

ANÁLISE FUNCIONAL E EVOLUTIVA DO CROMOSSOMO B EM  
*Astyanax paranae* (CHARACIFORMES, CHARACIDAE)

**DUÍLIO MAZZONI ZERBINATO DE ANDRADE SILVA**

UNIVERSIDADE ESTADUAL PAULISTA  
"Júlio de Mesquita Filho"

INSTITUTO DE BIOCIÊNCIAS DE BOTUCATU

ANÁLISE FUNCIONAL E EVOLUTIVA DO CROMOSSOMO B  
EM *Astyanax paranae* (CHARACIFORMES, CHARACIDAE)

**Aluno:** Duílio Mazzoni Zerbinato de Andrade Silva

**Orientador:** Prof. Dr. Fausto Foresti

Tese apresentada ao Instituto de  
Biociências, Campus de Botucatu,  
UNESP, para obtenção do título de Doutor  
no Programa de Pós-Graduação em  
Ciências Biológicas (Genética)

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.  
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP  
BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Silva, Duílio Mazzoni Zerbinato de Andrade.

Análise funcional e evolutiva do cromossomo B em *Astyanax paranae* (Characiformes, Characidae) / Duílio Mazzoni Zerbinato de Andrade Silva. - Botucatu, 2018

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu  
Orientador: Fausto Foresti  
Capes: 20200005

1. Cromossomos humanos. 2. Expressão gênica. 3. Genes. 4. DNA Mitocondrial. 5. Cordados. 6. Peixe - Genética.

Palavras-chave: Cordados; Cromossomos B; DNA mitocondrial; Expressão Gênica; Genes Codificadores de Proteínas.

A lei da mente é implacável.  
O que você pensa, você cria;  
o que você sente, você atrai;  
o que você acredita,  
torna-se realidade.

*Buda*

Dedico este trabalho à minha família.

Esta pesquisa foi financiada com recursos da  
Fundação de Amparo à Pesquisa do Estado de São Paulo  
FAPESP - Processo Número: 2013/24367-0



## **Agradecimentos**

Agradeço ao professor Fausto Foresti pela orientação e motivação para a realização deste trabalho e também pelos ensinamentos sobre a vida e as relações humanas.

Ao professor Cláudio de Oliveira, pela orientação e exemplo como pesquisador.

Ao colega Ricardo Utsunomia pela colaboração na execução deste trabalho e por acreditar em mim mais do que eu mesmo, sempre me motivando a ir além do trivial.

Aos colegas do Laboratório de Biologia e Genética de Peixes, em especial aos amigos do laboratório de citogenética.

Ao técnico Renato Devidé pelo trabalho empenhado neste projeto, ensinamentos e momentos de descontração.

Aos membros da Universidade de Granada, em especial ao professor Juan Pedro M. Camacho e ao colega Francisco J. Ruíz Ruano, pelos ensinamentos, discussões e colaboração.

Aos membros do Laboratório de Biologia do Músculo Estriado da UNESP de Botucatu, em especial ao professor Robson Francisco Carvalho e à colega Paula Pacielli Freire, pela colaboração.

Aos membros do Laboratório de Genética de Peixes da UNESP de Bauru, pela parceria, em especial ao professor Fábio Porto Foresti e ao colega Sandro Natal Daniel.

Aos membros do Laboratório de Biologia Molecular e Reprodutiva da UNESP de Botucatu, em especial ao professor Rafael Henrique Nóbrega.

Ao Instituto de Biociências e à UNESP, pela minha formação como Biólogo.

Ao programa de pós-graduação em Genética.

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

A minha família, por tudo, mas principalmente por me permitir escolher e desenvolver meus projetos em um ambiente familiar de amor, carinho e respeito.

Aos meus amigos, por serem a família que eu escolhi e à minha namorada Joyce por me apoiar nos momentos finais deste trabalho.

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

## Resumo

Cerca de 15% dos organismos eucariotos possuem elementos genômicos adicionais dispensáveis chamados cromossomos B. Na maioria dos casos a composição e os efeitos da presença destes cromossomos é desconhecida. Estes elementos estão presentes em diversas espécies de peixes do gênero *Astyanax*. Assim, os objetivos do presente trabalho foram caracterizar a composição do cromossomo B de *A. paranae*, bem como os efeitos da presença deste cromossomo nesta espécie e sua relação com cromossomos B de espécies próximas. Foram identificados sete genes codificadores de proteínas no cromossomo B de *A. paranae*, dos quais quatro mostraram expressão diferencial nos ovários de fêmeas 0B e 1B. Além disto, quatro destes genes estão presentes nos cromossomos B de *A. fasciatus* e de *A. bockmanni*. Desta forma, nós concluímos que o cromossomo B de *A. paranae* interfere na expressão dos genes que carrega em seu próprio benefício e possui uma origem comum com cromossomos B encontrados em outras espécies de *Astyanax*. Além deste objetivo principal do trabalho, também foi realizada uma análise do DNA mitocondrial de *A. paranae* e, por extensão, dos vertebrados em geral. Esta análise revelou inicialmente que o conteúdo de bases do DNA mitocondrial dos vertebrados tende a um padrão comum. Adicionalmente, foi possível distinguir grupos entre as classes de vertebrados com conteúdos semelhantes na composição de bases do DNA mitocondrial, indicando que a evolução do DNA destas organelas poderia estar relacionada às características biológicas de cada grupo, como modo de respiração, ocupação do meio terrestre ou aquático, modo de vida em água doce ou salgada e taxa metabólica dos organismos.



## Abstract

About 15% of eukaryotic organisms have additional and dispensable genomic elements called B chromosomes. In most cases the composition and effects of the presence of these chromosomes is unknown. These elements are present in several species of fish of the genus *Astyanax*. Thus, the objectives of the present work were to characterize the composition of the B chromosome of *A. paranae*, as well as the effects of the presence of this chromosome in this species and its relation with B chromosomes of nearby species. Seven protein-coding genes were identified on the B chromosome of *A. paranae*, of which four showed differential expression in the ovaries of females 0B and 1B. In addition, five of these genes are present on the B chromosomes of *A. fasciatus* and *A. bockmanni*. Thus, we conclude that the B chromosome of *A. paranae* interferes with the expression of the genes it carries for its own benefit and has a common origin with B chromosomes found in other species of *Astyanax*. In addition to this main objective of the work, an analysis of the mitochondrial DNA of *A. paranae* and, by extension, of the vertebrates in general was also performed. This analysis initially revealed that the base content of vertebrate mitochondrial DNA tends to a common pattern, and it was possible to distinguish groups among vertebrate classes with similar contents in mitochondrial DNA bases, indicating that the DNA evolution of these organelles could be related to the biological characteristics of each group, such as the type of respiration, occupation of the terrestrial or aquatic environment, freshwater or salt water livelihood and metabolic rates of organisms.

## Sumário

<b>1 INTRODUÇÃO</b>	<b>1</b>
<b>2 OBJETIVOS</b>	<b>5</b>
2.1 OBJETIVOS GERAIS	5
2.2 OBJETIVOS ESPECÍFICOS	5
2.3 OBJETIVOS ADICIONAIS	6
<b>3 MATERIAIS E MÉTODOS</b>	<b>7</b>
<b>3.1 MATERIAIS</b>	<b>7</b>
<b>3.2 MÉTODOS</b>	<b>8</b>
3.2.1 ESTIMULAÇÃO DE MITOSES	8
3.2.2 OBTENÇÃO DE CROMOSSOMOS METAFÁSICOS MITÓTICOS	8
3.2.3 COLORAÇÃO CONVENCIONAL COM GIEMSA	8
3.2.4 DETECÇÃO DA HETEROCROMATINA CONSTITUTIVA (BANDA C)	8
3.2.5 HIBRIDAÇÃO <i>IN SITU</i> FLUORESCENTE POR DEPOSIÇÃO CATALISADA DE REPÓRTER CARD-FISH	8
3.2.6 ESTUDOS CARIOTÍPICOS	9
3.2.7 EXTRAÇÃO DE DNA GENÔMICO	9
3.2.8 EXTRAÇÃO DE RNA E SÍNTESE DO cDNA	9
3.2.8 SEQUENCIAMENTO DE NOVA GERAÇÃO DE DNA – DNA-SEQ	10
3.2.9 SEQUENCIAMENTO DE NOVA GERAÇÃO DE RNA – RNA-SEQ	10
3.2.10 DETECÇÃO DE GENES POSSIVELMENTE PRESENTES NO CROMOSSOMO B	11
3.2.11 CONFIRMAÇÃO DOS GENES PRESENTES NO CROMOSSOMO B POR QPCR	13
3.2.12 CONFIRMAÇÃO DA EXPRESSÃO DOS GENES PRESENTES NO CROMOSSOMO B POR RT-QPCR	13
3.2.13 ANÁLISES ESTATÍSTICAS	14
<b>4 RESULTADOS E DISCUSSÃO</b>	<b>15</b>
<b>4.1 CAPÍTULO 1</b>	<b>16</b>
ANÁLISE FUNCIONAL E EVOLUTIVA DO CROMOSSOMO B EM <i>ASTYANAX PARANAE</i> (CHARACIFORMES, CHARACIDAE).	16
RESULTADOS	17
DISCUSSÃO	25
<b>4.2 CAPÍTULO 2</b>	<b>30</b>
TENDÊNCIA EVOLUTIVA NA COMPOSIÇÃO DE BASES DO DNA MITOCONDRIAL ASSOCIADA A FATORES BIOLÓGICOS DAS CLASSES DE CORDADOS.	30
LETTER	31
<b>5 CONSIDERAÇÕES GERAIS</b>	<b>38</b>
<b>6 REFERÊNCIAS</b>	<b>40</b>
<b>7 MATERIAL SUPLEMENTAR – CAPÍTULO 1</b>	<b>44</b>
<b>8 MATERIAL SUPLEMENTAR – CAPÍTULO 2</b>	<b>54</b>
<b>9 APÊNDICES</b>	<b>56</b>

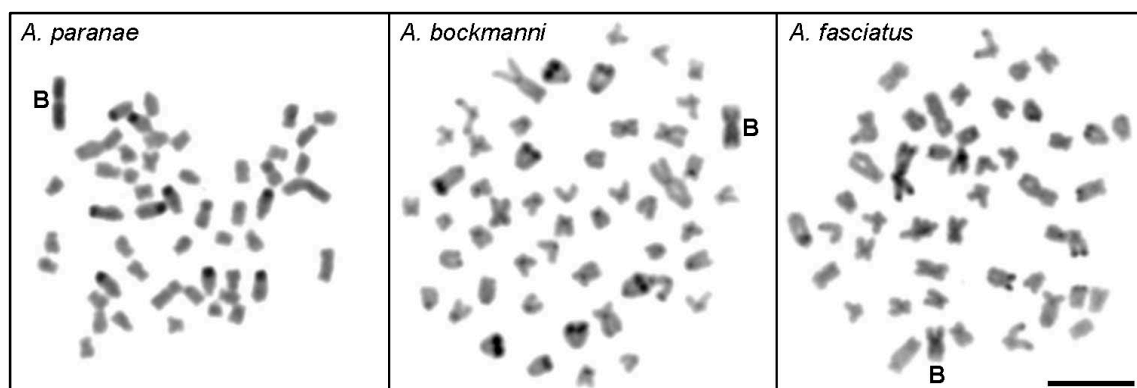
## 1 Introdução

Os organismos eucariotos possuem um lote padrão de cromossomos denominado complemento A. Além disso, cerca de 15% destes organismos possuem cromossomos adicionais denominados cromossomos B, ou supranumerários. Esses cromossomos são dispensáveis para a sobrevivência dos organismos e são encontrados em uma parte dos indivíduos de determinadas populações (Camacho et al. 2000). Sua morfologia é extremamente variada, até mesmo entre populações de uma mesma espécie, sendo encontrada grande diversidade em relação ao tamanho, sendo encontrados desde micros até grandes cromossomos B (Carvalho et al. 2008). Estes elementos podem se originar do próprio genoma A de uma dada espécie (Camacho 2005) ou podem ser adquiridos de espécies próximas através de processos de hibridação (Tosta et al. 2015).

Durante muito tempo acreditou-se que os cromossomos B eram elementos inertes e pobres em genes codificadores de proteínas (GCPs) (revisão em Camacho et al. 2005). De fato, até agora foram encontrados poucos GCPs nos cromossomos B analisados (revisão em Valente et al. 2016; Clark et al. 2016; Makunin et al. 2016; Navarro-Domínguez et al. 2017). No entanto, o conceito de que os cromossomos B são estruturas genéticas inertes vem sendo alterado pela descoberta e descrição de transcritos provenientes destes elementos (revisão em Valente et al. 2016; Navarro-Domínguez et al. 2017). Para exercer atividades funcionais e evoluir, os cromossomos B podem realizar modificações nos processos celulares normais, interagindo com o genoma dos hospedeiros em seu próprio benefício (Valente et al. 2016). Efeitos funcionais variados associados aos cromossomos B foram relatados em algumas espécies, como por exemplo, a ocorrência de resistência a antibióticos no fungo *Nectria haematococca* (Miao et al. 1991), a razão sexual paternal na mosca *Nasonia vitripennis* (Werren 1991), o aumento da viabilidade das sementes em alho (Plowman & Bougourd 1994), a participação no processo de determinação sexual em peixes ciclídeos (Yoshida et al.

2011), a influência na resposta imune e carga de parasitas no roedor *Apodemus flavicollis* (Adnađević et al. 2014) e a tolerância ao calor durante os estágios iniciais da esporogênese masculina em centeio (Pereira et al. 2017). Além disso, foi recentemente demonstrado que o cromossomo B de centeio é capaz de codificar uma proteína similar à argonauta AGO4B, com atividade “slicer” *in vitro* (Ma et al. 2016).

Neste sentido, consideramos interessante a investigação do conteúdo genético dos cromossomos B, bem como da expressão dos seus genes para uma melhor compreensão a respeito da origem e efeitos da presença destes elementos em organismos eucariotos. Assim, as espécies de peixes do gênero *Astyanax* surgem como um interessante modelo de investigação.



**Figura 1:** Metáfases de espécies de *Astyanax* após bandamento-C. Notar os diferentes tipos de cromossomos B presentes nas células destes organismos. Barra = 10  $\mu\text{m}$ .

O gênero *Astyanax* é representado por mais de 240 espécies de peixes (Eschmeyer 2017) e está distribuído do sul dos Estados Unidos ao centro da Argentina (Ornelas-García et al. 2008), num grupo complexo de organismos que se apresenta com espécies crípticas e complexos de espécies (Moreira-Filho & Bertollo 1991; Pansonato-Alves et al. 2013; Castro et al. 2014). Esta grande diversidade específica também está representada pela diversidade citogenética encontrada neste gênero, no qual os números diplóides variam entre  $2n=36$  em *A. schubarti* (Morelli et al. 1983) e  $2n=50$  em diversas outras espécies (Oliveira et al. 2008), bem como por inúmeros rearranjos macro/micro cromossômicos, uma alta dispersão de DNAs

repetitivos (Silva et al. 2017) e diversos tipos de cromossomos B (Oliveira et al. 2008) (Figura 1). Entre os cromossomos B encontrados no gênero *Astyanax*, a variante metacêntrica com tamanho similar ao do primeiro par do cariótipo normal da espécie é a mais comum (Oliveira et al. 2008). Foi postulado então que este tipo de elemento representaria a variante ancestral dos cromossomos B deste gênero e que sua presença nas diferentes espécies teria uma origem comum (Moreira-Filho et al. 2004).

Até o momento, somente sequências de DNA repetitivo haviam investigadas nos cromossomos B de *Astyanax*. Primeiramente, foi descrita a ocorrência do DNA satélite As-51 no cromossomo B de *A. scabripinnis* (Mestriner et al. 2000). Este DNA satélite mostrou-se distribuído ao longo dos dois braços deste cromossomo de forma simétrica, permitindo inferir de que se trataria de um isocromossomo, o que foi reforçado pela ocorrência de auto-pareamento deste cromossomo durante a meiose (Mestriner et al. 2000). Posteriormente, foram identificadas sequências de diversos tipos de DNA repetitivo (i.e. microsatélite (AC)<sub>15</sub>, DNAr 18S, satélite As-51 e histonas H1 e H3) nos cromossomos B de cinco espécies de *Astyanax* (Silva et al. 2014, 2016). A partir da identificação de sequências de DNAr 18S, responsáveis pela região organizadora de nucléolos – RON, nos cromossomos de *A. paranae* e *A. fasciatus*, foi possível testar a atividade destas sequências por meio de impregnação por nitrato de prata (Ag-RON). Isto revelou que somente as sequências de DNAr 18S presentes no cromossomo B de *A. fasciatus* estão ativas (Silva et al. 2016), demonstrando que este cromossomo B pode expressar seus genes e ser um elemento ativo.

Além disso, a presença de 35 DNAs satélites foi descrita nos cromossomos B de três espécies de *Astyanax* (Silva et al. 2017), sendo que os dados obtidos no estudo permitiram inferir que os cromossomos B de *A. paranae* e *A. bockmanni* permanecem perfeitamente isométricos, enquanto que o cromossomo B de *A. fasciatus* apresenta distribuição diferencial para alguns DNAs satélites entre seus braços, indicando que este cromossomo vem perdendo

isometria ao longo do tempo. Desta forma, os autores concluíram que o cromossomo B de *A. fasciatus* seria mais antigo que os cromossomos B das duas outras espécies. Porém, a discussão sobre a origem comum ou independente destes elementos ainda permanece aberta.

Quanto aos efeitos funcionais dos cromossomos B no gênero *Astyanax*, foi proposta uma possível participação no mecanismo de determinação sexual para os cromossomos B de *A. paranae*, uma vez que a presença destes cromossomos poderia estar relacionada à distorção da proporção sexual na população encontrada no Rio Capivara, Botucatu, SP (Silva et al. 2014). Esta hipótese é suportada pela ocorrência de cromossomos B restritos a indivíduos de apenas um dos sexos em diversas populações (Maistro et al. 1992; Rocon-Stange & Almeida-Toledo 1993; Mizoguchi & Martins-Santos 1997).

No presente trabalho, foi demonstrada a ocorrência de sete genes codificadores de proteínas no cromossomo B de *A. paranae*, dos quais quatro são compartilhados com os cromossomos B de *A. bockmanni* e de *A. fasciatus*. Assim, pode ser suposto que os cromossomos B destas três espécies possuem origem comum. Além disso, quatro genes presentes no cromossomo B de *A. paranae* mostraram expressão diferencial nos ovários de fêmeas 0B e 1B, demonstrando que o cromossomo B desta espécie deve exercer influência na expressão dos genes nos indivíduos portadores.

## 2 Objetivos

### 2.1 Objetivos Gerais

O presente projeto se insere 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. Assim, os objetivos gerais do presente trabalho foram:

- a) caracterizar o conteúdo genômico do cromossomo B de *A. paranae*;
- b) verificar os efeitos da presença do cromossomo B em *A. paranae*;
- c) testar se os cromossomos B do gênero *Astyanax* tiveram uma origem comum ou independente.

### 2.2 Objetivos Específicos

Visando-se atingir os objetivos gerais, os objetivos específicos deste trabalho foram os seguintes:

- a) identificar tipos de DNAs repetitivos presentes no cromossomo B de *A. paranae*;
- b) identificar genes codificadores de proteínas no cromossomo B de *A. paranae*;
- c) testar se a presença do cromossomo B altera a expressão dos genes que ele porta em fêmeas de *A. paranae* 0B e 1B;
- d) verificar se os genes presentes no cromossomo B de *A. paranae* estão presentes nos cromossomos B de outras espécies de *Astyanax*.

### 2.3 Objetivos Adicionais

O sequenciamento de Nova Geração de amostras de DNA total permite a obtenção não só de sequências do DNA nuclear, como também do DNA das mitocôndrias nos animais. A partir dos dados gerados é possível a obtenção da sequência completa do DNA destas organelas. Desta forma, os dados de NGS gerados neste trabalho também foram utilizados para se atingir os seguintes objetivos adicionais:

- a) descrever a sequência completa do DNA mitocondrial de *A. paranae*;
- b) caracterizar a composição de bases do DNA mitocondrial de diferentes classes de vertebrados, visando identificar padrões de composição e distribuição de bases ao longo da história evolutiva destes animais.



### 3 Materiais e Métodos

#### 3.1 Materiais

Foram analisados indivíduos das espécies de peixes caracídeos *A. paranae*, *A. fasciatus* e *A. bockmanni* (Tabela 1). Pesquisas anteriores mostraram a presença de cromossomos B em indivíduos de todas as localidades amostradas (Maistro et al. 1994; Daniel et al. 2012; Silva et al. 2016). As amostras foram coletadas de acordo com a legislação brasileira de proteção ambiental (permissão de coleta MMA/IBAMA/SISBIO-número 3245). Os procedimentos para amostragem, manutenção e análise das amostras foram realizados de acordo com diretrizes internacionais para o cuidado e uso de animais em pesquisa seguido pelo Colégio Brasileiro de Experimentação Animal (COBEA) e foram aprovados pelo Comitê de Ética da UNESP sobre Uso de Animais (CEUA) do Instituto de Biociências de Botucatu.

Após a retirada de material amostral e das análises, os indivíduos foram depositados na coleção do Laboratório de Biologia e Genética de Peixes (LBP) da UNESP, Botucatu, São Paulo, Brasil, sob os códigos LBP19572 (*A. paranae*) e LBP19573 (*A. fasciatus*). Os espécimes de *A. bockmanni* foram depositados na coleção do Laboratório de Genética de Peixes da UNESP, Bauru, São Paulo, Brasil.

**Tabela 1:** Localidades, número de indivíduos, número diplóide (2n) e características dos cromossomos B de *Astyanax* analisados.

Espécie	Localidade	Coordenadas	Indivíduos	2n	Tipo B	Nome B
<i>A. paranae</i>	Ribeirão Cascatinha	22°53'30"S 48°28'36"W	20	50	M grande	BpM
<i>A. fasciatus</i>	Córrego das Araras	22°22'43"S 47°25'37"W	25	48	M grande	BfMb
<i>A. bockmanni</i>	Rio Alambari	22°27'6"S 49°14'25"W	7	50	M grande	BbM

M = metacêntrico.

## **3.2 Métodos**

### **3.2.1 Estimulação de mitoses**

Com o objetivo de se obter uma maior frequência de células em divisão mitótica nas preparações cromossômicas, os animais foram mantidos em aquários aquecidos a 26°C durante 24 a 48 horas antes de serem preparados. Este método se mostrou eficiente para a estimulação de mitoses, sendo considerado menos invasivo que o método usualmente empregado com injeção de uma solução de fermento biológico (Cole & Levans 1971).

### **3.2.2 Obtenção de cromossomos metafásicos mitóticos**

Para obtenção dos cromossomos metafásicos foi utilizada a técnica proposta por (Foresti et al. 1981), com adaptações (Mat. Sup.). Esta metodologia envolve basicamente a inibição da polimerização dos microtúbulos pela ação da colchicina no momento da divisão celular, a hipotonização das células em suspensão por solução de KCl e a fixação celular pela mistura de metanol/ácido acético.

### **3.2.3 Coloração convencional com Giemsa**

As preparações cromossômicas fixadas nas lâminas foram coradas por 5 minutos com solução de Giemsa a 5 % em tampão fosfato (pH=6,7).

### **3.2.4 Detecção da heterocromatina constitutiva (banda C)**

Para obtenção de bandas-C foi utilizada a técnica descrita originalmente por (Sumner 1972), com adaptações (Mat. Sup.). Esta técnica baseia-se na ação sequencial de tratamentos com base, ácido e solução salina no material celular para eliminar seletivamente regiões do DNA cromossômico, permanecendo intactas somente regiões de heterocromatina.

### **3.2.5 Hibridação *in situ* fluorescente por deposição catalisada de repórter CARD-FISH**

Para a detecção do gene *nusap* no cromossomo B de *A. paranae* foi utilizada a técnica de CARD-FISH utilizando-se o kit HNPP Fluorescent Detection Set (Roche),

segundo instruções do fabricante. Esta técnica envolve a hibridação de uma sonda marcada com um anticorpo primário (digoxigenina-11-dUTP), o qual é detectado por um anticorpo secundário (anti-digoxigenina) associado a uma fosfatase alcalina. A amplificação do sinal em relação à técnica de FISH convencional é baseada na desfosforilação de múltiplas moléculas de HNPP (2-hydroxy-3-naphtoic acid-2'-phenylanilide phosphate) pela ação da fosfatase alcalina, gerando um precipitado de HNP-HNPP desfosforilado (composto altamente fluorescente) no sítio alvo. A sonda utilizada foi marcada por PCR (Mat. Sup.) incluindo-se na reação de amplificação o nucleotídeo digoxigenina-11-dUTP (Roche).

### **3.2.6 Estudos Cariotípicos**

As lâminas foram analisadas em fotomicroscópio óptico de fluorescência (Olympus BX61) e as imagens foram capturadas usando o programa Image Pro Plus 6.0 software (MediaCybernetics).

### **3.2.7 Extração de DNA genômico**

O DNA genômico dos exemplares analisados foi obtido a partir de amostras de fígado ou de músculo preservadas em etanol 95%, utilizando-se o Kit NucleoSpin® Tissue (Macherey Nagel) conforme instruções do fabricante. Os produtos gerados foram checados em gel de agarose 1%.

### **3.2.8 Extração de RNA e Síntese do cDNA**

Os indivíduos de *A. paranae* utilizados para a extração de RNA foram obtidos durante uma única coleta e foram mantidos sob as mesmas condições ambientais. Estes indivíduos foram sacrificados em solução de benzocáína no menor intervalo de tempo possível (período de 1 a 2 horas), uma vez que análises citogenéticas foram realizadas concomitantemente. Para extração de RNA, foram separadas amostras de tecidos de ovário e músculo, as quais foram imediatamente congeladas em nitrogênio líquido e armazenadas a -70 °C. O RNA foi extraído usando-se o kit TRIzol® (Invitrogen), conforme instruções

do fabricante. Posteriormente, as amostras foram tratadas com DNaseI e foram checadas em gel de agarose e pelo equipamento 2100 Bioanalyzer.

As amostras de RNA foram então separadas para síntese do cDNA e para a confecção das bibliotecas para RNA-seq. O cDNA correspondente a cada amostra foi sintetizado utilizando-se o kit High-Capacity cDNA Reverse Transcription (Applied Biosystems), conforme instruções do fabricante.

### **3.2.8 Sequenciamento de Nova Geração de DNA – DNA-seq**

Para a busca das sequências de DNA localizadas no cromossomo B de *A. paranae*, foi realizado o sequenciamento Illumina HiSeq2000 *paired-end* (2x101bp) utilizando-se DNA genômico de um indivíduo sem cromossomo B (0B) e de outro portador de um cromossomo B (1B), confirmados por bandamento-C. O sequenciamento massivo do genoma do indivíduo 0B gerou aproximadamente 341 milhões de *reads* (34 Gb), representando uma cobertura de 18,4 vezes do genoma da espécie (1,87 Gb) (Carvalho et al. 2002). Já o sequenciamento massivo do genoma do indivíduo 1B gerou aproximadamente 140 milhões de *reads* (14 Gb), representando uma cobertura de 7,5 vezes do genoma da espécie. As altas coberturas obtidas foram fundamentais para a detecção de genes codificadores de proteína diferencialmente representados entre os genomas 0B e 1B.

### **3.2.9 Sequenciamento de Nova Geração de RNA – RNA-seq**

O sequenciamento massivo das sequências de RNA transcritas em indivíduos sem e com cromossomos B (0B e 1B) foi realizado no sentido de melhor conhecer o conteúdo e efeitos destes cromossomos. A montagem dos transcriptomas pode ser considerada uma abordagem fundamental para auxiliar os estudos sobre o genoma de uma espécie, uma vez que, embora gerando uma quantidade menor de informações, apresenta-se consequentemente mais simples de se analisar, sendo útil para facilitar as análises iniciais

das sequências de DNA. Nesse sentido, foram sequenciadas amostras de gônadas e músculos de fêmeas de *A. paranae* com e sem cromossomos B, incluindo 3 fêmeas em cada grupo, totalizando 12 amostras. O sequenciamento massivo das amostras de RNA das gônadas gerou em torno de 33 milhões de *reads* (4,0 Gb) considerando os 6 exemplares amostrados. Para as 6 amostras de músculo foram obtidos em torno de 38 milhões de *reads* (4,7 Gb).

### 3.2.10 Detecção de genes possivelmente presentes no cromossomo B

A partir dos dados do RNA-seq, foi possível montar o transcriptoma de gônada e de músculo de *A. paranae* para serem utilizados como referência no mapeamento dos *reads* genômicos. Para isto, as 12 bibliotecas de RNA-seq foram concatenadas e uma normalização *in silico* com 50x de cobertura máxima foi realizada. Posteriormente, a montagem (*assemble*) foi executada com o programa Trinity (Haas et al. 2013) utilizando-se as opções padrão.

No presente projeto, um dos principais objetivos era identificar genes codificadores de proteínas presentes no cromossomo B de *A. paranae*. Desta forma, os CDSs maiores que 100 aminoácidos foram extraídos utilizando-se o programa Transdecoder (Haas et al. 2013); a redundância foi reduzida com o uso do programa CDHit-EST (Fu et al. 2012) através de alinhamento local e algoritmo *greedy*, sendo posteriormente agrupadas as sequências com 80% ou mais de similaridade em até 80% do comprimento (opções `-M 0 -aS 0.8 -c 0.8 -G 0 -g 1`).

Após a obtenção do referido transcriptoma de referência, foi realizado o mapeamento dos *reads* resultantes do sequenciamento dos transcriptomas e genomas dos indivíduos 0B e 1B utilizando o programa SSAHA2 (Ning et al. 2001), com pontuação mínima de alinhamento de 40 e 80% de identidade mínima. Além disso, foi utilizado um script customizado para contar o número de *reads* mapeados, como medida de abundância.

A abundância de cada CDS em cada biblioteca foi expressa como o número de cópias no genoma haploide usando a fórmula:

$$\text{Número de cópias} = \frac{\text{reads mapeados} \times \text{tamanho do read} \times \text{tamanho do genoma (nt)}}{\text{tamanho do CDS} \times \text{tamanho da biblioteca (nt)}}$$

O tamanho do genoma haploide de *A. paranae* foi considerado 1,87 Gb (Carvalho et al. 2002). Assim, a abundância média foi estimada em cada biblioteca genômica 0B e 1B separadamente. Posteriormente, os CDSs foram filtrados de acordo com sua abundância, sendo que os CDSs altamente representados no genoma 0B foram excluídos por serem, provavelmente, sequências repetitivas. O *fold change* (FC) por abundância de CDS foi calculado devido à presença do cromossomo B como log<sub>2</sub> do quociente entre o número de cópias por genoma haploide estimado no genoma 1B e no genoma 0B. Finalmente, os CDSs sobre-representados no genoma 1B foram anotados usando a *pipeline* Trinotate (<https://trinotate.github.io>) e a base de dados SWISS-PROT (Boeckmann et al. 2003).

A estimativa do nível de expressão dos CDSs anotados nas 12 bibliotecas de RNA-seq foi realizada através do cálculo da expressão gênica como RPKM segundo a seguinte equação:

$$\text{RPKM} = \frac{10^9 \times \text{reads mapeados no contig}}{\text{total de reads mapeados} \times \text{tamanho do contig (nt)}}$$

Para os genes sobre-representados no genoma 1B, análises adicionais foram realizadas. Primeiramente, verificou-se se as CDSs estavam completas no *contig* obtido; nos casos em que elas não estavam completas, um transcriptoma adicional com as bibliotecas 0B foi utilizado para buscar estas CDSs com o algoritmo BLASTN (Altschul et al. 1990). Sobre os *contigs* com os maiores CDSs, foi realizado um mapeamento com SSAHA2 com os *reads* das bibliotecas genômicas e transcriptômicas para estimar a cobertura média por posição ao longo do *contig*. Também foi calculado o desvio padrão

nas bibliotecas 0B e 1B e o número médio de cópias foi normalizado de acordo com o tamanho de cada biblioteca.

### 3.2.11 Confirmação dos genes presentes no cromossomo B por qPCR

Foram desenhados *primers* (Tabela S1) a partir das sequências dos genes possivelmente presentes no cromossomo B para quantificar a sua abundância nos genomas das espécies de *Astyanax* em indivíduos 0B e +B (três indivíduos de cada genótipo, checados previamente por citogenética) (Tabela 2). Excepcionalmente, os *primers* desenhados para o gene *msh4* não se mostraram eficientes para a quantificação por esta técnica.

A quantificação relativa (RQ) dos genes alvo foi obtida utilizando-se o método de  $2^{-\Delta Ct}$  (Nguyen et al. 2013) através da comparação com um gene autossômico de cópia única, *hypoxanthine phosphoribosyltransferase* (Hprt). Assim, se o gene alvo fosse autossômico, a razão do seu número de cópias comparada com o Hprt esperado seria de 1:1 ou se o gene alvo estivesse presente no cromossomo B, as razões esperadas estariam positivamente correlacionadas com o número de cromossomos B no indivíduo. A qPCR foi realizada com auxílio do equipamento StepOne Real-Time PCR Systems (Life Technologies). Os genes-alvos e o gene de referência foram analisados simultaneamente em triplicatas de três amostras independentes. As condições de programação do ciclo foram 95 °C por 10 min; 45 ciclos de 95 °C por 15 s, e 60 °C por 1 min. A especificidade dos produtos de PCR foi confirmada pela análise da curva de dissociação.

### 3.2.12 Confirmação da expressão dos genes presentes no cromossomo B por RT-qPCR

Os genes presentes no cromossomo B de *A. paranae* que mostraram padrões de expressão diferencial nas análises de RNA-seq foram selecionados para confirmação da sua expressão por RT-qPCR. Assim, foram desenhados *primers* (Tabela S2) em regiões de junção éxon-éxon das CDSs a partir das sequências de interesse destes genes. Os *primers* foram então utilizados para quantificar a sua expressão em gônadas de indivíduos 0B e 1B, sendo utilizados

ao menos três indivíduos de cada genótipo, checados previamente por análise citogenética (Tabela 2). A expressão relativa de cada gene foi obtida comparando-se a sua expressão com a expressão do gene autossômico Hprt utilizando-se o método  $2^{-\Delta\Delta Ct}$  (Nguyen et al. 2013). O equipamento e parâmetros utilizados para esta análise foram os mesmos utilizados no item 3.2.11.

### 3.2.13 Análises estatísticas

A correlação entre a quantidade relativa dos genes possivelmente presentes no cromossomo B e o número de cromossomos B foi verificada pelo teste de correlação de Spearman. A diferença entre a expressão gênica nos indivíduos 0B e 1B foi identificada pelo teste de Mann-Whitney.

**Tabela 2.** Estratégia experimental para análise dos genes presentes nos cromossomos B de espécies de *Astyanax*.

Experimento	Técnica	Spp	Sexo	Parte do corpo	Bs	N
<b>Conteúdo do cromossomo B</b>	Illumina DNA-seq	Apa	F	Fígado	0	1
					1	1
	qPCR (DNA)	Apa	F	Fígado	0	3
					1	3
					2	3
		Afa	?	Fígado	0	3
					1	3
		Abo	?	Fígado	0	4
	1				3	
	<b>Expressão no genoma com cromossomo B</b>	Illumina RNA-seq	Apa	F	Ovário	0
1						3
Músculo					0	3
					1	3
qPCR (cDNA)		Apa	F	Ovário	0	4
					1	6
			M	Testículo	0	4
					1	3

Spp = espécie, Apa = *A. paranae*, Afa = *A. fasciatus*, Abo = *A. bockmanni*, F = fêmea, M = macho, Bs = número de cromossomos B, N = número de indivíduos analisados.



## **4 Resultados e Discussão**

As informações obtidas dos estudos citogenéticos e moleculares a respeito dos cromossomos B realizados em espécies de peixes do gênero *Astyanax*, bem como as análises do genoma mitocondrial de Cordados são apresentadas a seguir na forma de capítulos. A colocação na forma de artigo científico foi feita no sentido de facilitar a delimitação das informações para posterior publicação.

As referências citadas nestes capítulos encontram-se relacionadas no final da tese.

### **Capítulo 1:**

**Análise funcional e evolutiva do cromossomo B em *Astyanax paranae* (Characiformes, Characidae).**

### **Capítulo 2:**

**Tendência evolutiva na composição de bases do DNA mitocondrial associada a fatores biológicos das classes de Cordados.**

## 4.1 Capítulo 1

### **Análise funcional e evolutiva do cromossomo B em *Astyanax paranae* (Characiformes, Characidae).**

#### **Resumo**

Cerca de 15% dos organismos eucariotos possuem elementos genômicos adicionais dispensáveis chamados cromossomos B. Contudo, em grande parte dos organismos portadores, a composição e os efeitos destes elementos no genoma ainda são desconhecidos. Estes elementos estão presentes em diversas espécies de peixes do gênero *Astyanax*. Assim, os objetivos do presente trabalho foram caracterizar a composição e estrutura do cromossomo B de *Astyanax paranae*, bem como identificar os possíveis efeitos da presença deste cromossomo nesta espécie e sua relação com cromossomos B de espécies próximas. As análises realizadas permitiram identificar sete genes codificadores de proteínas no cromossomo B de *A. paranae*, sendo que quatro mostraram expressão diferencial entre os ovários de fêmeas 0B e 1B. Além disto, quatro destes genes estão presentes nos cromossomos B de *A. fasciatus* e de *A. bockmanni*. Desta forma, pode ser estabelecido que o cromossomo B de *A. paranae* apresenta-se como um elemento genômico funcional e exerce interferência em seu próprio benefício na expressão dos genes que nele estão contidos. Além disso, os dados obtidos permitem supor a existência de uma origem comum dos cromossomos B envolvendo outras espécies no gênero *Astyanax*.

## Resultados

### *Identificação de genes codificadores de proteínas no cromossomo B de A. paranae*

Os *reads* das bibliotecas de DNA-seq e RNA-seq 0B e 1B foram mapeados sobre o transcriptoma de *A. paranae* de referência montado *de novo*, contando o número de *reads* de cada biblioteca mapeados em cada CDS e expressando os resultados normalizados como o número de cópias gênicas por genoma haploide, no caso do gDNA e RPKM para RNA.

Posteriormente, foram selecionadas 17.065 CDSs com pelo menos uma cópia e no máximo vinte cópias no genoma 0B, a fim de descartar CDSs pobremente mapeadas ou possivelmente representando sequências de DNA repetitivo. A seleção dos genes possivelmente presentes no cromossomo B foi feita com base no FC. Assim, as CDSs com  $FC > 1$  (i.e. genes com duas vezes maior número de cópias no genoma 1B) foram selecionadas, resultando no total de 55 CDSs. A anotação revelou que 18 destas CDSs eram ortólogas a 17 genes codificadores de proteínas do genoma de *A. mexicanus* (McGaugh et al. 2014). Posteriormente, as sequências obtidas destes genes foram analisadas para buscar segmentos de maior comprimento possível, incluindo algumas regiões UTR. Com isso, foram encontrados os transcritos completos para todos os genes analisados (Tabela 1).

Utilizando-se estes transcritos como referência, foram mapeados separadamente os *reads* das bibliotecas de gDNA para verificar a variação na cobertura por posição ao longo do comprimento do transcrito. Assim, foram verificados dois tipos de cobertura para os genes no genoma 1B (Tabela 2). Alguns genes mostraram cobertura uniforme (CU), enquanto outros apresentaram cobertura irregular (CI), sendo que alguns genes mostraram amplificação regional em algumas posições (CI-ar) e um gene mostrou uma parte da sequência repetida *in tandem* com padrão de DNA satélite (CI-sat) (Tabela 2) (Figuras 1, 2 e S1-5).

**Tabela 1.** Resultado do mapeamento dos *reads* de gDNA 0B e 1B nas CDSs do transcriptoma de *A. paranae*. Somente os *contigs* com mais de uma e menos de 20 cópias no genoma 0B e  $\log_1(1B/0B) \geq 1$  são mostrados.

CDS do Transcriptoma	Tamanho (bp)	<i>reads</i> mapeados no gDNA		Log2 (1B/0B)	Anotação	Ortólogos em <i>A. mexicanus</i>
		0B	1B	Gen		
c4_g1_i1 m.47835	360	13,55	1165,32	6,43	INPP	
c4_g1_i1 m.47835	360	13,55	1165,32	6,43	INPP	
c0_g6_i10 m.27466	450	1,47	72,89	5,63	G2E3	
c0_g6_i10 m.27464	906	3,15	119,70	5,25	G2E3	
c0_g2_i1 m.53260	3768	1,49	55,39	5,21	SBNO2	+
c2_g3_i1 m.38839	882	19,26	677,49	5,14	NUSAP	+
c0_g6_i10 m.27465	747	2,07	70,27	5,09	G2E3	
c2_g2_i1 m.46606	699	9,04	304,06	5,07	G2E3	
c2_g2_i1 m.46607	396	9,22	303,04	5,04	G2E3	
c0_g1_i1 m.45162	594	5,09	146,35	4,84	B2L14	
c2_g1_i2 m.32128	480	5,20	115,00	4,47	DJC25	
c2_g1_i14 m.32125	1323	2,78	57,04	4,36	DJC25	
c2_g1_i14 m.32126	594	1,51	25,89	4,10	DJC25	
c0_g6_i10 m.27463	1311	10,25	131,66	3,68	G2E3	
c0_g3_i1 m.45147	1818	10,28	90,17	3,13	TOP2A	
c1_g1_i5 m.20817	2277	0,99	8,35	3,08	NOBOX	
c1_g1_i3 m.29965	1785	1,29	7,67	2,57	ZN782	+
c2_g13_i2 m.38836	315	57,77	312,95	2,44	NUSAP	+
c0_g2_i4 m.7753	711	8,74	47,13	2,43	MPV17	+
c1_g1_i3 m.37124	903	1,13	5,63	2,32	TNR14	
c0_g1_i1 m.13342	408	2,30	11,02	2,26	RM44	
c0_g1_i1 m.13342	408	2,30	11,02	2,26	RM44	
c0_g1_i2 m.58557	1119	9,81	46,24	2,24	YTX2	
c0_g1_i3 m.55683	348	49,14	206,96	2,07	LORF2	
c0_g4_i1 m.31807	3246	1,21	5,05	2,06	RNF17	+
c3_g7_i1 m.45800	1563	1,01	4,19	2,05	SF3B2	+
c0_g1_i3 m.55682	501	69,29	239,47	1,79	LORF2	
c1_g1_i1 m.52930	1569	2,19	7,33	1,74	MDM2	+
c0_g2_i2 m.21307	1419	2,11	6,70	1,67	MOT1	+
c1_g3_i11 m.18423	339	1,36	4,29	1,66	CANX	+
c0_g1_i3 m.37006	2628	1,90	5,87	1,63	PACS1	
c0_g1_i4 m.8403	669	2,72	8,35	1,62	POL	
c0_g1_i1 m.33308	2160	1,42	4,30	1,60	SIMC1	
c8_g1_i1 m.56660	354	1,13	3,43	1,60	KPB1	
c0_g2_i1 m.44191	1200	36,44	109,74	1,59	RTJK	
c1_g1_i5 m.9080	777	1,13	3,33	1,56	GRID1	
c2_g3_i2 m.5480	1269	1,21	3,49	1,52	UBP42	

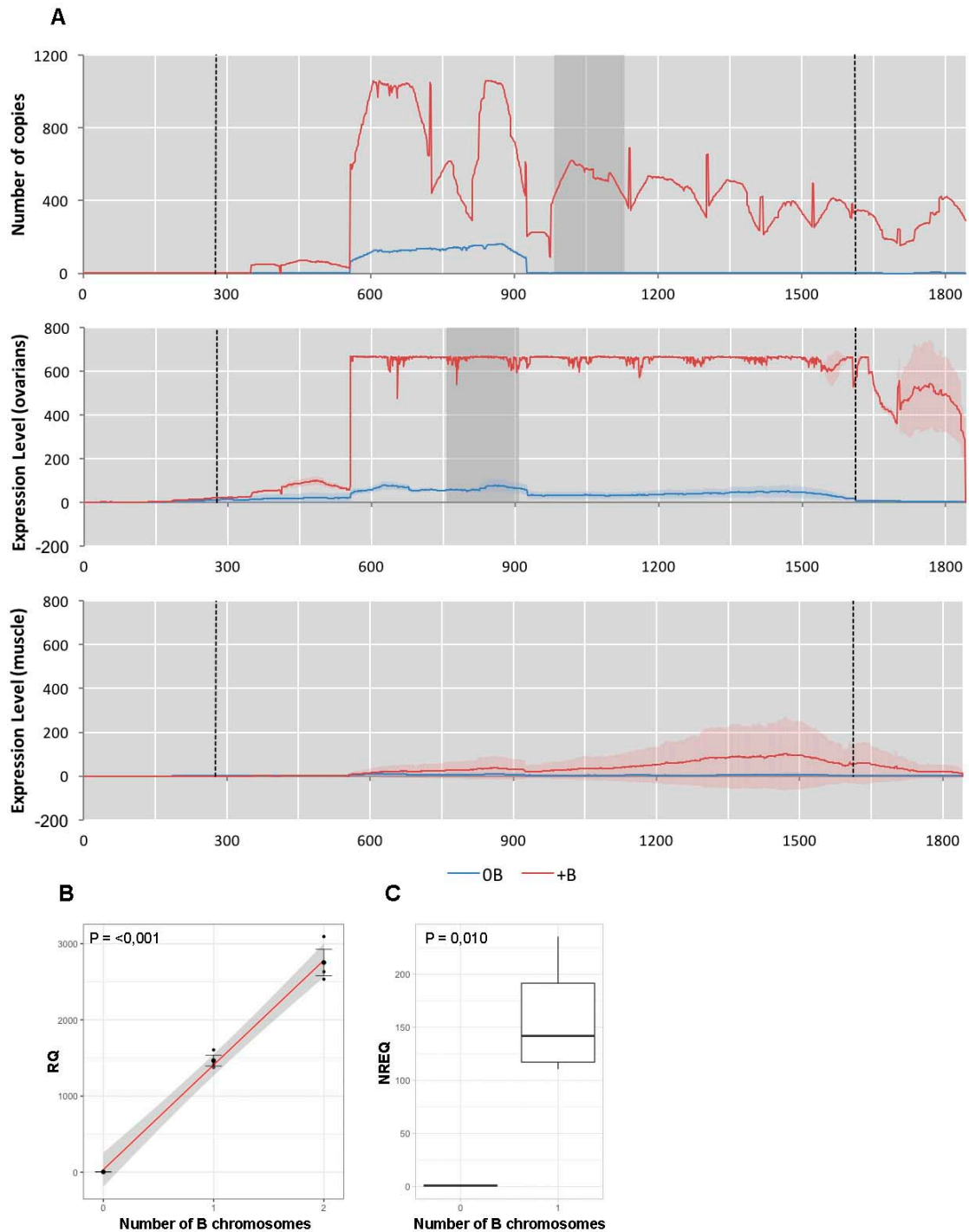
c2_g5_i3 m.64301	309	17,57	49,73	1,50	IGKC	
c0_g1_i1 m.58598	1434	1,93	5,44	1,49	BRAC	
c2_g7_i1 m.64302	324	12,15	32,90	1,44	IPLL5	+
c0_g2_i1 m.12772	2010	1,48	3,90	1,40	MSH4	+
c0_g2_i6 m.21986	432	5,29	13,52	1,35	STXB	
c0_g2_i1 m.27317	384	1,11	2,74	1,30	DOPDA	+
c0_g1_i14 m.46330	2475	1,15	2,83	1,30	CCPG1	+
c0_g1_i6 m.2997	1434	1,13	2,73	1,27	DD19A	+
c0_g1_i3 m.24839	306	1,09	2,62	1,26	SC11C	
c2_g2_i1 m.5548	2496	1,41	3,37	1,25	DHX30	
c1_g3_i1 m.5526	663	44,76	106,35	1,25	KIF2C	
c5_g1_i1 m.51178	462	9,93	23,36	1,23	RINI	
c0_g1_i4 m.18589	630	3,10	7,20	1,22	PVRL1	
c2_g1_i1 m.58366	324	2,10	4,82	1,20	MYPCL	
c3_g3_i1 m.21647	426	1,08	2,47	1,19	KAT5	+
c0_g1_i1 m.48115	624	1,44	3,23	1,16	CPNS1	+
c0_g10_i6 m.6826	672	5,12	11,04	1,11	STXA	
c1_g1_i4 m.45640	624	1,13	2,39	1,08	DCTP1	+

O sinal “+” indica que foram encontrados ortólogos para estes genes no genoma de *A. mexicanus*.

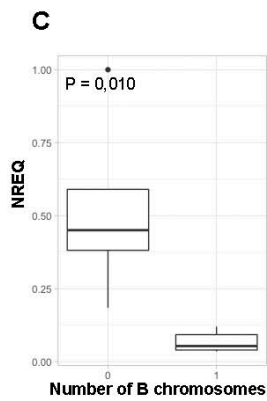
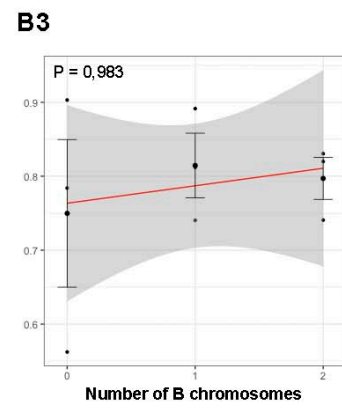
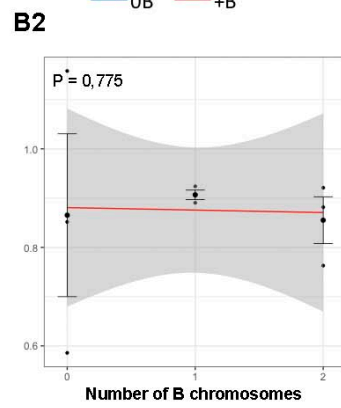
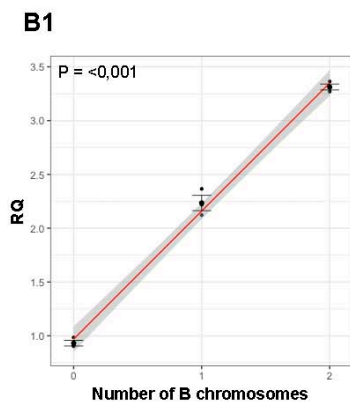
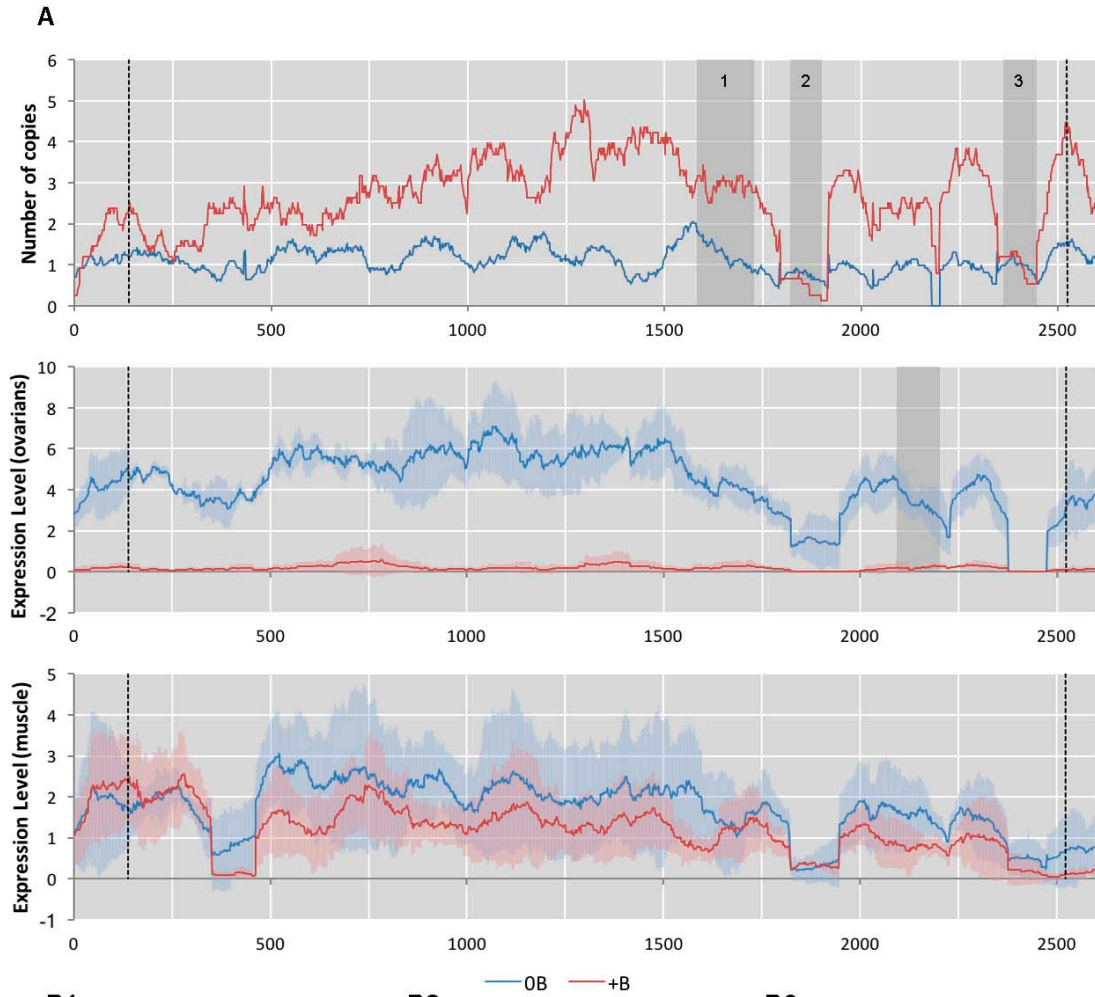
**Tabela 2.** Características dos genes presentes no cromossomo B de *A. paranae* confirmados por qPCR em gDNA desta espécie.

Acrônimo	Anotação	Tamanho CDS (bp)	Distribuição cópias no B	Log2(+B/0B)			RT-qPCR ovários
				Gen	RNAgon	RNAmus	
<i>sbno2</i>	Protein strawberry notch homolog 2	4992	CI-sat	5,21	-3,03	-0,83	sub
<i>nusap</i>	Nucleolar and spindle-associated protein 1	1313	CI-ar	5,14	6,73	3,80	super
<i>rnfl17</i>	RING finger protein 17	4750	CI- ar	2,06	1,11	1,58	sem dif.
<i>mdm2</i>	E3 ubiquitin-protein ligase Mdm2	1458	CU	1,74	0,46	-1,42	sem dif.
<i>mot1</i>	Monocarboxylate transporter 1	1414	CI- ar	1,67	3,97	-0,59	super
<i>msh4</i>	MutS protein homolog 4	2278	CU	1,40	-0,16	-2,29	NP
<i>ccpg1</i>	Cell cycle progression protein 1	2271	CI- ar	1,30	-4,56	-0,78	sub

CI-sat = Cobertura irregular semelhante a satélite. CI-ar = Cobertura irregular com amplificação regional. CU = Cobertura uniforme. Gen = Log2(1B/0B) do DNA genômico. RNAgon = Log2(1B/0B) do RNA dos ovários. RNAmus = Log2(1B/0B) do RNA do músculo. “super” e “sub” na coluna RT-qPCR ovários indicam que estes genes estão super ou sub-expressos nos ovários de indivíduos 1B comparando-se com indivíduos 0B. “sem dif.” Indica que não houve diferença na expressão entre indivíduos 0B e 1B.



**Figura 1:** Cobertura para o transcrito do gene *nusap* no gDNA (0B e 1B) e RNA (0B e 1B) de ovários e músculo a partir das bibliotecas Illumina de indivíduos de *A. paranae* (A), e qPCR no gDNA (B) e cDNA (C). Note que a cobertura é maior na biblioteca 1B ao longo de toda a CDS (delimitado pelas linhas pontilhadas verticais). Da mesma forma, note a maior cobertura para este transcrito na biblioteca de RNA de ovário 1B. As regiões sombreadas em A marcam as regiões amplificadas por qPCR. A qPCR no gDNA (B) mostrou que o número de cópias do gene *nusap* aumenta com o número de cromossomos B, seguindo um padrão dose-dependente, indicando sua presença no cromossomo B. A qPCR no cDNA (C) confirmou que o gene *nusap* está super-expresso no ovário das fêmeas 1B. RQ = Quantidade relativa. NREQ = Quantidade relativa de expressão normalizada. P = P-value para teste de correlação de Spearman (B) e teste de Mann-Whitney (C).



**Figura 2:** Cobertura para o transcrito do gene *ccpg1* no gDNA (0B e 1B) e RNA (0B e 1B) de ovários e músculo a partir das bibliotecas Illumina de indivíduos de *A. paranae* (A), e qPCR no gDNA (B) e cDNA (C). Note que a cobertura é maior na biblioteca 1B na maior parte da CDS (delimitado pelas linhas pontilhadas verticais). Entretanto, as regiões sombreadas 2 e 3 possuem coberturas similares entre as bibliotecas 0B e 1B, indicando que este gene está truncado no cromossomo B. As três regiões sombreadas foram selecionadas para amplificação por qPCR, a região 1 que parece estar presente no cromossomo B e as regiões 2 e 3 que parecem não estar neste cromossomo. A qPCR na região 1 mostrou que o número de cópias para este gene nesta região aumenta com o número de cromossomos B em um padrão dose-dependente (B1). Já para as regiões 2 e 3 o número de cópias independe do número de cromossomos B (B2 e B3). Da mesma forma, note a menor cobertura para este transcrito na biblioteca de RNA de ovário 1B. A qPCR no cDNA (C) confirmou que o gene *ccpg1* está sub-expresso no ovário das fêmeas 1B. RQ = Quantidade relativa. NREQ = Quantidade relativa de expressão normalizada. P = P-value para teste de correlação de Spearman (B) e teste de Mann-Whitney (C).

A técnica de qPCR mostrou que a abundância genômica aumenta linearmente com o número de cromossomos B para sete (*ccpg1*, *mdm2*, *mot1*, *msh4*, *nusap*, *rnf17*, *sbno2*) dos 17 genes candidatos a residirem no cromossomo B de *A. paranae* (Tabela 2, Figuras 1, 2 e S1-5). Assim, foi considerado que estes sete genes estão presentes neste cromossomo. Notadamente, a quantificação por qPCR dos genes truncados (*ccpg1*, *rnf17*, *sbno2*) mostrou esta relação apenas nas regiões que apresentam alta cobertura na biblioteca 1B, indicando a sua localização no cromossomo B. Por outro lado, nenhuma relação foi observada nas regiões de baixa cobertura, indicando sua ausência no cromossomo B e a existência de cópias fragmentadas destes genes neste cromossomo (Figuras 2 e S4, 5). No entanto, o gene *rnf17* mostrou-se uma exceção, apresentando uma das regiões de baixa cobertura com maior quantidade de cópias nos indivíduos +B (Figura S4).

A aplicação da técnica de CARD-FISH permitiu a detecção do gene *nusap* na região intermediária dos braços do cromossomo B de *A. paranae* (Figura 3).

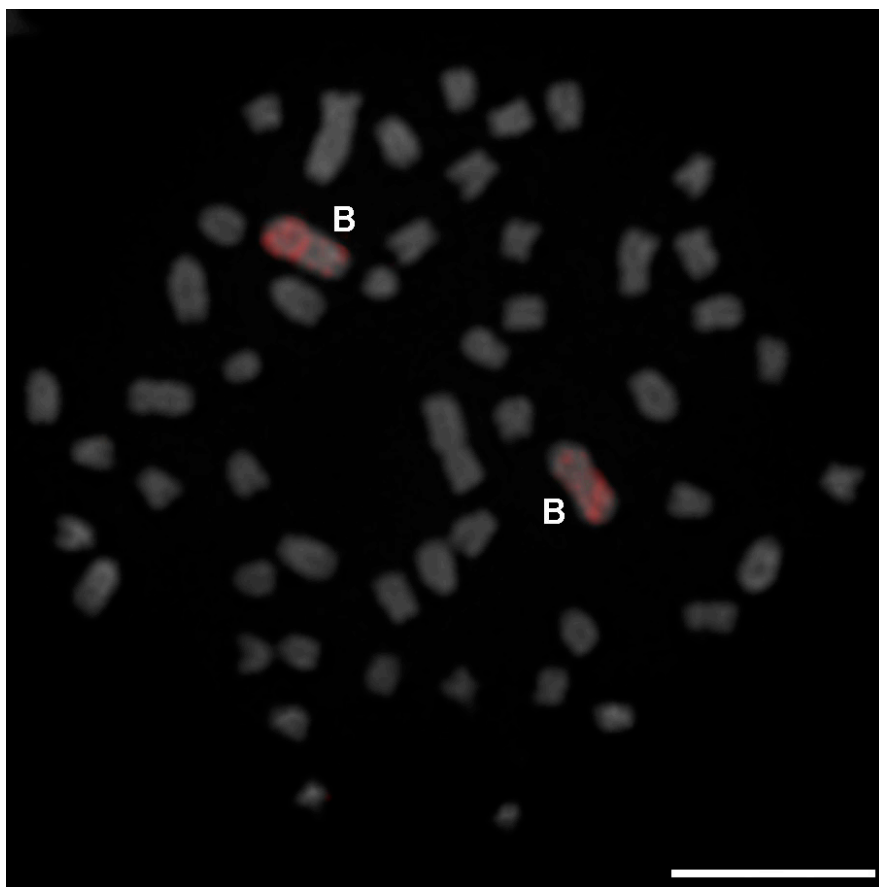
### ***Expressão dos genes presentes no cromossomo B de A. paranae***

Para determinar a expressão dos sete genes identificados e presentes no cromossomo B de *A. paranae* verificou-se o FC 1B/0B RNA das bibliotecas



estabelecidas para tecidos de ovários e músculo. Nos ovários foram identificados três genes com  $FC > 1$  (*mot1*, *nusap* e *rnf17*) e dois genes com  $FC < -1$  (*ccpg1* e *sbno2*), enquanto que nos músculos dois genes mostraram  $FC > 1$  (*nusap* e *rnf17*) e dois genes apresentaram  $FC < -1$  (*mdm2* e *msh4*) (Tabela 2). De fato, os genes *nusap* e *mot1* revelaram super-expressão nos ovários dos indivíduos 1B (Figuras 1 e S2). No entanto, o gene *rnf17* não mostrou expressão diferencial entre os ovários de indivíduos 0B e 1B (Figura S4). Já os genes *ccpg1* e *sbno2* estão, de fato, sub-expressos nos ovários 1B (Figuras 2 e S5).

A análise de expressão ao longo das CDSs nas amostras de músculo mostrou expressão diferencial apenas para o gene *rnf17* (Figura S4). Assim, optou-se por não realizar a técnica de RT-qPCR neste tecido.

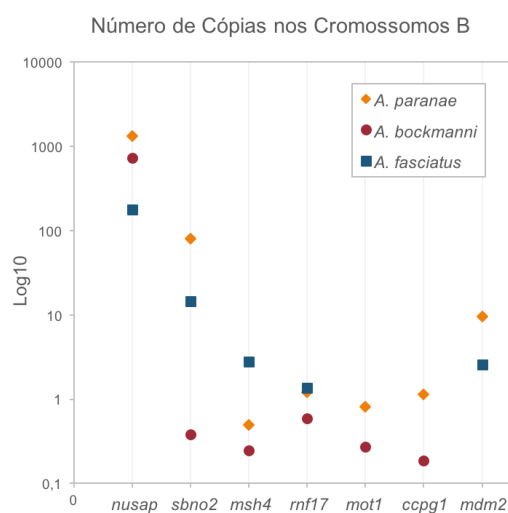


**Figura 3:** Metáfase de *A. paranae* após hibridação *in situ* fluorescente (CARD-FISH) com sonda do gene *nusap* (vermelho). Barra = 10  $\mu\text{m}$ .

### **Identificação dos genes codificadores de proteínas presentes no cromossomo B de *A. paranae* em cromossomos B de outras espécies de *Astyanax***

A análise por qPCR permitiu inferir que os genes *nusap*, *sbno2*, *msh4*, *rnf17*, *mot1* e *ccpg1* estão presentes no cromossomo B de *A. bockmanni*, enquanto os genes *nusap*, *sbno2*, *msh4*, *rnf17* e *mdm2* estão presentes no cromossomo B de *A. fasciatus*. Esta análise também mostrou que as regiões truncadas dos genes presentes no cromossomo B de *A. paranae* também estão ausentes nos cromossomos B de *A. bockmanni* e *A. fasciatus* (Figuras S6, 7), com exceção do gene *ccpg1* que possui uma região presente no cromossomo B de *A. bockmanni* que, contudo, não foi identificada no cromossomo B de *A. paranae* (Figuras S6).

A comparação da quantidade de cópias dos genes presentes nos cromossomos B das três espécies analisadas revelou que os genes *nusap*, *sbno2*, *mdm2*, *mot1* e *ccpg1* possuem maior número de cópias no cromossomo B de *A. paranae*, enquanto que os genes *msh4* e *rnf17* possuem maior número de cópias no cromossomo B de *A. fasciatus*. Todos os genes analisados mostraram menor número de cópias no cromossomo B de *A. bockmanni*, com exceção do gene *nusap* (Figura 4).



**Figura 4:** Comparação entre o número de cópias dos genes alocados nos cromossomos B de três espécies de *Astyanax*. Notar que os genes *mot1* e *ccpg1* não estão presentes no cromossomo B de *A. fasciatus* e o gene *mdm2* não está presente no cromossomo B de *A. bockmanni*.

## Discussão

Estudos recentes demonstraram a ocorrência de genes ativos codificadores de proteínas nos cromossomos B de diferentes espécies (revisão em Valente et al. 2016; Clark et al. 2016; Makunin et al. 2016; Navarro-Domínguez et al. 2017). A expressão diferencial de genes presentes no cromossomo B poderia indicar que este cromossomo exerceria influência no aumento dos níveis de expressão dos genes existentes nos cromossomos do complemento normal, estabelecendo, então, uma relação com o genoma A da espécie (Huang et al. 2017). No presente trabalho, a super-expressão dos genes completos identificados no cromossomo B nos indivíduos 1B (*mot1* e *nusap*) pode indicar que as cópias presentes neste cromossomo são recrutadas para uma expressão adicional destes genes no indivíduo. No entanto, não foi possível caracterizar a expressão de sequências exclusivas do cromossomo B. Além disso, fragmentos de genes presentes em um cromossomo B podem interferir na expressão das cópias canônicas destes genes presentes nos cromossomos do complemento A (Banaei-Moghaddam et al. 2013). Neste sentido, nós identificamos que dois genes truncados presentes no cromossomo B de *A. paranae* apresentaram expressão reduzida nos ovários 1B (*ccpg1* e *sbno2*).

Entre os genes identificados no cromossomo B de *A. paranae*, o gene *nusap* se destaca por estar super-expresso nos ovários dos indivíduos 1B (Figura 1). Este gene codifica uma proteína associada aos microtúbulos durante a divisão mitótica, podendo causar divisões deficientes em caso de super ou sub-expressão (Raemaekers et al. 2003). Assim, considerando-se que os cromossomos B são elementos genômicos adicionais, sugere-se que a presença destes elementos implicaria em um aumento da maquinaria relacionada ao processo de divisão celular, possibilitando a inclusão de mais microtúbulos para atuarem no posicionamento dos cromossomos. Desta forma, a

super-expressão do gene *nusap* nas gônadas de indivíduos portadores de cromossomos B poderia indicar uma adaptação destes genomas à presença deste cromossomo, seja com relação aos maiores níveis de expressão das cópias presentes no genoma A ou até mesmo das cópias presentes no próprio cromossomo B.

O gene *ccpg1* também chama a atenção por estar presente no cromossomo B, ainda que de forma incompleta, e por apresentar-se sub-expresso nos ovários 1B (Figura 2). Foi verificado que este gene está envolvido com o controle do ciclo e proliferação celular (Kostenko et al. 2006). Assim, a sub-expressão deste gene poderia estar ligada a um retardo no processo de divisão celular, o qual poderia ser causado pelo aumento do conteúdo de DNA nuclear resultante da presença dos cromossomos B adicionais (revisão em Bell & Burt 1990).

Outro gene interessante presente no cromossomo B é o *mot1*, que está super-expresso nos ovários dos indivíduos 1B (Figura S2). Este gene está envolvido com a regulação da transcrição (Muldrow et al. 1999). Desta forma, este gene poderia estar ativo diferencialmente nos indivíduos com o cromossomo B devido à necessidade de expressão diferencial de genes em favor da manutenção destes cromossomos.

O gene *sbno2* também mostrou uma sub-expressão nos ovários dos indivíduos 1B (Figura S5). Este gene pertence à família de co-repressores da *helicase Strawberry notch* e está envolvido em processos como homeostase óssea (Maruyama et al. 2013) e da cascata anti-inflamatória, podendo ter ambas as funções de co-ativador e co-repressor (Kasmi et al. 2007). Assim, acreditamos que a sub-expressão deste gene nos ovários dos indivíduos 1B poderia permitir a maior expressão dos demais genes necessários para a manutenção do cromossomo B.

Outros dois genes presentes no cromossomo B (*mdm2* e *rnfl7*) não mostraram expressão diferencial nos tecidos analisados. No entanto, isto não significa que estes

genes não possam ser atuantes e favorecerem o cromossomo B em outras etapas do desenvolvimento, e. g. durante o desenvolvimento embrionário, quando ocorre uma alta proliferação celular e o cromossomo B se mostraria uma carga adicional para as células do organismo.

De forma geral, os cromossomos B de diversos organismos possuem características semelhantes que permitem classificá-los dentro das suas particularidades de elementos extras do genoma (Camacho et al. 2000). No entanto, estes elementos possuem composições gênicas muito distintas nos diferentes grupos devido ao seu modo de origem independente em grupos distantes ou até mesmo entre espécies de um mesmo gênero (Pansonato-Alves et al. 2014). Neste sentido e até de modo surpreendente, um dos genes identificado no cromossomo B de *A. paranae*, *mdm2*, também está presente no cromossomo B do gafanhoto *Locusta migratoria* (Ruiz-Ruano, comunicação pessoal). Além disto, genes envolvidos com funções semelhantes como divisão celular, foram identificados em cromossomos B de diferentes espécies (Valente et al. 2014; Navarro-Domínguez et al. 2017). Este fato poderia indicar que, provavelmente, os cromossomos B podem portar e alterar a expressão de genes com funções específicas para sua própria manutenção, independentemente do grupo biológico ao qual estejam inseridos. Tal fato parece indicar que, apesar de estarem em indivíduos de grupos muito distantes, estes elementos ativam maquinaria celular semelhante para seu próprio benefício.

A origem dos cromossomos B dentro do gênero *Astyanax* vem sendo discutida intensamente nos últimos anos. Dados de sequenciamento genômico e de mapeamento de sequências de DNA repetitivo indicam que alguns destes cromossomos podem ter tido origem comum, enquanto que outros tipos parecem ter surgido independentemente em diferentes espécies (Silva et al. 2016). Contudo, o uso de DNAs repetitivos como

marcadores não se apresenta como uma ferramenta segura neste tipo de estudo, uma vez que este tipo de sequência de DNA evolui rapidamente, com dinâmica própria, podendo estar constantemente sendo inserido ou deletado dos cromossomos B (Clark et al. 2016; Silva et al. 2017).

Os resultados obtidos na presente análise demonstram que os cromossomos B de três espécies de *Astyanax* pertencentes ao clado 1 (Rossini et al. 2016) possuem pelo menos quatro genes codificadores de proteínas compartilhados. Tal fato leva à conclusão de que estes cromossomos possuem uma origem comum e que derivam de um cromossomo B ancestral, como proposto anteriormente por Moreira-Filho et al. (2004). Neste sentido, considera-se que seria pouco parcimonioso propor que os cromossomos B das três espécies teriam surgido em eventos independentes a partir de fragmentos com a mesma composição gênica ou que teriam sido colonizados pelos mesmos genes. Entretanto, a partir dos dados aqui obtidos não é possível elucidar se estes cromossomos B foram herdados de uma mesma espécie ancestral nas três espécies ou se foram adquiridos por eventos de hibridação com espécies próximas que portavam este cromossomo desde sua origem.

O número de cópias das sequências identificadas nos cromossomos B das três espécies mostrou-se diferente (Figura 4), indicando que estes genes sofrem eventos de amplificação/erosão particulares em cada cromossomo B. Mostra ainda que cada cromossomo B segue seu próprio caminho evolutivo em cada espécie. Tais processos fazem parte do processo de evolução dos cromossomos B (Banaei-Moghaddam et al. 2013).

Diversos mecanismos podem provocar o aumento no número de cópias de um dado gene em um cromossomo (ver Clark et al. 2016). Dentre eles, a recombinação meiótica pode ser um mecanismo importante, de forma que as sequências localizadas na

região intermediária dos braços cromossomos experimentariam uma maior taxa de recombinação e, então, uma maior variação no número de cópias que as sequências próximas ao centrômero (Clark et al. 2016). Corroborando com esta afirmação, no presente trabalho verificou-se que o gene *nusap* mostrou um alto número de cópias no cromossomo B de *A. paranae* e foi localizado por FISH na região intermediária dos braços deste cromossomo (Figura 3). Tal fato parece confirmar que sua localização pode contribuir de fato para o aumento do número de cópias observado entre as espécies.

## 4.2 Capítulo 2

### **Tendência evolutiva na composição de bases do DNA mitocondrial associada a fatores biológicos das classes de Cordados.**

#### **Resumo**

As mitocôndrias são estruturas celulares que possuem o seu próprio DNA, a despeito do DNA nuclear da célula da célula que as contêm. Em muitos organismos a composição de bases do DNA destas organelas viola a segunda lei de paridade de Chargaff (PR2), que presume que uma fita simples de DNA deva ter proporções similares dos nucleotídeos complementares, ou seja, as quantidades  $A = T$  e  $G = C$ . A deaminação das bases A e C durante a fase de fita simples no processo de replicação molecular é tida como o principal mecanismo capaz de alterar a *skew* das bases AT e GC no DNA mitocondrial, que é calculado da seguinte forma:  $AT\ skew = (A-T)/(A+T)$  e  $GC\ skew = (G-C)/(G+C)$ . Assim, em decorrência do processo de diferenciação estabelecido pelo fato das mitocôndrias violarem a lei de paridade PR2, padrões distintos de composição são observados entre os grupos de eucariotos. Embora estudos restritos tenham sido realizados dentro de grupos específicos, uma análise detalhada e comparativa da distribuição da composição das bases AT e GC no DNA mitocondrial entre as classes de animais ainda não foi realizada. Na presente investigação é realizada a análise comparativa da composição de bases do DNA mitocondrial descrita para componentes de 11 classes de animais através dos *skews* de AT e GC. A análise dos dados revelou tendência para a existência de um padrão evolutivo comum na composição de bases do DNA mitocondrial destas classes de organismos, sendo que as modificações ocorridas devem proporcionar ou estariam ligadas a vantagens adaptativas. O padrão evolutivo



identificado revela diferenciação específica em grupos de animais com características particulares quanto aos *skews* GC e AT. Assim, é proposto que as alterações na composição de bases do DNA mitocondrial ao longo do processo evolutivo poderiam ser resultado não só de mecanismos bioquímicos, mas também determinado por características biológicas particulares dos organismos, como modo de respiração, ocupação do meio terrestre ou aquático, de água doce ou salgada, entre outros.

### **Letter**

As mitocôndrias são as estruturas celulares responsáveis pela respiração celular nos organismos eucariotos. Elas possuem seu próprio DNA, que difere do DNA nuclear da célula e dos cloroplastos em vegetais. É aceito que estas estruturas evoluíram a partir de bactérias e foram incorporadas para a formação das células dos eucariotos envolvendo um ou dois eventos endosimbiontes (Margulis 1981; Martin 2005). O DNA mitocondrial é circular e formado, em geral, por 13 genes codificadores de proteínas, 22 tRNAs e 2 rRNAs alocados em uma cadeia pesada (H) e uma leve (L). A organização e composição destes genes difere entre os organismos, podendo ser observadas inversões e duplicações (Saccone et al. 2000).

A partir do Sequenciamento de Nova Geração de DNA total de seres eucariotos é possível obter também sequências do DNA mitocondrial, além das sequências do DNA nuclear. Desta forma, sequências completas do DNA mitocondrial têm sido obtidas para diversos organismos como resultado secundário de projetos de sequenciamento total do genoma de uma dada espécie. A montagem da sequência completa do DNA mitocondrial é facilitada pelo uso de softwares específicos, como no caso do programa MitoBim (Hahn et al. 2013).

A segunda lei de paridade de Chargaff (PR2) diz respeito à proporção de bases nucleotídicas em uma cadeia de DNA simples. Essa lei presume que:  $C=G$  e  $A=T$  (Rudner et al. 1968). Espera-se que esta lei seja mantida quando as taxas de mutação são similares para substituições de nucleotídeos complementares entre as duas cadeias (Sueoka 1995). Com relação ao DNA nuclear dos organismos eucariotos, eubactéria, archaeabacteria e na maioria dos casos de vírus, esta lei parece ser obedecida. No entanto, no DNA organelar ela é burlada (Mitchell & Bridge 2006). Em organismos mais simples como as archeobacterias, o DNA organelar mantém proporções iguais entre as bases, como previsto pela 2ª lei de paridade de Chargaff (ver Frank & Lobry 1999). Porém, em outros organismos, a proporção de bases tende a se diferenciar através de mecanismos ainda discutidos (revisão em Saccone et al. 2000).

Os estudos acerca dos desvios na composição de bases do DNA mitocondrial têm se limitado principalmente à observação dos animais superiores como um todo e relacionam somente fatores moleculares às alterações de conteúdo do DNA mitocondrial. No presente trabalho pretendeu-se investigar as proporções de bases do DNA mitocondrial dos animais superiores através dos AT e GC *skews*, observando as modificações ocorridas neste segmento genômico em relação às peculiaridades de cada grupo separadamente, uma vez que animais com características muito distintas são encontrados entre os grupos mais complexos.

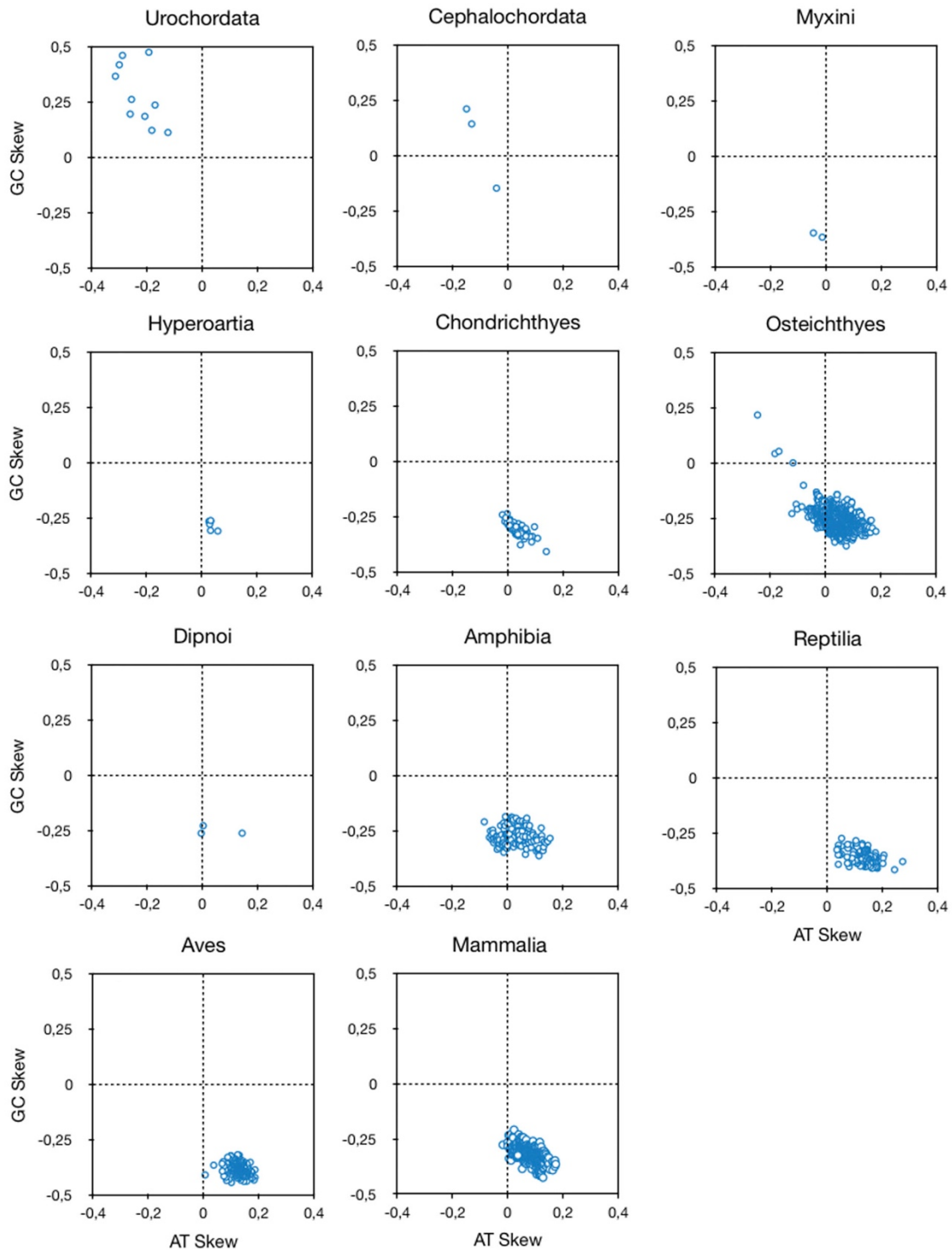
Para isto, foram utilizadas sequências de DNA mitocondrial completas prospectadas na base de dados GenBank do NCBI. Para realizar as comparações do DNA mitocondrial entre as classes de cordados, foi utilizada uma sequência representativa para cada gênero disponível. Algumas classes foram agrupadas por subfiló devido ao baixo número de representantes disponíveis. Foi então calculado o conteúdo nucleotídico (A, T, G, C, N) nas sequências através de um script em Python

([https://github.com/fjruiaruano/ngs-protocols/blob/master/count\\_acgtn.py](https://github.com/fjruiaruano/ngs-protocols/blob/master/count_acgtn.py)) e excluídas as sequências incompletas e que possuíam bases indeterminadas (N), a fim de se poder calcular as proporções exatas de nucleotídeos. Finalmente, foram calculados os *skews* de AT e GC para cada gênero dos grupos de animais através das fórmulas propostas por (Perna & Kocher 1995):  $AT\ skew = (A-T)/(A+T)$  e  $GC\ skew = (G-C)/(G+C)$ , onde A, T, G e C representam a ocorrência das quatro bases. Com base nos *skews* de AT e GC foram então calculadas as diferenças entre as classes com maior número de representantes (Amphibia, Aves, Chondrichthyes, Mammalia, Osteichthyes, Reptilia) através da Análise de Variância de dupla classificação para medidas repetidas (Two-way repeated measures ANOVA), com pós-teste de Bonferroni.

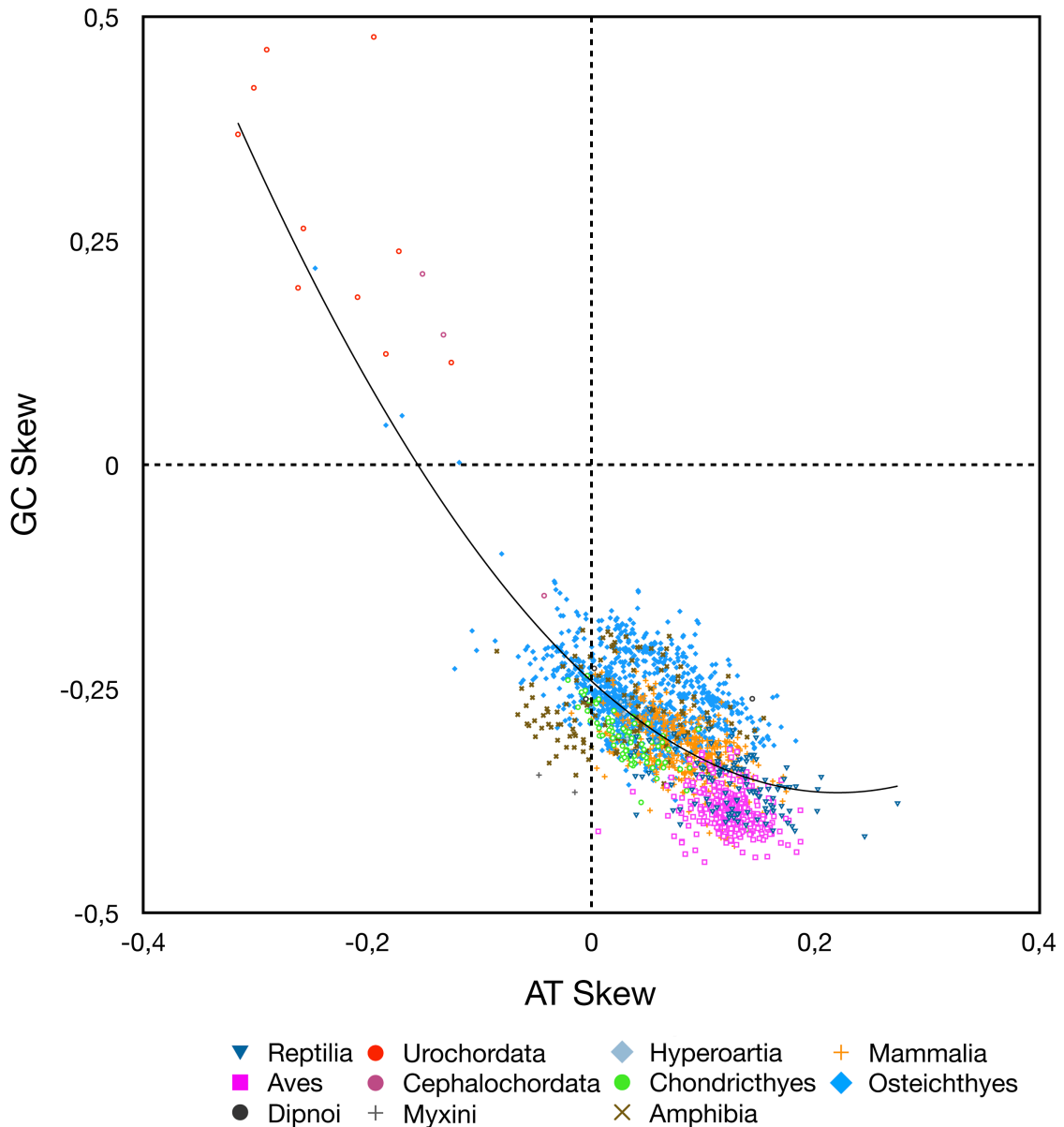
De modo geral, os cordados apresentaram *AT skew* positivo e *GC skew* negativo (Figura 1). No entanto, algumas exceções foram observadas em Chondrichthyes, Osteichthyes, Dipnoi, Amphibia e Mammalia, nos quais alguns gêneros apresentaram *AT* e *GC skew* negativo (Figura 1). Este último padrão também foi observado nos grupos ancestrais Cephalocordata e Myxini (Figura 1). Além disso, raras exceções foram observadas em Osteichthyes com *AT skew* negativo e *GC skews* positivo (Figura 1), sendo estas exceções causadas provavelmente por alterações do DNA mitocondrial já descritas (Fonseca et al. 2014). Padrões semelhantes a estas exceções também foram observado nos grupos ancestrais Urochordata e Cephalocordata (Figura 1).

Em conjunto, os dados obtidos para todos os grupos parecem seguir uma tendência que leva ao padrão observado nos grupos mais recentes (*AT skew* positivo e *GC skew* negativo) (Figura 2). Este padrão corresponde a um acúmulo de adenina e citosina no DNA mitocondrial frente às bases complementares timina e guanina, respectivamente. O acúmulo de A e C no DNA mitocondrial parece estar ligado ao processo de deaminação destas bases durante a fase de fita simples no processo de

replicação (Reyes et al. 1998). Além disso, de acordo com a hipótese de transcrição do uso de códons (THCU), o acúmulo de A no DNA mitocondrial favoreceria o processo de transcrição nestas organelas (Xia 1996).



**Figura 1:** Representação gráfica dos *skews* de AT e GC para grupos de Cordados.



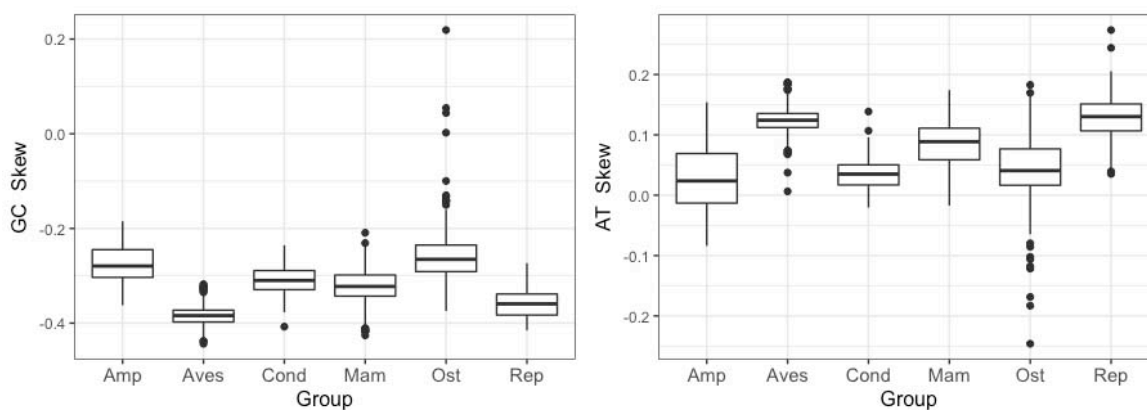
**Figura 2:** Tendência para os *skews* de AT e GC em Cordados.  $R^2=0,584$ .

As hipóteses supracitadas ajudam a explicar o padrão geral de evolução do DNA mitocondrial em vertebrados. Entretanto, a comparação dos *skews* de AT e GC entre as classes de vertebrados revelou que alguns destes grupos possuem características semelhantes, enquanto outras classes possuem características únicas (Tabela S3), indicando que a evolução do DNA mitocondrial pode estar relacionada a características biológicas dos grupos. Nesta análise, Reptilia, Amphibia e Mammalia formam um grupo com características similares, enquanto Chondrichthyes, Osteichthyes e Aves

possuem características próprias. Assim, pode ser considerado que as modificações ocorridas na evolução do DNA mitocondrial destes grupos estariam associadas ao processo de saída dos animais do meio aquático para o meio terrestre, o qual está também relacionado à evolução do sistema respiratório.

Neste sentido, o grupo formado por Reptilia, Amphibia e Mammalia teria como característica comum a ocupação do meio terrestre e o modo de respiração aéreo. Adicionalmente, podemos considerar que os Chondrichthyes possuem uma especialização do DNA mitocondrial relacionada à ocupação majoritária da água salgada, enquanto que Osteichthyes se diferenciam por ocupar o meio aquático de forma geral, tanto de água salgada quanto de água doce, enquanto o grupo das aves se diferenciaria pela ocupação do meio terrestre e respiração e metabolismo particulares devido ao voo.

Os dados revelaram que o grupo das Aves possui a menor variância nos *skews* de AT e GC (Figura 2), provavelmente devido à sua menor diversidade e especialização quanto ao modo de vida. Neste sentido, em Chondrichthyes também foram observados *skews* com pouca variância, o que poderia estar relacionado à menor diversidade de ocupação de ambientes. Por outro lado, Amphibia e Osteichthyes possuem uma alta variância nos *skews* observados (Figuras 2), em consonância com a maior diversidade no modo de vida encontrado em Amphibia e com a ocupação de diferentes meios aquáticos em Osteichthyes.



**Figura 3:** Comparação entre os *skews* de AT e GC entre os grupos de vertebrados.

Finalmente, consideramos que os mecanismos bioquímicos de evolução do DNA mitocondrial podem estar associados ao modo de vida dos organismos. Assim, estes mecanismos poderiam conduzir o DNA mitocondrial a uma especialização dentro de cada grupo de cordados, a qual resultaria em sistemas metabólicos mais eficientes, de acordo com características peculiares de cada grupo. Estes sistemas poderiam então levar ao estabelecimento de modificações relacionadas à economia de energia durante os processos de replicação do material genômico das mitocôndrias e dos processos respiratórios celulares desempenhados por estas organelas. Assim, sequências de DNA contendo maior quantidade de AT com ligações duplas de hidrogênio entre as bases formariam estruturas metabolicamente mais ágeis e econômicas nos processos energéticos das reações que as combinações GC, estabelecidas através de pontes triplas de ligações de hidrogênio.

A evolução do DNA mitocondrial já havia sido estudada anteriormente em cordados (Nikolaou & Almirantis 2006), abordando a existência de desvios nas proporções de bases AT e GC existente na cadeia do DNA simples desta organela celular. Contudo, no presente estudo os dados são analisados pela primeira vez em associação com características biológicas particulares dos grupos de organismos cordados. Considera-se que seria de interesse investigar como as características biológicas influenciam na evolução do DNA mitocondrial em cada grupo em particular, tendo em vista as diferenças e a diversidade no modo de vida encontrado dentro de cada grupo. Neste sentido, seria interessante elucidar como as características biológicas se relacionam com os processos bioquímicos, bem como são promovidas e se estabelecem as alterações da composição do material genômico das mitocôndrias.

## 5 Considerações Gerais

O desenvolvimento deste trabalho forneceu um panorama da composição genética dos cromossomos B no gênero *Astyanax*, bem como dos efeitos da presença deste cromossomo em *A. paranae* e da sua relação com outros cromossomos B de duas espécies de *Astyanax*.

Primeiramente, a descrição do satelitoma de *A. paranae* (Anexo 1) permitiu uma ampliação a respeito da biologia dos DNAs satélites de forma geral. O estudo realizado faz parte de uma série de estudos pioneiros envolvendo análise profunda de grande quantidade de dados para caracterização do conteúdo de DNAs satélites. Durante muitos anos, foram publicados diversos estudos citogenéticos envolvendo espécies do gênero *Astyanax* com base somente no mapeamento físico do DNA satélite As-51. Assim, a descrição de mais de 30 sequências de DNA satélite componentes do satelitoma de *A. paranae* (Anexo 1) fornece uma série de marcadores úteis para estudos futuros neste gênero. Além disso, a identificação de diversas sequências destes DNAs satélites nos cromossomos B de três espécies de *Astyanax* permitiu um melhor conhecimento a respeito da origem, evolução e distribuição destes elementos no genoma destas espécies.

Em seguida, um estudo mais detalhado de sequenciamento do genoma de *A. paranae* permitiu a identificação de genes codificadores de proteínas no cromossomo B desta espécie e abriu uma linha importante de investigação a respeito destes elementos (Capítulo 1). Inicialmente, foi possível observar os efeitos da presença dos cromossomos B nos ovários de fêmeas 1B, indicando que este cromossomo influencia na expressão dos genes que carrega, provavelmente em seu próprio benefício, corroborando hipóteses formuladas em estudos anteriores. Verificou-se ainda, que



quatro dos genes presentes no cromossomo B de *A. paranae* são compartilhados com os cromossomos B de *A. bockmanni* e de *A. fasciatus*, sendo esta constatação importante para corroborar a hipótese de que estes cromossomos possuem uma origem comum.

Além disso, após a obtenção do genoma mitocondrial completo de *A. paranae* (Anexo 2), foram observadas algumas incongruências com dados já publicados, de forma que decidiu-se revisar e analisar a estrutura e comportamento de todos os genomas mitocondriais disponíveis no banco de dados para peixes e, em seguida, para outras classes de cordados em relação ao conteúdo de bases que os compõe (Capítulo 2). Esta análise permitiu a identificação de um padrão evolutivo comum de modificação do DNA mitocondrial entre estes animais e que, além disso, permitiu observar que alguns grupos se apresentam com características e padrões particulares de acúmulo das diferenciações. Assim, estes padrões foram relacionados de modo ainda não estudado a características biológicas destas classes de organismos. Neste sentido, considera-se que as hipóteses levantadas abrem perspectivas para trabalhos a respeito da evolução do DNA mitocondrial de modo geral, envolvendo não apenas as modificações ocorridas nas sequências destes segmentos genômicos, mas também os possíveis mecanismos determinantes das diferenciações e suas relações com os aspectos biológicos particulares nestes grupos de animais.

Finalmente, consideramos que as informações aqui apresentadas fornecem uma base para novos estudos a respeito da origem, estrutura e função dos cromossomos B em peixes, bem como hipóteses interessantes a respeito da evolução do conteúdo de bases do DNA mitocondrial em cordados.

## 6 Referências

- Adnadević, T. et al., 2014. Possible Influence of B Chromosomes on Genes Included in Immune Response and Parasite Burden in *Apodemus flavicollis*. *PLoS ONE*, 9(11), pp.e112260–8.
- Altschul, S.F. et al., 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215, pp.403–410.
- Banaei-Moghaddam, A.M. et al., 2013. Formation and Expression of Pseudogenes on the B Chromosome of Rye. *The Plant Cell*, 25(7), pp.2536–2544.
- Bell, G. & Burt, A., 1990. B-chromosomes: Germ-line parasites which induce changes in host recombination. *Parasitology*, 100, Suppl:S19-26.
- Boeckmann, B., 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Research*, 31, pp.365–370.
- Camacho, J.P.M., 2005. B Chromosomes. In T. R. Gregory, ed. *The evolution of the genome*. San Diego, pp. 223–286.
- Camacho, J.P.M., Sharbel, T.F. & Beukeboom, L.W., 2000. B-chromosome evolution. *Phil. Trans. R. Soc. Lond. B*, 355, pp.163–178.
- Carvalho, M.L. et al., 2002. Nuclear DNA content determination in Characiformes fish (Teleostei, Ostariophysi) from the Neotropical region. *Genetics and Molecular Biology*, 25(2), pp.49–55.
- Carvalho, R.A., Martins-Santos, I.C. & Dias, A.L., 2008. B chromosomes: an update about their occurrence in freshwater Neotropical fishes (Teleostei). *Journal of Fish Biology*, 72(8), pp.1907–1932.
- Castro, J.P. et al., 2014. Diversity of the *Astyanax scabripinnis* species complex (Teleostei: Characidae) in the Atlantic Forest, Brazil: species limits and evolutionary inferences. *Reviews in Fish Biology and Fisheries*, 25(1), pp.231–244.
- Clark, F.E. et al., 2016. Dynamic Sequence Evolution of a Sex-Associated B Chromosome in Lake Malawi Cichlid Fish. *Journal of Heredity*, 108(1), pp.53–62.
- Cole, C.J. & Levans C.R., 1971. Chromosome preparations of amphibians and reptiles: improved technique. *Herpetol. Rev.*, 3(6), pp.102.
- Daniel, S.N. et al., 2012. Cytogenetic characterization of distinct B chromosomes in a population of the fish *Astyanax bockmanni* (Teleostei, Characiformes). *Caryologia*, 65(3), pp.229–233.
- Eschmeyer, W.N. & Fong, J.D., 2016. Catalog of fishes: Species by family/subfamily.
- Foresti et al., 1981. Polymorphic nature of the nucleolus organizer regions in fishes. *Cytogenet Cell Genet.*, 31, pp. 134-141.
- Fonseca, M.M., Harris, D.J. & Posada, D., 2014. The Inversion of the Control Region in Three Mitogenomes Provides Further Evidence for an Asymmetric Model of Vertebrate mtDNA Replication. D. Mishmar, ed. *PLoS ONE*, 9(9), pp.e106654–10.

- Frank, A.C. & Lobry, J.R., 1999. Asymmetric substitution patterns: a review of possible underlying mutational or selective mechanisms. *Gene*, 238, pp.65–77.
- Fu, L. et al., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28, pp.3150–3152.
- Haas, B.J. et al., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8), pp.1494–1512.
- Hahn, C., Bachmann, L. & Chevreur, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads--a baiting and iterative mapping approach. *Nucleic Acids Research*, 41(13), p.e129.
- Huang, W. et al., 2017. B chromosome contains active genes and impacts the transcription of A chromosomes in maize (*Zea mays* L.). *BMC Plant Biology*, pp.1–14.
- Kasmi, El, K.C. et al., 2007. Cutting Edge: A Transcriptional Repressor and Corepressor Induced by the STAT3-Regulated Anti-Inflammatory Signaling Pathway. *The Journal of Immunology*, 179(11), pp.7215–7219.
- Kostenko, E.V. et al., 2006. Ccpg1, a Novel Scaffold Protein That Regulates the Activity of the Rho Guanine Nucleotide Exchange Factor Dbs. *Molecular and Cellular Biology*, 26(23), pp.8964–8975.
- Ma, W. et al., 2016. Rye B chromosomes encode a functional Argonaute-like protein with in vitro slicer activities similar to its A chromosome paralog. *New Phytologist*, 213(2), pp.916–928.
- Maistro, E.L. et al., 1992. Occurrence of macro B chromosome in *Astyanax scabripinnis paranae* (Pisces, Characiformes, Characidae). *Genetica*, 87, pp.101–106.
- Maistro, E.L., Fausto, F. & Oliveira, C., 1994. New occurrence of a macro B-chromosome in *Astyanax scabripinnis paranae* (Pisces, Characiformes, Characidae). *Rev. Brasil. Genet.*, 17(2), pp.153–156.
- Makunin, A.I. et al., 2016. Contrasting origin of B chromosomes in two cervids (Siberian roe deer and grey brocket deer) unravelled by chromosome-specific DNA sequencing. *BMC Genomics*, pp.1–14.
- Margulis, L., 1981. Symbiosis in Cell Evolution, New York: Freeman.
- Martin, W., 2005. The missing link between hydrogenosomes and mitochondria. *Trends in Microbiology*, 13(10), pp.457–459.
- Maruyama, K. et al., 2013. Strawberry notch homologue 2 regulates osteoclast fusion by enhancing the expression of DC-STAMP. *The Journal of Experimental Medicine*, 210(10), pp.1947–1960.
- McGaugh, S.E. et al. 2014. The cavefish genome reveals candidate genes for eye loss. *Nature Communications*, 5, pp.1–10.
- Mestriner, C.A. et al., 2000. Structural and functional evidence that a B chromosome in the characid fish *Astyanax scabripinnis* is an isochromosome. *Heredity*, 85, pp.1–9.

- Miao, V.P., Covert, S.F. & VanEtten, H.D., 1991. A Fungal Gene for Antibiotic Resistance on a Dispensable (“B”) Chromosome. *Science*, 254, pp.1773–1776.
- Mitchell, D. & Bridge, R., 2006. A test of Chargaff’s second rule. *Biochemical and Biophysical Research Communications*, 340(1), pp.90–94.
- Mizoguchi, S.M.H.N. & Martins-Santos, I.C., 1997. Macro- and microchromosomes B in females of *Astyanax scabripinnis* (Pisces, Characidae). *Hereditas*, 127, pp.249–253.
- Moreira-Filho, O. & Bertollo, L.A.C., 1991. *Astyanax scabripinnis* (Pisces, Characidae): A species complex. *Rev. Brasil. Genet.*, 14(2), pp.331–357.
- Moreira-Filho, O., Galetti, P.M., Jr & Bertollo, L.A.C., 2004. B chromosomes in the fish *Astyanax scabripinnis* (Characidae, Tetragonopterinae): An overview in natural populations. *Cytogenetic and Genome Research*, 106(2-4), pp.230–234.
- Morelli, S. et al., 1983. Cytogenetic Considerations on the Genus *Astyanax* (Pisces, Characidae). I. Karyotypic Variability. *Caryologia*, 36(3), pp.235–244.
- Muldrow, T.A. et al., 1999. MOT1 Can Activate Basal Transcription In Vitro by Regulating the Distribution of TATA Binding Protein between Promoter and Nonpromoter Sites. *Molecular and Cellular Biology*, pp.2835–2845.
- Navarro-Domínguez, B. et al., 2017. Protein-coding genes in B chromosomes of the grasshopper *Eyprepocnemis plorans*. *Scientific Reports*, pp.1–12.
- Nguyen, P., 2013. Neo-sex chromosomes and adaptive potential in tortricid pests. *Proc Natl Acad Sci U S A*, 110, pp.6931–6936.
- Nikolaou, C. & Almirantis, Y., 2006. Deviations from Chargaff’s second parity rule in organellar DNA. *Gene*, 381, pp.34–41.
- Ning, Z., 2001. SSAHA: a fast search method for large DNA databases. *Genome Res.*, 11(10), pp.1725–9.
- Oliveira, C., Foresti, F. & Hilsdorf, A.W.S., 2008. Genetics of neotropical fish: from chromosomes to populations. *Fish Physiology and Biochemistry*, 35(1), pp.81–100.
- Ornelas-García, C. et al., 2008. Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evolutionary Biology*, 8(1), pp.340.
- Pansonato Alves, J.C. et al., 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(2-3), pp.133–142.
- Pansonato-Alves, J.C. et al., 2014. Single Origin of Sex Chromosomes and Multiple Origins of B Chromosomes in Fish Genus Characidium. *PLoS ONE*, 9(9), pp.e107169–10.
- Pereira, H.S. et al., 2017. Rye (*Secale cereale*) supernumerary (B) chromosomes associated with heat tolerance during early stages of male spermatogenesis. *Annals of Botany*, 119(3), pp.325–337.
- Perna, N.T. & Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution*, 41, pp.353–358.

- Plowman, A.B. & Bougourd, S.M., 1994. Selectively advantageous effects of B chromosomes on germination behaviour in *Allium schoenoprasum* L. *Heredity*, 72, pp.587–593.
- Raemaekers, T. et al., 2003. NuSAP, a novel microtubule-associated protein involved in mitotic spindle organization. *The Journal of Cell Biology*, 162(6), pp.1017–1029.
- Reyes, A. et al., 1998. Asymmetrical Directional Mutation Pressure in the Mitochondrial Genome of Mammals. *Molecular Biology and Evolution*, 15(8), pp.957–966.
- Rocon-Stange, E.A. & Almeida-Toledo, L.F.D., 1993. Supernumerary B chromosome restricted to males in *Astyanax scabripinnis* (Pisces, Characidae). *Rev. Brasil. Genet.*, 16(3), pp.601–615.
- Rossini, B.C. et al., 2016. Highlighting *Astyanax* Species Diversity through DNA Barcoding. *PLoS ONE*, 11(12), pp.e0167203–20.
- Rudner, R., Karkas, J.D. & Chargaff, E., 1968. Separation of *B. Subtilis* DNA into Complementary Strands, III. Direct Analysis. *Proc Natl Acad Sci USA*, 60(3), pp.921–922.
- Saccone, C. et al., 2000. Evolution of the mitochondrial genetic system: an overview. *Gene*, 261, pp.153–159.
- Silva, D.M.Z.A. et al., 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), pp.e94896–10.
- Silva, D.M.Z.A. et al., 2016. Origin of B chromosomes in the genus *Astyanax* (Characiformes, Characidae) and the limits of chromosome painting. *Molecular Genetics and Genomics*, 291(3), pp.1407–1418.
- Silva, D.M.Z.A. et al., 2017. High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*. *Scientific Reports*, pp.1–12.
- Sueoka, N., 1995. Intrastrand Parity Rules of DNA Base Composition and Usage Biases of Synonymous Codons. *Journal of Molecular Evolution*, 40, pp.318–325.
- Sumner, A.T., 1972. A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.*, 75, pp.304–306.
- Tosta, V.C. et al., 2015. Possible Introgression of B Chromosomes between Bee Species (Genus *Partamona*). *Cytogenetic and Genome Research*, 144(3), pp.220–226.
- Valente, G.T. et al., 2016. B chromosomes: from cytogenetics to systems biology. *Chromosoma*, pp.1–9.
- Valente, G.T. et al., 2014. Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Mol Biol Evol.*, 8, pp.2061–72.
- Werren, J.H., 1991. The Paternal-Sex-Ratio Chromosome of *Nasonia*. *Am. Nat.*, 137, pp.392–402.
- Xia, X., 1996. Maximizing Transcription Efficiency Causes Codon Usage Bias. *Genetics*, 144, pp.1309–1320.
- Yoshida, K. et al., 2011. B Chromosomes Have a Functional Effect on Female Sex Determination in Lake Victoria Cichlid Fishes. *PLoS Genetics*, 7(8), pp.e1002203–10.

## **7 Material suplementar – Capítulo 1**

### **7.1 Métodos**

#### **7.1.1 Obtenção de cromossomos metafásicos mitóticos**

Para obtenção dos cromossomos metafásicos foi utilizada a técnica proposta por (Fausto et al. 1981), com adaptações. O procedimento utilizado consiste em:

- 1) injetar, na região intra-abdominal, solução aquosa de colchicina (0,025%) na proporção de aproximadamente 1 ml / 100 g de peso do animal;
- 2) deixar o peixe em aquário bem aerado, por um período de 40 minutos;
- 3) sacrificar o animal, retirando o rim anterior. Transferir o material para uma placa de Petri, contendo 7 ml de solução hipotônica de KCl (0,075M);
- 4) dissociar o material com o auxílio de pinças de dissecação, complementando esse processo com o auxílio de uma pipeta Pasteur, até obter uma solução aquosa homogênea;
- 5) transferir a suspensão obtida para um tubo de centrifuga e deixar o tubo no interior de uma estufa a 37 °C por 20 minutos;

#### **7.1.2 Detecção da heterocromatina constitutiva (banda C)**

Para obtenção de bandas C foi utilizada a técnica descrita originalmente por (Sumner 1972). O procedimento consiste em:

- 1) secar as lâminas recém pingadas ao ar por uma noite;
- 2) hidrolisar as lâminas por 30 minutos em HCl 0,2N à temperatura ambiente e lavar em água destilada;
- 3) passar por uma solução de Ba(OH)<sub>2</sub> por cerca de 12 segundos e lavar em água destilada;
- 4) lavar em HCl 1N a 60 °C e, em seguida, em água destilada;

- 5) incubar por 30 minutos em 2xSSC (pH = 6,8) a 60 °C;
- 6) corar por 8 minutos com solução de Giemsa a 5% em tampão fosfato (pH = 6,7).
- 7) lavar deixar secar ao ar.

### 7.1.3 Confeção da sonda para FISH

As sequências foram sintetizadas por PCR (Polymerase Chain Reaction) a partir do DNA total de *Astyanax paranae* usando os primers relacionados na Tabela S1.

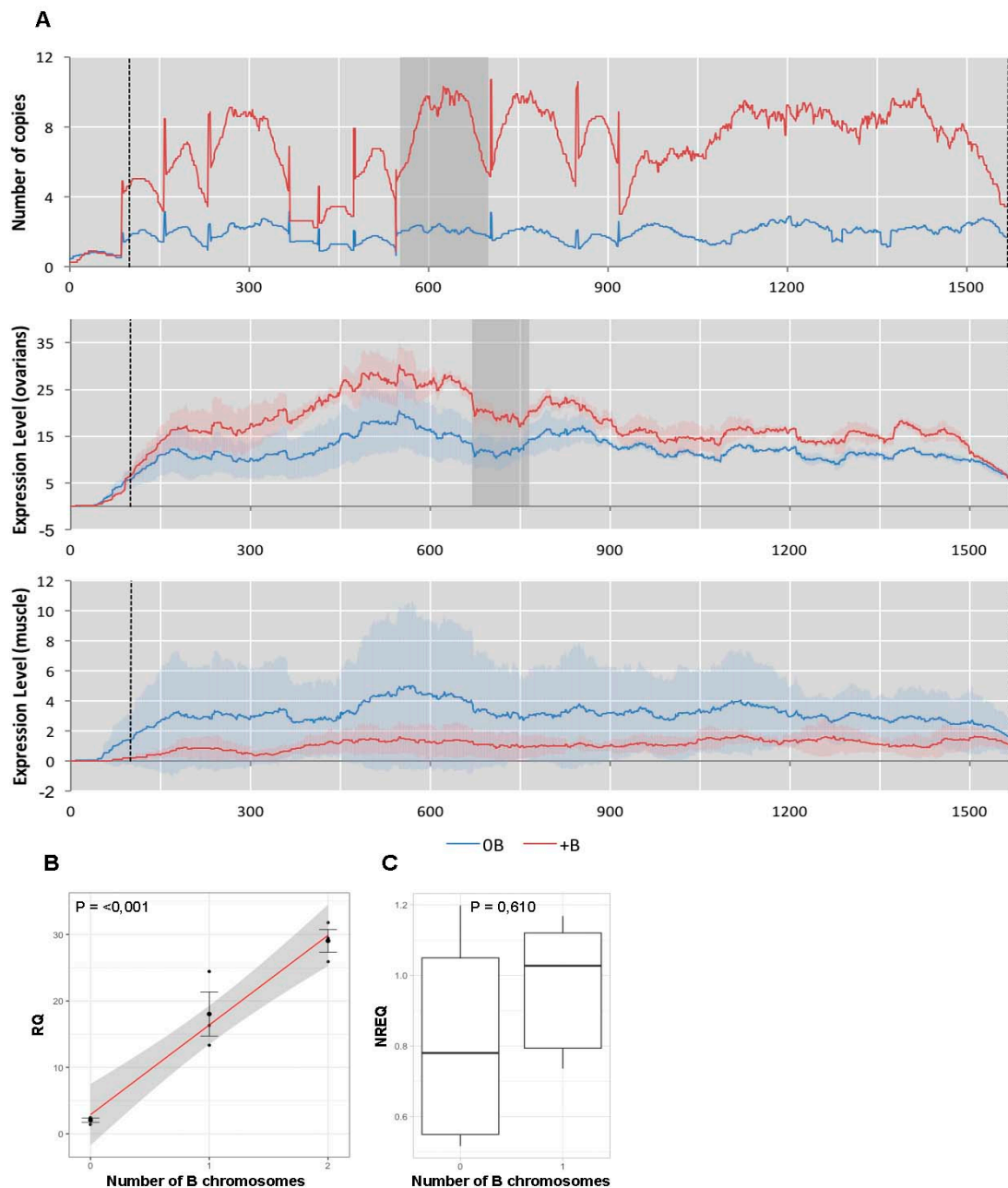
As reações foram realizadas com volume final de 25 µL, contendo 200 µM de cada dNTP (dATP, dCTP, dGTP, dTTP, digoxigenina-11-dUTP), 1,5 mM de MgCl<sub>2</sub>, 1 µL de Taq DNA tampão (20 mM Tris-HCl, pH 8,4 e 50 mM KCl), 0.5 U de Taq polymerase (Invitrogen), 0,1 µM de cada primer e 50 ng de DNA genômico, utilizando-se o seguinte programa:

```
95 °C ---- 5 m
95 °C ---- 1 m
54 °C ---- 45 s
72 °C ---- 1 m
72 °C ---- 10 m
12 °C ---- manutenção
```

} 30 x

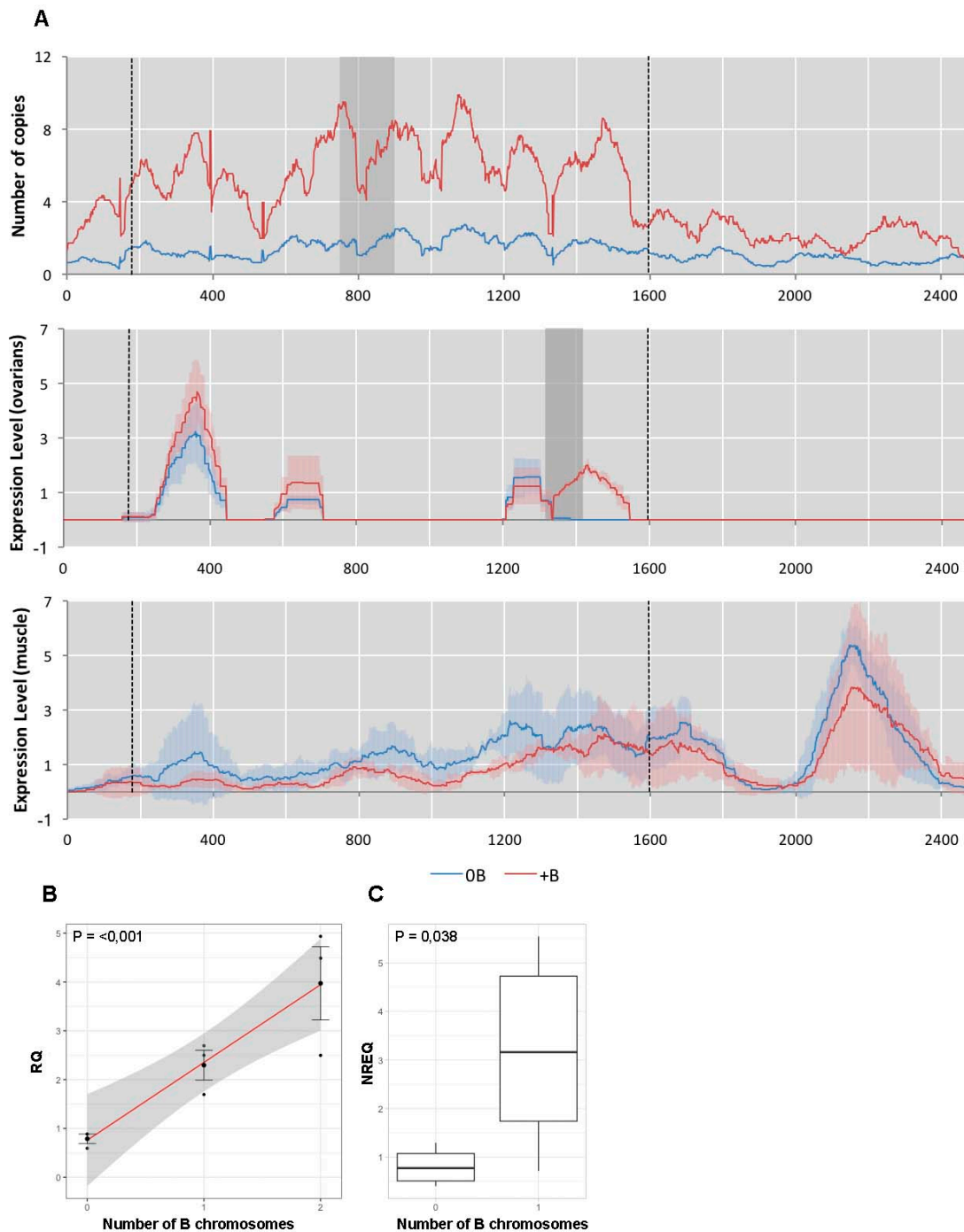
Os produtos gerados foram visualizados em gel de agarose 1%.

## 7.2 Figuras

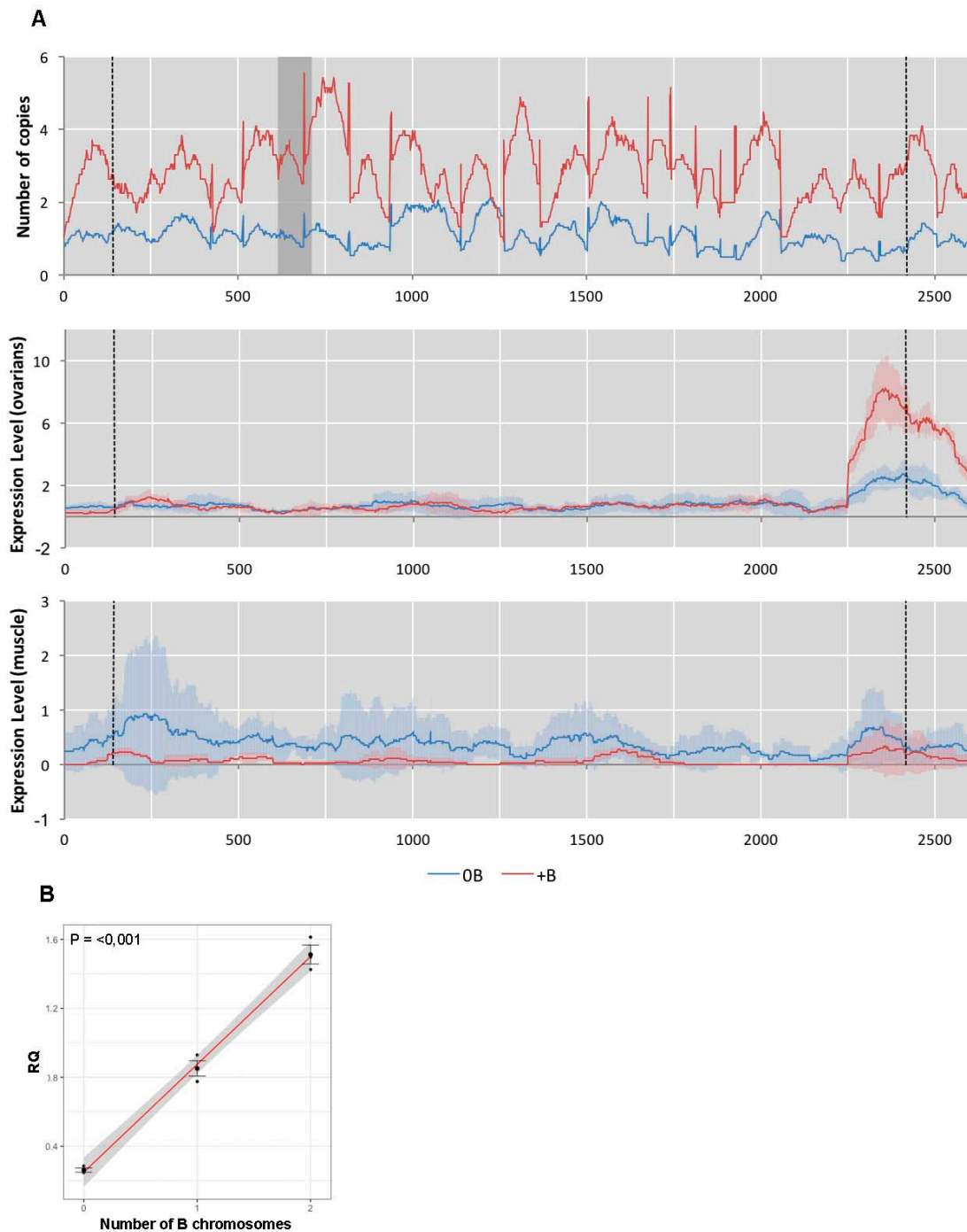


**Figura S1:** Cobertura para o transcrito do gene *mdm2* no gDNA (0B e 1B) e RNA (0B e 1B) de ovários e músculo a partir das bibliotecas Illumina de indivíduos de *A. paranae* (A), e qPCR no gDNA (B) e cDNA (C). Note que a cobertura é maior na biblioteca 1B ao longo de toda a CDS (delimitado pelas linhas pontilhadas verticais). Da mesma forma, note a maior cobertura para este transcrito na biblioteca de RNA de ovário 1B. As regiões sombreadas em A marcam as regiões amplificadas por qPCR. A qPCR no gDNA (B) mostrou que o número de cópias do gene *mdm2* aumenta com o número de cromossomos B, seguindo um padrão dose-dependente. Por outro lado, a qPCR no cDNA (C) mostrou que o gene *mdm2* não está diferencialmente expresso entre as fêmeas 0B e 1B. RQ = Quantidade relativa. NREQ = Quantidade relativa de expressão normalizada. P = P-value para teste de correlação de Spearman (B) e teste T de Student (C).

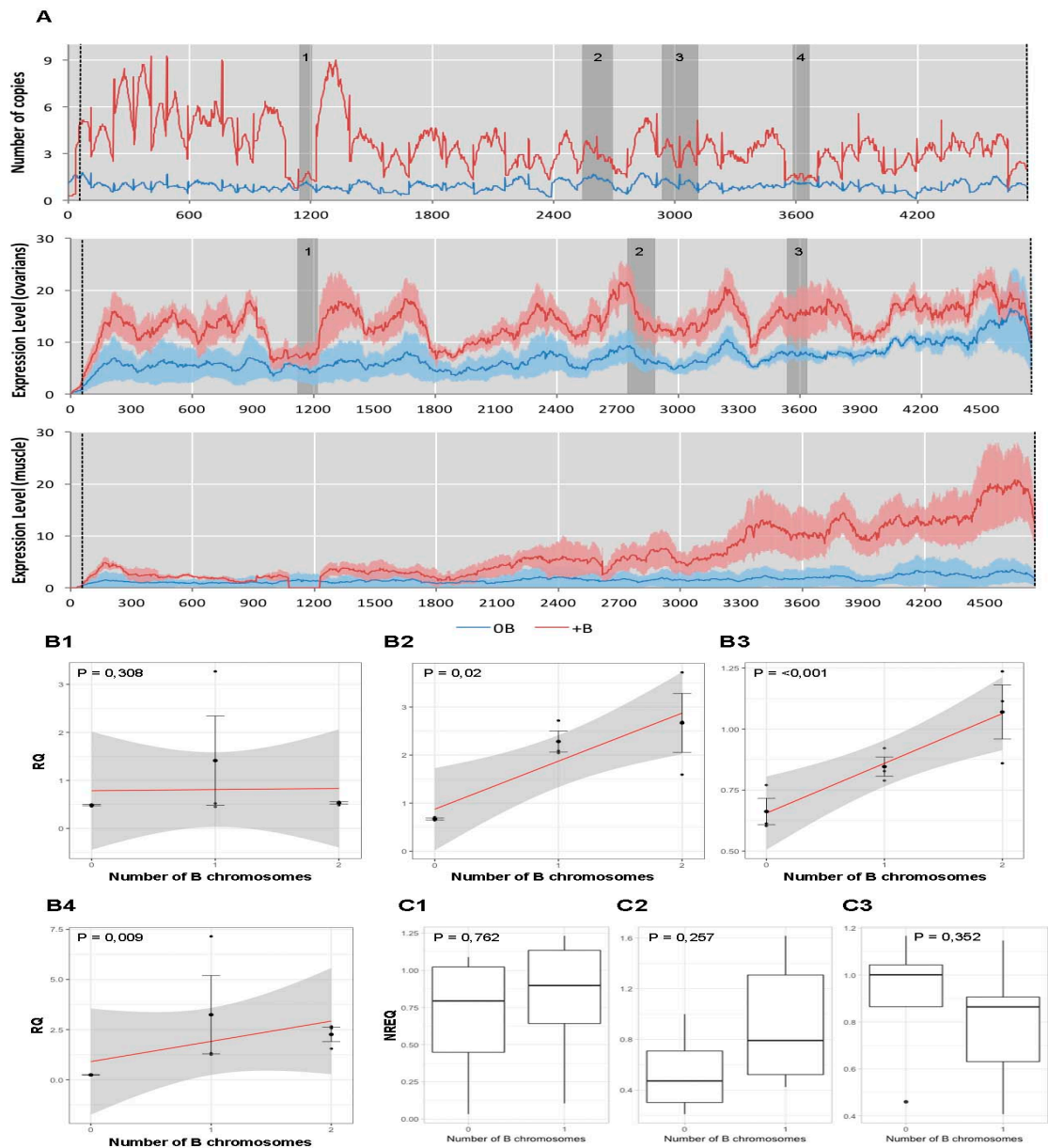




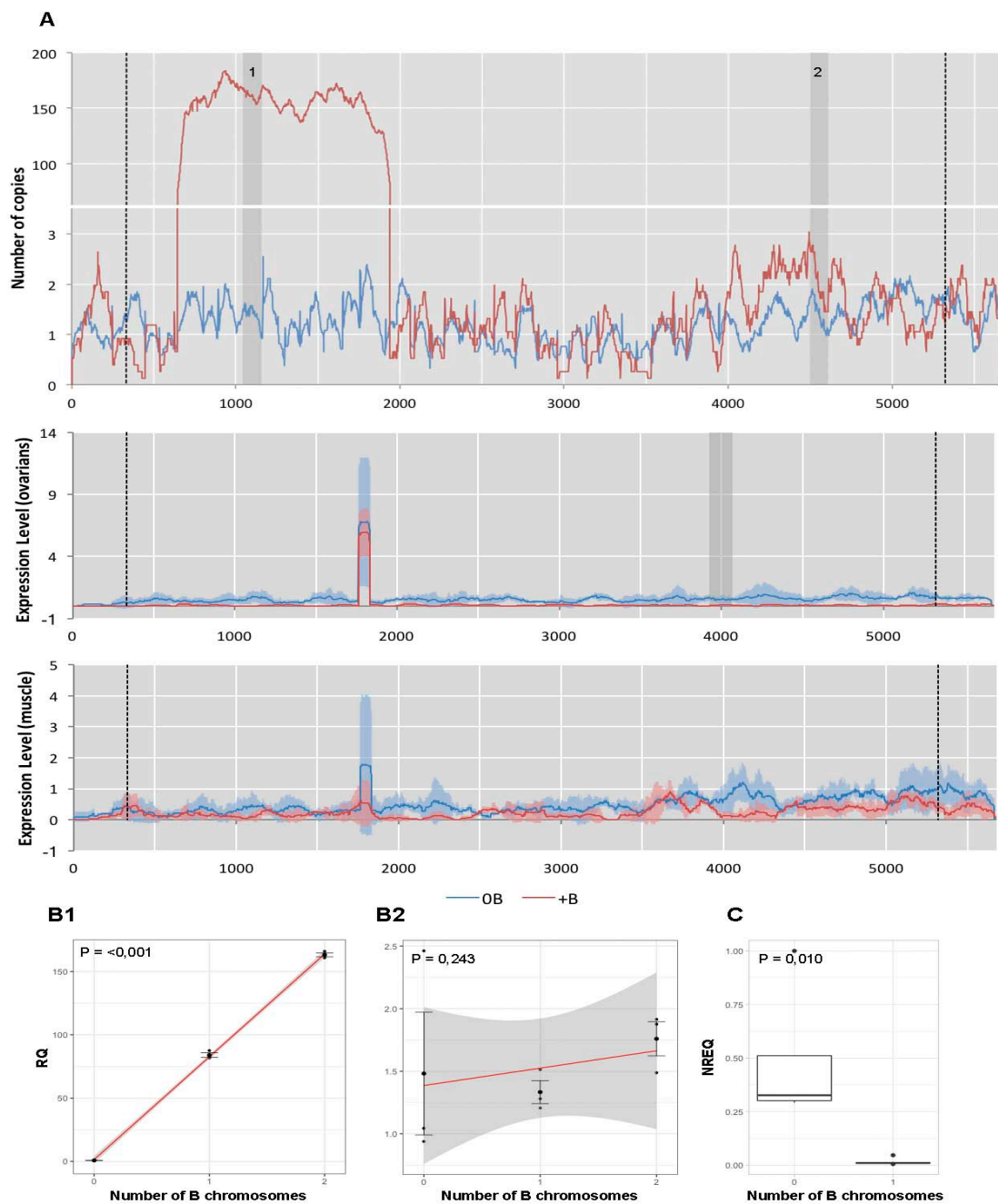
**Figura S2:** Cobertura para o transcrito do gene *mot1* no gDNA (0B e 1B) e RNA (0B e 1B) de ovários e músculo a partir das bibliotecas Illumina de indivíduos de *A. paranae* (A), e qPCR no gDNA (B) e cDNA (C). Note que a cobertura é maior na biblioteca 1B ao longo de toda a CDS (delimitado pelas linhas pontilhadas verticais). Da mesma forma, note a maior cobertura para este transcrito na biblioteca de RNA de ovário 1B. As regiões sombreadas em A marcam as regiões amplificadas por qPCR. A qPCR no gDNA (B) mostrou que o número de cópias do gene *mot1* aumenta com o número de cromossomos B, seguindo um padrão dose-dependente. A qPCR no cDNA (C) confirmou que o gene *mot1* está super-expresso no ovário das fêmeas 1B. RQ = Quantidade relativa. NREQ = Quantidade relativa de expressão normalizada. P = P-value para teste de correlação de Spearman (B) e teste de Mann-Whitney (C).



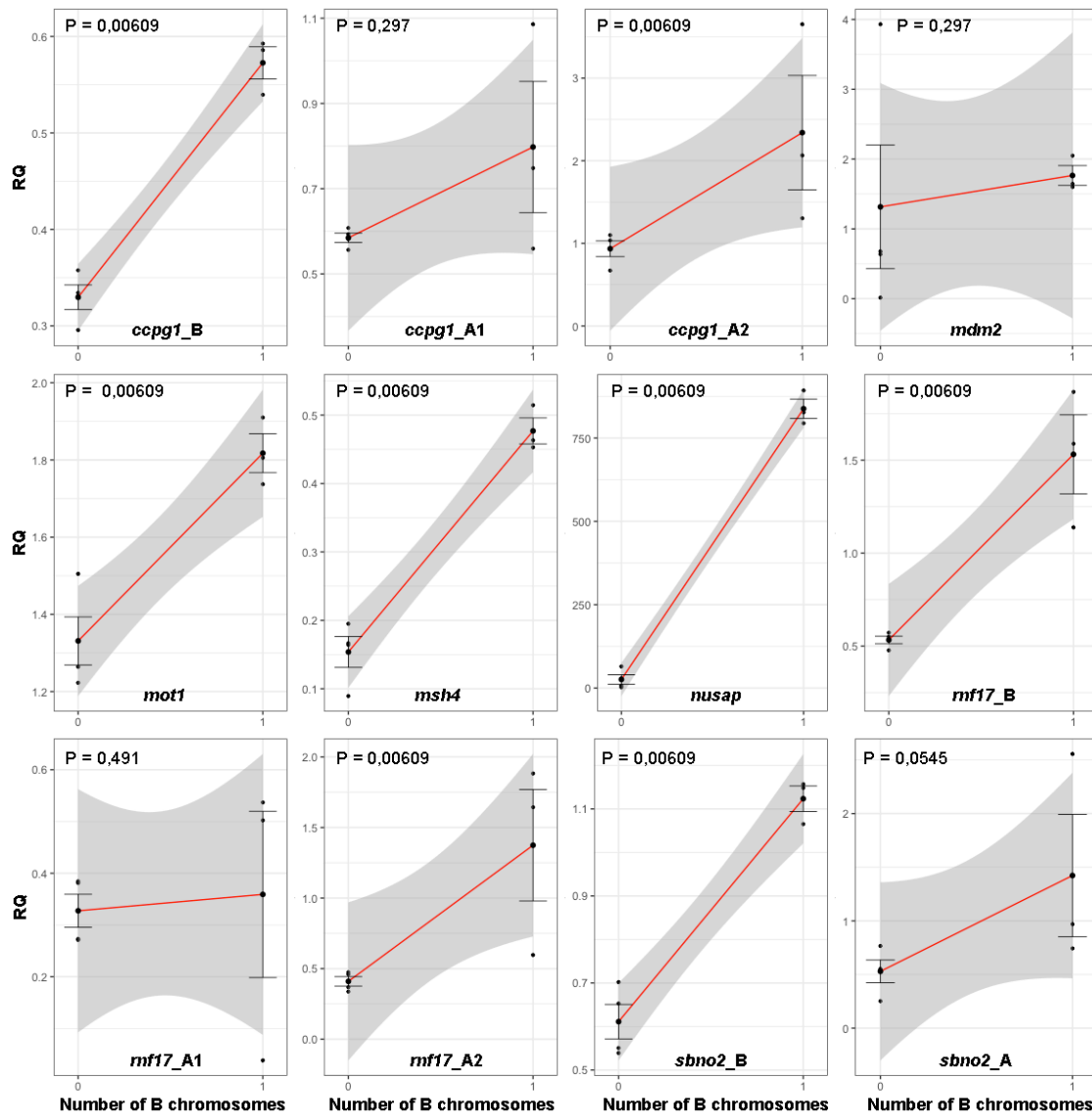
**Figura S3:** Cobertura para o transcrito do gene *msh4* no gDNA (0B e 1B) e RNA (0B e 1B) de ovários e músculo a partir das bibliotecas Illumina de indivíduos de *A. paranae* (A), e qPCR no gDNA (B) e cDNA (C). Note que a cobertura é maior na biblioteca 1B ao longo de toda a CDS (delimitado pelas linhas pontilhadas verticais). Na biblioteca de RNA dos ovários a cobertura é maior para os indivíduos 1B apenas em uma região no final da CDS. A qPCR no gDNA (B) mostrou que o número de cópias do gene *msh4* aumenta com o número de cromossomos B, seguindo um padrão dose-dependente. Não foi possível amplificar a região com expressão diferencial nos ovários por qPCR a partir do cDNA. RQ = Quantidade relativa. P = P-value para teste de correlação de Spearman (B).



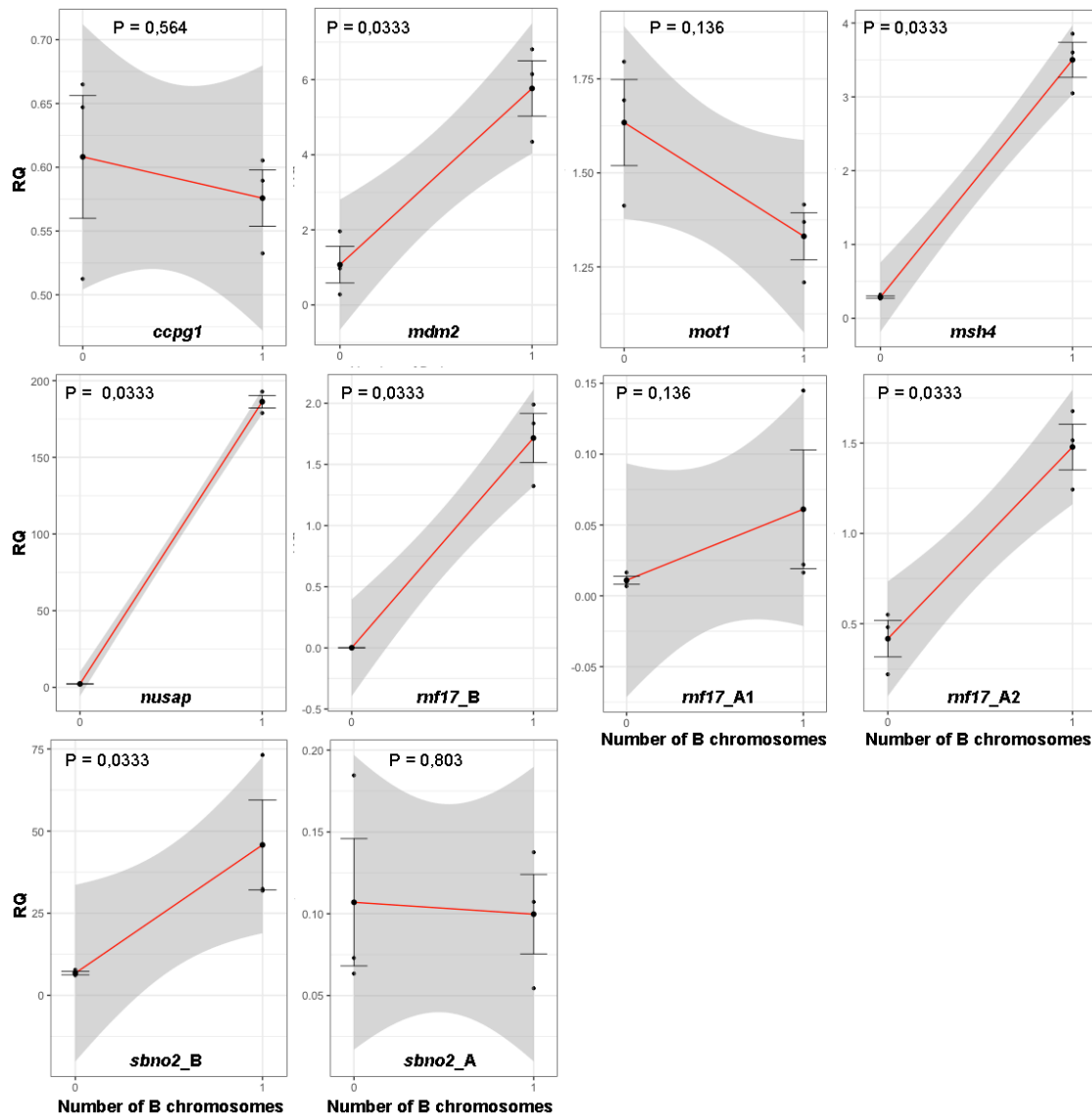
**Figura S4:** Cobertura para o transcrito do gene *rnf17* no gDNA (0B e 1B) e RNA (0B e 1B) de ovários e músculo a partir das bibliotecas Illumina de indivíduos de *A. paranae* (A), e qPCR no gDNA (B) e cDNA (C). Note que a cobertura é maior na biblioteca 1B na maior parte da CDS (delimitado pelas linhas pontilhadas verticais). Entretanto, as regiões sombreadas 1 e 4 possuem coberturas similares entre as bibliotecas 0B e 1B, indicando que este gene está truncado no cromossomo B. As três regiões sombreadas foram selecionadas para amplificação por qPCR, as regiões 1 e 4 que parecem não estar neste cromossomo e as regiões 2 e 3 que parecem estar presentes no cromossomo B. A qPCR na região 1 mostrou que o número de cópias independe do número de cromossomos B (B1). Já para as regiões 2 e 3, a qPCR mostrou que o número de cópias para este gene nestas regiões aumenta com o número de cromossomos B em um padrão dose-dependente (B2 e B3). Por outro lado, a qPCR na região 4 mostrou que o número de cópias está correlacionado ao número de cromossomos B, porém, não como um padrão dose-dependente, indicando que esta região não está presente no cromossomo B (B4). Da mesma forma, note a cobertura irregular para este transcrito na biblioteca de RNA de ovário 1B. No entanto, a qPCR no cDNA (C) não evidenciou expressão diferencial em nenhuma das regiões testadas. RQ = Quantidade relativa. NREQ = Quantidade relativa de expressão normalizada. P = P-value para teste de correlação de Spearman (B) e teste T de Student (C).



**Figura S5:** Cobertura para o transcrito do gene *sbno2* no gDNA (0B e 1B) e RNA (0B e 1B) de ovários e músculo a partir das bibliotecas Illumina de indivíduos de *A. paranae* (A), e qPCR no gDNA (B) e cDNA (C). Note que a cobertura é maior na biblioteca 1B apenas em uma parte da CDS (delimitado pelas linhas pontilhadas verticais). Na maior parte a CDS, as bibliotecas apresentam cobertura semelhante, indicando que este gene possui um padrão de amplificação regional. As duas regiões sombreadas foram selecionadas para amplificação por qPCR, a região 1 que parece estar presente no cromossomo B e a região 2 que parece não estar neste cromossomo. A qPCR na região 1 mostrou que o número de cópias para este gene nesta região aumenta com o número de cromossomos B em um padrão dose-dependente (B1). Já para a região 2 o número de cópias independe do número de cromossomos B (B2). Da mesma forma, note a menor cobertura para este transcrito na biblioteca de RNA de ovário 1B. A qPCR no cDNA (C) confirmou que o gene *sbno2* está sub-expresso no ovário das fêmeas 1B. RQ = Quantidade relativa. NREQ = Quantidade relativa de expressão normalizada. P = P-value para teste de correlação de Spearman (B) e teste de Mann-Whitney (C).



**Figura S6:** Quantificação por qPCR dos genes presentes no cromossomo B de *A. paranae* em gDNA de *A. bockmanni*. RQ = Quantificação relativa. P = P-value para teste de correlação de Spearman. As letras "B" e "A (1 ou 2)" após os nomes dos genes indicam a quantificação nas regiões presentes ou ausentes no cromossomo B de *A. paranae*, respectivamente.



**Figura S7:** Quantificação por qPCR dos genes presentes no cromossomo B de *A. paranae* em gDNA de *A. fasciatus*. RQ = Quantificação relativa. P = P-value para teste de correlação de Spearman. As letras “B” e “A (1 ou 2)” após os nomes dos genes indicam a quantificação nas regiões presentes ou ausentes no cromossomo B de *A. paranae*, respectivamente.

### 7.3 Tabelas

**Tabela S1.** Primers utilizados nos experimentos de qPCR em gDNA de espécies de *Astyanax*.

Nome	Sequencia	Nome	Sequencia
cc_A1_F	GAAAAGCGAGACATGGTGTGT	cc_A1_R	TCTTTTGCTCTGAGATGCTGGT
cc_A2_F	GGTTGGCTGAGGATTGGACA	cc_A2_R	GGGAGTGTTGAACAGCTGCT
cc_F	GAACCGTGCTCTGAAGACCA	cc_R	GCGTTGGTTCTCAGACAGGA
hprt_F	GGCCAGGGAGATCATGAAGG	hprt_R	TGGAGCGGTCACTATTTCCG
mdm2_F	TTCTCAGTGGAGGAGGAAAGC	mdm2_R	TGCTGCGTGAGTCTGAAGAC
mot1_F	GCTGGATCTCTCATGAGGCC	mot1_R	AAGGAAGCCACGGTGTTTGA
ms_F	GGAATTCACAGACTGCATGGC	ms_R	AAGTGAACGGTCGCTCATCA
nusap_F	CTGCCAGTCGTGCCCCATT	nusap_R	TTTGTGACTGAAAGGACAGCTG
rnf_A1_F	TGGAGTCCAAGCGCTTACAG	rnf_A1_R	CCTTTTGTGGCTTACGGCTC
rnf_A2_F	AGCACATTAAGACGCTCCCC	rnf_A2_R	CATGTCCGAGCCCATGACAG
rnf_B_F	CCTGGATACTGCTGTGTGGG	rnf_B_R	CTCAGGTCCTTAAGGCAGGC
rnf_B2_F	ACACGGTTAAGCTCCACATCA	rnf_B2_R	GAGGCAACACTCCAGCAGTA
sbno_A_F	CCACACACTGAAACAGCAGC	sbno_A_R	CCGAGCTGCTATCTGACTGG
sbno_B_F	AGAGACCAGCACTCTCTCCA	sbno_B_R	GATGGCCTCAAGCTGTAGGG

**Tabela S2.** Primers utilizados nos experimentos de RT-qPCR em cDNA de *A. paranae*.

Nome	Sequencia	Nome	Sequencia
cc_ex_F	CAAGACGAGGGTGGTGTATGG	cc_ex_R	CAGGATGCACTTGTGAGCG
hprt_ex_F	CTCACGGACTCATCTTGGACAGGAC	hprt_ex_R	GCAGATCTGCAAAGAAGCTGTAGCC
md_ex_F	GCAGCATATCGGATGTGGGT	md_ex_R	CTCGCTGTACGTGTCCGAAT
mo_ex_F	TCGGTCCTCCTTTACTGGGA	mo_ex_R	GAGGAACAAGCTTCCGACCA
nusap_ex_F	AAGATTCAGCCTAAGCCGCC	nusap_ex_R	CTGAAGGGGACACTGTGCTT
rnf_ex_F	TGATAACCAACTGAAGAGGGT	rnf_ex_R	TGCACACCTTTCCTCTCTCC
rnf_exA1_F	ACAGAGAGAGAAGGAGTGGTCT	rnf_exA1_R	GTAGAAGAGGTGTCCTGTCTT
rnf_exA2_F	CCACAGGCTGAGTGTGAGTT	rnf_exA2_R	CAAGCACACTCTTCCAGTTGT
sbno_ex_F	CAGAACAGCAAGCAAGTTGGTAT	sbno_ex_R	CTGCTCCTGTATGCGCTTCT

## 8 Material suplementar – Capítulo 2

**Tabela S3.** Análise de Variância de Dupla Classificação para Medidas Repetidas para variáveis associadas ao conteúdo de bases do DNA mitocondrial em classes de Cordados.

Comparisons for factor: Group				
Comparison	Diff of Means	t	P	P<0,050
ost vs. cond	0,0286	11,12	<0,001	Yes
ost vs. aves	0,0224	16,762	<0,001	Yes
ost vs. amp	0,014	7,039	<0,001	Yes
ost vs. mamma	0,0105	8,624	<0,001	Yes
ost vs. rep	0,00627	2,979	0,044	Yes
rep vs. cond	0,0223	6,998	<0,001	Yes
rep vs. aves	0,0161	6,974	<0,001	Yes
rep vs. amp	0,00772	2,813	0,074	No
rep vs. mamma	0,00427	1,896	0,872	Do Not Test
mamma vs. cond	0,0181	6,712	<0,001	Yes
mamma vs. aves	0,0119	7,636	<0,001	Yes
mamma vs. amp	0,00345	1,611	1	Do Not Test
amp vs. cond	0,0146	4,695	<0,001	Yes
amp vs. aves	0,00843	3,818	0,002	Yes
aves vs. cond	0,0062	2,256	0,363	No
Comparisons for factor: Group within GC				
Comparison	Diff of Means	t	P	P<0,05
ost vs. aves	0,123	49,214	<0,001	Yes
ost vs. rep	0,0959	24,299	<0,001	Yes
ost vs. mamma	0,061	26,644	<0,001	Yes
ost vs. cond	0,0478	9,905	<0,001	Yes
ost vs. amp	0,0133	3,568	0,005	Yes
amp vs. aves	0,11	26,6	<0,001	Yes
amp vs. rep	0,0826	16,073	<0,001	Yes
amp vs. mamma	0,0477	11,902	<0,001	Yes
amp vs. cond	0,0345	5,905	<0,001	Yes
cond vs. aves	0,0756	14,68	<0,001	Yes
cond vs. rep	0,0481	8,042	<0,001	Yes
cond vs. mamma	0,0133	2,626	0,13	No
mamma vs. aves	0,0623	21,379	<0,001	Yes
mamma vs. rep	0,0349	8,266	<0,001	Yes
rep vs. aves	0,0275	6,33	<0,001	Yes
Comparisons for factor: Group within AT				
Comparison	Diff of Means	t	P	P<0,05
rep vs. amp	0,098	19,075	<0,001	Yes
rep vs. cond	0,0928	15,512	<0,001	Yes
rep vs. ost	0,0833	21,119	<0,001	Yes
rep vs. mamma	0,0434	10,29	<0,001	Yes
rep vs. aves	0,00483	1,113	1	No



aves vs. amp	0,0932	22,524	<0,001	Yes
aves vs. cond	0,088	17,087	<0,001	Yes
aves vs. ost	0,0785	31,323	<0,001	Yes
aves vs. mamma	0,0386	13,229	<0,001	Yes
mamma vs. amp	0,0546	13,622	<0,001	Yes
mamma vs. cond	0,0494	9,79	<0,001	Yes
mamma vs. ost	0,0399	17,439	<0,001	Yes
ost vs. amp	0,0147	3,946	0,001	Yes
ost vs. cond	0,00947	1,964	0,744	No
cond vs. amp	0,00522	0,894	1	No

ost = Osteichthyes; cond = Chondrichthyes; amp = Amphibia; mamma = Mammalia; rep = Reptilia.

## 9 Apêndices

A seguir, apresentamos nossos artigos publicados sobre cromossomos B do gênero *Astyanax* e a caracterização do DNA mitocondrial de *A. paranae* produzidos como parte desta tese.

Silva, D.M.Z.A. et al., 2016. Origin of B chromosomes in the genus *Astyanax* (Characiformes, Characidae) and the limits of chromosome painting. *Molecular Genetics and Genomics*, 291(3), pp.1407–1418.

Silva, D.M.Z.A. et al., 2017. The complete mitochondrial genome sequence of *Astyanax paranae* (Teleostei: characiformes). *Mitochondrial DNA Part B*, 1(1), pp., 586–587.

Silva, D.M.Z.A. et al., 2017. High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*. *Scientific Reports*, pp.1–12.