



UNIVERSIDADE ESTADUAL PAULISTA
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU
Departamento de Morfologia

Laboratório de Biologia e Genética de Peixes

**MICRODISSECÇÃO E SEQUENCIAMENTO COMPLETO DE
CROMOSSOMOS B EM PEIXES APLICADOS EM ESTUDOS
EVOLUTIVOS E FUNCIONAIS**

RICARDO UTSUNOMIA

**Botucatu/SP
Fevereiro/2016**



UNIVERSIDADE ESTADUAL PAULISTA
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU
Departamento de Morfologia

Laboratório de Biologia e Genética de Peixes

RICARDO UTSUNOMIA

**MICRODISSECÇÃO E SEQUENCIAMENTO COMPLETO DE
CROMOSSOMOS B EM PEIXES APLICADOS EM ESTUDOS
EVOLUTIVOS E FUNCIONAIS**

Tese apresentada ao Programa de Pós-graduação em Ciências Biológicas (Genética) – Instituto de Biociências, UNESP, *campus* de Botucatu, como parte dos requisitos para obtenção do título de Doutor.

Orientador: Prof. Dr. Fausto Foresti

Co-orientador: Dr. José Carlos Pansonato Alves

**Botucatu/SP
Fevereiro/2016**

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

Utsunomia, Ricardo.

Microdissecção e sequenciamento completo de cromossomos B em peixes aplicados em estudos evolutivos e funcionais / Ricardo Utsunomia. - Botucatu, 2016

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

Orientador: Fausto Foresti

Coorientador: José Carlos Pansonato Alves

Capes: 20204000

1. Citogenética. 2. Peixe - Evolução. 3. Peixe - Populações. 4. Characideo. 5. Microdissecção. 6. Cromossomos B.

Palavras-chave: Characidae; Citogenética; DNA repetitivo; Moenkhausia; Sequenciamento de nova geração.

AGRADECIMENTOS

Expresso aqui meus agradecimentos aos que, de alguma forma, contribuíram e auxiliaram na realização deste trabalho

Ao professor Dr. Fausto Foresti, pelo exemplo como pessoa e pesquisador, pelos incentivos e por abrir as portas de seu Laboratório desde o nosso primeiro encontro em 2008.

Ao professor Dr. Cláudio de Oliveira pela agradável convivência diária e exemplo de pesquisador e dedicação ao trabalho.

Ao professor Dr. Juan Pedro Martinez Camacho, pelos valiosos ensinamentos e contribuições para a realização deste trabalho, pelo exemplo como pesquisador e de amor à ciência.

Ao professor Dr. Vladimir A. Trifonov, por me receber de portas abertas em seu Laboratório e pela agradável convivência diária durante o meu período em Novosibirsk.

Ao técnico Renato Devidé, pela convivência diária e disposição em ajudar nas coletas e preparações citogenéticas desde a minha chegada ao Laboratório.

Ao amigo e co-orientador Dr. José Carlos Pansonato Alves por todos os ensinamentos, paciência, companheirismo e dedicação. Muito do que sei e me tornei devo a você.

Aos colegas Francisco J. Ruiz-Ruano e Alexey Makunin pelo esforço despendido nas diferentes análises bioinformáticas realizadas neste trabalho.

Aos amigos da ala citogenética do Laboratório pelo companheirismo diário: Cristiano, Fabilene, Patrícia Elda, Silvana e Viviani. Em especial, Priscilla e Duílio por todo o auxílio na realização deste trabalho, pela grande amizade e convivência diária.

Aos amigos Maria Lígia, Érica, André e Diogo pela amizade e por tudo o que fizeram por mim durante todos estes anos.

Aos amigos Sâmia e Caio pela grande amizade, companheirismo, risadas e ensinamentos sobre a vida.

À todos os alunos e ex-alunos integrantes do Laboratório de Biologia e Genética de Peixes da UNESP de Botucatu e Laboratório de Genética de Peixes da UNESP de Bauru que sempre estiveram dispostos a ajudar e ensinar.

À Universidade Estadual Paulista e ao Instituto de Biociências de Botucatu por proporcionar toda a estrutura necessária para a realização deste trabalho. Em especial a todos os funcionários do Departamento de Morfologia e da Seção da Pós-Graduação que sempre nos atendem prontamente.

À FAPESP pelo apoio financeiro.

À toda minha família, em especial os meus pais Neide e Mitsuo e irmã Karen, que sempre me apoiaram e confiaram nas minhas decisões, amo vocês!

Resumo

No presente estudo, duas populações da espécie de peixe *Moenkhausia sanctaefilomenae*, uma coletada no rio Batalha, Bauru-SP (bacia do rio Tietê) e outra coletada no rio Novo, Ocaçu-SP (bacia do rio Paranapanema), foram analisadas através de técnicas citogenéticas básicas e moleculares. Além disso, o sequenciamento na plataforma Illumina HiSeq2000 de ambos os genomas e de uma variante do cromossomo B microdissecado também foram realizados. Os resultados obtidos revelaram que todos os indivíduos da população do rio Batalha apresentam de 0 a 6 cromossomos B por célula, enquanto os indivíduos da população do rio Novo não mostram elementos supranumerários. Deve-se destacar que os cromossomos B encontrados nos indivíduos do rio Batalha podem ser divididos em 2 variantes, sendo uma eucromática (B₁) e altamente frequente na população, e outra heterocromática (B₂) e com uma baixa prevalência. A hibridação *in situ* fluorescente (FISH) de DNAs repetitivos evidenciou que ambas as variantes apresentam *clusters* de DNAr 18S, histona H3 e satDNAs MS3 e MS7, com diferenças apenas na abundância do DNAr 18S. A análise das sequências presentes nos cromossomos B revelam que a variante B₁ é mais antiga que a variante B₂, embora ambas pareçam ter derivado de um mesmo cromossomo. Além disso, a seleção parece estar relaxada em ambas as variantes para o gene de histona H3. Para um melhor conhecimento do conteúdo genômico do cromossomo B, uma biblioteca microdissecada da variante B₁ foi sequenciada e os dados obtidos revelam um intenso acúmulo de elementos transponíveis neste cromossomo supranumerário, além de possíveis genes de cópia única, que estão presentes em diferentes grupos de ligação no genoma de *Astyanax mexicanus*. Estes resultados serão utilizados como base para estudos evolutivos e funcionais deste sistema de cromossomo B. A integralização de abordagens cromossômicas e moleculares foi utilizada na tentativa de melhor compreender questões relacionadas à biologia evolutiva de cromossomos supranumerários em peixes Neotropicais.

Abstract

In the present study, two populations of the fish species *Moenkhausia sanctaefilomenae*, one collected at the Batalha river, Bauru-SP (Tietê river basin) and the other collected at the Novo river, Ocaçu-SP (Paranapanema river basin), were analyzed through classical and molecular cytogenetic techniques. Besides, Illumina HiSeq2000 sequencing was performed for both genomes and for one microdissected B-variant. Results obtained revealed that all the individuals from the Batalha river show from 0 to 6 B chromosomes per cell, while the specimens from Novo river did not show any supernumerary chromosomes. One must note that the B chromosomes found in the Batalha river might be classified into 2 variants, one euchromatic (B₁) and present in all the population, and another heterochromatic (B₂) with lower prevalence. Fluorescence in situ hybridization (FISH) using repetitive DNA probes revealed that both B-types showed 18S rDNA, H3 histone and MS3 and MS7 satDNAs. Interestingly, the only between-variant difference was the abundance of 18S rDNA. Sequence analysis revealed that the B₁ variant is older than B₂, although both variants seem to be derived from the same autosomal pair. Notably, selection seems to be relaxed for H3.3 histone genes located on the B chromosomes. For a better knowledge about the genomic content of B chromosomes, a microdissected library of B₁ variant was sequenced and data obtained revealed an intense accumulation of transposable elements in this supernumerary, in addition to single copy genes, that are present in different linkage groups of *Astyanax mexicanus* genome. These results will be used for future evolutionary and functional studies about B chromosomes. The combined approach of chromosomal and molecular methods was used to better understand evolutionary issues of supernumerary chromosomes in Neotropical fish.

Sumário

Introdução.....	1
1.1 Considerações sobre cromossomos B.....	2
1.2 Isolamento de cromossomos e suas aplicações em estudos citogenéticos	2
1.3 <i>Moenkhausia sanctaefilomenae</i> como modelo de estudo de cromossomos B	3
2 Objetivos.....	5
3 Materiais e Métodos	6
3.1 Materiais	6
3.1.1 Espécies estudadas e locais de coleta	6
3.2 Métodos	6
3.2.1 Estimulação de mitoses	6
3.2.2 Obtenção de cromossomos metafásicos	7
3.2.3 Identificação de heterocromatina constitutiva	8
3.2.4 Extração de DNA genômico.....	8
3.2.5 Preparação das sondas de DNA para hibridação <i>in situ</i> fluorescente (FISH)	8
3.2.5.1 Preparação das sondas de DNAr 18S e 5S	8
3.2.5.2 Preparação da sonda para snDNA U2	9
3.2.5.3 Preparação da sonda para histona H3	10
3.2.6 Microdissecção	11
3.2.7 Kit GenomePlex Single Cell Whole Genome Amplification (WGA4).....	11
3.2.8 WGA3 – Reamplificação e marcação	12
3.2.9 Hibridação <i>in situ</i> fluorescente (FISH).....	13
3.2.10 Sequenciamento Illumina	14
3.2.11 Amplificações do material microdissecado.....	14
3.2.12 Purificação dos fragmentos de DNA.....	14
3.2.13 Clonagem.....	15
3.2.13.1 Identificação dos clones de interesse.....	15
3.2.14 Sequenciamento nucleotídico (Sanger)	16
3.2.15 Análises bioinformáticas	16
3.2.15.1 RepeatExplorer	17
3.2.15.2 Análise de cromossomo microdissecado.....	17
3.2.15.3 Extração dos DNAs satélites MS3 e MS7 dos <i>reads</i> Illumina	18
3.2.15.4 Análises nucleotídicas	18
3.3 Estudos cariotípicos.....	19
3.4 Análises nucleotídicas	19
4 Resultados e Discussão	20
4.1 Capítulo 1 – Uncovering the ancestry of B chromosomes in <i>Moenkhausia sanctaefilomenae</i> (Teleostei, Characidae)	21
4.2 Capítulo 2 – Microdissecção e sequenciamento completo do cromossomo B ₁ de <i>Moenkhausia sanctaefilomenae</i> (Teleostei, Characidae)	54
5 Considerações Finais	72
6 Referências Bibliográficas	74

1. Introdução

1.1 Considerações sobre cromossomos B

Cromossomos B, também referidos como supranumerários, são elementos genômicos adicionais e dispensáveis que estão presentes em alguns indivíduos de determinadas populações de algumas espécies e, geralmente, surgiram dos cromossomos do conjunto A. No entanto, estes elementos genômicos geralmente seguem seu próprio caminho evolutivo (Camacho et al. 2000). Sua ocorrência e distribuição é conhecida para um grande conjunto de organismos, como fungos, plantas e animais (Navas-Castillo et al. 1986; Alfenito e Birchler 1993; Covert 1998). Como propriedades básicas, estes cromossomos não se recombinam com os cromossomos do conjunto padrão A e são elementos dispensáveis para o ciclo do genoma hospedeiro (Camacho et al. 2000). De forma geral, estudos realizados nestes cromossomos buscam informações a respeito da possível função destes segmentos, incluindo desde a análise da distribuição nas espécies, passando por investigações sobre sua estrutura, origem, transmissão, dinâmica populacional, evolução, até a integração destes cromossomos com os do conjunto genômico A (Camacho et al. 2000; Bugrov et al. 2007).

Após a identificação da presença de cromossomos B em determinada população de uma espécie, uma caracterização citogenética extensa destas amostras se torna importante para um melhor conhecimento sobre a distribuição, modos de herança e da dinâmica populacional destes elementos extras (Hewitt e East 1978; Pardo et al. 1995; Beukeboom et al. 1996). Adicionalmente, o conhecimento sobre a presença/ausência de cromossomos B em uma determinada população permite o desenvolvimento de marcadores moleculares que atestarão a condição dos indivíduos para esta característica, sem a necessidade de obter preparações citogenéticas, permitindo uma melhor compreensão da dinâmica evolutiva destes cromossomos e da sua distribuição (Tosta et al. 2007; Muñoz-Pajares et al. 2011) da espécie em estudo. No entanto, para maiores esclarecimentos das questões relacionadas à origem e evolução destes cromossomos, informações sobre a estrutura molecular e composição nucleotídica destes elementos são necessárias.

Alguns estudos demonstraram que em determinadas espécies os cromossomos B podem compartilhar sequências com os cromossomos A ou apresentar uma estrutura molecular distinta (Page et al. 2001; Cheng e Lin 2003; Bugrov et al. 2007). Comumente, os cromossomos B são compostos principalmente por sequências de DNA repetitivo, entre os quais o DNA ribossômico, DNA centromérico e telomérico, além de elementos móveis, que, frequentemente são encontrados nos elementos supranumerários de uma variedade elevada de organismos (Camacho 2005). No entanto, sequências codificantes também já foram

encontradas nestes cromossomos (Miao et al. 1991; Graphodatsky et al. 2005; Teruel et al. 2009a; Valente et al. 2014; Makunin et al. 2014), podendo, inclusive, estarem ativas (Ruiz-Estévez et al. 2012). Tais informações tendem a aumentar à medida que novas tecnologias de análise surgem.

Com o intuito de investigar a biologia de cromossomos B, diversos autores focaram no estudo da porção repetitiva presente em abundância nos cromossomos supranumerários. Para deduzir a origem de um cromossomo B presente em uma espécie de gafanhoto, López-León et al. (1994) utilizaram a técnica de FISH com sondas de genes ribossômicos e de um DNA satélite isolado do genoma da espécie *Eyprepocnemis plorans*. Os resultados apontaram para uma origem intraespecífica a partir do cromossomo X desta espécie, basicamente pelos padrões de hibridação similares das duas sequências referidas nos cromossomos X e B. Por outro lado, em uma espécie de vespa nomeada *Nasonia vitripennis* foi possível verificar que a origem dos cromossomos supranumerários existentes estaria, provavelmente, relacionada a um evento interespecífico (McAllister e Werren 1997), principalmente após os autores verificarem que algumas sequências repetitivas presentes nos supranumerários desta espécie eram mais relacionadas a outras espécies do que a ela mesma.

A introdução da técnica de microdissecção cromossômica, realizada pela extração seletiva de cromossomos, no início da década de 1980, permitiu uma abordagem mais direta sobre a estrutura e composição molecular dos cromossomos B. Até o momento, a utilização destes cromossomos como sonda em experimentos de pintura cromossômica é o método mais empregado em estudos evolutivos, com resultados obtidos em organismos das mais distintas ordens (McQuade et al. 1994; Sharbel et al. 1998; Trifonov et al. 2002; Long et al. 2008; Teruel et al. 2009b; Machado et al. 2011). No entanto, métodos que permitam acessar o conteúdo nucleotídico propriamente dito destes cromossomos poderiam resultar em análises evolutivas mais significativas. Nesse sentido, a utilização do material genético proveniente de cromossomos B microdissecados para a amplificação de determinadas sequências permite abordagens distintas de estudo, como a comparação contra sequências genômicas em nível intra e interespecíficas e a verificação do grau de degeneração destas sequências. Assim, inferências sobre o tipo de origem destes elementos supranumerários, bem como tempo do surgimento e suas possíveis funções podem ser realizadas, tal como demonstrado por Teruel et al. (2009a).

1.2 Isolamento de cromossomos e suas aplicações em estudos citogenéticos

A capacidade de isolar os cromossomos, seja por microdissecção ou por outros métodos (ex. citometria de fluxo) torna possível conhecer a composição nucleotídica de qualquer cromossomo sem precedentes (ex. amplificação de sequências utilizando o DNA do supranumerário como molde). Uma primeira abordagem foi demonstrada por Scalenghe et al. (1981) que, utilizando o processo de clonagem de cromossomos politênicos microdissecados, conseguiram isolar determinadas sequências presentes em diferentes cromossomos. No entanto, a aleatoriedade e dificuldade de integrar as sequências de cada um dos clones obtidos se torna um problema na caracterização destes cromossomos. Aplicando esta técnica para construir uma biblioteca do cromossomo sexual W da espécie de peixe *Cynoglossus semilaevis*, Wang et al. (2009) identificaram diversas sequências descritas em outros peixes, comprovando uma certa eficácia deste método, no entanto, tal ferramenta permitia uma cobertura de apenas 1% de todo o cromossomo W, o que acaba por limitar a aplicação desta ferramenta.

Além de providenciar uma melhor compreensão sobre a organização dos genomas, o desenvolvimento de tecnologias de sequenciamento de nova geração (NGS) possibilitou avanços no conhecimento da genômica funcional, estrutura da cromatina, classificação de espécies, descobrimento de genes, dentre outras aplicações (Metzker 2010). Nesse sentido, a utilização de plataformas NGS para o sequenciamento de cromossomos, ou regiões, microdissecadas se torna uma ferramenta de interesse para a citogenética. Em estudo pioneiro, Weise et al. (2009) demonstra a viabilidade de capturar regiões cromossômicas para NGS. Neste trabalho, os autores evidenciam que utilizando apenas seis cromossomos microdissecados e posteriormente amplificados por DOP-PCR foi possível obter material suficiente e de qualidade para sequenciamento na plataforma 454 (Roche). Nesse sentido, a combinação das técnicas de microdissecção, amplificação do material e sequenciamento cromossômico por NGS se torna uma ferramenta poderosa em estudos evolutivos e funcionais, o que irá permitir, principalmente, o aumento do conhecimento sobre cromossomos sexuais e supranumerários.

1.3 *Moenkhausia sanctaefilomenae* como modelo de estudo de cromossomos B

O gênero de peixes *Moenkhausia* se insere na família Characidae e está amplamente distribuído pela região Neotropical (Lima et al. 2003). Sua taxonomia é bastante complexa e diversos estudos evidenciam que os aspectos filogenéticos devem ser mais bem estudados (Benine et al. 2009). Estudos citogenéticos realizados em poucas espécies deste gênero evidenciaram algumas características particulares. Embora apresente uma macroestrutura

cariotípica conservada com $2n=50$ cromossomos, variações na distribuição das RONS, bem como da localização de blocos de heterocromatina constitutiva nas regiões pericentroméricas dos cromossomos já forma descritas anteriormente. No entanto, o alvo de estudo mais frequente neste grupo se centra na ocorrência de distintos cromossomos B no genoma desta espécie (Foresti et al. 1989; Portela-Castro et al. 2001; Dantas et al. 2007; Hashimoto et al. 2012; Scudeler et al. 2015).

Foresti et al. (1989) descreveram a presença de diversas variantes de cromossomos B presentes nas células de indivíduos de *M. sanctaefilomenae*. Neste estudo, os autores relatam a ocorrência de variação intra e interindividual em relação à distribuição dos elementos supranumerários (um a nove cromossomos B), o que pode estar relacionado à ocorrência de não disjunção somática destes cromossomos. Além disso, o padrão de heterocromatina constitutiva revelado após bandamento C demonstra que cada um dos cromossomos B apresenta características distintas quanto à sua composição molecular. Em estudo posterior de outra população, uma situação distinta foi detectada, sendo que apenas indivíduos machos apresentavam cromossomos supranumerários, apontando para a ocorrência de um mecanismo celular que resulta na eliminação destes cromossomos nas fêmeas (Portela-Castro et al. 2001). Por fim, em uma terceira população desta espécie, a presença de RONS ativas em alguns cromossomos supranumerários foi constatada e evidenciou uma possível função destes cromossomos (Hashimoto et al. 2012).

Embora estes estudos tenham fornecido dados sobre a distribuição e dinâmica populacional destes cromossomos B, poucas informações acerca da origem e função foram geradas, principalmente se levarmos em conta a falta de conhecimento sobre a composição destes cromossomos (Foresti et al. 1989; Hashimoto et al. 2012). Nesse sentido, a realização de um mapeamento físico com diversas sondas compostas por DNA repetitivo irá fornecer, além de um panorama geral de distribuição destas sequências no complemento A de cromossomos, também evidências físicas da possível presença destas sequências nos cromossomos B, gerando indícios sobre a origem e composição de determinados elementos supranumerários. Além disso, o sequenciamento de dois cromossomos B inteiros microdissecados, um heterocromático e um eucromático, fornecerá informações a respeito das diferenças nucleotídicas que esta condição induz e também dados sobre a presença de genes, ativos ou não, microssatélites e elementos transponíveis. Considerando os aspectos descritos acima, os dados obtidos poderão fornecer evidências sobre a origem, função e evolução destes cromossomos no genoma de *M. sanctaefilomenae*.

2. Objetivos

O presente projeto se inseriu no programa geral de estudos citogenéticos de peixes Neotropicais que vêm sendo desenvolvidos no Laboratório de Biologia e Genética de Peixes (IB/UNESP/Botucatu), visando a um melhor conhecimento da estrutura cromossômica das células eucariontes e em particular dos peixes.

Conforme destacado, os peixes do gênero *Moenkhausia* apresentam distintas variantes de cromossomos B, com a existência de elementos supranumerários eucromáticos e heterocromáticos, característica que torna esse grupo de peixes um bom modelo em estudos cromossômicos. Nesse contexto, os objetivos do presente trabalho consistiram em identificar os aspectos evolutivos e funcionais do sistema de cromossomos supranumerários em *M. sanctaefilomenae*, bem como estudar o possível mecanismo de origem destes elementos com a aplicação de técnicas citogenéticas e moleculares. Dessa forma, o presente estudo teve como objetivos específicos:

- a) Expandir o conhecimento sobre a estrutura cromossômica neste grupo de peixes, focando na presença de cromossomos B existentes, por meio de análises cromossômicas convencionais, que incluíram a caracterização cariotípica por coloração convencional e localização da heterocromatina constitutiva por bandamento-C;
- b) Ampliar o conhecimento sobre a organização e distribuição genômica de sequências repetitivas nesta espécie, com especial atenção à presença de determinadas sequências nos elementos supranumerários;
- c) Obter informações sobre a estrutura, composição, origem e evolução dos cromossomos B em *Moenkhausia sanctaefilomenae*, através dos dados gerados pelo sequenciamento total de cromossomos B;
- d) Verificar diferenças estruturais existentes entre os cromossomos B eu e heterocromáticos de *Moenkhausia sanctaefilomenae*;

3 Materiais e Métodos

3.1 Materiais

3.1.1 Espécies estudadas e locais de coleta

No presente trabalho foi realizada a análise citogenética de duas populações da espécie de peixes *Moenkhausia sanctaefilomenae* (Teleostei, Characiformes) coletadas em diferentes bacias hidrográficas brasileiras indicadas na Tabela 1.

De cada indivíduo foram retirados fragmentos de tecido renal para estudos citogenéticos e também fragmentos do fígado, que foram fixados em etanol 95% para estudos moleculares. Os exemplares utilizados na amostragem foram fixados em formol 5%, conservados em álcool 70% e após identificação foram depositados na coleção de Peixes do Laboratório de Biologia e Genética de Peixes do Departamento de Morfologia, UNESP de Botucatu-SP.

Tabela 1. Exemplares de peixes da espécie *M. sanctaefilomenae* analisados. Destaque para presença ausência de cromossomos supranumerários.

Espécie	LBP	Localidade	Cromossomos B	Coordenadas
<i>M. sanctaefilomenae</i>	19830	Rio Batalha	Presentes 0-6	22°24'23.65"S, 49°05'51.38"W
<i>M. sanctaefilomenae</i>	19831	Rio Novo	Ausentes	22°28'13.32"S, 49°55'26.17"W

3.2 Métodos

Para a análise citogenética convencional foram utilizados os métodos de estimulação de mitoses (Cole e Leavens 1971), preparação direta de células renais (Foresti et al. 1981) e identificação da heterocromatina constitutiva (Sumner 1972).

Para a análise citogenética molecular foi utilizada a técnica de FISH (Pinkel et al. 1986) para a localização cromossômica das sequências de DNAr 5S e 18S, snDNA U2, histona H3, satDNAs MS3 e MS7 e para a pintura dos cromossomos B₁ e B₂.

3.2.1 Estimulação de Mitoses

Para obtenção de maior número de células em mitose foi utilizada a técnica de estimulação de divisão celular por meio da inoculação intramuscular de suspensão de células

de fermento biológico, conforme técnica descrita por Cole e Leavans (1971) para anfíbios e répteis e adaptada para peixes por Oliveira et al. (1988).

O procedimento utilizado consiste em:

- 1) preparar uma solução de fermento biológico (Fleischmann) na seguinte proporção: 0,5 g de fermento, 0,5 g de açúcar e 7 ml de água destilada;
- 2) incubar a solução em uma estufa (37°C) por cerca de 10 minutos;
- 3) injetar a solução dorso-lateralmente no peixe na proporção de 1 ml por 100 g de peso do animal;
- 4) deixar o animal em aquário aquecido e bem aerado, por cerca de 48 horas.

Fragments de tecidos de rim e brânquias foram utilizados para a obtenção de preparações cromossômicas visando o estudo cariotípico das espécies.

3.2.2 Obtenção de Cromossomos Metafásicos

A técnica utilizada para obtenção de cromossomos metafásicos seguiu a descrição de Foresti et al. (1981), com algumas adaptações. Essa metodologia envolve basicamente, a inibição da polimerização dos microtúbulos pela colchicina e a hipotonização das células em suspensão *in vitro*, seguindo-se a fixação celular pela mistura de metanol/ácido acético. O procedimento consiste em:

- 1) anestesiá o animal em solução de benzocaína a 0,01%;
- 2) sacrificar o animal, retirando fragmentos de tecido do rim anterior. Transferir o material para uma pequena placa de vidro, contendo 10 ml meio de cultura RPMI previamente aquecido em estufa a 37°C;
- 3) dissociar o material com o auxílio de pinças de dissecação, complementando esse processo com o auxílio de uma pipeta Pasteur até a obtenção de suspensão homogênea;
- 4) adicionar com uma seringa de 1ml, 2 gotas de solução de colchicina a 0,01% e levar à estufa a 37°C por 30 minutos;
- 5) centrifugar (1000 rpm) por 10 minutos e descartar o sobrenadante;
- 6) acrescentar 10 ml de solução hipotônica de KCl (0,075M) previamente aquecida a 37°C e levar a estufa 37°C por 21 minutos;
- 7) retirar da estufa, colocando 10 gotas de fixador gelado (metanol e ácido acético na proporção de 3:1, respectivamente); agitar levemente a mistura com uma pipeta Pasteur e deixar repousar por 5 minutos a temperatura ambiente;

- 8) adicionar cerca de 10 ml de fixador e novamente agitar a mistura; levar à centrífuga (1000 rpm) por 10 minutos;
- 9) gotear 2 gotas do material em lâminas de vidro;
- 10) deixar secar ao ar.

As lâminas podem ser guardadas no congelador durante muito tempo, servindo, assim, para aplicação de várias técnicas de bandamento cromossômico e FISH.

3.2.3 Identificação da Heterocromatina Constitutiva

Essa técnica caracteriza-se pela ação sequencial de base, ácido e solução salina para eliminar seletivamente o DNA cromossômico, permanecendo intactas somente regiões de heterocromatina, mais compactadas. O procedimento foi seguido conforme descrito por Sumner (1972), com adaptações e consiste em:

- 1) deixar a lâmina submersa em HCl 0,2N por 25 minutos, à temperatura ambiente;
- 2) lavar com água destilada, à temperatura ambiente;
- 3) incubar a lâmina em solução de hidróxido de Bário 5%, a 60°C, por 5 a 15 segundos;
- 4) lavar rapidamente em HCl 1 N, a 60°C, por alguns segundos;
- 5) incubar a lâmina em 2XSSC, a 60°C, por 25 minutos;
- 6) lavar com água destilada à temperatura ambiente;
- 7) corar com Giemsa 5% por 10 minutos.

3.2.4 Extração de DNA genômico

O DNA genômico foi obtido a partir de amostras de tecido de fígado ou músculo preservadas em etanol, utilizando-se o Kit Wizard Genomic DNA Purification (PROMEGA) conforme instruções do fabricante.

3.2.5 Preparação das sondas de DNA para hibridação fluorescente *in situ* (FISH)

3.2.5.1 Preparação das sondas de DNAr 18S e 5S

A sonda de DNAr 18S foi obtida por PCR (Polymerase Chain Reaction) a partir do DNA total de *Moenkhausia sanctaefilomenae* usando os primers 18S F (5'CCG CTT TGG TGA CTC TTG AT 3') e 18S R (5'CCG AGG ACC TCA CTA AAC CA 3') (White et al. 1990). Esta sonda foi marcada com Digoxigenina-11-dUTP (Roche Applied Science) por PCR e a detecção do sinal de hibridação foi realizada usando anti-digoxigenina-rodamina (Roche Applied Science). Os parâmetros para marcação compreenderam 31 µl de água Milli-Q, 5 µl de tampão da enzima Taq polimerase (10X), 5 µl de MgCl₂ (25mM), 1 µl de dATP,

dCTP e dGTP e 0,7 µl de dTTP (2mM cada), 0,8 µl de Digoxigenina-11-dUTP, 1µl de cada primer (10mM), 0,5 µl de Taq polimerase (5U/µL) e 2 µl de DNA molde. Total da reação: 50 µl.

O programa de PCR utilizado apresentou as seguintes condições:

34 ciclos	{	94°C ---- 2 min
		95°C ---- 45 seg
		52,1°C ---- 45 seg
		72°C ---- 1 min e 30 seg
		72°C ---- 5 min
		12°C ---- manutenção

Os produtos gerados foram checados em gel de agarose 1%.

A sonda de DNAr 5S foi obtida por PCR (Polymerase Chain Reaction) a partir do DNA total de *Moenkhausia sanctaefilomenae* usando os primers 5S F (5'TAC GCC CGA TCT CGT CCG ATC 3') e 5S R (5'CAG GCT GGT ATG GCC GTA ACG 3') (Pendás et al. 1994). Esta sonda foi marcada com Biotina-16-dUTP (Roche Applied Science) por PCR e a detecção do sinal de hibridação foi realizada usando Avidina-fluoresceína conjugada (FITC) e os parâmetros para marcação foram: 33 µl de água Mili-Q, 2 µl de tampão da enzima Taq polimerase (10X), 5 µl de MgCl₂ (25mM), 1 µl de dATP, dCTP e dGTP e 0,5 µl de dTTP (2mM cada), 2 µl de Biotina-16-dUTP, 1 µl de cada primer (10mM), 0,5µl de Taq polimerase (5U/µL) e 2 µl de DNA molde. Total da reação: 50 µl.

O programa de PCR utilizado apresentou as seguintes condições:

34 ciclos	{	95°C ---- 5 min
		95°C ---- 1 min
		55°C ---- 45 seg
		72°C ---- 1 min
		72°C ---- 7 min
		12°C ---- manutenção

Os produtos gerados foram checados em gel de agarose 1%.

3.2.5.2 Preparação da sonda para snDNA U2

A sonda utilizada para detecção de sequências snDNA U2 foi obtida por PCR (Polymerase Chain Reaction) a partir do DNA total de *Moenkhausia sanctaefilomenae*, usando os primers U2F (5'-ATC GCT TCT CGG CCT TAT G-3') e U2R (5'-TCC CGG

CGG TAC TGC AAT A-3') (Bueno et al. 2013). Esta sonda foi marcada com Digoxigenina-11-dUTP (Roche Applied Science) por PCR e a detecção do sinal de hibridação foi realizada usando anti-digoxigenina-rodamina (Roche Applied Science). Os parâmetros para marcação foram: 31 µl de água Mili-Q, 5 µl de tampão da enzima Taq polimerase (10X), 5 µl de MgCl₂ (25mM), 1 µl de dATP, dCTP e dGTP e 0,7 µl de dTTP (2mM cada), 0,8 µl de Digoxigenina-11-dUTP, 1µl de cada primer (10mM), 0,5 µl de Taq polimerase (5U/µL) e 2 µl de DNA molde. Total da reação: 50 µl.

O programa de PCR utilizado apresentou as seguintes condições:

34 ciclos {

- 94°C ---- 2 min
- 95°C ---- 45 seg
- 52,1°C ---- 45 seg
- 72°C ---- 1 min e 30 seg
- 72°C ---- 5 min
- 12°C ---- manutenção

Os produtos gerados foram checados em gel de agarose 1%.

3.2.5.3 Preparação da sonda para histona H3

A sonda utilizada para detecção de sequências histônicas foi amplificada e marcada por PCR a partir do DNA total de *Moenkhausia sanctaefilomenae*, usando os primers H3F (5'-ATGGCTCGTACCAAGCAGACVGC-3') e H3F (5'-ATATCCTTRGGCATRATRG TGAC-3') (Colgan et al. 1998) e os parâmetros para marcação foram: 33,9 µl de água Mili-Q, 5 µl de tampão da enzima Taq polimerase (10X), 2 µl de MgCl₂ (25mM), 1 µl de dATP, dCTP e dGTP e 0,7 µl de dTTP (2mM cada), 1 µl Digoxigenina-11-dUTP, 1,5 µl de cada primer (10mM), 0,4 µl de Taq polimerase (5U/µL) e 1 µl de DNA molde. Total da reação: 50 µl.

O programa de PCR utilizado apresentou as seguintes condições:

34 ciclos {

- 95°C ---- 10 min
- 94°C ---- 15 seg
- 48°C ---- 15 seg
- 68°C ---- 15 seg
- 68°C ---- 7 min
- 12°C ---- manutenção

Os produtos gerados foram checados em gel de agarose 1%.

3.2.6 Microdissecção

Os procedimentos para a microdissecção cromossômica consistiram em (Henning et al. 2008):

1. Diluir uma alíquota da suspensão celular final no fixador (metanol: ácido acético na proporção 3:1), respectivamente;
2. Pingar as suspensões celulares em lamínulas previamente limpas e esterilizadas, esperar secar ao ar.
3. Lavar a lamínula em uma solução de PBS 1X por 1 minuto.
4. Incubar a lamínula em uma solução de tripsina (1% de tripsina, PBS 1X) por 20 segundos.
5. Lavar em PBS 1X e corar com Giemsa 5% em PBS por 9 minutos.
6. Colocar a lamínula contendo a preparação celular num microscópio invertido (Axiovert 100 – Zeiss) e proceder à microdissecção dos cromossomos com o auxílio de uma agulha de vidro acoplada a um micromanipulador mecânico (Eppendorf) e uma pipeta silanizada (método de silanização descrito em Henning et al. 2008).
7. Cada cromossomo microdissecado foi transferido para uma micropipeta contendo uma solução de coleção (1.5ug/ul proteinase K, 0.1% SDS, 0.1% Triton X-100, 1mM EDTA, 10mM Tris-HCl, pH 8.0, 10mM NaCl) e mantido a 60°C por 1h em uma câmara úmida.
8. Quebrar a ponta das micropipetas em um microtubo de 0.2ml contendo 5ul de água miliQ estéril e amplificar o material utilizando o kit GenomePlex Single Cell Whole Genome Amplification Kit (WGA-Sigma).

Nota: Todos os experimentos de microdissecção foram realizados utilizando cópias únicas de cromossomos.

3.2.7 Kit GenomePlex® Single Cell Whole Genome Amplification (WGA4 – Sigma) (Gribble et al. 2004)

Esta reação consistiu na fragmentação aleatória não enzimática do DNA genômico seguida pela ligação de adaptadores com os dois locais de iniciação de PCR específicos para ambas as extremidades de cada fragmento conforme proposto por Marshall et al. (2004).

Passo 1 - Fragmentação: Adicionar 1 µl da solução Lysis and Fragmentation Buffer (0,5 µl de proteinase K + 8 µl de tampão 10x) ao tubo contendo os cromossomos microdissecados.

Incubar o mix de DNA a 50°C por 1 hora, aquecer por exatos 4 minutos a 99°C, manter no gelo e centrifugar brevemente antes de preparar a biblioteca.

Passo 2 – Preparação da Biblioteca: Para a preparação da biblioteca foram adicionados ao tubo de PCR 2µl de Single Cell Library Preparation e 1µl de Library Stabilization Solution. A amostra foi colocada em termociclador a 95°C por 2 minutos. Em seguida adicionou-se 1µl de Library Preparation Enzyme e a amostra foi colocada em termociclador e submetida as seguintes condições:

16°C ----- 20 min
24°C ----- 20 min
37°C ----- 20 min
75°C ----- 5 min
12°C --- manutenção

Passo 3 – Amplificação: Na sequência, foi realizada uma reação de amplificação, adicionando 48,5 µl de água ultrapura, 7,5 µl de 10x Amplification Master Mix e 5,0 µl de WGA DNA polimerase. A amostra foi submetida ao seguinte programa:

25 ciclos { 95°C ---- 3 min
94°C ---- 30 seg
65°C ---- 5 min
12°C ---- manutenção

O produto da amplificação foi visualizado em gel de agarose 1%.

3.2.8 WGA3 – Re-amplificação e Marcação.

Para a reação de reamplificação foi preparado os parâmetros foram: 15µl de água livre de nuclease, 2,5µl de Amplification Master Mix (10X), 1,7µl de WGA DNA polimerase, 1µl de dNTP (2mM cada), 0,3µl de dTTP (2mM), 0,7µl Digoxigenina-11-dUTP (1mM), 2µl de DNA molde. A amostra foi submetida ao seguinte programa:

14 ciclos { 95°C ---- 3 min
94°C ---- 15 seg
65°C ---- 5 min
12°C ---- manutenção

O produto da amplificação foi visualizado em gel de agarose 1%.

3.2.9 Hibridação *in situ* Fluorescente (FISH)

O mapeamento das sondas composta por DNA repetitivo e da sonda do cromossomo W produzida por microdissecção em cromossomos metafásicos foi realizado por hibridação fluorescente *in situ* de acordo com o procedimento estabelecido por Pinkel et al. (1986), com modificações conforme descrito abaixo:

Etapa 1: Preparação das lâminas

As lâminas foram lavadas em PBS 1x por 5 minutos à temperatura ambiente e desidratadas em série de etanol a 70%, 85% e 100% a -20°C por 5 minutos em cada banho. Posteriormente, as lâminas foram tratadas com solução de RNase (100 µg/ml) por 1 hora em câmara úmida a 37°C, lavadas 2 vezes em 2xSSC por 10 minutos e uma vez em PBS 1x por 5 minutos. Em seguida, as lâminas foram tratadas com pepsina 0,005%/HCl 10mM por 10 minutos a 37°C, lavadas em PBS 1x à temperatura ambiente por 5 minutos e fixadas com formaldeído 1%/PBS 1x/MgCl₂ 50mM por 10 minutos a temperatura ambiente. Depois de fixadas, as lâminas foram lavadas em PBS 1x por 5 minutos e desidratadas em série de etanol a 70%, 85% e 100% a -20°C por 5 minutos em cada banho. Após a secagem, as lâminas foram desnaturadas em formamida 70% a 70°C por 5 minutos e desidratadas em série de etanol a 70%, 85% e 100% a -20°C por 5 minutos em cada banho. Simultaneamente à desnaturação das lâminas, desnaturou-se a solução de hibridação (10% sulfato dextrato, 50% formamida e 20xSSC e água) contendo 3-6 µL de sonda marcada por 10 minutos a 100°C. Posteriormente, 15 µL de solução de hibridação foram colocados sobre a lâmina, que foi coberta com lamínula e mantida em câmara úmida a 37°C *overnight*.

Etapa 2: Detecção dos sinais correspondentes

Após a hibridação as lâminas foram lavadas 2 vezes em solução de formamida 15%/0,2xSSC a 42°C por 10 minutos cada. Em seguida, foram lavadas 3 vezes em solução de 0,1xSSC a 60°C por 5 minutos cada e uma vez em solução Tween 0,5%/4xSSC à temperatura ambiente por 5 minutos. Após as lavagens, as lâminas foram cobertas com 50 µL de solução de bloqueio (leite em pó, 20xSSC, e água) por 15 minutos e lavadas 2 vezes em Tween 0,5%/4xSSC a temperatura ambiente. Em seguida, as lâminas foram incubadas com 30 µL de solução de bloqueio com anticorpo (24 µL de solução de bloqueio e 6µL de antidigoxi-rodamina) por 1 hora em câmara úmida e escura a 37°C. Posteriormente, foram lavadas 3 vezes em Tween 0,5%/4xSSC por 5 minutos, desidratadas em série de etanol 70%, 85% e

100% a -20°C por 5 minutos em cada banho, secas e cobertas com 22 µl de Vectashield® Mounting Medium com DAPI (Vector). Todas as lâminas foram analisadas em um fotomicroscópio óptico de fluorescência (Olympus BX61) e as imagens metafásicas foram capturadas usando o programa Image Pro Plus, 6.0 software (MediaCybernetics).

3.2.10 Sequenciamento Illumina

O DNA extraído de um exemplar de cada uma das populações analisadas e uma biblioteca de cromossomo microdissecado foram utilizados para sequenciamento na plataforma Illumina HiSeq2000 pela empresa Macrogen Inc. (Coreia do Sul). O serviço requerido e utilizado foi o sequenciamento de baixa cobertura *paired-ends reads* 2x101bp (retorno de aproximadamente 1Gb).

3.2.11 Amplificações do material microdissecado

A partir das bibliotecas de cromossomos B₁ e B₂ de *M. sanctaefilomenae* microdissecados, foram amplificados fragmentos correspondentes à região parcial codificante do gene histona H3 e os DNAs satélites MS3 e MS7 para posterior clonagem e sequenciamento. Os produtos foram submetidos à eletroforese em gel de agarose.

3.2.12 Purificação dos fragmentos de DNA

Os fragmentos de DNA de interesse foram recortados do gel de agarose e purificados utilizando o “GFX™ PCR and Gel Band Purification Kit” (Amersham Biosciences), de acordo com as orientações do fabricante. Assim, deve-se recortar as bandas do gel de agarose e colocá-las num tubo de 0,5ml. Adicionar 10µl de capture buffer para cada 10mg de gel. Misturar no vortex e incubar a 60°C, mexendo a cada 5 minutos, até que a agarose se dissolva totalmente. Centrifugar brevemente a agarose dissolvida e transferir o conteúdo para a coluna GFX. Incubar por um minuto em temperatura ambiente e, em seguida, centrifugar por 1 minuto a 16.000rpm. Descartar o líquido que passou pela coluna para o tubo coletor e recolocar a coluna GFX novamente no mesmo tubo coletor. Adicionar 500µl de wash buffer à coluna GFX e centrifugar por 1 minuto a 16.000rpm. Descartar o tubo coletor e transferir a coluna GFX para um tubo de microcentrífuga de 1,5ml estéril. Secar a coluna GFX para evitar a presença de álcool. Deixar 5 minutos à temperatura ambiente. Adicionar sobre a coluna 50µl de água MilliQ autoclavada. Incubar por 2 minutos em temperatura ambiente e, em seguida, centrifugar por 1 minuto a 16.000rpm para recolher o DNA purificado. Descartar a coluna GFX e armazenar no freezer os DNAs recuperados.

3.2.13 Clonagem

As amostras de DNA isoladas do gel de agarose foram ligadas ao plasmídeo pGEM®-T (Promega) para posterior transformação de células competentes de *Escherichia coli* DH5 α (Gibco.Brl Life Technologies), segundo as instruções do fabricante. Em um tubo de 0,5ml, adicionar 3 μ l do inserto de interesse (fragmento de DNA), 1 μ l de T4 DNA ligase, 5 μ l de tampão de reação 2x, 1 μ l do plasmídeo pGEM-T (54ng/ μ l). Misturar cuidadosamente com uma micropipeta e incubar a reação a 4°C durante 12-16 horas.

Descongelar as células competentes em gelo, aliquotar 50 μ l destas células competentes e transferir para um tubo de microcentrífuga de 1,5 ml, mantendo em gelo. Recolocar o restante das células competentes em freezer -70°C. Adicionar 4 μ l do produto de ligação (inserto-plasmídeo) ao tubo contendo os 50 μ l de células competentes, misturando cuidadosamente com uma micropipeta. Manter o tubo em gelo por 30 minutos. Aplicar um choque térmico de 20 segundos em banho-maria a 37°C. Colocar o tubo imediatamente em gelo e manter por 2 minutos. Adicionar ao tubo 950 μ l de meio líquido LB (peptona 1%, NaCl 0,17M, extrato de levedura 0,5%) à temperatura ambiente e agitar a 225 rpm por uma hora a 37°C.

O produto deve, então, ser centrifugado por 10 segundos a 14.000 rpm e descartada parte do sobrenadante. Espalhar 150 μ l do produto de transformação em placas de Petri estéreis contendo meio sólido LB (peptona 1%, NaCl 0,17M, extrato de levedura 0,5%, ágar 1,5%, ph 7,5), contendo 2 μ l de ampicilina (50 mg/ml) por mililitro de meio LB e 50 μ l de X-Gal (5-bromo-4-cloro-3-indolil- β -D-galactoside) (50 mg/ml). Incubar as placas com o meio voltado para baixo *overnight* a 37°C. As colônias brancas foram transferidas para tubos de 15 ml contendo 4 ml de meio LB líquido acrescido de ampicilina com ajuda de palitos ou ponteiras de micropipeta autoclavados.

Os tubos com as bactérias devem ser mantidos *overnight* sob agitação de 300rpm a 37°C. Transferir 0,8 ml do meio com as bactérias para tubos de microcentrífuga de 1,5 ml contendo aproximadamente 200 μ l de glicerol estéril. O material foi homogeneizado e armazenado a -70°C.

3.2.13.1 Identificação dos clones de interesse

Os clones de interesse foram identificados através de reações de amplificação utilizando *primers* específicos para o plasmídeo. A reação foi feita num volume total de 50 μ l (50 pmol de

cada *primer* - M13F (5'- 3') AGC GGA TAA CAA TTT CAC ACA GG e M13R (5'- 3') CCC AGT CAC GAC GTT GTA AAA CG -, tampão de reação 1x (KCl 50 mM; MgCl₂ 1,5mM; Tris-HCl 10 mM), 2 mM de cada dNTP e 0,5 U de *Taq* DNA polimerase (Amersham Biosciences).

Encostar um palito estéril sobre a colônia recombinante e mergulhar o palito no tubo de 0,5 ml contendo as soluções para amplificação. As reações de amplificação foram realizadas num termociclador PTC-100TM (MJ Research), de acordo com as seguintes condições: um ciclo de 3 minutos a 95°C seguido de 35 ciclos de 30 segundos a 95°C, 1 minuto a 60°C, 2 minutos a 72°C e um ciclo final de 5 minutos a 72°C. Para verificar a efetividade da reação e a amplificação de fragmentos dentro do tamanho esperado, checar o produto de PCR em gel de agarose 2%.

3.2.14 Sequenciamento nucleotídico (Sanger)

O sequenciamento nucleotídico foi realizado utilizando o kit BIG DY Terminator k.3.1 (Applied Biosystems), segundo as instruções do fabricante. Para cada amostra a ser sequenciada foi preparada a seguinte reação: 2µl de pré-mix (kit), 1µl de primer (3,3mM), 40-100ng DNA (volume máximo de 5µl), água estéril (volume necessário para 10µl de reação final). O programa de PCR de sequenciamento consistiu de uma etapa inicial de 2 minutos a 96°C, seguida de 36 ciclos de 45 segundos a 96°C, 30 segundos a 50°C e 4 minutos a 60°C e de uma etapa final a 4°C. Os produtos dessa reação foram purificados para eliminação de nucleotídeos terminadores não incorporados. Adicionar 2µl de acetato de sódio 3M e 80µl de etanol 95% em temperatura ambiente em cada tubo contendo o produto da reação de PCR de sequenciamento. Misturar no vortex e centrifugar a 20°C por 30 minutos a 14.000g. Remover cuidadosamente o sobrenadante por aspiração. Adicionar 400µl de etanol 70% em temperatura ambiente e, em seguida, centrifugar a 20°C por 5 minutos a 14.000g. Descartar cuidadosamente o sobrenadante por aspiração e deixar o pellet secar por uma hora em temperatura ambiente protegido da luz. O pellet poderá ficar guardado por um mês a 4°C protegido da luz. Ressuspender o pellet em 4µl de tampão de carregamento (Formamida:blue Dextran, na proporção de 5:2). Passar os tubos pelo vortex, desnaturar as amostras por 5 minutos a 95°C e colocá-las imediatamente no gelo após a desnaturação. Utilizar até 2µl de amostra na corrida de sequenciamento. As sequências foram lidas em um sequenciador automático ABI 3130 Genetic Analyzer (Applied Biosystems).

3.2.15 Análises bioinformáticas

As análises bioinformáticas apresentadas aqui foram realizadas em parceria com diferentes instituições de ensino e pesquisa, nomeadamente, Universidad de Granada, sob orientação do prof. Dr. Juan Pedro Martinez Camacho e Institute of Molecular and Cellular Biology, sob orientação do prof. Dr. Vladimir A. Trifonov.

3.2.15.1 RepeatExplorer

Inicialmente, um passo de filtro de qualidade dos *reads* Illumina foi realizado (retirada de *reads* com menos de 90% de bases com qualidade inferior a Q20) com o *software* Trimmomatic (Bolger et al. 2014). Após a etapa de filtragem, 400.000 pares de *reads* (200.000 *reads* de cada população) foram selecionados para a etapa de clusterização com o *software* RepeatExplorer (Novák et al. 2013) considerando *reads paired-end*, e cobertura da clusterização e montagem com tamanhos iguais a 55 e 40pb, respectivamente. Além disso, uma base de dados customizada de sequências repetitivas também foi construída rodando o programa RepeatModeler (Smit e Hubley 2010) no genoma montado de *A. mexicanus* (GenBank accession number APWO00000000.1) como um complemento para a anotação das repetições do RepBase. Esta base de dados customizada gerou 1.243 sequências, constituídas por 589.736pb e N50=613pb. Após a etapa de clusterização, nós buscamos por *clusters* de famílias de DNAs satélites. Posteriormente, nós extraímos e processamos manualmente os contigs montados utilizando o *software* Geneious Pro v8.04 utilizando a opção *High Sensivity/Slow* para visualizar os gráficos dotplots visando detectar repetições *in tandem*. Finalmente, os *contigs* extraídos foram quebrados em monômeros e alinhados para obter a sequência consenso do DNA satélite para desenho de *primers*.

3.2.15.2 Análise de cromossomo microdissecado

Para a análise do cromossomo B microdissecado, as três bibliotecas de *M. sanctaefilomenae* (1- cromossomo B₁ microdissecado, 2- 0B gDNA, 3- 6B gDNA) foram processadas com a *pipeline* “DOPseq analyzer (https://github.com/ilyakichigin/DOPseq_analyzer).” desenvolvida por MSc Alexey Makunin. De forma geral, esta *pipeline* retira os adaptadores Illumina e os *primers* WGA GenomePlex e alinha os *reads* das bibliotecas ao genoma de *A. mexicanus* (GCA_000372685.1) e ao genoma humano (hg19, baixado do *UCSC Genome Browser*) com o *software* bowtie2. Desta forma, os *reads* que se alinham perfeitamente com o genoma humano são descartados como contaminação putativa. Adicionalmente, os *reads* também foram alinhados ao genoma de *D. rerio* (GCA_000002035.3), uma espécie de peixe mais distante de *M. sanctaefilomenae*,

porém com um genoma mais completo e melhor montado. Além disso, esta *pipeline* identifica regiões cromossomo-específicas baseadas em densidade de *reads*, bem como realiza uma busca por regiões variantes (*variant calling*) dos genes contra a referência utilizada.

O método escolhido para a análise de cobertura diferencial foi semelhante ao utilizado por Valente et al. (2014), buscando-se regiões com 4x de diferença de densidade de *reads* entre as duas bibliotecas genômicas. Para a análise da porção repetitiva deste cromossomo, os *reads* de *midib* foram alinhados contra os *contigs* gerados após análise pelo *software* RepeatExplorer (Novák et al. 2013) com os dados de genoma total 0B e 6B através do *software* RepeatMasker (Smit et al. 2010) para estimar a abundância de cada um dos *clusters* na biblioteca *midib*.

3.2.15.3 Extração dos DNAs satélites MS3 e MS7 dos *reads* Illumina

Para obter um *score* mais confiável e detalhado sobre a abundância haplotípica dos satDNAs MS3 e MS7 nas bibliotecas genômicas, nós extraímos os monômeros de cada um dos satélites diretamente dos *reads* Illumina. Uma vez que os *reads* são menores que o tamanho dos monômeros de ambos os satélites, nós juntamos os *reads paired-ends* utilizando *fastq-join* (<https://code.google.com/p/ea-utils/wiki/FastqJoin>) com um mínimo de sobreposição de sequências de 6pb.

Para o satDNA MS7, nós alhamos os *reads* contra um dímero de sequência deste satDNA com o *software* RepeatMasker (Smit et al. 2010) e, com um *script* Python (https://github.com/fjruiaruano/ngs-protocols/blob/master/rm_getseq.py); nós utilizamos a informação do alinhamento para extrair apenas a região alinhada. Posteriormente, nós mapeamos estas sequências com o Geneious Pro v8.04 contra um dímero e extraímos a região central do alinhamento. Para o MS3, após o mapeamento com Geneious, nós cortamos a ponta esquerda do mapeamento ao monômero referência e juntamos este corte ao outro lado da sequência, utilizando um outro *script* Python (https://github.com/fjruiaruano/ngs-protocols/blob/master/sat_cutter.py).

3.2.15.4 Análises nucleotídicas

Todas as sequências foram inicialmente analisadas no programa Geneious Pro v8.05. O alinhamento das sequências foi feito com a versão *online* do programa Mafft (Katoh e Standley, 2013). As análises de diversidade foram realizadas com o programa DnaSP (Librado e Rozas, 2009). Para a montagem de uma árvore de extensão mínima (*minimum spanning tree*), nós selecionamos os blocos mais conservados do alinhamento utilizando o

programa Gblocks (Castresana, 2000) a partir dos haplótipos gerados no programa DNAsp. A árvore de extensão mínima foram geradas no programa Arlequin v3.5 (Excoffier e Lischer, 2010) considerando somente os haplótipos correspondentes a mais de uma sequência. Todas as sequências foram submetidas a uma busca na base de dados do NCBI (<http://www.ncbi.nlm.nih.gov/blast>) com a ferramenta BLASTn (Altschul et al. 1990) para confirmar sua identidade.

Comparações de substituições sinônimas por sítio sinônimo (dS) e não sinônimas por sítios não sinônimos (dN) da histona H3.2 de cada uma das bibliotecas (B₁, B₂ e 0B gDNA) foram realizadas utilizando o teste não-paramétrico de Kruskal-Wallis seguido por pós-testes de Dunn, considerando $\alpha=0,05$.

3.3 Estudos Cariotípicos

As preparações cromossômicas convencionais, com fluorocromos e com hibridação *in situ* foram analisadas em fotomicroscópio óptico (Olympus BX61), estabelecendo-se o número diploide modal (contagem mínima de 23 metáfases cromossômicas por espécime) presente em cada população amostrada no presente trabalho. As melhores metáfases foram capturadas com *software* Image Pro Plus, 6.0 (MediaCybernetics). Os fluorocromos foram capturados separadamente (utilizando filtros UV específicos), coloridos e sobrepostos no mesmo *software* Image Pro Plus, 6.0 (MediaCybernetics).

Os cromossomos foram recortados usando o *software* Adobe Photoshop versão 12.0 - Adobe System e organizados em pares de homólogos em classes e ordem decrescente de tamanho para a montagem dos cariótipos, sendo classificados em metacêntricos (m), submetacêntricos (sm), subtelocêntricos (st) e acrocêntricos (a), com base no trabalho de Levan *et al.* (1964).

4 Resultados e Discussão

As informações obtidas dos estudos citogenéticos e moleculares efetuados em duas populações de *M. sanctaefilomenae* foram organizados na forma de artigos científicos, que são apresentados a seguir. O Capítulo 1 já foi submetido para publicação e o Capítulo 2 ainda é um estudo incipiente desses elementos e necessita de experimentos adicionais.

Capítulo 1

Uncovering the ancestry of B chromosomes in *Moenkhausia sanctaefilomenae* (Teleostei, Characidae)

Capítulo 2

Microdissecção e sequenciamento completo do cromossomo B₁ de *Moenkhausia sanctaefilomenae* (Teleostei, Characidae)

Capítulo 1

Uncovering the ancestry of B chromosomes in *Moenkhausia sanctaefilomenae* (Teleostei, Characidae)

Abstract

B chromosomes constitute a heterogeneous mixture of genomic parasites sometimes derived intraspecifically from the standard genome of the harboring species but, in other cases, they result from interspecific hybridization. The mode of origin determines DNA content in the B showing high similarity with the A genome, in the first case, but higher similarity with a different species in the second. The characid fish *Moenkhausia sanctaefilomenae* harbors highly invasive B chromosomes being present in all populations hitherto analyzed in the Parana and Tietê rivers. In order to investigate the origin of these B chromosomes, we analyzed here two natural populations: one carrying B chromosomes and the other lacking them, using a combination of molecular cytogenetic techniques, nucleotide sequence analysis and high-throughput sequencing. Our results have shown that i) B chromosomes have not yet reached the Paranapanema River basin, ii) B chromosomes are mitotically unstable, iii) there are two types of B chromosomes, the most frequent being lightly C-banded (similar to euchromatin in A chromosomes) (B₁) and the other being darkly C-banded (heterochromatin-like) (B₂), iv) the two B-types contain the same tandem repeat DNA sequences (18S ribosomal DNA, H3 histone genes, the MS3 and MS7 satellite DNAs), with higher content of 18S rDNA in the heterochromatic variant, v) all these repetitive DNAs are present together only in the paracentromeric region of the autosome pair no. 6, suggesting that B chromosomes are derived from this A chromosome, vi) the two B chromosome variants show MS3 sequences highly diverged between them and from the 0B genome, although the B₂ derived sequences show higher similarity with the latter (this suggests the independent origin of both

B variants, the least frequent B₂ presumably being younger), and vii) The dN/dS ratio for the H3.2 histone gene was almost 4-6 times higher for B chromosome than for A chromosome sequences, suggesting that purifying selection is relaxed for the DNA sequences located in the B chromosomes, presumably because they are mostly inactive.

Introduction

B chromosomes are dispensable genomic elements that are present in about 15% of eukaryotes. These chromosomes exhibit a parasitic nature, and their interaction with the host genome determines population frequencies being highly dynamic as a consequence of “the arms race” between A and B chromosomes (Camacho et al. 2000; Jones, 1985; Nur, 1966). Since B chromosomes do not always occur in pairs, their segregation does not conform to a Mendelian system, which may facilitate these chromosomes to show transmission rates higher than 0.5, resulting in transmission advantages collectively named drive (Camacho et al. 2000; Hewitt, 1976).

The origin of B chromosomes has been investigated in different organisms and, basically, they may arise from the A chromosomes of their current host species (intraspecific origin), or else interspecifically through hybridization (Camacho et al. 2000). The intraspecific origin of B chromosomes has been demonstrated, for instance, in maize (Lamb et al. 2005; Peng and Cheng, 2011), the migratory locust (Teruel et al. 2010), rye (Martis et al. 2012), and the fish *Astyanax paranae* (Silva et al. 2014). However, examples of B chromosomes arisen through interspecific hybridization have been reported in plants genus *Coix* (Sapre and Deshpande, 1987), the fish *Poecilia formosa* (Schartl et al. 1995), and the wasp *Nasonia vitripennis* (McAllister and Werren, 1997; Perfectti and Werren, 2001).

Moenkhausia (Teleostei, Characidae) is a small freshwater characid fish genus widely distributed in South American river basins and comprising about 86 species (Eschmeyer and

Fong, 2015). *M. sanctaefilomenae* shows a conserved karyotype in respect to the standard genome ($2n= 50$ biarmed chromosomes) but it also carried micro-B chromosomes in all populations hitherto analyzed (Dantas et al. 2007; de Brito Portela-Castro et al. 2000; Foresti et al. 1989; Hashimoto et al. 2012; Scudeler et al. 2015). Specifically, several populations collected in the Tietê River basin showed euchromatic, partially or totally heterochromatic B chromosomes (Dantas et al. 2007; Foresti et al. 1989; Hashimoto et al. 2012; Scudeler et al. 2015), whereas one population collected at the Paraná River showed only the euchromatic variant, which was restricted to males (de Brito Portela-Castro et al. 2000).

Recently, after karyotyping three populations from the Tietê River basin and performing chromosome painting using a B-specific probe, Scudeler et al. (Scudeler et al. 2015) concluded that the heterochromatic B-type had an intraspecific origin, due to sharing DNA sequences with several A chromosomes, and it arose independently from the euchromatic B chromosome, since a painting probe produced from the former B-type did not paint the latter variant. Trying to go further in these conclusions, we analyzed here a natural population carrying the euchromatic and heterochromatic variants, by means of a combination of cytogenetic (C-banding, microdissection, chromosome painting, FISH mapping, silver staining), molecular (PCR amplification, cloning, Sanger DNA sequencing and Illumina sequencing), phylogenetic, and bioinformatics techniques, also including the first report for a B-lacking population in this species.

Materials and Methods

Ethics Statement

Sampling was carried out on private lands and the owners gave permission to conduct this study. The animals were captured using nets, transported to the Laboratory, kept in a fish tank and were anesthetized before the analyses. The animals were collected in accordance

with Brazilian environmental protection legislation (Collection Permission MMA/IBAMA/SISBIO—number 3245) and the procedures for sampling, maintenance and analysis of the fishes were performed in compliance with the Brazilian College of Animal Experimentation (COBEA) and was approved (protocol 504) by the BIOSCIENCE INSTITUTE/UNESP ETHICS COMMITTEE ON USE OF ANIMALS (CEUA).

Sampling, chromosome banding and DNA extraction

Individuals of *M. sanctaefilomenae* were sampled in 2 rivers from the Paraná River system, 23 specimens (11 females and 12 males) from the Batalha River, belonging to the Tietê River basin (Bauru, SP; 22°24'23.65'S, 49°05'51.38"W) and 16 individuals (11 females and 5 males) from the Novo River, at the Paranapanema River basin (Ocaçu, SP; 22°28'13.32"S, 49°55'26.17"W). These rivers are separated from each other by approximately 170 Km, as the crow flies, and by several hundred kilometers through the rivers connecting them. After analysis, all specimens were deposited at the fish collection of the Laboratório de Biologia e Genética de Peixes (LBP) at UNESP, Botucatu, São Paulo, Brazil, under the voucher numbers LBP19830 (Batalha River) and LBP19831 (Novo River).

Mitotic chromosomes were obtained from cell suspensions of the anterior kidney, according to Foresti et al. (Foresti et al. 1981). C-banding was carried out according to Sumner (Sumner, 1972) and the active nucleolar organizer regions (NORs) were revealed according to Howell and Black (Howell and Black, 1980). The chromosomes were classified as metacentric (m), submetacentric (sm), subtelocentric (st) and acrocentric (a) according to Levan et al. (Levan et al. 1964). Genomic DNA (gDNA) was obtained from liver cells using the Promega Wizard Genomic DNA Purification Kit according to the manufacturer's instructions.

Chromosome microdissection

Cell suspensions were dropped onto clean coverslips. The coverslips were washed in 1x PBS solution for 1 minute, incubated in a trypsin solution (1% Trypsin, 1x PBS) for 20 seconds and washed again in 1x PBS. Afterwards, the preparations were stained with Giemsa 5% in PBS for 5 minutes. For the microdissection of B₁, we used the cell suspensions of the individual 69693, which presented only this variant (Table 1). To microdissect the B₂ variant, C-banding was performed to allow a better identification of the heterochromatic B chromosome in the metaphase spread. In addition, one autosome from a B-carrying individual was also microdissected (pair No. 1, henceforth called A₁). One must note that all microdissection experiments were carried out using single copy chromosomes.

Each microdissected chromosome was transferred to a micropipette containing a collection solution (1.5ug/ul proteinase K, 0.1% SDS, 0.1% Triton X-100, 1mM EDTA, 10mM Tris-HCl, pH 8.0, 10mM NaCl) and placed for 1h at 60°C in a moist chamber. After that, pipette tips were broken in a 0.2ml microtube containing 5ul of sterile miliQ water and were then amplified using the GenomePlex Single Cell Whole Genome Amplification Kit (WGA-Sigma).

Whole Genome Sequencing (WGS) and Satellite DNA identification

For a deeper search of satDNAs in the *M. sanctaefilomenae* genome, we sequenced gDNA from two individuals collected at the Batalha River (carrying up to 6B) and Novo River (0B) in the Illumina HiSeq2000 platform yielding 2x10¹bp paired-end reads. After a quality trimming step (filtering out reads with less than a 90% of bases with quality lower than Q20) with Trimmomatic (Bolger et al. 2014), we sampled 200.000 pairs of reads (100.000 reads from each population) for clustering using RepeatExplorer (Novák et al. 2013) considering paired-end reads, with clustering and assembly overlap length equal to 55 and

40bp, respectively. For this analysis, we also built a custom database of repeated sequences by running RepeatModeler (Smit and Hubley, 2010) on the *A. mexicanus* assembled genome (GenBank accession number APWO00000000.1) as a complement to RepBase for cluster annotation. This custom database resulted in 1,243 sequences, constituted by 589,736bp and N50= 613bp. We then searched for clusters of satDNA families, i. e., unannotated, with sphere or ring shapes, and graph density higher than 0.1 After that, we manually processed the assembled contigs using Geneious Pro v8.04 using the High Sensivity/Slow option to visualize the dotplot graphics in order to detect tandem repetitions. We then split them into monomers, aligned them and got a consensus sequence of the monomeric units.

Repetitive DNA probes

5S rDNA, U2 snDNA and H3 histone gene probes were obtained by PCR directly from the genome of *M. sanctaefilomenae* using primers described previously (Bueno et al. 2013; Colgan et al. 1998; Pendás et al. 1994) and others described here, as follows. In order to yield an optimum-sized probe of 18S rDNA (~600bp), a new set of primers was designed, from sequences of the organisms available in GenBank, as 18S6F (5'-CTCTTTCGAGGCCCTGTAAT-3') and 18S6R (5'-CAGCTTTGCAACCATACTCC-3'). Probes of the two satDNAs were obtained using divergent primers as follows: MS3F (5'-TGGTTCCCAATTTGCAATCAAG-3') and MS3R (5'-ATCGGACCTTTCTTCGCTTTACA-3'); MS7F (5'-CACAAGCCTTATGTTACCATGA-3') and MS7R (5'-GTACAGTAAAGTTGTAAGTGGT-3').

FISH

Prior to FISH, all probes were labeled with digoxigenin-11-dUTP or biotin-16-dUTP. The painting probes were labeled using the GenomePlex (WGA3 Reamplification Kit-Sigma) following the manufacturer's protocol and the repetitive DNA probes were labeled by PCR.

FISH was performed under high stringency conditions using the method described by Pinkel et al. (Pinkel et al. 1986). The pre-hybridization conditions were different accordingly to the probes used. Thus, slides probed with repetitive sequences were incubated with RNase (50µg/ml) for 1 h at 37°C and the chromosomal DNA was denatured in 70% formamide/2xSSC for 5 min at 70°C. For each slide, 30µl of hybridization solution (containing 200ng of each labelled probe, 50% formamide, 2xSSC and 10% dextran sulphate) was denatured for 10 minutes at 95°C, dropped on the slides and hybridized overnight at 37°C in a 2xSSC moist chamber. Slides probed with whole chromosome paints were incubated with 0,005% pepsin/10mM HCl for 10 min and the chromosomal DNA was denatured in 70% formamide/2xSSC for 3 min at 70°C. For each slide, 30ul of hybridization solution (containing 200ng of labelled probe, 50% formamide, 2xSSC, 10% dextran sulphate and 3ug of salmon sperm DNA) was denatured for 10 minutes at 85°C and allowed to pre-hybridize for 30min at 37°C, dropped on the slides, sealed with rubber cement and hybridized at 37°C in a 2xSSC moist chamber for 36h. Post hybridization, all the slides were washed in 0.2xSSC/15% formamide for 20 min at 42°C, followed by a second wash in 0.1xSSC for 15 min at 60°C and a final wash at room temperature in 4xSSC, 0.5% Tween for 10 min. Probe detection was carried out with Avidin-FITC (Sigma) or anti-digoxigenin-rhodamine (Roche), the chromosomes were counterstained with DAPI (4',6-diamidino-2-phenylindole, Vector Laboratories) and analyzed using an optical photomicroscope (Olympus BX61). The images were captured using the Image Pro plus 6.0 software (Media Cybernetics). From each individual, a minimum of 10 cells was analyzed to confirm the FISH results and estimate the number of the B chromosomes per cel.

DNA amplification, cloning and sequencing

We amplified, cloned and sequenced partial H3 histone genes and both satDNAs from different samples, including a microdissected euchromatic B chromosome (B₁), a microdissected heterochromatic B chromosome (B₂), gDNA 0B from *M. sanctaefilomenae*-Novo River (0B gDNA) and 0B gDNA from *A. fasciatus*. In addition, MS3 and MS7 satDNA sequences were also obtained from the longest autosome pair (A₁). The reactions were performed in 1x PCR buffer, 1.5mM of MgCl₂, 200μM of each dNTP, 0.5U of *Taq* polymerase (Invitrogen), 0.1μM of each primer and 50ng of DNA. The cycle to amplify these regions consisted of an initial denaturation at 95°C for 5min followed by 32 cycles at 95°C for 45s, 56°C for 30s and 72°C for 1min and a final extension of 72°C for 15min. PCR products were visualized in a 2% agarose gel and the fragment obtained from each sample was extracted from the gel, cloned into the pGEM-T Easy Vector (Promega, Madison, Wisconsin, USA). DNA sequencing was performed with Big Dye TM Terminator v3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems), following manufacturer's instructions. Although different strategies were adopted (i.e. testing different primer pairs and adding DMSO to the PCR), we failed to amplify any region of the 45S rDNA from B chromosomes, probably because GC-rich sequences are usually under amplified in WGA steps (Teruel et al. 2009).

Extraction of MS3 and MS7 from Illumina reads

In order to obtain a detailed and reliable score of haplotype abundance for MS3 and MS7 satDNAs sequences from genomic libraries, we extracted the monomers directly from the Illumina reads. Since reads are smaller than the monomer size of both satellites, we joined

the paired-reads using fastq-join (<https://code.google.com/p/ea-utils/wiki/FastqJoin>) with a minimum overlapping size of 6bp.

For MS7, we aligned the joined reads against a dimer sequence of the satDNA with the RepeatMasker software (Smit et al. 2010) and, by a custom Python script (https://github.com/fjruiizuano/ngs-protocols/blob/master/rm_getseq.py); we used the alignment information from the output file (with extension .out) to extract only the aligned region. After that, we mapped these sequences with Geneious Pro v8.04 against the dimer and extracted the central region corresponding with one monomer, by manually deleting those sequences that did not cover a whole monomer. For MS3, after Geneious mapping, we cut the ends mapped out of the reference monomer and added them to the other end to get sequences starting and ending at the same positions. We did this by another custom Python script (https://github.com/fjruiizuano/ngs-protocols/blob/master/sat_cutter.py).

Sequence analysis

Consensus sequences from forward and reverse strands were obtained using Geneious Pro v8.04 and alignments were generated using the Muscle algorithm (Edgar, 2004) under the default parameters. DNA diversity analyses, considering indels, were performed with the DnaSP v5.05 (Librado and Rozas, 2009). Minimum spanning trees were built on the basis of pairwise differences, using ARLEQUIN v3.5.1.3 (Excoffier and Lischer, 2010) and were visualized with HAPSTAR v0.7 (Teacher and Griffiths, 2011).

Comparisons of synonymous substitutions per synonymous site (dS) and non-synonymous substitutions per non-synonymous site (dN) of H3.2 from each library (B₁, B₂, A₁ and 0B gDNA) were carried out using the nonparametric Kruskal-Wallis followed by Dunn's multiple post-hoc tests, considering $\alpha=0,05$.

Results

Individuals from the two analyzed populations showed similar standard karyotype, all exhibiting $2n= 50$ biarmed chromosomes ($6m + 16sm + 28st$) with no sex-related chromosomal dimorphism. Additionally, mitotically unstable B chromosomes were observed in the genomes of all studied specimens collected at the Batalha River (BR) population, as manifested by the intraindividual variation in B number (0 – 6). However, all individuals from the Novo River (NR) population lacked B chromosomes.

In the BR population, there were two types of B chromosomes on the basis of C-banding pattern and population frequency. The most frequent variant (named B_1) shows a light C-banding pattern (similar to euchromatin in the A chromosomes), whereas the least frequent variant (B_2) shows a dark C-banding pattern (similar to heterochromatin in the A chromosomes) (Fig 1). While B_1 was found in all 23 individuals analyzed from the BR population (100% prevalence), B_2 was present only in four of them (17% prevalence) (Table 1). Both B types were found in males and females.

Similar DNA content in both B chromosome variants

Whole Chromosome Painting (WCP) with the B_1 and B_2 probes showed similar hybridization signals for both probes on both B-variants, in addition to the pericentromeric regions of about one third of the A chromosomes (Fig 1a-d). This was also apparent when double WCP with both B-derived probes was performed on metaphase cells from B-lacking NR individuals, with both hybridization signals co-locating in most cases (Fig 1e, f).

FISH mapping revealed remarkable differences between the two populations for the number of H3 histone gene clusters (all showing pericentromeric location), with only two chromosome pairs (no. 6 and 22) carrying them in the B-lacking population (NR) (Fig 2a) and seven (no. 3, 6, 14, 21, 22, 24, and 25) in the B-carrying one (BR) (Fig 2b). The FISH

analysis with the 5S rDNA probe also revealed between-population differences but, in this case, the NR population showed higher number of clusters (at pericentromeric location on chromosomes no. 1, 2, 5, 7, 17, and 18) than BR (only in chromosomes no. 1, 2, 5 and 7). On the contrary, both populations showed exactly the same pattern for the U2 snDNA, which was located in the pericentromeric region of chromosomal pairs no. 20 and 21.

The most extreme difference between the two populations was found for 18S rDNA which, in NR, was located only in the short arm of chromosome pair no. 6 whereas, in BR, distal clusters appeared on many chromosomes in homozygous (13, 14, 15, 17) or heterozygous (2, 12, 20, 22, 24, 25) state, in addition to the cluster on chromosome 6. Sequential FISH and silver impregnation (the latter indicating active rDNA clusters) showed that the rDNA cluster on chromosome 6 was always active, while most other clusters on other chromosomes (including B chromosomes) in the BR population were inactive (Table S1, Fig S1). Considering the B chromosomes, NOR activity was observed only for B₁ in about 10% of the cells analyzed in a single individual, indicating that B chromosome contribution to rRNA synthesis does not seem to play an important role in the cellular physiology of B-carrying *M. sanctaefilomenae* specimens.

Out of these tandem repeat gene families assayed, both types of B chromosomes carried only 18S rDNA and H3 histone gene sites, but the heterochromatic variant (B₂) carried a larger 18S rDNA cluster than the euchromatic one (B₁) (Figs 2 and 3).

To improve our knowledge of DNA content in B chromosomes, we searched for satellite DNA (satDNA) tandem repeats in two sets of Illumina Hiseq2000 Paired-End reads obtained from whole *M. sanctaefilomenae* genome sequencing runs from a B-lacking individual from the NR population and a B-carrying one from the BR population. Due to the fact that all individuals analyzed from the BR population carried B chromosomes, we had to analyze a B-lacking genome from the NR population. It was, however, unfortunate that the

two repeat families being present in the B chromosomes (18S rDNA and H3 histone genes) showed extensive spread across A chromosomes in the B-carrying population and thus impeded to detect changes in DNA repeat coverage between the B-carrying and B-lacking genomes. For this reason, we used the Illumina reads to search for satellite DNAs that might be useful as additional B-specific FISH markers.

Sequence clustering analysis resulted in 11,149 clusters constituting a genome proportion of 25.6% and 24.4% for the +B and 0B individuals, respectively. We designed primer pairs for seven putative satellite DNAs, PCR amplified them and selected those with a ladder-like pattern in the agarose gels. We then generated DNA probes for FISH and mapped them to A and B chromosomes (data not shown), and selected the two satDNA families showing a clustered distribution which were present in both A and B chromosomes (Fig 3e, f), henceforth named MS3 (CL27) and MS7 (CL96). The MS3 satDNA showed a consensus sequence of 186 bp with a cluster density of 0.16 (Fig S2) and did not show any similarity with the custom database. Notably, this cluster was 2x more abundant in the BR library (0.236%) than in the NR library (0.117%). The MS7 satDNA showed a consensus sequence of 100 bp with a cluster density of 0.55 (Fig S2) and yielded similarity hits with DNA/TcMar-Tc1 (54% of hits) with the *A. mexicanus* database. In terms of abundance in different libraries, there was almost no difference for this cluster between the two populations analyzed (0.0285% for BR and 0.0295% for NR). FISH analysis corroborated the abundance data and revealed that MS3 is located in the pericentromeric region of 13 chromosome pairs in the BR population and only 6 pairs in the NR population. Therefore, the higher abundance of MS3 in the B-carrying population was not due to B chromosomes (even though they actually carry this satellite) but to its presence in a twice as much number of A chromosomes. Conversely, MS7 was located in the telomeric regions of 15 pairs in both analyzed populations.

Taken together, these results indicate that both B variants contain essentially the same DNA repeats (H3 histone genes, 18S rDNA and MS3 and MS7 satDNAs). The fact that the autosome pair no. 6 is the only A chromosome carrying all these repeat families strongly points to the possibility that both B chromosomes derived from the pericentromeric region of this chromosome.

The two B chromosome variants show similar degree of mitotic instability

Since the number of B chromosomes varied among cells within a same individual, we performed an analysis of the degree of mitotic instability causing this variation. For this purpose, we used a mitotic instability index, previously developed in the migratory locust (Pardo et al. 1995), which is based on the assumption that the median number of B chromosomes in the adult represents the number of B chromosomes in the zygote stage. This mitotic instability index (MI) measures the sum of deviations in B number in a sample of cells, in respect to the median, normalized per B chromosome.

The fact that both B chromosome types contained H3 histone genes and 18S rDNA helped us to identify the two B types in 657 mitotic metaphase cells submitted to double FISH and subsequent C-banding, in 23 individuals from the BR population (mean= 29 cells per individual, SD= 6), to accurately scoring the number and type of B chromosomes (Fig 4). In each individual, we calculated the mean number of B chromosomes per cell, the median number of Bs and the mitotic instability index (MI). The results revealed that B₁ and B₂ show a similar degree of MI, but B₁ was almost nine fold more frequent than B₂ (Table 1).

A comparison of the mean number of B chromosomes and MI per individual between the present data and those previously reported for the Tiete River by (Foresti et al. 1989) and (Dantas et al. 2007), by means of the Kruskal-Wallis test, revealed significant differences for both mean (H= 21.16, df= 2, N= 46, p< 0.0001) and MI (H= 15.4, df= 2, N= 46, P= 0.0005).

Given that the three samples were taken in the same river but at different times, these results suggest a tendency of MI to increase across years, but we found an inverted V-shaped variation in the mean number of B chromosomes per individual (Fig 4). This might suggest that B chromosomes have increased their mitotic instability across years but the number of Bs per individual has reached a maximum, in consistency with the existence of a tolerance threshold.

DNA sequence analysis suggests different age for the two B variants

We PCR amplified three out of these repetitive sequences (H3 histone, and MS3 and MS7 satDNAs) from four sources (0B gDNA from NR, microdissected B₁, B₂ and A₁ chromosomes from BR). Notably, we could not amplify any fragment of the major ribosomal sites from microdissected libraries as well as histone H3 genes from the A₁ chromosome. These PCR products were cloned and sequenced in both directions. After discarding primer regions, a total of 318-328bp, 183-186bp and 93-106bp from the H3 histone, MS3 and MS7, respectively, were obtained from several clones and Illumina reads (Table 2). In order to minimize the impact of possible PCR/sequencing errors, we discarded the singletons for subsequent analyses.

The H3 histone gene sequences obtained from the B-types showed two distinct isotypes: H3.2 (28 clones, 11 of which showed a 10bp deletion, and 15 were polymorphic in respect to *Danio rerio* H3.2 amino acids sequence; UniProt accession number Q4RF4) and H3.3 (6 clones, all of which were non-defective and identical to *Danio rerio* H3.3 amino acids sequence; UniProt accession number Q6PI20). In the case of the 0B gDNA, however, we obtained only the H3.2 isotype (19 clones, 2 of which were defective showing a 10bp deletion and 1 polymorphic in relation to *Danio rerio* H3.2 amino acids sequence).

Calculation of dN for H3.2 sequences in each group revealed significant differences between the 0B gDNA sequences and those from B₁ and B₂ (H= 161.7, dF= 2, N= 47, p<0.0001). Post-hoc comparisons (not shown) failed to show significant differences between both B chromosome types, but these latter showed significantly higher dN values than the H3.2 sequences coming from the 0B genome (Table 3). Conversely, dS showed significant differences between 0B, B₁ and B₂ sequences (H= 40.12, dF= 2, N= 47, p<0.0001), and post-hoc comparisons revealed significant differences in all cases (not shown). It was remarkable that dS was almost twice as much higher in B₁ than B₂, suggesting that the former is probably older and has had longer time to accumulate synonymous changes. In addition, the dN/dS ratio in the B₁ and B₂ sequences was higher than that in the 0B gDNA, suggesting that purifying selection is relaxed in the B chromosomes. The minimum spanning tree built with H3.2 haplotypes showed a certain degree of differentiation of the B-derived sequences (Fig 5).

In the case of the MS3 and MS7 satDNAs, we obtained 62 and 57 PCR clones, respectively. In addition, 193 and 916 monomers were extracted from the Illumina reads for MS3 and MS7 satDNAs, respectively (Table 2). We aligned the Illumina reads for each satDNA and built minimum spanning trees considering haplotype relative abundance, on which we traced the PCR sequences obtained from the different libraries (Fig 6). It is expected that Illumina reads provide accurate estimates about haplotype abundance without the bias of PCR amplification. In both cases, the satDNAs amplified from *A. fasciatus* showed the highest divergence (as inspected in the nucleotide alignment), as expected for DNA sequences submitted to concerted evolution. The most abundant haplotypes for MS3 and MS7 in *M. sanctaefilomenae*, found in the Illumina reads (Fig 6), were present in the 0B and 6B genomes. However, in MS3, this haplotype was found only in the B₂ chromosome and not in the B₁ chromosome or the 0B genomic DNA. The minimum spanning trees showed higher conservatism for MS7 (Fig 6b) than MS3 (Fig 6a). Remarkably, the tree for the MS3 satDNA

showed that the most abundant haplotype in the 0B and 6B Illumina-sequenced genomes was present in the PCR sequences obtained from the B₂ chromosome but not in those coming from the B₁ chromosome. This difference should not be expected if one B-type would have derived from the other. In addition, the absence of the most common haplotype in B₁ but its presence in B₂ would be consistent with an independent and more recent origin for the latter (conceivably from the same A chromosome). The tree for the MS7 satDNA was consistent with the former conclusion since all DNA sequences found in the B₂ chromosome corresponded to the most abundant haplotype in 0B and 6B genomes, whereas only 6% of the DNA sequences obtained from the B₁ chromosome corresponded to the former haplotype and the remainder belonged to a different haplotype showing one mutational difference.

Discussion

To date, all studied populations of *M. sanctaefilomenae* harbored B chromosomes (Dantas et al. 2007; de Brito Portela-Castro et al. 2000; Foresti et al. 1989; Hashimoto et al. 2012; Scudeler et al. 2015). Since the first description, several authors have proposed the occurrence of three B-types in this species, distinguishable by the C-banding patterns (Foresti et al. 1989; Hashimoto et al. 2012). Recently, it was reported that some B chromosomes in this species carry 18S rDNA clusters, while others do not (Scudeler et al. 2015). In our study, C-banding and FISH allowed us to clearly identify only two different B-types, which showed differences in C-banding patterns, abundance of 18S rDNA and population frequency. In the present paper, we also analyzed a novel population collected at the Paranapanema river basin and, for the first time in this species, no B chromosome-bearing individuals were found. However, the question whether B chromosomes ever existed but were eliminated from this specific population, or if B chromosomes originated after the split of these populations, and have not yet reached the NR population, remains open.

Individuals carrying different variants of supernumerary chromosomes have been reported in several species. In fish, *Prochilodus lineatus*, *Astyanax scabripinnis*, *A. goyacensis* and *M. sanctaefilomenae* are the most studied examples (Artoni et al. 2006; dos Santos et al. 2013; Foresti et al. 1989; Salvador and Moreira-Filho, 1992), while the grasshopper *Eyprepocnemis plorans* constitutes the most complete example of B chromosome diversification (Cabrero et al. 1999; López-León et al. 1993). In this latter species, several studies showed the existence of numerous B variants (more than 40) that probably arose from a common ancestor B chromosome (Cabrero et al. 1999; Cabrero et al. 2014). In the B chromosome system of the fish *M. sanctaefilomenae*, the B₁ and B₂ variants showed a similar degree of mitotic instability, whereas the frequency of the euchromatic variant (B₁) was almost nine fold higher than that of the heterochromatic one (B₂). Since mitotic instability is a frequent drive mechanism (Camacho, 2005), we can infer that it does not make the difference between B₁ and B₂, although other conceivable drive mechanisms (e.g. meiotic) should be analyzed in future experiments. Notwithstanding, our sequence analysis of the MS3 and MS7 satellite DNAs suggests that B₁ is older than B₂, and this would be consistent with the higher frequency of the former in the BR population. If B₂ would actually be younger, we should expect that it would increase in frequency during next years, most likely at the expense of B₁ since B frequency appears to have reached a maximum in this population. This population thus provides the opportunity for witnessing the possible replacement of a B variant for another, in resemblance with that previously reported in the grasshopper *Eyprepocnemis plorans* (Zurita et al. 1998).

Sharing repetitive DNAs between A and B chromosomes is a common feature, as evidenced in different animals, like fish, grasshoppers and mammals (Artoni et al. 2006; Graphodatsky et al. 2005; Mestriner et al. 2000; Silva et al. 2014; Teruel et al. 2010). The composition of B chromosomes has been used to identify the probable ancestral chromosome

in the host species (Artoni et al. 2006; Bueno et al. 2013; Silva et al. 2014; Teruel et al. 2010). Our WCP and FISH mapping results have shown that the B₁ and B₂ chromosomes in *M. sanctaefilomenae* are composed of the same repetitive DNA sequences, suggesting their common intraspecific origin. Recently, Scudeler et al. (Scudeler et al. 2015) suggested an intraspecific origin of B chromosomes in this species, based on WCP results indicating the presence of DNA sequences shared between A and B chromosomes. Our present results are consistent with this conclusion. However, they also suggested the independent origin for different B variants in this species, because they were not painted with the only B-probe employed. We cannot rule out the presence of other B variants in these authors' samples, but our present analysis has revealed that all observed B chromosomes in the 23 individuals analyzed in the BR population contained H3 genes and 18S rDNA.

Since the autosomal pair No. 6 is the only one in the A karyotype that exhibits co-located histone, 18S rDNA, MS3 and MS7 sites (i.e. the repetitive DNA sequences contained in the B chromosomes), we suggest that this pair might be the probable B chromosome ancestor. Remarkably, the minimum spanning trees obtained from MS3 and MS7 satDNAs sequences also supported the hypothesis of intraspecific origin due to the high similarity, and even shared haplotypes, between the sequences allocated in the A and B chromosomes of *M. sanctaefilomenae*. However, the fact that 52% of MS3 and 100% of MS7 sequences observed in B₂ corresponded to the most frequent haplotype found in the Illumina reads from the B-lacking and B-carrying populations, whereas these figures were 0 and 6%, respectively, for B₁, suggests that the two B chromosome arose independently from autosome 6, so that the most recent B type (B₂) still conserves many satDNA repeats of the commonest haplotype in the A genome found in even distant populations. The observed differences in satellite DNA sequences between the two B variants also suggest that concerted evolution might act separately for each B type. Remarkably, nucleotide diversity for MS3 and MS7 is lower in B₁,

perhaps because its higher chance for sequence homogenization due to its higher age and population frequency, the latter allowing the presence of two or more Bs in the germ line which could form bivalents during meiosis thus facilitating unequal crossovers and sequence homogenization. The fact that B₁ shows almost twice dS value for H3.2 histone genes than B₂ reinforces the conclusion that B₁ is older.

B chromosomes containing 18S rRNA genes have already been described in different fish species (Poletto et al. 2010; Silva et al. 2014), but their activity has only been shown in the euchromatic variant (B₁) in *M. sanctaefilomenae* (Hashimoto et al. 2012). It is interesting to note that the main structural difference between the eu- (B₁) and heterochromatic (B₂) variants is the higher content of rDNA in the latter. Similarly, in the A chromosomes of different salmonid species, rDNA loci with smaller FISH signals show faint C-band heterochromatin, while larger clusters are coincident with strongly positive C-bands (Fujiwara et al. 1998; Pendás et al. 1993). Such condition is probably related with the interspersed organization of the rDNA within the repeated DNA sequences of the heterochromatin, although the possibility that the rDNA itself contributes to heterochromatin formation cannot be discarded (Pendás et al. 1993). In this context, the B chromosomes of *M. sanctaefilomenae* thus provide a model to test the relationship between rDNA and heterochromatin content and their roles in determining the C-banding pattern.

Other multigene families have also been found in B chromosomes of several species (Menezes-de-Carvalho et al. 2015; Ruíz-Estévez et al. 2013; Silva et al. 2014; Teruel et al. 2010). Regarding histone genes, only the H3.2 subtype has hitherto been reported in B chromosomes (Silva et al. 2014; Teruel et al. 2010). In *M. sanctaefilomenae*, two non-defective (thus potentially active) H3 histone types (H3.2 and H3.3) were found in the eu- and heterochromatic B chromosome variants. Remarkably, while H3.2 was represented by several haplotypes in both B-types, the six H3.3 clones analyzed from the two B chromosome

variants showed exactly the same DNA sequence. Interestingly, the coding nature of histone genes provides additional inferences about the DNA sequences contained in the B chromosomes, in terms of the dN/dS ratio. In general, higher dN/dS ratios are expected for coding DNA sequences residing in B chromosomes, compared to the same sequences in the A chromosomes, since selection is assumed to be relaxed in B chromosomes due to their dispensable nature. This has been reported for the H3 and H4 histone genes in the grasshopper *L. migratoria*, with dN/dS ratios being 2,23 and 1,72 higher in the B chromosome, respectively (Teruel et al. 2010), and for the H1 histone gene in the fish *A. paranae*, with dN/dS ratio being 3 times higher in the B chromosome (Silva et al. 2014). Our present results are consistent with these previous observations since they showed 5.77 (B₁) and 3.78 (B₂) higher dN/dS ratios in the B chromosomes compared to the 0B genome, indicating that purifying selection is relaxed for the H3.2 genes located in the B chromosomes. This suggests that H3 genes in the B chromosomes are most likely inactive.

In conclusion, B chromosome content for several types of repetitive DNA sequences (18S rDNA, H3 histone genes and two satellite DNAs), and its comparison with that in A chromosomes by means of FISH mapping, chromosome painting, and DNA sequencing (by Sanger and Illumina methods) has revealed that the Neotropical fish *M. sanctafilomenae* harbors two B chromosome variants differing in C-banding pattern, frequency and abundance of 18S rDNA. Both B variants presumably derived independently from the same A chromosome (autosome no. 6) but the heterochromatic variant shows signs of being younger than the euchromatic one. Finally, both B variants showed higher dN/dS ratios for the H3.2 histone gene, suggesting that purifying selection is relaxed for the B-sequences, as expected if they were mostly inactive.

Acknowledgements

We thank to Renato Devide for technical assistance.

References

Corresponding references are listed in the end of the thesis

Supporting information Captions

Figure S1: Metaphase plates of *M. sanctaefilomenae* from Batalha River after FISH with major ribosomal probes (**a, c, e**) and sequential Ag-NOR staining (**b, d, f**). Note that active NOR sites were observed only in the B₁ variant.

Figure S2: Graph layouts (clusters CL27 and CL96) identified by NGS.

Table 1. Intra- and interindividual variation in the number of B₁ and B₂ chromosomes in the BR population.

Id no.	Sex	Cells with 0-6 B ₁									Cells with 0-2 B ₂						Total	
		0	1	2	3	4	5	6	Mean	Median	MI	0	1	2	Mean	Median		MI
69616	M	2	2	12	3	2			2.05	2	0.31	21			0	0		21
69618	M	0	1	1	5	7			3.29	3.5	0.20	14			0	0		14
69626	F	3	13	13	2				1.45	1	0.65	31			0	0		31
69628	F	3	14	8	2	3			1.60	1	0.80	0	12	18	1.60	2	0.20	30
69676	M	0	4	21	13				2.24	2	0.22	38			0	0		38
69678	F	8	15	6					0.93	1	0.48	2	16	11	1.31	1	0.45	29
69679	M	4	0	8	18	2			2.44	3	0.23	3	22	7	1.13	1	0.31	32
69680	F	2	4	14	4				1.83	2	0.25	24			0	0		24
69682	M	4	14	4					1.00	1	0.36	22			0	0		22
69637	M	2	3	10	10				2.12	2	0.34	25			0	0		25
69693	F	0	15	14	3				1.63	2	0.28	32			0	0		32
69695	M	3	20	5					1.07	1	0.29	28			0	0		28
69696	F	0	4	24	5				2.03	2	0.14	33			0	0		33
69710	F	0	2	8	18	2			2.67	3	0.16	30			0	0		30
69713	M	1	5	13	9				2.07	2	0.29	28			0	0		28
69714	F	3	26	4					1.03	1	0.21	33			0	0		33
69729	M	0	6	9	8	6	2	1	2.75	3	0.35	32			0	0		32
69730	F	0	0	4	11	11			3.27	3	0.19	26			0	0		26
69731	M	11	18	10					0.97	1	0.54	39			0	0		39
69739	M	0	13	16	2				1.65	2	0.24	0	29	2	1.06	1	0.06	31
69740	M	0	5	4	9				2.22	2.5	0.31	18			0	0		18
69741	F	1	5	10	12				2.18	2	0.34	28			0	0		28
69743	F	0	1	10	20	3			2.74	3	0.15	34			0	0		34
Mean									1.97	2	0.32				0.22	0.22	0.26	
SE									0.15	0.17	0.03				0.25	0.26	0.08	

MI= mitotic instability index, F= female, M= male, SE= standard error.

Table 2. Genetic variation found in MS1, MS2, H3.2 and H3.3 sequences obtained through microdissection and PCR amplification, and directly from reads.

	Sample	N	Length	Hap	Hd	π	Def	
MS3	0B reads	129	183-184	38	0.8975	0.06334	-	
	6B reads	64	183-184	28	0.9127	0.07408	-	
	Pair No.1	19	183-185	2	0.3509	0	-	
	B ₁	15	184-185	6	0.8762	0.02174	-	
	B ₂	17	183-184	5	0.7059	0.04243	-	
	0B gDNA	6	184-186	3	0.800	0.05448	-	
	<i>A. fasciatus</i>	5	184	5	1	0.02391	-	
MS7	Sample	N	Length	Hap	Hd	π	Def	
	0B reads	394	96-100	110	0.9425	0.02061	-	
	6B reads	522	93-100	129	0.9266	0.02041	-	
	Pair No.1	13	98	1	0	0	-	
	B ₁	16	100	2	0.1250	0.00125	-	
	B ₂	18	100	1	0	0	-	
	0B gDNA	6	100	3	0.7333	0.01133	-	
<i>A. fasciatus</i>	4	106	2	0.6667	0.00629	-		
H3.2	Sample	N	Length	Hap	Hd	π	Def	Pol
	B ₁	10	318-328	5	0.8222	0.01726	2	7
	B ₂	18	318-328	7	0.8627	0.01247	9	8
	0B gDNA	19	318-328	6	0.6608	0.00376	2	1
H3.3	Sample	N	Length	Hap	Hd	π	Def	
	B ₁	1	328	1	0	0	0	
	B ₂	5	328	1	0	0	0	

N= Number of clones, Hap= Number of haplotypes, Hd= Haplotype diversity, π = Nucleotide diversity per site, Def= Defective copies (Stop codons or gapped sequences), Pol = Polymorphic sequences (sequences with different amino acids in relation to *Danio rerio* H3.2).

Table 3: Number of synonymous and non-synonymous substitutions per synonymous (dS) and non-synonymous (dN) site, respectively, observed in the DNA sequences of H3.2 histone genes.

H3.2				
Sample	N	dN	dS	dN/dS
B ₁	10	0,00623±0,004615	0,04031±0,02544	0,1545
B ₂	18	0,00626±0,00399	0,02641±0,03286	0,2370
0B gDNA	19	0,00045±0,00132	0,00699±0,00845	0,0643

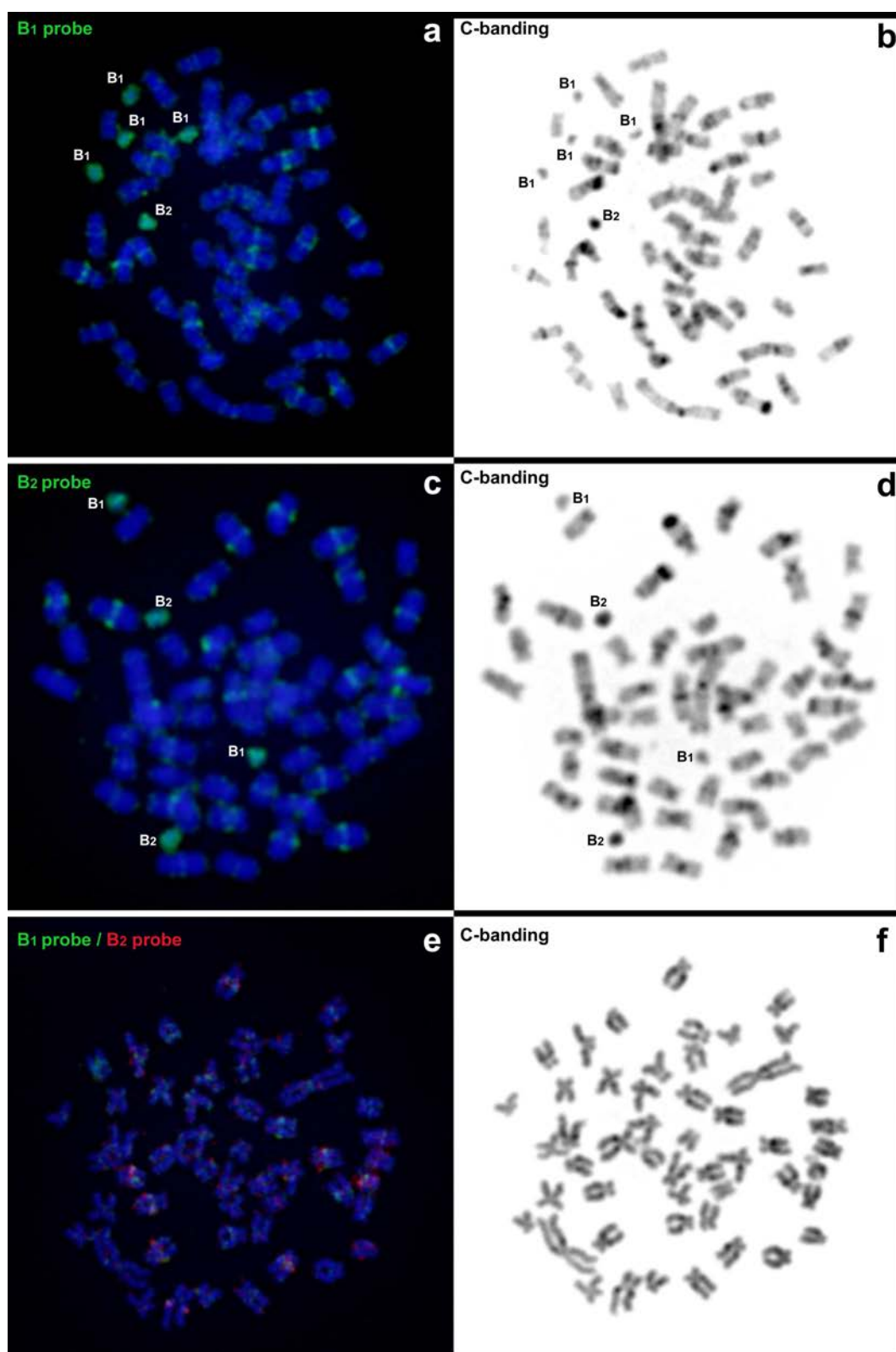


Fig 1: Metaphase plates of *M. sanctaefilomenae* after whole chromosome painting with B₁ and B₂ probes and their respective sequential C-banding. (a-d) represent *M. sanctaefilomenae* from Batalha River. (e-f) represent *M. sanctaefilomenae* from Novo River.



Fig 2: *M. sanctaefilomenae* karyotypes constructed from mitotic metaphase cells submitted to FISH with different repetitive DNA probes. a) Novo River population. b) Batalha River population.

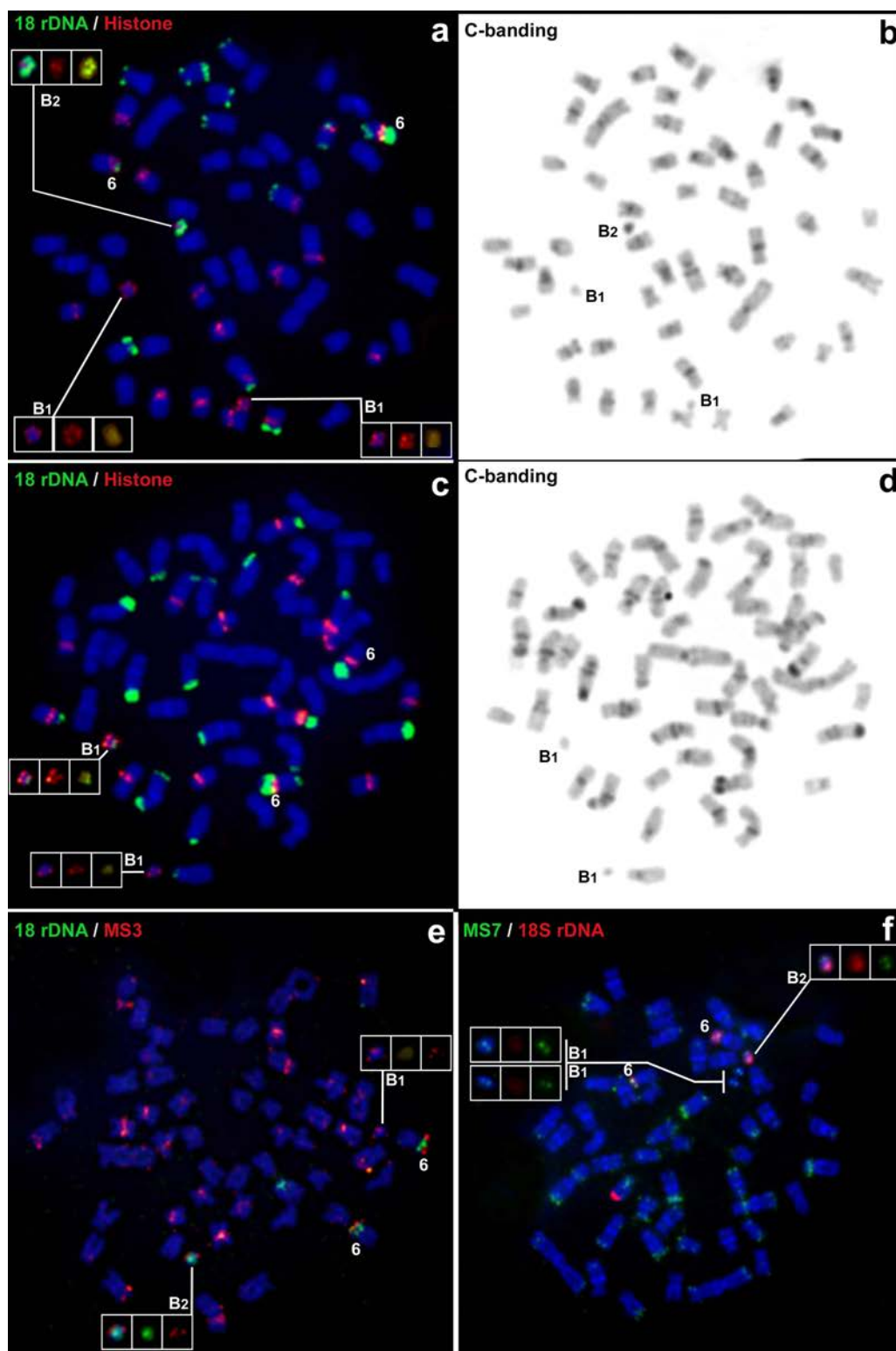


Fig 3: Metaphase plates of *M. sanctaefilomenae* from Batalha River after FISH with different repetitive probes and sequential C-banding to show the clustering of 18S rDNA, H3 histone and satDNAs in the different B chromosomes. Note that B₁ is euchromatic showing small blocks of 18S rDNA and B₂ is heterochromatic showing larger amounts of 18S rDNA.

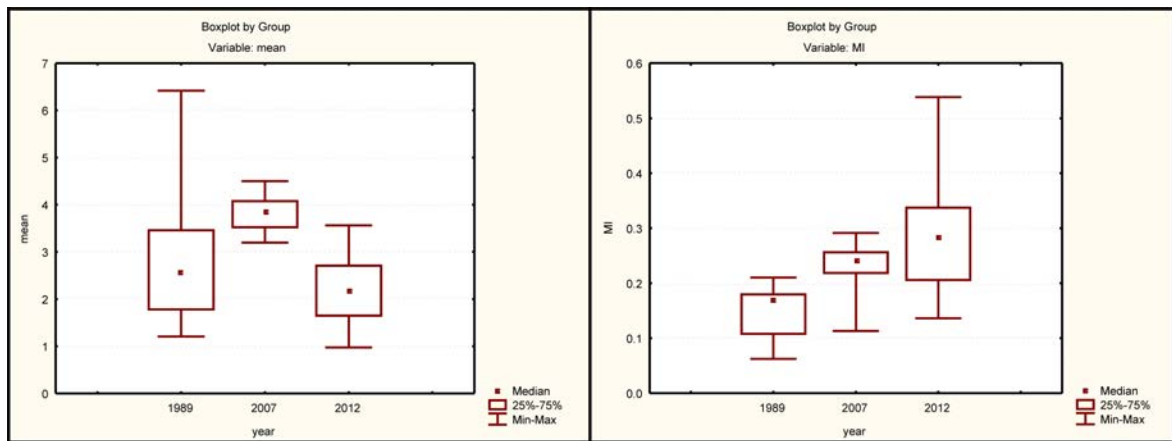


Fig 4: Comparisons of the mean number of B chromosomes and mitotic instability index (MI) per individual between the present data and those previously reported for the Tiete River by Foresti et al. (1998) and Dantas et al. (2007).

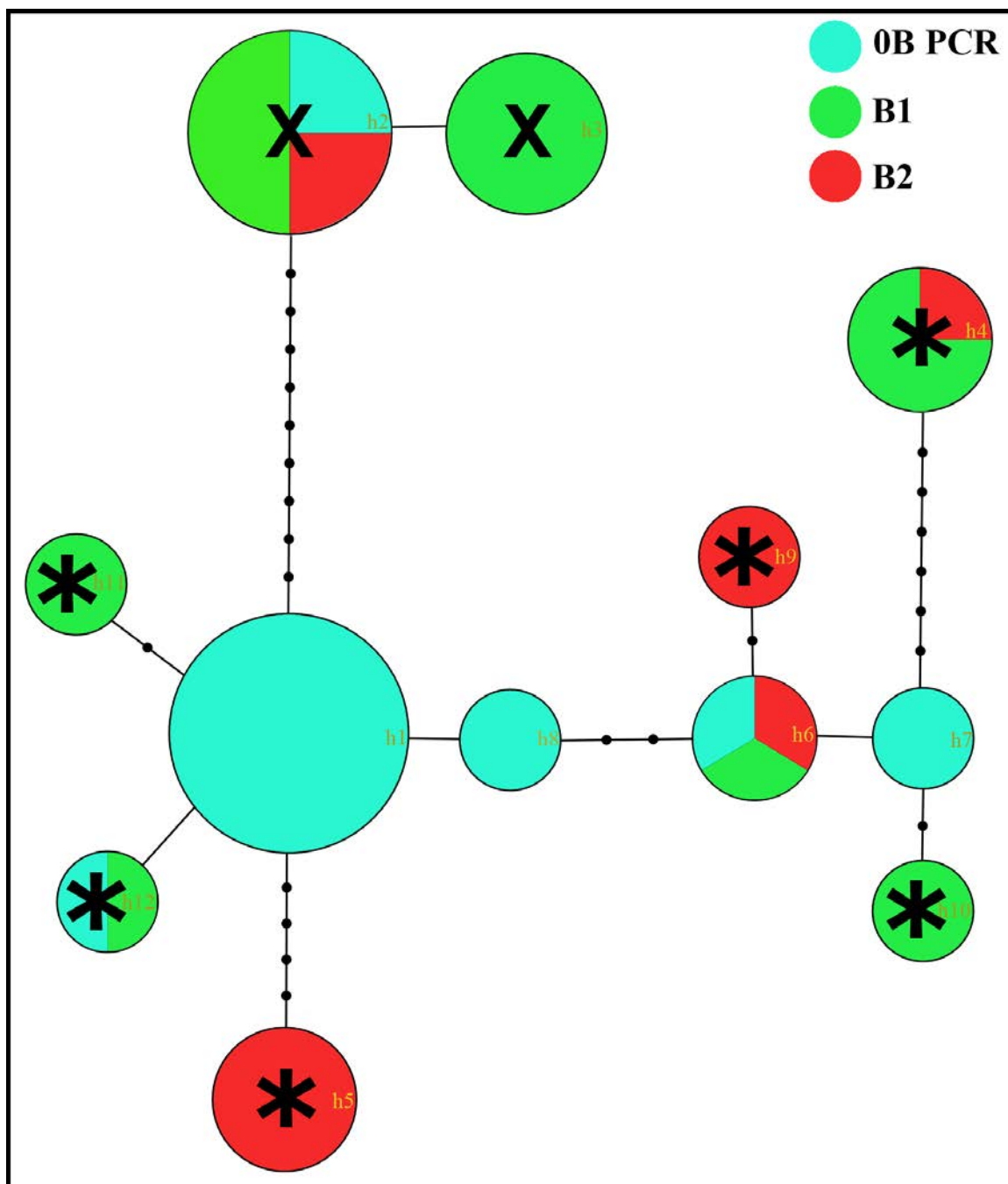


Fig 5: Minimum-spanning trees (MSTs) showing the relationships between the different haplotypes of H3.2 obtained from different libraries. Colored circles represent haplotypes and each black dot represents a mutational step. “X” inside the circles indicates defective copies (a 10 bp deletion), and “*” inside the circles indicates polymorphic amino acids in comparison with *Danio rerio* H3.2 amino acids sequence (UniProt accession number Q4RF4).

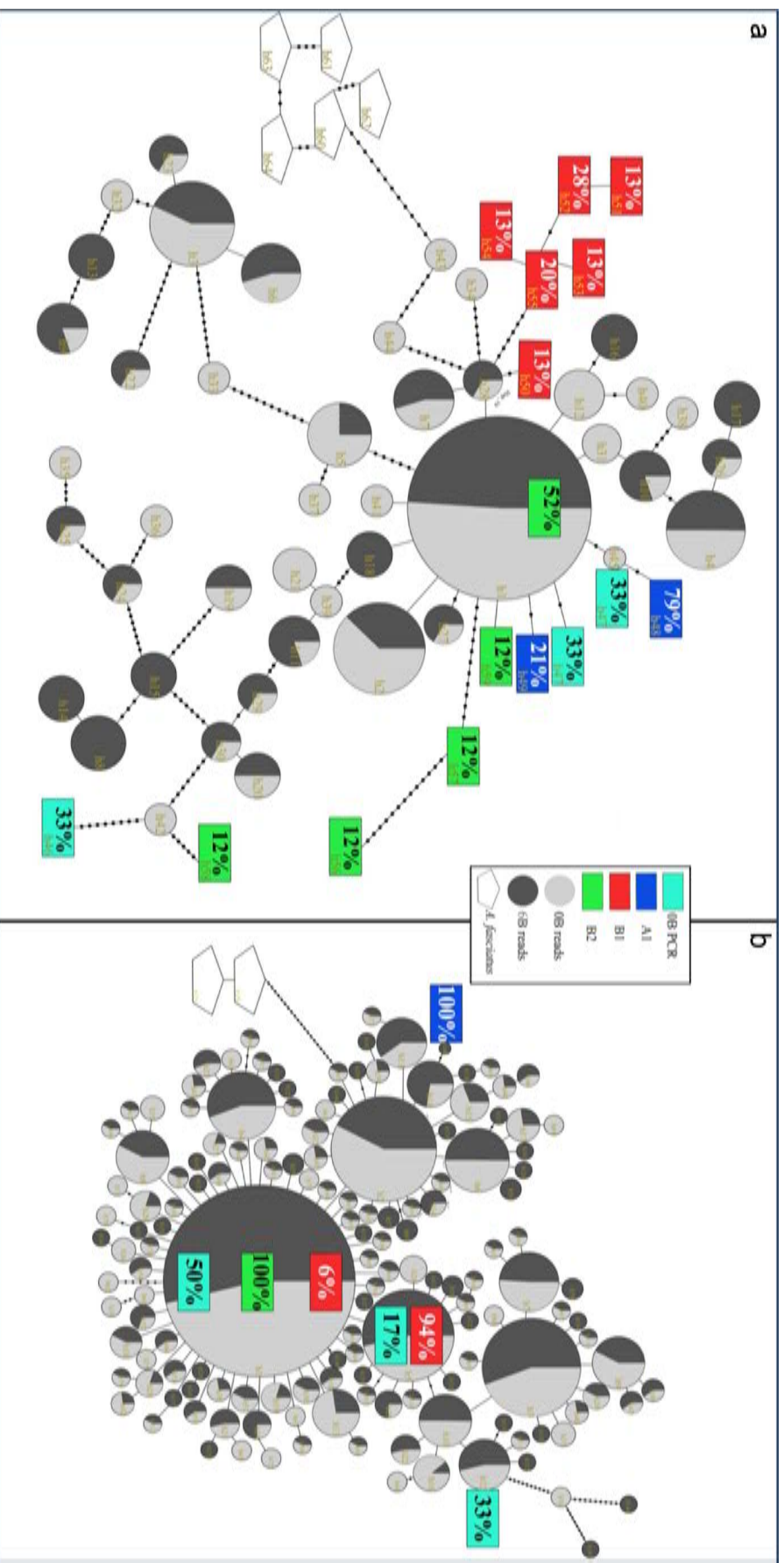


Fig 6: Minimum spanning trees (MST) showing the relationships between the different haplotypes of MS3 and MS7 satDNAs obtained from distinct libraries. Haplotypes retrieved directly from Illumina reads are represented by light/dark gray circles, and their circle diameter is proportional to their abundance, whereas PCR-amplified haplotypes are represented by colored rectangles with the percentage of clones corresponding to each haplotype. Each black dot represents a mutational step.

§Table S1: Intrapopulational polymorphism of 18S rDNA location in *M. sanctaeflorenae*. Gold shaded, the constant Ag-NOR pair No. 6.

Samples	Chromosomes																									Total	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
69615	-	-	-	p*	q	p	q*	p	-	-	p*	pq*	q*	-	q*	-	-	-	-	-	-	-	-	-	-	14	
69616	-	-	-	-	-	p	-	q*	-	-	-	q*	-	-	q*	p*	p	-	-	-	-	q*	-	-	p*	q*	11
69618	-	-	-	q*	-	p	-	-	-	-	-	q*	-	-	-	-	-	-	-	-	-	-	-	-	-	4	
69626	-	p*	-	-	-	p	-	-	-	-	p*	q*	-	q*	q*	q*	-	-	-	p*	-	-	q*	-	-	9	
69628	-	-	q*	-	-	p	-	-	-	-	q	-	-	q*	p	-	-	-	-	-	-	-	p*	-	-	9	
69676	-	-	q*	-	q*	p	-	-	-	-	q*	q*	-	q	-	-	-	-	-	-	-	-	-	-	-	8	
69678	-	-	-	-	-	p	-	-	-	-	-	q*	-	-	-	-	-	-	-	-	-	-	-	-	-	3	
69679	-	pq*	q*	-	-	p	-	-	-	-	pq*	q*	-	q*	-	-	-	q*	-	-	-	-	-	-	-	8	
69680	-	-	-	-	-	p	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	
69682	-	-	-	-	q*	p	-	-	-	-	-	q*	-	q	-	-	-	-	-	-	-	-	-	-	-	6	
69636	-	pq*	-	q	pq	p	-	-	-	-	-	q*	p*	-	pq*	-	q*	-	p*	-	p*	-	-	-	-	13	
69637	-	-	-	q	-	p	-	q*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	
69693	-	-	q*	-	-	p	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	
69694	-	-	q*	-	-	pq	-	p*	-	p	-	q*	-	q*	-	q*	-	-	-	-	-	-	-	-	-	8	
69695	-	-	-	-	-	p	-	-	-	-	-	q*	-	-	-	-	-	-	-	-	-	-	-	-	-	3	
69696	-	-	-	-	-	p	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	
69710	q*	p*	-	pq*	-	p	-	q*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	
69711	-	p*	-	q*	-	p	-	q*	-	-	q*	q*	-	p*	-	-	-	q	-	-	-	-	-	-	-	9	
69712	-	-	-	q*	-	p	-	-	-	-	pq*	q*	-	p*	-	-	-	-	-	-	-	-	-	-	-	5	
69713	-	-	-	-	p	p	-	-	-	-	-	q*	p*	-	-	-	-	-	-	-	-	-	-	-	-	6	
69714	-	-	q*	-	-	p	-	-	-	-	p*	q*	-	q*	-	q*	-	-	-	-	-	-	-	-	-	6	
69729	-	-	-	-	p*	p	-	-	-	-	-	q	-	-	q*	p*	-	-	-	-	-	-	-	-	-	7	
69730	-	-	-	q*	-	p	-	-	-	-	pq*	-	p*	-	-	-	-	-	-	-	-	-	-	-	-	5	
69731	-	p*	-	-	-	p	-	-	-	-	p*	q	q	q	q	p	q*	-	q*	-	p*	-	p*	q*	-	19	
69732	-	-	q*	-	-	p	-	-	-	-	q*	-	p*	q*	p*	-	-	-	p*	-	-	-	-	-	-	8	
69739	-	-	q*	p*	-	pq*	-	-	-	-	p*	-	pq	-	p*	-	-	-	-	-	-	-	-	p	-	10	
69740	-	q*	q*	pq*	-	pq*	-	-	-	-	q	q	p*	-	p*	-	-	-	-	-	-	-	-	-	-	12	
69741	-	-	-	-	q*	p	-	-	-	-	p	-	-	-	p*	-	-	-	-	-	-	-	-	-	-	6	
69743	-	-	-	q*	-	pq*	q*	-	-	-	-	-	-	p	-	p*	-	-	-	-	-	-	-	-	-	7	

p: short arms, q: long arms, pq: both arms, *: heteromorphic pair

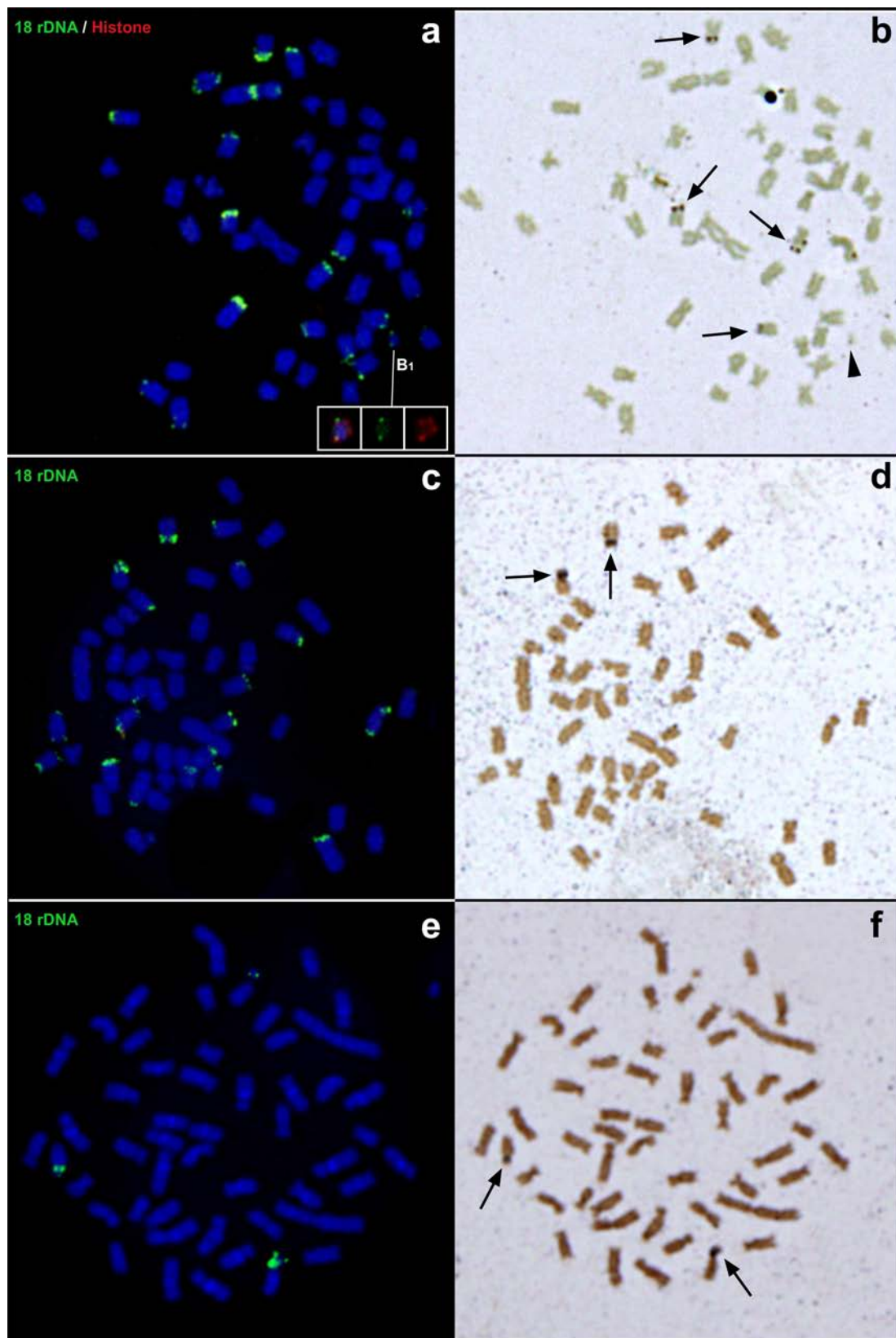


Fig S1: Metaphase plates of *M. sanctaefilomenae* from Batalha River after FISH with major ribosomal probes (a, c, e) and sequential Ag-NOR staining (b, d, f). Note that active NOR sites were observed only in the B₁ variant.

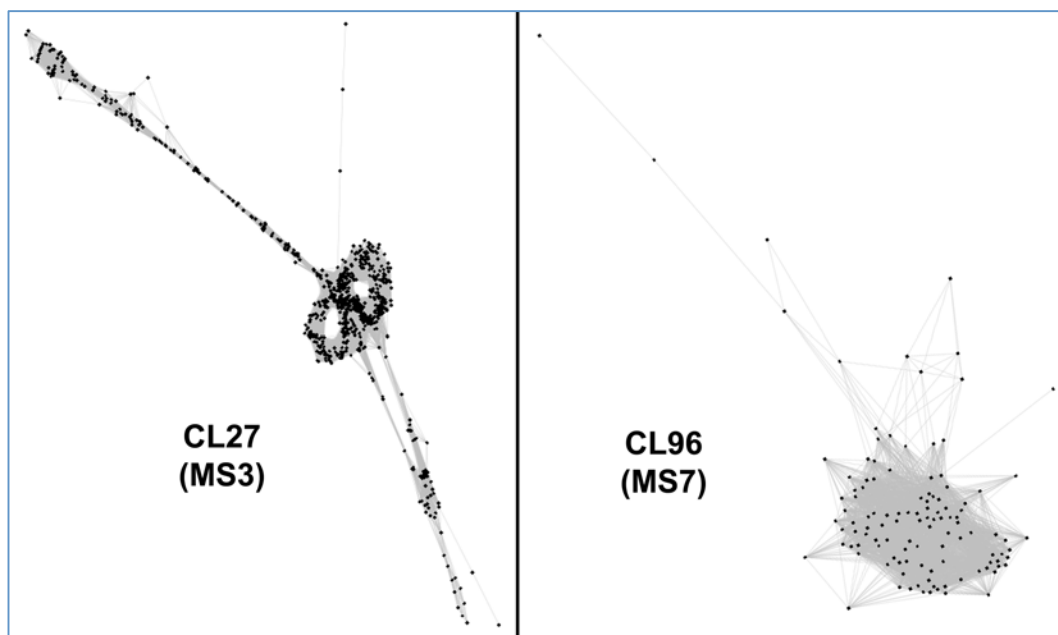


Fig S2: Graph layouts (clusters CL27 and CL96) identified by NGS.

Capítulo 2

Microdissecção e sequenciamento completo do cromossomo B₁ de *Moenkhausia sanctaefilomenae* (Teleostei, Characidae).

Resumo

Cromossomos B ou supranumerários estão presentes em aproximadamente 15% dos organismos eucariotos e, de forma geral, não afetam o desenvolvimento normal dos organismos. Sua natureza não recombinante acarreta processos degenerativos semelhantes aos envolvidos na evolução de cromossomos sexuais e incluem processos de degeneração gênica, heterocromatinização e acúmulo de sequências de DNA repetitivo. Por esta razão, a grande maioria dos estudos realizados têm tido como foco a porção repetitiva dos cromossomos B, enquanto o conhecimento sobre a porção gênica destes elementos genômicos é ainda restrito a poucas espécies. No presente estudo, o sequenciamento completo do cromossomo B₁ eucromático da espécie de peixe *Moenkhausia sanctaefilomenae* foi realizado visando a uma identificação e caracterização global inicial deste elemento genômico. Os resultados preliminares mostraram que alguns *scaffolds*, genes e elementos repetitivos estão putativamente localizados neste cromossomo. Deve-se destacar que os diferentes *scaffolds* mapeados em B₁ correspondem a diferentes grupos de ligação no genoma de *Astyanax mexicanus*, indicando que este elemento supranumerário é um depósito de sequências de diferentes regiões genômicas.

Introdução

Cromossomos B ou supranumerários são elementos extras ao complemento padrão A e não são essenciais para o desenvolvimento e/ou crescimento normal dos organismos que os

portam (Jones e Rees 1982; Shaw e Hewitt 1990). Estes elementos apresentam comumente um comportamento mitótico e meiótico irregular que permite o seu acúmulo nas células germinativas, evidenciando uma herança não mendeliana com taxas de transmissão acima de 0.5 (Camacho et al. 2000; Camacho 2005). Por apresentarem tais características os cromossomos B são considerados parasitas intracelulares que, durante o seu ciclo de vida, estão em uma “corrida armamentista” constante contra os elementos do genoma A das espécies hospedeiras (Camacho et al. 2000; Houben et al. 2013).

A origem destes cromossomos pode ocorrer de diferentes formas, aparecendo como derivação de cromossomos autossômicos ou sexuais, de forma inter- ou intraespecífica (revisão em Camacho et al. 2000). Subsequentemente, por apresentarem taxas de recombinação reduzidas, a evolução molecular e diferenciação dos cromossomos B ocorrem por processos análogos à “Catraca de Muller” (*Muller's ratchet*), resultando em silenciamento gênico, processos de heterocromatinização e acúmulo de sequências de DNA repetitivo (Green 1990), em um processo evolutivo que se assemelha à evolução dos cromossomos sexuais (Camacho et al. 2000; Houben et al. 2013). No entanto, nos últimos anos, um conjunto crescente de evidências vem apontando também para a existência de diversos genes nos cromossomos B de diferentes espécies, muitos deles apresentando-se com alta integridade e até mesmo funcionais (Yoshida et al. 2011; Martis et al. 2012; Banaei-Moghaddam et al. 2013; Trifonov et al. 2013; Makunin et al. 2014; Valente et al. 2014).

Em peixes Neotropicais, uma grande diversidade de sistemas de cromossomos B já foi reportada, podendo ser encontrados em espécies de diferentes famílias (Oliveira et al. 2009). Entretanto, deve-se destacar que a maioria dos estudos realizados é de caráter descritivo, enquanto análises mais refinadas sobre a origem e evolução destes elementos, que incluem pintura cromossômica e/ou análises de sequências repetitivas são ainda restritas a poucas espécies (Mestriner et al. 2000; Vicari et al. 2011; Silva et al. 2014; Sampaio et al. 2015;

Capítulo 1). Nesse contexto, o conhecimento sobre a existência de possíveis genes de cópia única presentes nestes cromossomos em peixes é ainda mais escasso, ou até mesmo inexistente.

A espécie *Moenkhausia sanctaefilomenae* (Teleostei, Characidae) está amplamente distribuída pelos rios da bacia do Rio Paraná. Notavelmente, todas as populações analisadas citogeneticamente desta espécie apresentam microcromossomos B em suas células, exceto a população coletada em componente da bacia do Rio Paranapanema (Foresti et al. 1989; Portela-Castro et al. 2001; Dantas et al. 2007; Hashimoto et al. 2012; Scudeler et al. 2015; Capítulo 1). Em estudo recente, Utsunomia et al. (Capítulo 1) observaram a ocorrência de duas variantes de cromossomos B com origens distintas em exemplares coletados em componentes da bacia do Rio Tietê, sendo uma eucromática (B₁) e outra heterocromática (B₂), ambas portadoras das famílias multigênicas DNAr 18S e histona H3 e dos DNAs satélites MS3 e MS7. Assim, visando aprofundar o conhecimento sobre o conteúdo genômico do cromossomo supranumerário eucromático de *M. sanctaefilomenae*, o sequenciamento completo da variante B1 microdissecada através da plataforma Illumina HiSeq foi realizado.

Material e Métodos

Amostragem, obtenção de cromossomos mitóticos e extração de DNA

No presente trabalho, foram coletados 23 exemplares da espécie de peixe *M. sanctaefilomenae* no rio Batalha (22°7.02'S, 49°16.01W), bacia do Rio Tietê, próximo à cidade de Bauru-SP (vide Anexo I). Do mesmo modo, 16 exemplares desta espécie também foram coletados no Rio Novo (22°28'13.32"S, 49°55'26.17"W), componente da bacia do Rio Paranapanema, próximo à cidade de Ocaçu-SP. Os procedimentos para coleta, manutenção e análise dos animais foram realizados de acordo com as regras internacionais em

experimentação animal, seguidas pela Universidade Estadual Paulista. De cada indivíduo analisado foram retirados fragmentos de tecido renal para estudos citogenéticos e também fragmentos de tecido do fígado, que foram fixados em etanol 95% para estudos moleculares posteriores. Os exemplares utilizados na amostragem foram fixados em formol 10%, conservados em álcool 70% e, após identificação, foram depositados na coleção de peixes do Laboratório de Biologia e Genética de Peixes do Departamento de Morfologia, UNESP de Botucatu-SP.

Os cromossomos mitóticos foram obtidos de suspensões celulares da porção anterior do rim de acordo com Foresti et al. (1981). O DNA genômico (gDNA) foi obtido dos fragmentos de fígado utilizando o kit *Promega Wizard Genomic DNA Purification Kit*, de acordo com as instruções do fabricante.

Microdissecção cromossômica

As suspensões celulares foram pingadas em lamínulas limpas, previamente lavadas em PBS 1x por 1 minuto, incubadas em uma solução de tripsina (1% de tripsina em PBS 1x) por 20 segundos e lavadas novamente em PBS 1x. Posteriormente, as preparações foram coradas com solução de Giemsa 5% em PBS 1x por 5 minutos. Para a microdissecção da variante B₁ (eucromática), nós utilizamos as suspensões celulares do indivíduo 69693, que só apresentava essa variante em suas células.

O cromossomo microdissecado foi transferido para uma micropipeta contendo uma solução de coleção (1,5ug/ul de proteinase K, 0,1% SDS, 0,1% Triton X-100, 1mM EDTA, 10mM Tris-HCl, pH 8,0, 10mM NaCl) e colocado em uma câmara úmida por 1 hora a 60°C. Posteriormente, as pontas das pipetas contendo o material foram quebradas em uma microtubo de 0,2ml contendo 5ul de água miliQ estéril, sendo, então, realizada a amplificação

utilizando o kit GenomePlex Single Cell Whole Genome Amplification (WGA-Sigma), gerando a biblioteca denominada *midib*.

Sequenciamento Illumina

O sequenciamento de três bibliotecas distintas (1- cromossomo B₁ microdissecado, 2- 0B gDNA, 3- 6B gDNA) foi realizado na plataforma Illumina HiSeq2000 pela empresa Macrogen Inc. As bibliotecas genômicas foram preparadas com o kit *TruSeq DNA PCR-free kit*, enquanto a biblioteca do cromossomo B microdissecado foi preparada com o kit *TruSeq Nano kit*, pulando a etapa de fragmentação das amostras, uma vez que o método de WGA utilizado já resulta em amostras fragmentadas.

Análises bioinformáticas

Para a análise estrutural do cromossomo B microdissecado, as três bibliotecas de *M. sanctaefilomenae* relacionadas acima foram processadas com a *pipeline* “*DOPseq analyzer* (https://github.com/ilyakichigin/DOPseq_analyzer).” desenvolvida por MSc Alexey Makunin. De forma geral, esta *pipeline* retira os adaptadores Illumina e os *primers* WGA GenomePlex e alinha os *reads* das bibliotecas ao genoma de *A. mexicanus* (GCA_000372685.1) e ao genoma humano (hg19, baixado do *UCSC Genome Browser*) com o *software* bowtie2. Assim, os *reads* que se alinham perfeitamente com o genoma humano são descartados como contaminação putativa. Adicionalmente, os *reads* também foram alinhados ao genoma de *D. rerio* (GCA_000002035.3), uma espécie de peixe mais distante de *M. sanctaefilomenae*, porém com um genoma mais completo e melhor montado. Finalmente, esta *pipeline* também identifica regiões cromossomo-específicas baseadas em densidade de *reads*, bem como realiza uma busca por regiões variantes (*variant calling*) dos genes contra a referência utilizada.

O método escolhido para a análise de cobertura diferencial foi semelhante ao utilizado por Valente et al. (2014), buscando-se regiões com 4x de diferença de densidade de *reads* entre as duas bibliotecas genômicas. Para a análise da porção repetitiva deste cromossomo, os *reads* de *midib* foram alinhados contra os *contigs* gerados após análise pelo *software* RepeatExplorer (Novák et al. 2013) com os dados de genoma total 0B e 6B através do *software* RepeatMasker (Smit et al. 2010) para estimar a abundância de cada um dos *clusters* na biblioteca *midib*.

Resultados e Discussão

O sequenciamento Illumina HiSeq das bibliotecas gerou um total de 15.299.574 *reads* para o indivíduo 0B, 12.638.474 *reads* para o indivíduo 6B e 4.899.386 *reads* para o cromossomo B₁ microdissecado. Informações adicionais sobre as quantidades de *reads* mapeados e/ou contaminados estão sumarizados na Tabela 1. Inicialmente, as análises bioinformáticas revelaram maiores taxas de alinhamento dos *reads* do cromossomo B₁ (*midib*) contra o genoma humano (40,25%) do que em relação ao genoma de *A. mexicanus* (27,95%). No entanto, a porção final de *reads* putativamente contaminados foi baixa (5% dos *reads*), o que sugere a existência de regiões de baixa qualidade de montagem do genoma de *A. mexicanus*, justamente nas regiões genômicas compartilhadas entre *A. mexicanus* e o genoma humano. Paralelamente, as mesmas análises foram realizadas contra o genoma do peixe *D. rerio* como referência, mas os resultados também não foram satisfatórios, provavelmente devido à grande distância filogenética entre *D. rerio* e *M. sanctaefilomenae* (dados não mostrados).

Nos últimos anos, diferentes abordagens têm sido utilizadas para a caracterização global de sequências presentes nos cromossomos B. Destas, podem ser destacadas a análise

do índice de cobertura relativa de genomas –B e +B (Valente et al. 2014); o *screening* diferencial em bibliotecas genômicas +B e –B (Yoshida et al. 2011); e o sequenciamento pela plataforma 454 de cromossomos *flow-sorted* (Martis et al. 2012) ou microdissecados (Valente et al. 2014). Uma vez que na população de *M. sanctaefilomenae* do Rio Tietê apenas indivíduos portadores de cromossomos B foram encontrados, a técnica de sequenciamento de cromossomo microdissecado foi escolhida para este estudo. De forma geral, nossas análises revelaram que embora a biblioteca microdissecada apresente uma quantidade relativamente alta de contaminação, muitas sequências putativas do próprio cromossomo B foram identificadas, comprovando a eficácia deste método como uma abordagem inicial para a caracterização massiva destes cromossomos.

A busca de regiões-alvo com base na densidade de *reads* de *midB* revelou várias regiões extensas, com pelo menos 20 posições e média “pd” menor que 15kbp (pd = distância entre posições cobertas pelos *reads*: quanto menor, maior a densidade de *reads* e maior a chance de se caracterizar como uma região-alvo, ou seja, ser uma sequência localizada no B). A princípio, estas regiões-alvo podem ser tratadas como pertencentes ao cromossomo B₁, embora dois pontos devam ser levados em consideração: *i*) o valor de corte para a média “pd” é escolhido arbitrariamente, ou seja, regiões com 12-14kb estão inclusas, enquanto regiões de 15,6-16kb foram excluídas; e *ii*) os resultados com menos de 20 posições de alinhamento foram ignorados; apesar destes dados também serem informativos, os resultados finais se tornariam altamente caóticos e difíceis de se interpretar.

Em geral, as regiões-alvo identificadas cobriram os *scaffolds* de forma integral. Apenas em três casos apenas uma parte dos *scaffolds* foi coberta com maior densidade, ou seja, valores de “pd” menores (Tabela 2, Figura 2). Deve-se destacar que estas últimas regiões se caracterizam como mais interessantes, pois este padrão de mapeamento é característico da ocorrência de translocações (geralmente espera-se que as margens das regiões B-específicas

não correspondam ao fim dos *scaffolds* dos cromossomos A). A análise comparativa destes blocos B-específicos contra o genoma de *A. mexicanus* revelou que estas regiões pertencem a diferentes grupos de ligação (LG) nesta espécie (Tabela 2), indicando que diferentes regiões genômicas parecem contribuir para a composição do cromossomo B. Além disso, os genes que estão alocados em cada uma das regiões-alvo foram identificados e as respectivas quantidades de genes estão sumarizadas na Tabela 2.

As análises de *variant calling* nas regiões putativamente B-específicas retornaram nove regiões com variantes de genes altamente destrutivas em relação ao genoma de *A. mexicanus*, tais como inserções/deleções e substituições na cadeia de aminoácidos (Tabela 3). Deve-se notar que todas as 9 regiões/blocos correspondem ao LG21 de *A. mexicanus*, o que poderia indicar que o grupo sintênico ao LG21 de *A. mexicanus* em *M. sanctaefilomenae* poderia ser o proto-cromossomo B₁ ou um dos principais doadores de sequências para o cromossomo B₁ desta espécie. De fato, alguns estudos têm demonstrado que, após a sua origem, os cromossomos B recebem sequências de DNA de diversos cromossomos, inclusive DNA organelar (Alfenito e Birchler 1993; Houben et al. 2001; Peng et al. 2005; Valente et al. 2014).

Em estudo recente, Valente *et al.* (2014) reportaram uma série de genes com alta integridade relacionados com os processos de divisão celular no cromossomo B de *Astatotilapia latifasciata* e sugeriram que estes genes poderiam desempenhar um papel importante na continuidade da transmissão deste cromossomo B nesta espécie.. Aparentemente, nossos dados não indicam que os diversos genes encontrados no cromossomo B₁ sejam fundamentais para as etapas do ciclo celular (Tabela 3). No entanto, uma maior cobertura de sequenciamento dos cromossomos B de *M. sanctaefilomenae* se faz necessária para determinar, de fato, quais os genes que constituem estes elementos fundamentais nesta espécie.

De forma geral, as regiões-alvo encontradas na biblioteca *midiB* não apresentaram correspondência nas análises de cobertura diferencial realizadas com as bibliotecas genômicas 0B (*wg0B*) e 6B (*wg6B*). Inicialmente, regiões com 4x de diferença de densidade de *reads* entre as duas bibliotecas genômicas foram buscadas. Para a *wg6B*, 3 regiões foram encontradas mas nenhuma delas estava entre as regiões-alvo encontradas em *midiB*. De forma oposta, *wg0B* apresentou 12 regiões com densidade diferencial em relação à *wg6B*, sendo que apenas uma região era correspondente à região-alvo de *midiB*. Embora as bibliotecas *wg0B* e *wg6B* correspondam à diferentes populações de uma mesma espécie, as diferenças de densidade positiva para *wg0B* não eram esperadas. Nesse sentido, as supostas “duplicações” encontradas no genoma 0B poderiam ser explicadas pela má montagem do genoma de *A. mexicanus*, baixa cobertura de sequenciamento dos genomas 0B e 6B de *M. sanctaefilomenae* ou pela ocorrência de CNV (do inglês: *copy number variations*) entre as populações desta espécie, fenômeno amplamente relatado em humanos (Sudmant et al. 2000; Sebat et al. 2004; Freeman et al. 2006).

A busca por elementos repetitivos no cromossomo B de *M. sanctaefilomenae* resultou em um total de aproximadamente 9% de sequências mapeadas sobre os *clusters* gerados pelo *software* RepeatExplorer, totalizando 198 sequências de DNA repetitivo putativos presentes neste cromossomo (Tabela 4). Com o limite mínimo arbitrário estabelecido de 0.1% de *reads* alinhados contra os *clusters* do *software* RepeatExplorer, 25 *clusters* finais foram listados, sendo representados em sua maioria por famílias de elementos transponíveis, principalmente da superfamília Tc1/mariner. De fato, a invasão dos cromossomos B por elementos repetitivos é amplamente conhecida em diversos organismos, sendo um dos principais eventos da evolução molecular destes elementos (revisão em Camacho et al. 2000). Deve-se ressaltar que as proporções genômicas obtidas pelo *software* RepeatExplorer devem ser interpretadas com cautela, pois a expansão/retração destes DNAs repetitivos em diferentes populações é fato

conhecido e bem documentado (Charlesworth et al. 1994; Schlötterer 2000) e podem estar associadas à dinâmica destes elementos nos próprios cromossomos do complemento A, como já observado ocorrer nesta espécie (Capítulo 1). Além disso, como esta biblioteca já passou por um processo de amplificação inicial pelo método de WGA, as proporções podem estar enviesadas.

Os resultados obtidos aqui revelam que diferentes classes de sequências compõem o pequeno cromossomo B₁ de *M. sanctaefilomenae*, incluindo múltiplas famílias de DNA repetitivo, notavelmente elementos transponíveis. Estes dados sugerem que este cromossomo B acumula sequências de DNA repetitivo de diferentes origens, existindo como um depósito de elementos transponíveis, ocasionado principalmente pela ausência de recombinação destes cromossomos (Portela-Castro et al. 2001). Além disso, nossos dados também revelam que diferentes grupos de ligação contribuíram para a composição do cromossomo B, provavelmente através de um processo de duplicação gênica mediada por elementos móveis ou por retro-inserções. Nesse sentido, nós propomos que o cromossomo B de *M. sanctaefilomenae* se origina a partir de uma duplicação segmental, provavelmente do par No. 6 (Capítulo 1), formando o proto-cromossomo B. Subsequentemente, este cromossomo é invadido por diversas sequências oriundas de diferentes regiões genômicas, incluindo blocos que contêm genes de cópia única, elementos transponíveis e DNAs satélites, em um processo semelhante ao descrito ocorrer em *Astatotilapia latifasciata* (Valente et al. 2014).

Considera-se que os resultados aqui obtidos poderão ser utilizados como base para a realização de diferentes experimentos. Para os genes de cópia única putativamente localizados no cromossomo B, *primers* serão desenhados: *i*) com a porção 3' do *primer* anelando exatamente na posição da substituição/deleção; *ii*) ao redor da substituição/deleção para posterior sequenciamento e detecção dos polimorfismos; *iii*) ancorando em regiões conservadas nos genomas +B e -B para realizar experimentos de qPCR. Além disso, *primers*

para os elementos transponíveis de diferentes famílias que foram obtidos na presente análise também serão desenhados e testados. Finalmente, nossos dados evidenciaram a eficácia de um método simples e barato para uma caracterização inicial global de sequências presentes nos cromossomos B.

Referências

As referências correspondentes a este capítulo encontram-se relacionadas no final desta tese.

Tabela 1: Informações gerais relativas ao número de *reads* mapeados em cada uma das bibliotecas usadas

Biblioteca	Reads	Reads ca	Contaminação	Reads mapeados	Posições
MiDi B - <i>A. mexicanus</i>	4899386	4840782	25669	501577	34530
MiDi B - <i>D. rerio</i>	4899386	4840782	25669	79217	19300
WGS 0B	15299574	15299228	929	3824346	1972669
WGS 6B	12638474	12638196	766	3088360	1747279

Reads ca: Número de reads utilizados na análise após filtro de qualidade

Tabela 2: Tabela contendo as informações de todos os *scaffolds* mapeados com os *reads* Illumina de *midib*. Grifado em amarelo, os *scaffolds* cuja cobertura foi apenas parcial.

Scaffold	Começo da região	Fim da região	Posições	Média segmento	Média "pd"	Genes	LG - Amex
KB882136.1	713431	922707	60	6.4719	3383	-	LG19
KB871630.1	17063	5846906	878	7.2308	6545	62	LG21
KB872020.1	6987	289578	40	7.198	7150	1	x
KB872138.1	8739	233907	32	7.5363	7183	3	x
KB871605.1	7903	1129504	141	7.4017	7915	14	LG21
KB872023.1	133775	298337	20	8.0971	8604	-	x
KB871813.1	882	681225	78	7.1967	8741	3	x
KB882155.1	2143915	2552139	46	7.3495	8805	3	LG3
KB871667.1	22568	907281	96	7.4549	9235	17	LG21
KB871992.1	150	396269	42	7.5269	9568	7	LG21
KB871725.1	26468	768512	75	7.871	9942	9	x
KB871731.1	22776	875349	84	7.9664	10168	18	LG21
KB871978.1	59508	360012	30	6.9897	10262	2	x
KB882169.1	10232	2380128	230	7.7986	10262	30	LG21
KB871932.1	30273	361000	32	7.7307	10615	11	LG13
KB871865.1	9584	560530	51	7.8169	10917	1	x
KB871785.1	21036	631510	50	8.0841	12380	7	LG5
KB871970.1	17870	343294	24	8.2259	14067	9	LG5
KB882087.1	2563154	2930783	34	7.202	10714	1	x

LG – Amex: Localização do *scaffold* no genoma de *A. mexicanus*.

Tabela 3: Informações das nove regiões com variantes de genes altamente destrutivas em relação ao genoma de *A. mexicanus* e a localização de cada um dos *scaffolds* no genoma de *A. mexicanus*.

Scaffold	Gene	GO	SNP	De	Para	LG - Amex
KB871630.1	ATP-binding cassette sub-family A member 1-like	ATPase activity	2080348	CGT	C	LG21
KB871630.1	leucine aminopeptidase 3	Aminopeptidase activity	5154507	C	CAA	LG21
KB871630.1	leucine aminopeptidase 3	Aminopeptidase activity	5154514	AGC	A	LG21
KB882169.1	PHD finger protein 3	Zinc ion binding	982536	G	GGC	LG21
KB882169.1	PHD finger protein 3	Zinc ion binding	982538	A	AAAATGTG	LG21
KB871605.1	FAT atypical cadherin 1a	Adesão celular	498730	AT	A	LG21
KB871605.1	LIM and calponin homology domains 1a	Zinc ion binding	694695	T	G	LG21
KB871667.1	Rho guanine nucleotide exchange factor (GEF) 40	Regulação do sinal de transdução da proteína Rho	414620	GTA	G	LG21
KB871667.1	Rho guanine nucleotide exchange factor (GEF) 40	Regulação do sinal de transdução da proteína Rho	414623	C	CGT	LG21

GO: Função anotada no *Gene Ontologies*. SNP: Posição do SNP no *Scaffold* corresponde de *A. mexicanus*. LG – Amex: Localização do *scaffold* no genoma de *A. mexicanus*.

Tabela 4: Informações sobre os *clusters* que apresentaram *reads* alinhados acima do limite estabelecido de 0,1% do total de *reads*. Nas últimas duas colunas à direita, as proporções genômicas obtidas na análise do RepeatExplorer retirado de Utsunomia et al. (Capítulo 1).

<i>Cluster - RE</i>	% de <i>reads</i> alinhados	Anotação	0B - RE	6B - RE
CL48	0,8848%	TcMar Tc1	0.104%	0.074%
CL15	0,7468%	RC.Helitron	0.186%	0.369%
CL9	0,5114%	Unknown	0.424%	0.455%
CL40	0,4713%	TcMar Tc1	0.113%	0.119%
CL11	0,3860%	DNA.CMC.EnSpm	0.392%	0.398%
CL7	0,3749%	Unknown	0.496%	0.512%
CL6	0,3504%	TcMar Tc1	0.511%	0.506%
CL8	0,3305%	TcMar Tc1	0.531%	0.442%
CL204	0,2895%	Unknown	0.001	0.0195%
CL89	0,2212%	Unknown	0.0265%	0.0405%
CL1	0,2204%	RC.Helitron	0.638%	0.719%
CL2	0,1954%	Satellite	0.621%	0.679%
CL4	0,1679%	Satellite	0.494%	0.561%
CL10	0,1532%	TcMar Tc1	0.416%	0.433%
CL17	0,1405%	DNAr 5S	0.248%	0.258%
CL3	0,1312%	TcMar Tc1	0.623%	0.66%
CL21	0,1286%	TcMar Tc1	0.236%	0.241%
CL26	0,1179%	DNA.hAT.Ac	0.188%	0.178%
CL16	0,1121%	TcMar Tc1	0.246%	0.282%
CL169	0,1113%	TcMar Tc1	0.017%	0.0095%
CL12	0,1105%	TcMar Tc1	0.326%	0.355%
CL24	0,1077%	Unknown	0.186%	0.215%
CL84	0,1011%	Simple repeat	0%	0.072%
CL27	0,1010%	SatDNA MS3	0.117%	0.236%
CL96	0,1000%	SatDNA MS7	0.0295%	0.0285%

RE: RepeatExplorer

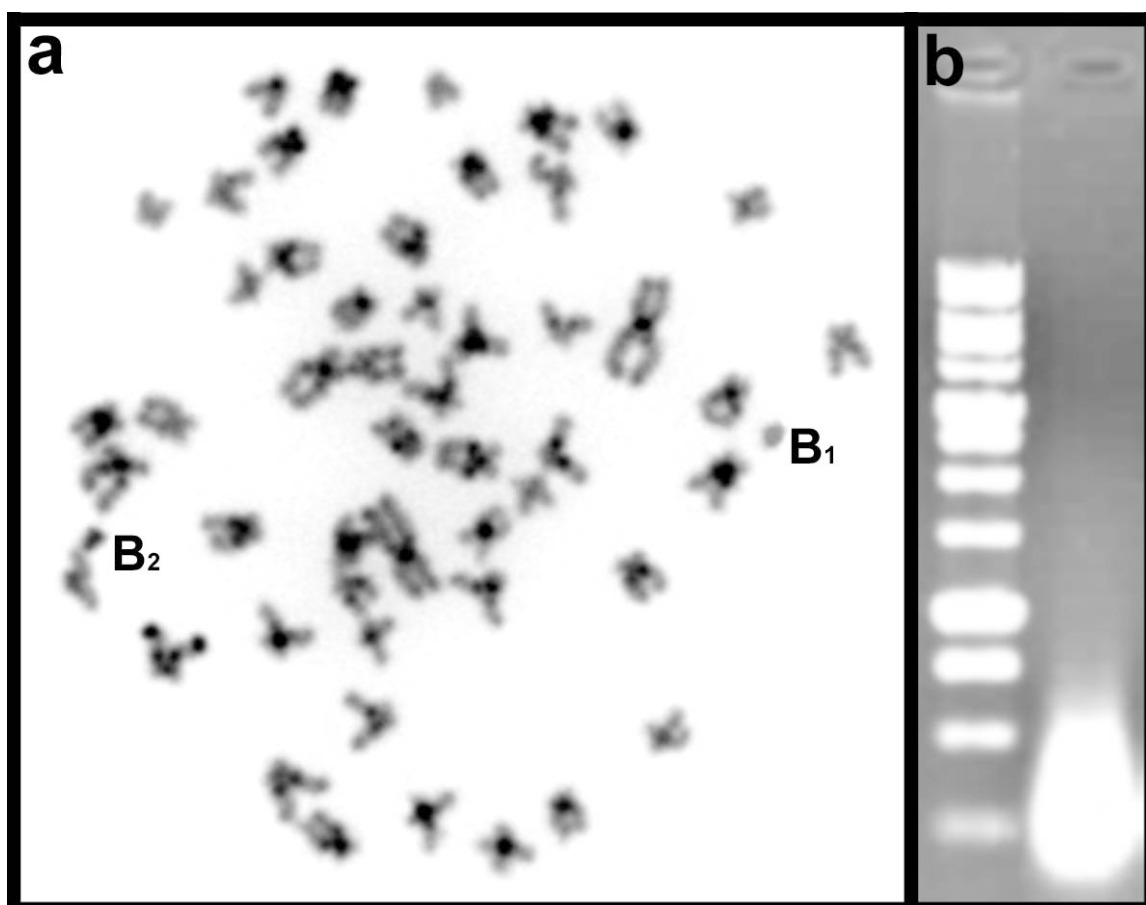


Figura 1: Metáfase de exemplar de *Moenkhausia sanctaefilomenae* coletado na bacia do rio Tietê após bandamento-C, evidenciando dois tipos de cromossomos B (B₁ e B₂) (a). Em (b), gel de agarose após microdissecção e amplificação com GenomePlex .

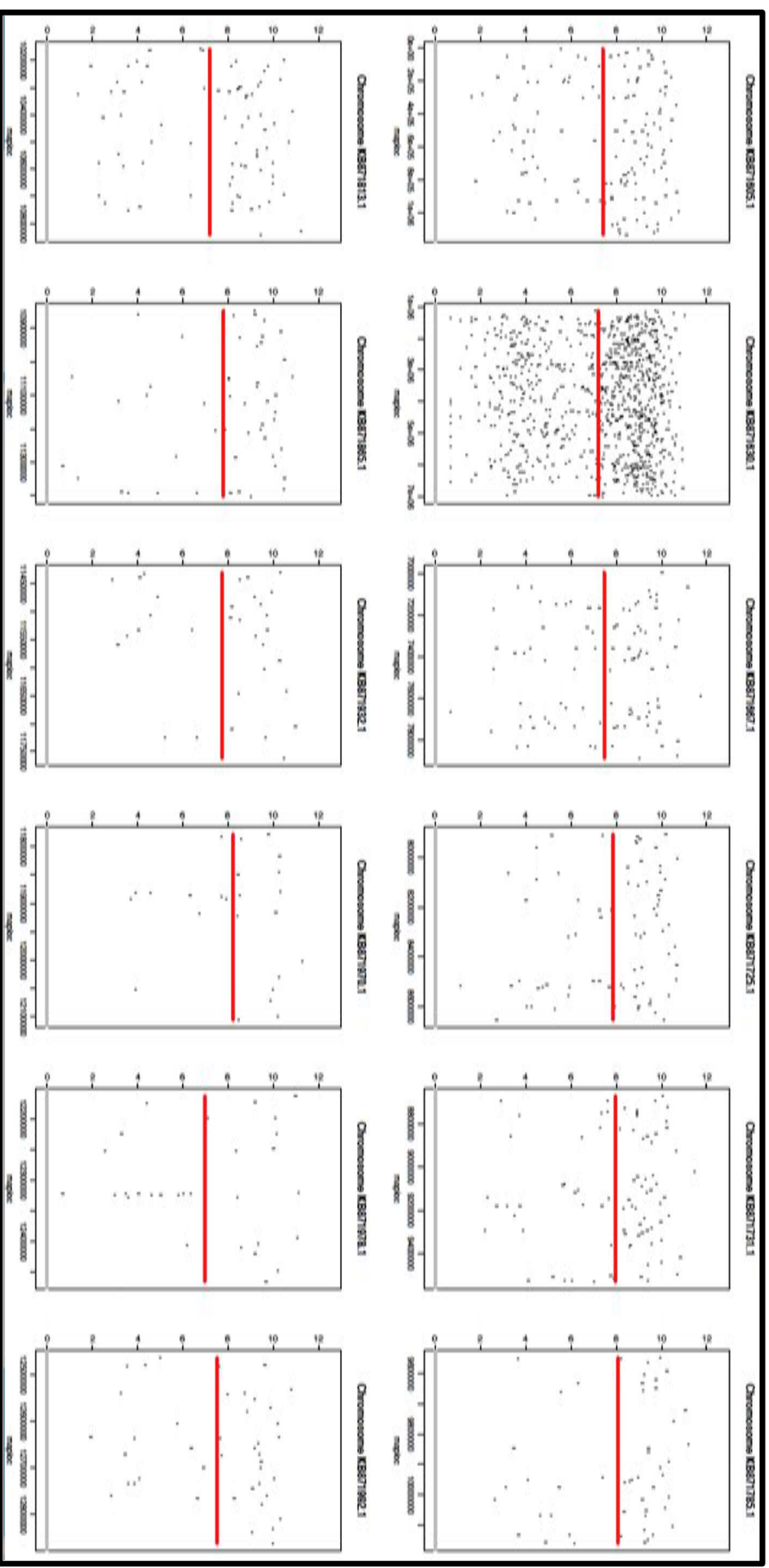


Figura 2: Gráficos mostrando o mapeamento dos *reads* de *midB* ao longo de diferentes *scaffolds* de *A. mexicanus*. No eixo X estão representadas as coordenadas cumulativas ao longo dos *scaffolds* mapeados. Notar que todos estão representados pelo mesmo tamanho virtual, embora os tamanhos reais sejam diferentes. No eixo Y estão representados os valores de log de “pd” (pd = distância entre posições cobertas pelos *reads*: quanto menor, maior a densidade de *reads* e maior a chance de se caracterizar como uma região-alvo, ou seja, ser uma sequência localizada no B.

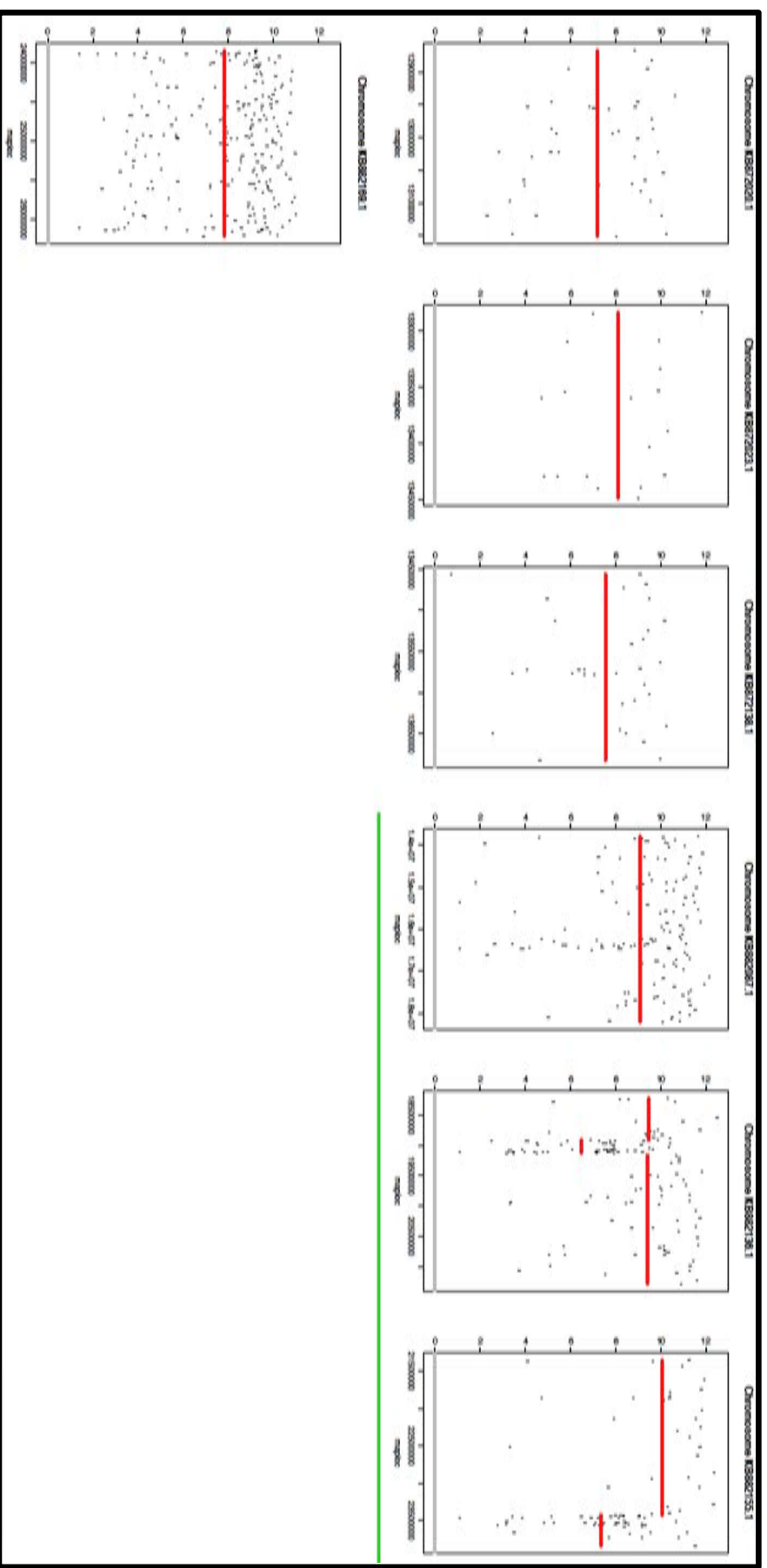


Figura 3: Gráficos mostrando o mapeamento dos *reads* de *miB* ao longo de diferentes *scaffolds* de *A. mexicanus*. No eixo X estão representadas as coordenadas cumulativas ao longo dos *scaffolds* mapeados. Notar que todos estão representados pelo mesmo tamanho virtual, embora os tamanhos reais sejam diferentes. No eixo Y estão representados os valores de \log de “pd” (pd = distância entre posições cobertas pelos *reads*: quanto menor, maior a densidade de *reads* e maior a chance de se caracterizar como uma região-alvo, ou seja, ser uma sequência localizada no B. Indicado em verde, os três *scaffolds* com maior densidade de *reads* em uma região específica do *scaffold*, indicativo de translocações.

5. Considerações Finais

O presente estudo se inseriu no programa geral de estudos citogenéticos de peixes Neotropicais que vem sendo desenvolvido no Laboratório de Biologia e Genética de Peixes (IBB/UNESP/Botucatu) e visou aumentar o conhecimento sobre a estrutura física e molecular de cromossomos supranumerários em peixes Neotropicais, bem como sobre a origem e os aspectos evolutivos deste sistema na espécie *Moenkhausia sanctaefilomenae*.

Os estudos citogenéticos e moleculares desenvolvidos em diferentes populações da espécie *M. sanctaefilomenae* trouxeram informações significativas relacionadas à origem de distintas variantes de cromossomos B nesta espécie. Nesse sentido, destaca-se a importância das análises nucleotídicas realizadas, que permitiram uma melhor delimitação da origem de cada uma das variantes observadas aqui, o que não seria possível apenas com a análise estrutural destes elementos (e.g. dados de FISH). Desta forma, apesar das variantes B₁ e B₂ presentes nas células de *M. sanctaefilomenae* coletadas no rio Batalha compartilharem os mesmos DNAs repetitivos, as informações relativas às suas frequências na população e aos haplótipos de dois DNAs satélites distintos, revelaram que uma das variantes (B₁-eucromática) é mais antiga que a outra (B₂-heterocromática). Nesse sentido, a análise de sequências de DNA repetitivo e as informações sobre a sua localização sugerem que o cromossomo B possa ter se originado em um evento intraespecífico a partir do par No.6.

A análise focada em sequências repetitivas também revelou que os cromossomos B de *M. sanctaefilomenae* são elementos que apresentam seleção relaxada para o gene da histona H3.3, além de uma atividade insignificativa das regiões organizadoras de nucléolo (RONs) presentes nos Bs para a síntese de RNAr nesta espécie. Este tipo de informação está de acordo com alguns estudos realizados em diferentes organismos e indicam que a contribuição dos cromossomos Bs para a fisiologia da célula é mínima. No entanto, alguns estudos contemporâneos têm revelado, principalmente com dados de sequenciamento de nova geração, que os cromossomos B apresentam diversos genes de cópia única, muitos deles com alta integridade. Nesse sentido, é cada vez mais claro que cada sistema de cromossomo B é único em suas respectivas linhagens, podendo desempenhar funções específicas, ou serem apenas elementos estritamente parasíticos e inertes.

A cada ano, novos sistemas de cromossomos Bs são descritos em diferentes peixes Neotropicais. De forma geral, estes sistemas se mostram extremamente variáveis, sendo que os cromossomos supranumerários podem ser de tamanho grande ou pequeno; heterocromáticos ou eucromáticos; compartilhar sequências com cromossomos A,

cromossomos sexuais ou com nenhum elemento do complemento A; portar genes conhecidos, notadamente famílias multigênicas; estar restritos a um único sexo ou ambos; estar limitados a algumas populações de determinadas espécies; ser mitoticamente estáveis ou instáveis; e estar em fase de invasão ou extinção na população. Todas estas características foram/vêm sendo traçadas através do uso de ferramentas citogenéticas clássicas. Nesse sentido, a aplicação de técnicas mais diretas para o conhecimento sobre a composição destes cromossomos se faz necessária para uma melhor compreensão sobre os mecanismos de origem e manutenção que os cromossomos B apresentam, principalmente no nível gênico.

Os resultados apresentados relativos ao sequenciamento do cromossomo B₁ microdissecado revelaram que apesar do seu tamanho reduzido, este cromossomo apresenta numerosas classes de DNA repetitivo, composto em sua maioria por elementos transponíveis e até mesmo genes de cópia única. Estes dados evidenciaram a eficácia de um método simples e barato para uma caracterização inicial global de sequências presentes nos cromossomos B.

Finalmente, deve-se destacar que a continuidade dos estudos apresentados neste trabalho será de fundamental importância para aumentar o conhecimento sobre a biologia evolutiva dos cromossomos supranumerários em geral.

6. Referências bibliográficas

- Alfenito MR, Birchler JA. 1993. Molecular characterization of a maize B chromosome centric sequence. *Genetics*, 135: 589-597.
- Artoni RF, Vicari MR, Endler AL, Cavallaro ZI, de Jesus CM, de Almeida MC, Moreira-Filho O, Bertollo LAC. 2006. Banding pattern of A and B chromosomes of *Prochilodus lineatus* (Characiformes, Prochilodontidae), with comments on B chromosomes evolution. *Genetica* 127, 277-284.
- Banaei-Moghaddam AM, Meier K, Karimi-Ashtiyani R, Houben A (2013) Formation and expression of pseudogenes on the B chromosome of rye. *Plant Cell* 25(7): 2536 - 2544
- Benine RC, Mariguela TC, Oliveira C. 2009. New species of *Moenkhausia* Eigenmann, 1903 (Characiformes: Characidae) with comments on the *Moenkhausia oligolepis* species complex. *Neotropical Ichthyology*, 7: 161–168.
- Beukeboom LW, Seif M, Mettenmeyer T, Plowman AB, Michiels NK. 1996. Paternal inheritance of B chromosomes in a parthenogenic hermaphrodite. *Heredity*, 77: 646-654.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, btu170.
- Bugrov AG, Karamysheva TV, Perepelov EA, *et al.* 2007. DNA content of the B chromosomes in grasshopper *Podisma kanoi* Storozh. (Orthoptera, Acrididae). *Chromosome Research* 15:315-325.
- Bueno D, Palacios-Gimenez OM, Cabral-de-Mello DC. 2013. Chromosomal mapping of repetitive DNAs in the grasshopper *Abracris flavolineata* reveal possible ancestry of the B chromosome and H3 histone spreading. *PLoS One* 8(6): e66532.
- Cabral-de-Mello DC, Valente GT, Nakajima RT, Martins C. 2012. Genomic organization and comparative chromosome mapping of the U1 snRNA gene in cichlid fish, with an emphasis in *Oreochromis niloticus*. *Chromosome Research* 20: 279–292
- Cabrero J, López-León M, Bakkali M, Camacho JPM. 1999. Common origin of B chromosome variants in the grasshopper *Eyprepocnemis plorans*. *Heredity* 83: 435-439.
- Cabrero J, López-León M, Ruíz-Estévez M, Gómez, R, Petitpierre E, Rufas J, Massa B, Kamel Ben Halima M, Camacho JPM. 2014. B1 was the ancestor B chromosome

- variant in the western Mediterranean area in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research* 142: 54-58.
- Camacho JPM, Sharbel TF, Beukeboom LW. 2000. B chromosome evolution. *Philosophical Transactions of the Royal Society B Biological Sciences* 355: 163-178.
- Camacho JPM. 2005. B chromosomes. In: Gregory TR editor. *The Evolution of the Genome*. Elsevier, San Diego. p 223-286.
- Charlesworth B, Snlegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371: 215-220.
- Cheng YM, Lin BY. 2003. Cloning and characterization of maize B chromosome sequences derived from microdissection. *Genetics* 164: 299-310.
- Cole CJ, Leavens C. 1971. Chromosome preparations of amphibians and reptiles: improved technique. *Herpetology* 3(6): 102.
- Colgan DJ, McLauchlan A, Wilson GDF. 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology* 46: 419-437.
- Covert SF. 1998. Supernumerary chromosomes in filamentous fungi. *Current Genetics* 33: 311-319.
- Dantas ESO, Vicari MR, Souza IL, Moreira-Filho O, Bertollo LAC, Artoni RF. 2007. Cytotaxonomy and karyotype evolution in *Moenkhausia* Eigenmann, 1903 (Teleostei, Characidae). *The Nucleus* 50: 505-518.
- Edgar RC. 2004. Muscle: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5: 113.
- Eschmeyer W, Fong J. 2015. Species by family/subfamily in the Catalog of Fishes.
- Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567.
- Foresti F, Toledo-Filho SA, Almeida-Toledo LF. 1981. Polymorphic nature of the nucleolus organizer regions in fishes. *Cytogenetic and Cell Genetics* 31: 134-141.
- Foresti F, Almeida-Toledo LF, Toledo SA. 1989. Supranumerary chromosome system C-banding pattern characterization and multiple Nucleolus Organizing Region in *Moenkhausia sanctaefilomenae* (Pisces, Characidae). *Genetica* 79: 107-114.
- Freeman JL, Perry GH, Feuk L, et al. 2006. Copy number variation: New insights in genome diversity. *Genome Research* 16: 949 – 961.

- Fujiwara A, Abe S, Yamaha E, Yamazaki F, Yoshida, MC. 1998. Chromosomal localization and heterochromatin association of ribosomal RNA genes loci and silver stained nucleolar organizer regions in salmonid fishes. *Chromosome Research* 6: 463-471.
- Graphodatsky AS, Kukekova AV, Yudkin DV, Trifonov VA, Vorobieva NV, Beklemisheva VR, Perelman PL, Graphodatskaya DA, Trut LN, Yang F. 2005. The proto-oncogene C-KIT maps to canid B-chromosomes. *Chromosome Research* 13: 113-122.
- Green DM. 1990. Muller's ratchet and the evolution of supernumerary chromosomes. *Genome* 33: 818-824.
- Gribble S, Ng BL, Prigmore E, Burford DC, Carter NP. 2004. Chromosome paints from single copies of chromosomes. *Chromosome Research* 12: 143-151.
- Hashimoto DT, Fergusson-Smith MA, Rens W, Foresti F, Porto-Foresti F. 2011. Chromosome mapping of H1 histone and 5S rRNA genes clusters in three species of *Astyanax* (Teleostei, Characiformes). *Cytogenetic and Genome Research* 134: 64-71.
- Hashimoto DT, Voltolin TA, Paes ADNVA, Foresti F, Bortolozzi J, Porto-Foresti F. 2012. Cytogenetic analysis of B chromosome in one population of the fish *Moenkhausia sanctaefilomenae* (Steindachner, 1907) (Teleostei, Characiformes). *Comparative Cytogenetics* 6: 141-151
- Henning F, Trifonov V, Ferguson-Smith MA, Almeida-Toledo LF. 2008. Non-homologous sex chromosomes in two species of the genus *Eigenmannia* (Teleostei:Gymnotiformes). *Cytogenetic and Genome Research* 121(1): 55-58.
- Hewitt GM. 1976. Meiotic drive for B-chromosomes in the primary oocytes of *Myrmekotettix maculatus* (Orthoptera: Acrididae). *Chromosoma* 56: 381-391.
- Hewitt GM, East TM. 1978. Effects of B chromosomes on development in grasshopper embryos. *Heredity* 41: 347-356.
- Houben A, Verlin D, Leach CR, Timmis JN. 2001. The genomic complexity of micro B chromosomes of *Brachycome dichromosomatica*. *Chromosoma* 110(7): 451-459
- Höckner M, Erdel M, Spreiz A, Utermann G, Kotzot D. 2009. Whole genome amplification from microdissected chromosomes. *Cytogenetic and Genome Research* 125: 98-102.
- Houben A, Banaei-Moghaddam AM, Klemme S, Timmis JN. 2013. Evolution and biology of supernumerary B chromosomes. *Cellular and Molecular Life Sciences* 71: 467 - 478.

- Howel WM, Black DA. 1980. Controlled silver-staining of nucleolus organizer regions with a protective colloidal developer: a 1-step method. *Experientia* 36: 1014-1015.
- Jones RN, Rees H. B chromosomes. Academic Press, 1982.
- Jones RN. 1985. Are B chromosomes selfish? In *The evolution of genome size* C.S.T, ed. (London: Wiley), 397-425.
- Lamb JC, Kato A, Birchler JA. 2005. Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. *Chromosoma* 113: 337-349.
- Leaver MJ. 2001. A Family of Tc1-like transposons from the genomes of fishes and frogs: evidence for horizontal transmission. *Gene* 271: 203-214.
- Levan A, Fredga K, Sandberg AA. 1964. Nomenclature for centromeric position of chromosomes. *Hereditas* 52: 201–220.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
- Lima FCT, Malabarba LR, Buckup PA, Silva JFP, Vari RP, Harold A, Benine R, Oyakawa OT, Pavanelli CS, Menezes NA. 2003. Characidae – genera *incertae sedis* in Characidae. In: Reis RE, Kullander SO, Ferraris Jr CJ (Eds). *Check List of the Freshwater fishes of the South and Central America*. EDIPUCRS. Porto Alegre, 106–169.
- Long H, Qi ZX, Sun XM *et al.* 2008. Characters of DNA constitution in the rye B chromosome. *Journal of Integrative Plant Biology* 50: 183-189.
- López-León M, Cabrero J, Pardo M, Viseras E, Camacho JPM, Santos J. 1993. Generating high variability of B chromosomes in *Eyprepocnemis plorans* (grasshopper). *Heredity* 71: 352-352.
- López-León MD, Neves N, Schwarzacher T, Heslop-Harrison JS, Hewitt GM, Camacho JPM. 1994. Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Research* 2: 87-92.
- Machado TC, Pansonato-Alves JC, Pucci MB, Nogaroto V, Almeida MC, Oliveira C, Foresti F, Bertollo LAC, Moreira-Filho O, Artoni RF, Vicari MR. 2011. Chromosomal painting and ZW sex chromosomes differentiation in *Characidium* (Characiformes, Crenuchidae). *BMC Genetics* 12: 65.
- Makunin AI, Dementyeva PV, Graphodatsky AS, Volobouev VT, Kukekova AV, Trifonov VA. 2014. Genes on B chromosomes of vertebrates. *Molecular Cytogenetics* 7:99
- Marshall OJ. 2004. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* 20: 2471–72.

- Martins C, Galetti Jr. PM. 1999. Chromosomal localization of 5S rDNA genes in *Leporinus* fish (Anostomidae, Characiformes). *Chromosome Research* 7: 363-367.
- Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Uwe S, Gundlach H, Wicker T, Simková H, et al. 2012. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences* 109: 13343-13346.
- McAllister BF, Werren JH. 1997. Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Chromosoma* 106: 243-253.
- McQuade L, Hill R, Francis D. 1994. B chromosome systems in the greater glider, *Petauroides volans* (Marsupialia, Pseudocheiridae). 2. Investigation of B chromosome DNA sequences isolated by micromanipulation and PCR. *Cytogenetic and Cell Genetics* 66: 155-161.
- Menezes-de-Carvalho NZ, Palacios-Gimenez OM, Milani D, Cabral-de-Mello DC. 2015. High similarity of U2 snDNA sequence between A and B chromosomes in the grasshopper *Abracris flavolineata*. *Molecular Genetics and Genomics* 290: 1787-1792.
- Merlo MA, Cross I, Chairi H, Manchado M, Rebordinos L. 2010. Analysis of three multigene families as useful tools in species characterization of two closely-related species, *Dicentrarchus labrax*, *Dicentrarchus punctatus* and their hybrids. *Genes Genetics and Systematics* 85: 341-349.
- Mestriner CA, Galetti PM, Valentini SR, Ruiz IR, Abel LD, Moreira-Filho O, Camacho JPM. 2000. Structural and functional evidence that a B chromosome in the characid fish *Astyanax scabripinnis* is an isochromosome. *Heredity* 85: 1-9.
- Metzker ML. 2010. Sequencing Technologies – the next generation. *Nature Reviews Genetics* 11: 31-46.
- Miao VP, Covert SF, VanEtten HD. 1991. A fungal gene antibiotic resistance on a dispensable (“B”) chromosome. *Science* 254: 1773-1776.
- Muñoz-Pajares AJ, Martínez-Rodríguez L, Teruel M, Cabrero J, Camacho JPM, Perfectti F. 2011. A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics* 187: 853-863.
- Navas-Castilho J, Cabrero J, Camacho JPM. 1986. C-Banding response of seven supernumerary heterochromatic segments in grasshoppers. *Cytobios* 47: 107-113.

- Novák P, Neumann P, Pech J, Steinhaisl J, and Macas J. 2013. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29: 792-793.
- Nur U. 1966. Harmful B chromosomes in a mealy bug population *Genetics* 54: 1225-1238.
- Oliveira C, Almeida-Toledo LF, Foresti F, Britski HA, Toledo-Filho SA. 1988. Chromosome formulae of Neotropical freshwater fishes. *Revista Brasileira de Genética* 11: 577-624.
- Oliveira C, Foresti F, Hilsdorf AWS. 2009. Genetics of neotropical fish: from chromosomes to populations. *Fish Physiology and Biochemistry* 35: 81 – 100
- Page BT, Wanous MK, Birchler JA. 2001. Characterization of a Maize chromosome 4 centromeric sequence: Evidence for an evolutionary relationship with the B chromosome centromere. *Genetics* 159: 291-302.
- Pardo MC, López-León MD, Viseras E, Cabrero J, Camacho JPM. 1995. Mitotic instability of B chromosomes during embryo development in *Locusta migratoria*. *Heredity* 74: 164-169.
- Pendás A, Morán P, Garcia-Vazquez E. 1993. Ribosomal RNA genes are interspersed throughout a heterochromatic chromosome arm in Atlantic salmon. *Cytogenetic and Genome Research* 63: 128-130.
- Pendás AM, Moran P, Freije JP, Garcia-Vazquez E. 1994. Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon 5S rDNA. *Cytogenetic and Genome Research* 67(1): 31–36.
- Peng SF, Lin YP, Lin BY. 2005. Characterization of AFLP sequences from regions of maize B chromosome defined by 12 B-10L translocations. *Genetics* 169:375–388.
- Peng SF, Cheng YM. 2011. Characterization of satellite CentC repeats from heterochromatic regions on the long arm of maize B-chromosome. *Chromosome research* 19: 183-191.
- Perfectti F, Werren JH. 2001. The interspecific origin of B chromosomes: experimental evidence. *Evolution* 55: 1069-1073.
- Pineau P, Henry M, Suspene R, Marchio A, Dettai A, Debruyne R, Petit T, Lécu A, Moisson P, Dejean A, Wain-Hobson S, Vartanian JP. 2005. A universal primer set for PCR amplification of nuclear histone H4 genes from all animal species. *Molecular Biology and Evolution* 22: 582-588.

- Pinkel D, Straume T, Gray JW. 1986. Cytogenetic analysis using quantitative, high sensitivity fluorescence hybridization. *Proceedings of the National Academy of Sciences* 83: 2934–2938.
- Poletto AB, Ferreira IA, Martins C. 2010. The B chromosomes of the African cichlid fish *Haplochromis obliquidens* harbour 18S rRNA gene copies. *BMC Genetics* 11: 1.
- Portela-Castro ALBS, Julio HF, Nishiyama PB. 2001. New occurrence of microchromosomes B in *Moenkhausia sanctaefilomenae* (Pisces, Characidae) from the Paraná River of Brazil: analysis of the synaptonemal complex. *Genetica* 110: 277–283.
- Portela-Castro ALB, Júlio-Júnior HF. 2001. Karyotype relationships among species of the subfamily Tetragonopterinae (Pisces, Characidae): Cytotaxonomic and evolution aspects. *Cytologia* 67: 329–336.
- Ruíz-Estévez M, López-León MD, Cabrero J, Camacho JPM. 2013. Ribosomal DNA is active in different B chromosome variants of the grasshopper *Eyprepocnemis plorans*. *Genetica* 141: 337-345.
- Salvador LB, Moreira-Filho O. 1992. B chromosomes in *Astyanax scabripinnis* (Pisces, Characidae). *Heredity* 69: 50-56.
- Sampaio TR, Gouveia JG, da Silva CRM, Dias AL, Rosa R. 2015. Molecular analysis of the B microchromosome in *Steindachnerina insculpta* (Characiformes: Curimatidae) by microdissection. *Cytogenetic and Genome Research* 146: 51 – 57.
- dos Santos LP, Castro JP, Francisco CM, Vicari MR, de Almeida MC, Goll LG, Morelli S, Artoni RF. 2013. Cytogenetic analysis in the neotropical fish *Astyanax goyacensis* Eigenmann, 1908 (Characidae, incertae sedis): karyotype description and occurrence of B microchromosomes. *Molecular Cytogenetics* 6: 48.
- Sapre A, Deshpande D. 1987. Origin of B chromosomes in *Coix L.* through spontaneous interspecific hybridization. *Journal of Heredity* 78: 191-196.
- Scalenghe F, Turco F, Edstrom JE, Pirrota V, Melli M. 1981. Microdissection and cloning of DNA from specific region of *Drosophila melanogaster* polytene chromosomes. *Chromosoma* 82: 205-216.
- Schartl M, Nanda I, Schlupp I, Wilde B, Eppelen JT, Schmid M, Parzefall J. 1995. Incorporation of subgenomic amounts of DNA as compensation for mutational load in a gynogenetic fish. *Nature* 373: 68-71.
- Scudeler PES, Diniz D, Wasko AP, Oliveira C, Foresti F. 2015. Whole chromosome painting of B chromosomes of the red-eye tetra *Moenkhausia sanctaefilomenae* (Teleostei, Characidae). *Comparative Cytogenetics* 9: 661.

- Schloötterer C. 2000. Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109: 365–371.
- Sebat J, Lakshmi B, Troge J, et al. 2004. Large-scale copy number polymorphism in the human genome. *Science* 305: 525 – 528
- Sharbel TF, Green DM, Houben A. 1998. B-chromosome origin in the endemic New Zealand frog *Leiopelma hochstetteri* through sex chromosome devolution. *Genome* 41: 14-22.
- Shaw MW, Hewitt GM. 1990. B chromosomes, selfish DNA and theoretical models: where next? In: *Oxford Surveys in Evolutionary Biology*, vol. 7 (ed. D. Futuyma & J. Antonovics), pp. 197-223. Oxford University Press.
- Silva DMZA, Pansonato-Alves JC, Utsunomia R, Araya-Jaime C, Ruiz-Ruano FJ, Daniel SN, Hashimoto DT, Oliveira C, Camacho JPM, Porto-Foresti F, Foresti F. 2014. Delimiting the origin of a B chromosome by FISH mapping, chromosome painting and DNA sequence analysis in *Astyanax paranae* (Teleostei, Characiformes). *PLoS ONE* 9:e94896
- Smit A, Hubley R. 2010. RepeatModeler Open-1.0. Repeat Masker Website.
- Smit A, Hubley R, Green P. 2010. RepeatMasker Open-3.0. See [http://www repeatmasker org](http://www.repeatmasker.org).
- Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A, Sampas N, Bruhn L, Shendure J, 1000 Genomes Project, et al. 2010. Diversity of human copy number variation and multicopy *genes*. *Science* 330: 641–646.
- Sumner AT. 1972. A simple technique for demonstrating centromeric heterochromatin. *Experimental Cell Research* 75: 304-306.
- Teacher AGF, Griffiths DJ. 2011. HapStar: automated haplotype network layout and visualization. *Molecular Ecology Resources* 11: 151-153.
- Teruel M, Cabrero J, Perfectti F, Camacho JPM. 2009. B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma* 119: 217-225.
- Teruel M, Cabrero J, Montiel EE, Acosta MJ, Sánchez A, Camacho JPM. 2009. Microdissection and chromosome painting of X and B chromosomes in *Locusta migratoria*. *Chromosome Research* 17: 11-18.
- Tosta VC, Tavares MG, Fernandes-Salomão TM, Barros EG, Campos LAO, Camacho JPM. 2007. Development of a SCAR marker for the analysis of B chromosomes presence in *Partamona helleri* (Hymenoptera, Apidae). *Cytogenetic and Genome Research* 116: 127-129.

- Trifonov VA, Perelman PL, Kawada SI, Iwasa MA, Oda SI, Graphodatsky AS. 2002. Complex structure of B-chromosomes in two mammalian species: *Apodemus peninsulae* (Rodentia) and *Nyctereutes procyonoides* (Carnivora). *Chromosome Research* 10: 109-116.
- Trifonov VA, Dementyeva PV, Larkin DM, O'Brien PC, Perelman PL, Yang F, Ferguson-Smith MA, Graphodatsky AS. 2013. Transcription of a protein-coding gene on B chromosomes of the Siberian roe deer (*Capreolus pygargus*). *BMC Biology* 11:1–11.
- Valente GT, Conte MA, Fantinatti BE, Cabral-de-Mello DC, Carvalho RF, Vicari MR, Kocher TD, Martins C. 2014. Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Molecular Biology and Evolution* 31:2061–2072.
- Vicari MR, Pistune HFM, Castro JP, Almeida MC, Bertollo LAC, et al. 2011. New insights on the origin of B chromosomes in *Astyanax scabripinnis* obtained by chromosome painting and FISH. *Genetica* 139: 1073–1081.
- Volff JN, Körting C, Sweeney K, Scharl M. 1999. The non-LTR retrotransposon Rex3 from the fish *Xiphophorus* is widespread among teleosts. *Molecular Biology and Evolution* 16: 1427–1438.
- Volff JN, Körting C, Scharl M. 2000. Multiple lineages of the non-LTR retrotransposon Rex1 with varying success in invading fish genomes. *Molecular Biology and Evolution* 17: 1673–1684.
- Volff JN, Körting C, Froschauer A, Sweeney K, Scharl M. 2001. Non-LTR retrotransposons encoding a restriction enzyme-like endonuclease in vertebrates. *Journal of Molecular Evolution* 52: 351-360.
- Wang X, Zhang Q, Ren J, Jiang Z, Wang C, Zhuang W, Zhai T. 2009. The preparation of sex-chromosome-specific painting probes and construction of sex chromosome DNA library in half-smooth tongue sole (*Cynoglossus semilaevis*). *Aquaculture* 297: 78-84.
- Weise A, Timmermann B, Grabherr M, Werber M, Heyn P, Kosyakova N, Liehr T, Neitzel H, Konrat K, Bommer C, Dietrich C, Rajab A, Reinhart R, Mundlos S, Lindner TH, Hoffman K. 2009. High-throughput sequencing of microdissected chromosomal regions. *European Journal of Human Genetics* 18: 457-462.

- White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ. PCR protocols: a guide to methods and applications. 315-322.
- Yoshida K, Terai Y, Mizoiri S, Aibara M, Nishihara H, Watanabe M, Kuroiwa A, Hirai H, Hirai Y, Matsuda Y, Okada N. 2011. B Chromosomes have a functional effect on female sex determination in lake victoria cichlid fishes. PLoS Genetics 7(8): e1002203.
- Zurita S, Cabrero J, López-León M, Camacho JPM. 1998. Polymorphism regeneration for a neutralized selfish B chromosome. Evolution 52: 274-277.

ANEXO 1: Produção Acadêmica e Científica no Doutorado: 2013-2015**Artigos completos publicados em periódicos**

1. Silva DMZA, Pansonato-Alves JC, **Utsunomia R**, Daniel SN, Hashimoto D, Oliveira C, Porto-Foresti F, Foresti F. (2013). Chromosomal organization of repetitive DNA sequences in *Astyanax bockmanni* (Teleostei, Characiformes): dispersive location, association and co-localization in the genome. *Genetica*, 141: 329-336
2. Pansonato-Alves JC, Serrano EA, **Utsunomia R**, Scacchetti PC, Oliveira C, Foresti F. (2013). Mapping five repetitive DNA classes in sympatric species of *Hypostomus* (Teleostei: Siluriformes: Loricariidae): analysis of chromosomal variability. *Reviews in Fish Biology and Fisheries*, 23: 477-489.
3. Pansonato-Alves JC, Hilsdorf AWS, **Utsunomia R**, Silva DMZA, Oliveira C, Foresti F. (2013). Chromosomal mapping of repetitive DNA and Cytochrome C oxidase I sequence analysis reveal differentiation among sympatric samples of *Astyanax fasciatus* (Characiformes, Characidae). *Cytogenetic and Genome Research*, 141: 133-142.
4. **Utsunomia R**, Scacchetti PC, Pansonato-Alves JC, Oliveira C, Foresti F. (2014). Comparative chromosome mapping of U2 snRNA and 5S rRNA genes in *Gymnotus* species (Gymnotiformes, Gymnotidae): evolutionary dynamics and sex chromosome linkage in *G. Pantanal*. *Cytogenetic and Genome Research*, *online first*.
5. **Utsunomia R**, Pansonato-Alves JC, Scacchetti PC, Oliveira C, Foresti F. (2014). Scattered organization of histone multigene Family and transposable elements in *Synbranchus*. *Genetics and Molecular Biology*, 37: 30-36.
6. Silva DMZA, Pansonato-Alves JC, **Utsunomia R**, Araya-Jaime C, Ruiz-Ruano FJ, Daniel SN, Oliveira C, Camacho JPM, Porto-Foresti F, Foresti F. (2014). Delimiting the origin of a B chromosome by FISH mapping, chromosome painting and DNA sequence analysis in *Astyanax paranae* (Teleostei, Characiformes). *PLoS one*, 9(4): e94896.
7. Scacchetti PC, **Utsunomia R**, Pansonato-Alves JC, Costa-Silva GJ, Oliveira C, Foresti F. (2015) Extensive spreading of interstitial telomeric sites on the chromosomes of *Characidium* (Teleostei, Characiformes). *Genetica*, *online first*.

8. Sene VF, Pansonato-Alves JC, **Utsunomia R**, Oliveira C, Foresti F. (2014) Karyotype diversity and patterns of chromosomal evolution in *Eigenmannia* (Teleostei, Gymnotiformes, Sternopygidae). *Comparative Cytogenetics*, 8: 301-311
9. **Utsunomia R**, Pansonato-Alves JC, Paiva LRS, Costa-Silva GJ, Oliveira C, Bertollo LAC, Foresti F. (2014) Genetic differentiation among distinct karyomorphs of the wolf fish *Hoplias malabaricus* species complex (Characiformes, Erythrinidae) and report of unusual hybridization with natural triploidy. *Journal of Fish Biology*, 85:1682-1692.
10. **Utsunomia R**, Pansonato-Alves JC, Costa-Silva GJ, Mendonça FF, Scacchetti PC, Oliveira C, Foresti F. (2014) Molecular and cytogenetic analyses of cryptic species within the *Synbranchus marmoratus* Bloch, 1795 (Synbranchiformes: Synbranchidae) grouping: species delimitations, karyotypic evolution and intraspecific diversification. *Neotropical Ichthyology*, 12: 903-911.
11. Pansonato-Alves JC, Serrano EA, **Utsunomia R**, Camacho JPM, Costa-Silva GJ, Vicari MR, Artoni RF, Oliveira C, Foresti F. (2014) Single origin of sex chromosomes and multiple origins of B chromosomes in fish genus *Characidium*. *PLoS One*, 9: e107169.
12. Scacchetti PC, **Utsunomia R**, Pansonato-Alves JC, Vicari MR, Artoni RF, Oliveira C, Foresti F. (2015) Chromosomal mapping of repetitive DNAs in *Characidium* (Teleostei, Characiformes): Genomic organization and diversification of ZW sex chromosomes. *Cytogenetic and Genome Research*, 146: 136-143.
13. Silva DMZA, **Utsunomia R**, Pansonato-Alves JC, Oliveira C, Foresti F. (2015) Chromosomal mapping of repetitive DNA sequences in five species of *Astyanax* (Characiformes, Characidae) reveals independent location of U1 and U2 snRNA sites and association of U1 snRNA and 5S rDNA. *Cytogenetic and Genome Research*, 146: 144-152.
14. Scacchetti PC*, **Utsunomia R***, Pansonato-Alves JC, Costa-Silva GJ, Vicari MR, Artoni RF, Oliveira C, Foresti F. (2015) Repetitive DNA sequences and evolution of ZZ/ZW sex chromosomes in *Characidium* (Teleostei: Characiformes). *PLoS ONE*, 10(9): e0137231. * Contribuíram igualmente para o trabalho
15. Acosta AA, Caffara M, Fioravanti ML, **Utsunomia R**, Zago AC, Franceschini L, Silva RJ. (2015) Morphological and molecular characterization of *Clinostomum detruncatum* metacercariae infecting *Synbranchus marmoratus*. *Journal of Parasitology*, *online first*.

16. Sene VF, Pansonato-Alves JC, Ferreira DC, Utsunomia R, Oliveira C, Foresti F. (2015) Mapping of the retrotransposable elements *Rex1* and *Rex3* in chromosomes of *Eigenmannia* (Teleostei, Gymnotiformes, Sternopygidae). *Cytogenetic and Genome Research*, 146: 319-324.

Artigos submetidos para publicação

1. Oliveira MLM, **Utsunomia R**, Pansonato-Alves JC, Scacchetti PC, Primo CC, Vicari MR, Artoni RF, Centofante L, Moreira-Filho O, Oliveira C, Foresti F. (2016) Microstructural chromosome reorganization in the genus *Trichomycterus* (Siluriformes, Trichomycteridae). *Neotropical Ichthyology*.
2. Silva DMZA, **Utsunomia R**, Pansonato-Alves JC, Daniel SN, Penitente M, Hashimoto DT, Ruiz-Ruano FJ, Camacho JPM, Oliveira C, Porto-Foresti F, Foresti F. (2016) Independent origin of B chromosomes in the genus *Astyanax* (Characiformes, Characidae) and the limits of chromosome painting. *Molecular Genetics and Genomics*.
3. **Utsunomia R**, Silva DMZA, Ruiz-Ruano FJ, Araya-Jaime C, Pansonato-Alves JC, Hashimoto DT, Oliveira C, Trifonov VA, Porto-Foresti F, Camacho JPM, Foresti F. (2016) Uncovering the ancestry of B chromosomes in *Moenkhausia sanctaefilomenae* (Teleostei, Characidae). *PLoS ONE*.

Artigos em preparação

1. Scacchetti PC, **Utsunomia R**, Hermida M, Bekkaert M, Mank J, Taggart J, Foresti F, Martinez P. (2016) Evolution of sex chromosome differentiation in the genus *Characidium*: identification, location and transpecific conservation of sex-associated RAD-markers from *Characidium gomesi*. *Heredity*.

Capítulos de livros publicados

1. **Utsunomia R**, Pansonato-Alves JC, Scacchetti PC, Silva DMZA, Oliveira C, Foresti F. (2015) Citogenética Molecular: Fundamentos Básicos, Métodos e Aplicações. Cross-species/ZOO-FISH, Reverse-FISH, Reciprocal FISH. In: Rodrigo R. Resende; Carlos R. Soccol (Org.). *Biotecnologia Aplicada à Saúde, Fundamentos e Aplicações*. 1ed, São Paulo: Blucher, pp. 38-420.

Estágio no exterior

1. **Institute of Molecular and Cellular Biology, Novosibirsk, Rússia.** Treinamento em microdissecção e preparação de bibliotecas cromossômicas para produção de sondas cromossômicas e material para sequenciamento *next-generation*. Supervisor: Dr. Vladimir A. Trifonov

Prêmios e títulos

1. 2013 - Menção honrosa – Prêmio Pós-graduação. 3ª Reunião Brasileira de Citogenética e IV Simpósio Latino Americano de Citogenética e Evolução.
2. 2014 - Melhor trabalho - Prêmio Sílvio de Almeida Toledo. 60º Congresso Brasileiro de Genética.
3. 2014 - Menção honrosa – Prêmio Sílvio de Almeida Toledo. 60º Congresso Brasileiro de Genética.
4. 2015 – Melhor trabalho – Prêmio de Melhor trabalho de citogenética. XVI Simpósio de Citogenética e Genética de Peixes.