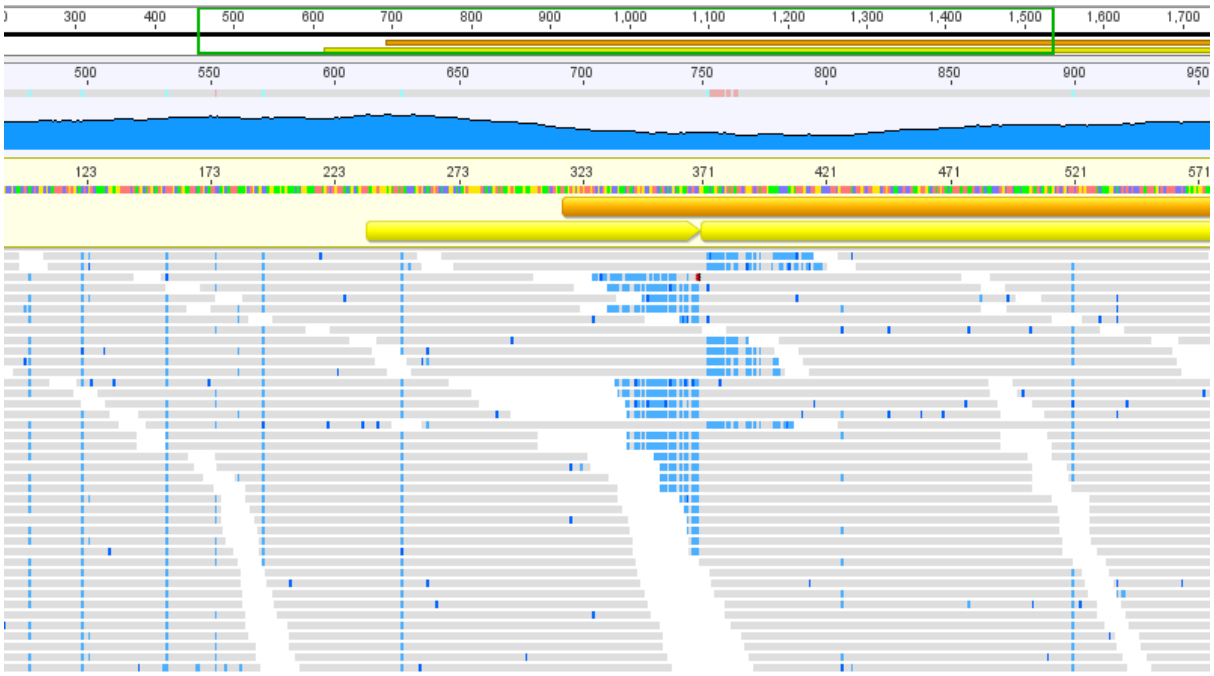


Figura S12. Análise de *reads* curtos de *Psalidodon fasciatus variante BfMa* de indivíduos 1B do gene *simc1*. Notar ausência de leitura entre as junções éxon-éxon (junção das anotações em amarelo) do gene *simc1* na região alta cobertura anotada em laranja.