

RESSALVA

Atendendo solicitação do(a)
autor(a), o texto completo desta tese
será disponibilizado somente a partir
de 16/05/2021.

Universidade Estadual Paulista “Júlio de Mesquita Filho”

Instituto de Biociências de Botucatu

Departamento de Morfologia

Análises em larga escala de proteínas e construção de redes
biológicas com foco em estudos de cromossomos B

Rafael Takahiro Nakajima

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

Botucatu/SP

Maio de 2019

Instituto de Biociências - Seção Técnica de Pós-Graduação

Distrito de Rubião Júnior s/n CEP 18618-970 Cx Postal 510 Botucatu-SP Brasil

Tel (14) 3880-0780 posgraduacao@ibb.unesp.br

Universidade Estadual Paulista “Júlio de Mesquita Filho”

Instituto de Biociências de Botucatu

Departamento de Morfologia

Análises em larga escala de proteínas e construção de redes
biológicas com foco em estudos de cromossomos B

Rafael Takahiro Nakajima

Orientador: Prof. Dr. Cesar Martins

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

Botucatu/SP

Maio de 2019

Instituto de Biociências - Seção Técnica de Pós-Graduação

Distrito de Rubião Júnior s/n CEP 18618-970 Cx Postal 510 Botucatu-SP Brasil

Tel (14) 3880-0780 posgraduacao@ibb.unesp.br

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

Nakajima, Rafael Takahiro.

Análises em larga escala de proteínas e construção de redes biológicas com foco em estudos de cromossomos B / Rafael Takahiro Nakajima. - Botucatu, 2019

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

Orientador: Cesar Martins

Capes: 20204000

1. Proteômica. 2. Cromossomos. 3. RNA mensageiro. 4. Sequenciamento de nucleotídeos em larga escala. 5. Ciclídeos.

Palavras-chave: RNA codificante; RNA mensageiro; ciclídeo africano; cromossomo extra; proteomica.

Agradecimentos

Meus sinceros agradecimentos:

- Ao meu orientador Prof. Dr. Cesar Martins pela oportunidade, confiança, incentivo e aprendizagem ao longo desses quatro anos de trabalho.
- Aos membros do Laboratório de Genômica Integrativa.
- Aos membros do Laboratório SBGL e CEVAP.
- Ao José Eduardo por sempre nos ajudar e pelo amigo que se tornou.
- Aos meus pais por me apoiarem nas minhas decisões, sempre estando ao meu lado em momentos bons e ruins. São exemplos que desejo seguir.
- À Talita Aleixo, por me ajudar e me aguentar nos maus e bons momentos, sempre me apoiando e com um bom humor.
- À nova família que ganhei que sempre acreditaram em nosso potencial e capacidade.
- Ao meus amigos: Akira, Chokito, Érica, Gigio, Costinha, Jones, Patty e Rancheira pelo apoio e descontrações.

Apoio financeiro

- À Fundação de Amparo à Pesquisa do Estado de São Paulo – processo 2015/11485-0 (bolsa), 2013/04533-3, 2015/016661-1 e 2018/24610-6
- Ao CAPES (bolsa – 1 ano) e CNPq 474684/2013-0

Resumo

Os cromossomos B ocorrem em cerca de 2.828 espécies de diferentes táxons, sendo basicamente heterocromáticos e compostos de DNAs repetitivos. Recentemente, análises genômicas em larga escala estão sendo utilizadas para elucidar questões acerca dos cromossomos supranumerários. Os peixes ciclídeos recebem grande interesse científico, uma vez que muitas espécies passaram por um rápido e extenso processo de radiação adaptativa. Em algumas espécies do grupo, como *Astatotilapia latifasciata*, foi descrita a presença de cromossomos B. Neste trabalho foi caracterizado o perfil de expressão proteico em tecidos específicos na *A. latifasciata* e realizada análise funcional da presença do cromossomo B nesta espécie de teleósteo, elucidando a influência que este pode acarretar em vias metabólicas específicas. Além disso, esses dados foram integrados com os resultados de RNA-Seq dessa espécie, e construídas sub redes de co-expressão e interação proteína-proteína. Também foi calculada a entropia de *Shannon*, a qual não apresentou diversidade na expressão dos transcritos em cada biblioteca comparada. Além disso, foi analisada a expressão diferencial de RNAm em cada tecido em relação a presença do cromossomo B e ao sexo. Dentre os transcritos diferencialmente expressos, a análise de enriquecimento funcional apresentou processos relacionados ao ciclo celular, resposta imune e resposta ao estresse. Na maioria dos casos analisados entre a expressão de proteínas, os transcritos *up regulated* e as sub redes, apareceram os genes Aurora quinase, tubulina e proteína de cinetócoro que estão relacionados ao controle do ciclo celular. Os resultados obtidos, de forma geral, não apresentaram grande alteração no perfil proteico ou gênico relacionado a presença do cromossomo B, apesar de identificar proteínas e genes putativos para as vias do GnRH (Gonadotropin-Releasing Hormone), do ciclo celular e do estresse. Contudo, essa alteração no perfil de expressão sugere regulação mais sutil e pontual. Portanto, o cromossomo B pode atuar na regulação de processos vitais para sobrevivência da célula e facilitar sua manutenção e transmissão para as próximas gerações.

Palavras-chave: proteômica, cromossomo extra, RNA mensageiro, RNA codificante, ciclídeo africano

Abstract

The B chromosomes occur in about 2,828 species of different taxa, being heterochromatic and composed of repetitive DNA. Recently, large-scale genomic analyzes are being used to elucidate questions about supernumerary chromosomes. Cichlid fish are of great scientific interest, since many species have gone through a rapid and extensive process of adaptive radiation. In some species of the group, such as *Astatotilapia latifasciata*, the presence of B chromosomes was described. In this work, we characterize the profile of protein expression in specific tissues in *A. latifasciata* and performed a functional analysis of the presence of the B chromosome in this species of teleost, elucidating the influence that it can cause in specific metabolic pathways. In addition, we integrate these data with the RNA-Seq results of this species, and construct sub-networks of co-expression and protein-protein interaction. The Shannon entropy was also calculated, which did not show diversity in the expression of the transcripts in each library. In addition, differential expression analysis was performed on each tissue separately and the relationship between the presence of B chromosome and sex chromosome was analyzed. Among the differentially expressed transcripts, functional enrichment analysis presented processes related to the cell cycle, immune response and stress response. In relation to abundance of proteins, the up regulated transcripts and the sub-networks we identified genes like the Aurora kinase, tubulin and kinetochore protein genes that are related to cell cycle control. The results obtained, in general, did not show a great change in the protein or gene profile related to the presence of the B chromosome, although it identified putative proteins and genes for the GnRH (Gonadotropin-Releasing Hormone), cell cycle and stress pathways. However, this change in the expression profile suggests subtler and punctual regulation. Therefore, the B chromosome can act in the regulation of vital processes for the survival of the cell and facilitate its maintenance and transmission for the next generations.

Keywords: proteomic, extra chromosome, messenger RNA, coding RNA, African cichlids

Lista de Ilustrações

Figura 1. Hipótese da origem do cromossomo B.	17
Figura 2. Ciclídeo africano <i>Astatotilapia latifasciata</i>	18
Figura 3. Cariótipo do ciclídeo <i>Astatotilapia latifasciata</i>	19
Figura 4. Exemplo de grafos direcionados, não direcionados e com peso.....	24
Figura 5. Representação ilustrativa do aquário em que foram mantidos os peixes para a extração proteica.	27
Figura 6. Corte histológico das gônadas femininas e masculinas.	37
Figura 7. Análise quantitativa da extração proteica.....	38
Figura 8. Mapa do ciclo celular.	40
Figura 9. Mapa biológico da sinalização do GnHR.....	41
Figura 10. Mapa biológico da fosforilação oxidativa.	42
Figura 11. Mapa biológico da sinalização MAPK.....	44
Figura 12. Mapa biológico da sinalização do GnHR.....	45
Figura 13. Mapa biológico do transporte de RNA.	46
Figura 14. Análise de entropia dos transcritos.	47
Figura 15. Análise das bibliotecas de encéfalo de macho.....	49
Figura 16. Análise das bibliotecas de encéfalo de fêmea.....	51
Figura 17. Análise das bibliotecas de músculo de macho.....	53
Figura 18. Análise das bibliotecas de músculo de fêmea.....	55
Figura 19. Análise das bibliotecas de gônada de macho.....	56
Figura 20. Análise de cobertura genômica.....	58
Figura 21. Análise das bibliotecas de gônada de fêmea.	59
Figura 22. Representação gráfica da análise das sub redes up regulated extraídas da rede de união Co-PPI-net que apresentam genes envolvidos no ciclo celular.....	71

Lista de Tabelas

Tabela 1. Genes envolvidos com ciclo celular e presentes no cromossomo B.	20
Tabela 2. Comparações realizadas para abundância das proteínas.....	30
Tabela 3. Mapas biológicos com a presença de proteínas com abundância alterada.....	39

Tabela 4. Número de acesso das enzimas encontradas para amostras de encéfalo de fêmeas B+.....	41
Tabela 5. Mapas biológicos com a presença de proteínas com abundância alterada.....	43
Tabela 6. Mapas biológicos com a presença de proteínas com abundância alterada.....	45
Tabela 7. Transcritos diferencialmente expresso em gônadas de fêmea.....	60
Tabela 8. Fisher's exact test dos genes em comum up regulated entre amostras de fêmea e macho de encéfalo.....	60
Tabela 9. Fisher's exact test dos genes em comum down regulated entre amostras de fêmea e macho de encéfalo.	61
Tabela 10. Fisher's exact test dos genes em comum up regulated entre amostras de fêmea e macho de músculo.....	62
Tabela 11. Termos ontológicos idênticos encontrados no transcriptoma e proteoma. ..	63
Tabela 12. Termos ontológicos idênticos encontrados no transcriptoma e proteoma. ..	64
Tabela 13. Termos ontológicos idênticos encontrados no transcriptoma e proteoma. ..	65
Tabela 14. Análise das sub redes.	69
Tabela 15. Genes relacionados ao ciclo celular entre os 10 maiores valores para degree e betweenness centrality.	70

Lista de abreviaturas e siglas

PCA	Análise da componente principal
B+	Indivíduos com cromossomo B
B-	Indivíduos sem cromossomo B
DNA	Ácido desoxirribonucleico
hnRNP Q-like	<i>Heterogeneous nuclear ribonucleoprotein Q-like</i>
TUBB1	<i>Tubulin beta 1</i>
KIF11	<i>Kinesin-like protein KIF11</i>
miRNAs	MicroRNA

RNA	Ácido ribonucleico
RNA _m	RNA mensageiro
RNA-Seq	Sequenciamento de ácido ribonucleico
CDKs	<i>Cyclin-dependent kinases</i>
MPS	<i>Monopolar spindle 1 kinase</i>
SAC	<i>Spindle assembly checkpoint</i>
APC/C	<i>Anaphase-promoting complex</i>
NDC80	<i>Kinetochore protein NDC80 homolog</i>
Nuf2	<i>Kinetochore complex component</i>
Spc24 e Spc25	<i>NDC80 kinetochore complex component</i>
PCR	Reação em cadeia da polimerase
ACN	Acetonitrila
DDT	<i>Dithiothreitol</i>
EDTA	Ácido etilenodiamino tetra-acético
BSA	Proteica de albumina bovina
rDNA	DNA ribossomal
tRNA	RNA transportador
SAMP	<i>Single-Assembler Multiple-Parameter</i>
PPI	Interação proteína-proteína
PPI-net	Rede biológica de interação proteína-proteína
ORFs	<i>openreadingframes</i>

Sumário

1	Introdução	15
1.1	Cromossomos B	15
1.2	Origem do cromossomo B	16
1.3	<i>Astatotilapia latifasciata</i> (Perciformes, Cichlidae) como modelo para análises de cromossomos B	17
1.4	Análises genômicas em larga escala de cromossomos B em <i>Astatotilapia latifasciata</i>	19
1.5	Análises funcionais: Proteômica	21
1.6	Integração de dados	22
1.6.1	Redes biológicas	22
1.7	Ciclo celular	24
2	Hipótese	25
3	Proposta	26
4	Objetivos	26
4.1	Objetivo Geral	26
4.2	Objetivos Específicos	26
5	Material e Métodos	26
5.1	Animais, extração de DNA e genotipagem de animais B+ e B-	26
5.2	Estudo morfológico das gônadas	27
5.3	Extração de proteínas	28
5.4	Quantificação, digestão e espectrometria de massa	28
5.5	Análise dos dados - Proteômica	29
5.6	Reanálise do transcriptoma (RNAm)	30
5.6.1	Filtragem das <i>reads</i>	31
5.6.2	Montagem <i>de novo</i> e filtragem	31

5.6.3	Entropia de <i>Shannon</i>	33
5.6.4	Expressão diferencial de RNAm	34
5.6.5	Análise comparativa	34
5.7	Construção das redes.....	34
6	Resultados.....	36
6.1	Análise histológica.....	36
6.2	Proteômica	38
6.3	Abundancia proteica e mapas biológicos	38
6.3.1	Encéfalo	39
6.3.2	Músculo.....	43
6.3.3	Gônada	44
6.4	Transcriptoma	46
6.4.1	Entropia de <i>Shannon</i>	47
6.4.2	Expressão diferencial.....	48
6.5	Análise comparativa: Transcriptoma	60
6.5.1	Relação macho e fêmea	60
6.6	Análise comparativa: Transcriptoma e Proteoma	62
6.6.1	Encéfalo	62
6.6.2	Músculo.....	63
6.6.3	Gônada	64
6.7	Redes Biológicas.....	65
7	Discussão	71
7.1	Mapas biológicos: Proteoma.....	71
7.2	Transcriptoma	73
7.3	Análise comparativa	74
7.4	Redes biológicas	75
8	Conclusão	75

9	Materiais Suplementares.....	76
10	Referências	76

1 INTRODUÇÃO

1.1 Cromossomos B

Cromossomos B equivalem a uma classe de elementos supranumerários, descritos como os primeiros “elementos egoístas” (Wilson, 1907) e definidos por Jones e Rees (1982) como cromossomos não homólogos aos cromossomos normais do complemento (denominados de A), os quais não apresentam capacidade de recombinação com os mesmos. Inicialmente foram considerados elementos dispensáveis, sendo classificados como não funcionais por não conterem genes essenciais e, conseqüentemente, por apresentar seu próprio caminho evolutivo (Beukeboom, 1994).

Aproximadamente, 2.828 espécies de táxons diferentes apresentam cromossomos B (D’Ambrosi *et al.*, 2017), os quais estão presentes em apenas uma parcela da população, podendo variar em tamanho e número e alterar o cariótipo do hospedeiro (*Picea glauca* número diploide $2n = 24$ com indivíduos podendo apresentar de 0 a 6 cromossomos Bs ($2n = 24 \text{ As} + 0 - 6 \text{ Bs}$); *Vulpes vulpes* $2n = 43 \text{ As} + 0 - 8 \text{ Bs}$; *Rattus rattus* $2n = 42 \text{ As} + 0 - 5 \text{ Bs}$) (Miao *et al.*, 1991; Carlson, 2009; Jones *et al.*, 2008; Ploskaya-Chaibi *et al.*, 2015). Os cromossomos B, em relação à sua composição genômica, em geral são heterocromáticos, compostos basicamente por elementos repetitivos de DNA (Loreto *et al.*, 2008; Fantinatti *et al.*, 2011; López-León *et al.*, 2008; Teruel *et al.*, 2010; Nur *et al.*, 1988; Jamilena *et al.*, 1994; Gutknecht *et al.*, 1995; Puertas, 2002), seqüências de DNA de organela (Martis *et al.*, 2012), pseudogenes (Banaei-Moghaddam *et al.*, 2014) e genes de cópia única (Miao *et al.*, 1991; Graphodatsky *et al.*, 2005; Rajcic *et al.*, 2015; Martis *et al.*, 2012; Valente *et al.*, 2014; Navarro-Domínguez *et al.*, 2017a; Clark *et al.*, 2017).

Em geral, não seguem a segregação mendeliana, apresentando descendentes em proporções diferentes de 0,5 (presença/ausência de B, B+ / B-). Portanto, o sucesso de sua manutenção depende de uma transmissão preferencial por um mecanismo de *drive* durante a divisão celular. O *drive* pode ocorrer por meio de dois tipos de mecanismos: i) envolve a não-disjunção dos cromossomos homólogos ou cromátides irmãs (mitose e meiose) e pela segregação preferencial (mitose e meiose) (Clark *et al.*, 2017; Jones, 2018). Outro ponto interessante é que o cromossomo B pode apresentar importante papel biológico na origem de cromossomos sexuais, determinação sexual e possível envolvimento no controle do ciclo celular (revisado em Martins *et al.*, 2011; Yoshida *et*

al., 2011; Valente *et al.*, 2014; Helleu *et al.*, 2016; Clark *et al.*, 2017), reforçando a relevância de sua manutenção dentro da população.

Em peixes, a presença de cromossomos B tem sido relatada em distintas populações de diversas espécies, sendo variáveis em tamanho, frequência e composição. Entretanto, os estudos estão, basicamente, restritos a análises citogenéticas convencionais envolvendo descrição da presença e frequência deste polimorfismo e, em alguns casos, sua natureza eucromática/heterocromática. Em relação à análise de cromossomos B em representantes da família Cichlidae, os mesmos foram descritos em sete espécies pertencentes à subfamília sul-americana (Cichlinae) (Feldeberg & Bertollo, 1984; Martins-Santos *et al.*, 1995; Feldberg *et al.*, 2004) e por volta de 20 espécies da subfamília africana Pseudocrenilabrinae, incluindo a espécie *Astatotilapia latifasciata* (Poletto *et al.*, 2010b; Fantinatti *et al.*, 2011; Yoshida *et al.*, 2011). Análises mais detalhadas realizadas por Yoshida e colaboradores (2011) no ciclídeo *Lithochormis rubripinnis* oriundo do Lago Victoria, Leste da África, sugeriram que nesta espécie, a presença do cromossomo B pode ter um efeito na determinação do sexo, influenciando o desenvolvimento de fêmeas. Entretanto, em outras doze espécies desse mesmo lago, incluindo a espécie-modelo desse projeto, *A. latifasciata* (Poletto *et al.*, 2010a, 2010b; Yoshida *et al.*, 2011; Kuroiwa *et al.*, 2014), ambos os sexos podem apresentar o cromossomo B. Recentemente, foi identificado a presença do cromossomo B somente em fêmeas de sete espécies de ciclídeos do Lago Malawi, Leste da África (*M. lombardoi*, *Metriaclima zebra* “Boadzul,” *M. zebra* “Nkhata Bay”, *Metriaclima greshakei*, *Metriaclima mbenji*, *Labeotropheus trewavasae*, *Melanochromis auratus*) (Clark *et al.*, 2017).

1.2 Origem do cromossomo B

Em relação à sua origem, o cromossomo B não possui uma história comum entre espécies, apresentando variações em sua estrutura e propriedades da cromatina (Houben, 2017). Zhou e colaboradores (2012) sugerem que a origem do cromossomo B em *Drosophila albomicans* ocorreu pela fusão entre o cromossomo 3 e um cromossomo sexual ancestral. Martins e colaboradores (2012) propuseram que a origem do cromossomo B em *Secale cereale* advém da translocação e segregação desigual após segmentação ou duplicação gênica da espécie. Já no ciclídeo *A. latifasciata*, a duplicação de um segmento do cromossomo A (incluindo o centrômero) deu origem a um proto-

cromossomo (proto-B). Esse proto-B foi submetido a eventos de duplicação, formando um isocromossomo, e, em seguida, houve um enriquecimento de sequências provenientes do complemento A (Figura 1) (Valente *et al.*, 2014).

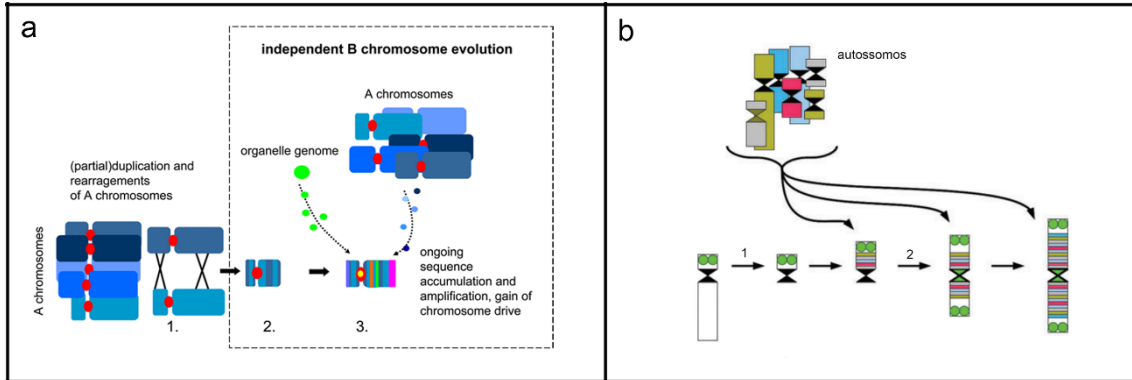


Figura 1. Hipótese da origem do cromossomo B. Em (a) o surgimento de um proto-B após a recombinação e um enriquecimento por sequências de DNA de organelas e DNAs repetitivos do genoma A (Houben *et al.*, 2013). Em (b) um resquício de cromossomo (com centrômero) duplicado origina a um proto-B. Esse proto-B foi submetido a eventos de duplicação formando um isocromossomo que sofreu enriquecimento de sequências provenientes do complemento A, originando o cromossomo B (Valente *et al.*, 2014).

1.3 *Astatotilapia latifasciata* (Perciformes, Cichlidae) como modelo para análises de cromossomos B

A espécie *Astatotilapia latifasciata* (Figura 2) pertence à família Cichlidae (Perciformes), que, por sua vez, é uma das famílias de peixes com maior número de espécies (Nelson, 2006), com cerca de 3.000 representantes distribuídos pela América Central e do Sul, Madagascar, Sudeste da Índia e África (Kocher, 2004), sendo a maior diversidade ocorrente nos grandes lagos africanos (Trewavas, 1983). Este grupo de peixes apresenta uma rápida radiação adaptativa nos grandes lagos do leste da África, onde quase 2.000 espécies diversificaram-se nos últimos 10 milhões de anos (Kocher, 2004). Devido a essas características biológicas mencionadas, os ciclídeos são considerados excelentes modelos experimentais para estudos em diversas áreas da genética (Chang *et al.*, 2005; Ijiri *et al.*, 2007; Wang *et al.*, 2007; Cnaani *et al.*, 2008; Ferreira e Martins, 2008; Valente *et al.*, 2009; Poletto *et al.*, 2010a,b; Brawand *et al.*, 2014;).



Figura 2. Ciclídeo africano *Astatotilapia latifasciata*. Nessa espécie é possível observar o dimorfismo sexual, os machos apresentam uma coloração amarelada e vermelha e as fêmeas são acinzentadas. Parte inferior da foto um exemplar macho e na porção superior uma fêmea (foto: Adachi).

Estudos estruturais e funcionais neste grupo de peixes tem se mostrado promissores em análises de diversificação de espécies, diferenciação sexual, mecanismos evolutivos relacionados diretamente aos seus genomas e cariótipos e análises de processos metabólicos. O sequenciamento completo dos genomas e transcriptomas das espécies de ciclídeos (Brawand *et al.*, 2014) estão disponíveis no BouillaBase (bouillabase.org) e SaciBase (sacibase.ibb.unesp.br), o que possibilitará abrir novas fronteiras ao conhecimento estrutural, funcional e evolutivo neste grupo.

Além disso, *A. latifasciata* apresenta de 0 a 2 cromossomos B heterocromáticos, com tamanho relativamente grande, similar ao primeiro par cromossômico do complemento A (Figura 3). Estes cromossomos extras não seguem a segregação mendeliana e apresentam sequências de DNAs repetitivos, DNAs ribossomais e genes de cópia única (Poletto *et al.*, 2010b; Fantinatti *et al.*, 2011; Valente *et al.*, 2014). Recentemente, Carmello e colaboradores (2017) demonstraram a expressão diferencial de uma variante do gene hnRNP Q-like em fêmeas portadoras do cromossomo extra nesta espécie, evidenciando a relação entre cromossomo B e determinação sexual nos ciclídeos,

o que pode estar associado ao *drive* meiótico em fêmeas (Ramos *et al.*, 2016; Clark *et al.*, 2017). Adicionalmente, Venturelli e colaboradores (em preparação) analisaram cruzamentos direcionados de indivíduos com Bs (B+) e sem Bs (B-) de *A. latifasciata* e observaram uma frequência de aproximadamente 75% de proles B+ do cruzamento entre fêmeas B+ e machos B- e B+, evidenciando uma possível influência desse cromossomo em fêmeas (Yoshida *et al.*, 2011; Clark *et al.*, 2017).

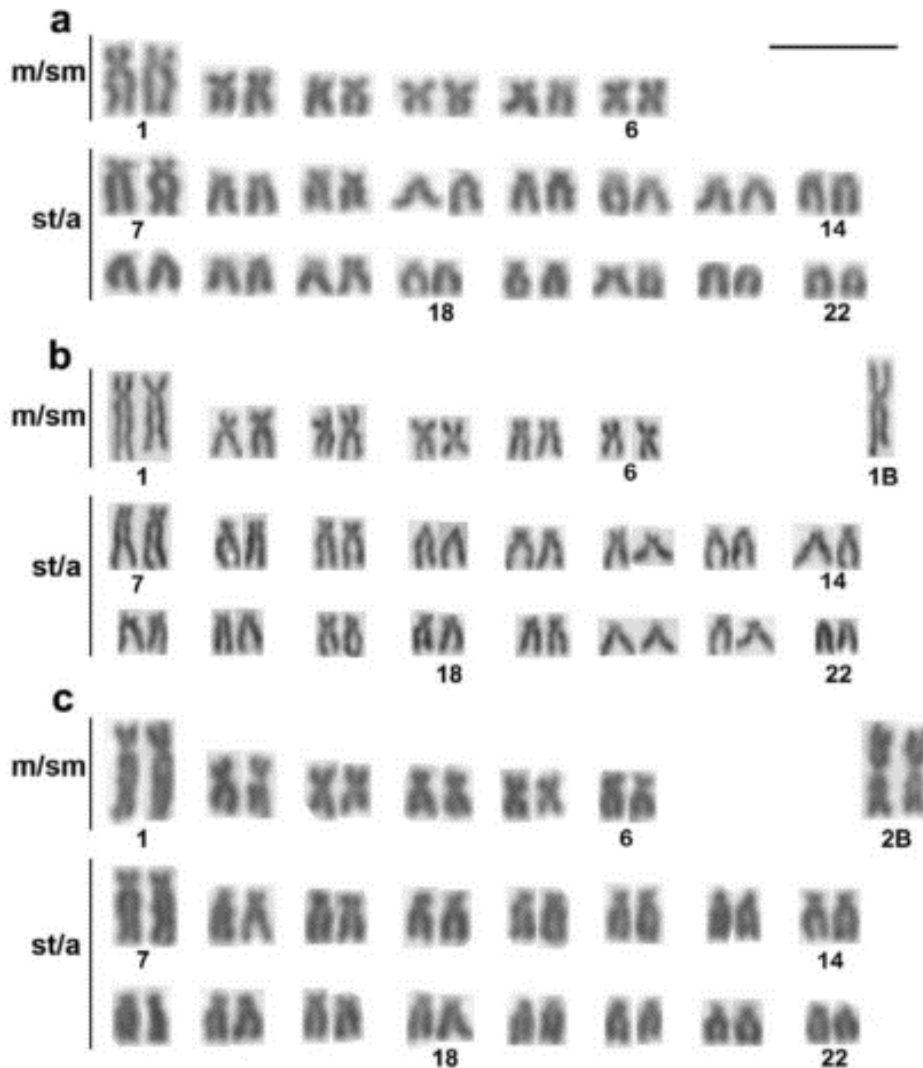


Figura 3. Cariótipo do ciclídeo *Astatotilapia latifasciata*. Cariótipo de indivíduos sem cromossomo B (a), com um (b) e dois (c) cromossomos supranumerários. Coloração em giemsa. (Poletto *et al.*, 2010).

1.4 Análises genômicas em larga escala de cromossomos B em *Astatotilapia latifasciata*

Junto a estudos citogenéticos, abordagens baseadas em análises genômicas em larga escala estão sendo utilizadas para elucidar questões acerca dos cromossomos

supranumerários. Valente e colaboradores (2014) analisaram os genomas de *A. latifasciata* com a presença e ausência do cromossomo B (B+ e B-), utilizando sequenciamento em larga escala. Neste estudo foram realizados procedimentos para identificação de regiões diferencialmente cobertas entre os dois genomas, permitindo a identificação de regiões “enriquecidas” nos genomas B+ e, possivelmente, presentes no cromossomo B.

Embora a maioria dos genes/sequências presentes no B sejam fragmentados, um pequeno número deles apresenta alta integridade. Entre os genes com alta integridade detectados, destaca-se um conjunto de genes envolvidos com controle da divisão celular. Assim, acredita-se que a presença de genes envolvidos com ciclo celular (Tabela 1) venha desempenhando papel importante na transmissão e manutenção do cromossomo B em *A. latifasciata* (Valente *et al.*, 2014).

Da mesma forma, o transcriptoma de *A. latifasciata* mostrou a presença de genes envolvidos com controle da divisão celular sendo transcritos pelo cromossomo B, o que também foi encontrado para o ciclídeo *Pundamilia nyererei* (Valente *et al.* 2014). As análises sugerem um possível papel funcional destes genes na manutenção dos cromossomos B nas células. Dessa forma, buscas realizadas em transcriptomas de *P. nyererei* (disponíveis em Bouillabase.org) revelaram que esta espécie contém transcritos B-específicos para os genes Separina, TUBB1 e KIF11 em diversos tecidos (Tabela 1) (Valente *et al.*, 2014). A partir destas informações, torna-se evidente que a presença de cromossomos B deve contribuir com novos transcritos ou mesmo alterar os padrões de expressão geral da célula.

Tabela 1. Genes envolvidos com ciclo celular e presentes no cromossomo B, de acordo com Valente *et al.* (2014). Os genes destacados em negrito apresentam evidências de atividade transcricional.

Nível de integridade (%)	No. de Acesso GenBank	Descrição gênica
70.93	XM_004570740.1	DNA repair protein XRCC2-like, mRNA
70.97	XM_004570733.1	Synaptonemal complex protein (SYCP) 2-like, mRNA
73.92	XM_004575607.1	Centromere-associated protein (CENP) E-like, mRNA
75.71	XM_004545355.1	Separin-like, mRNA

77.16	XM_004547754.1	Aurora kinase (AURK) A-B-like, mRNA
96.14	XM_004570623.1	Kinesin-like protein KIF11-like , mRNA
97.50	XM_004572890.1	Tubulin beta-5 (TUBB5) chain-like, mRNA
100	XM_004547671.1	Tubulin beta-1 (TUBB1) chain-like, mRNA
100	XM_004556260.1	Spindle and kinetochore-associated (SKA) protein 1-like, mRNA
100	XM_004576295.1	Regulator of telomere elongation helicase (RTEL) 1-like, mRNA

Legenda: Primeira coluna apresenta o nível de integridade dos genes encontrados. Segunda coluna, os números de acessos dos genes encontrados de acordo com o *GenBank*. Terceira coluna, a descrição dos respectivos genes. Os genes destacados em negrito apresentam evidências de atividade transcricional.

Para os miRNAs, com base nas mesmas amostras processadas para transcriptomas de *A. latifasciata*, foram construídas bibliotecas de microRNAs a fim de avaliar o perfil de expressão dos mesmos (Fantinatti, 2015). Em uma visão geral dos perfis de expressão de microRNAs, com base no banco de dados mirbase (<http://www.mirbase.org>), foi constatado que um total de 162 microRNAs apresentaram um padrão diferenciado entre as várias amostras considerando o sexo e a presença e ausência do cromossomo B.

1.5 Análises funcionais: Proteômica

Como mencionado acima, o avanço nas tecnologias de análise de DNA e RNA propiciou um melhor entendimento da composição genômica e expressão gênica nos organismos. Entretanto, estas informações ainda não são suficientes para elucidar precisamente os níveis de abundância proteica nas diferentes condições celulares, como observado por Vogel e Marcotte (2012), que demonstraram uma baixa correlação entre abundância de RNAm e proteínas em ratos ($R^2=0,4$). Além disso, nem sempre os bancos de dados fornecem informações quantitativas e qualitativas que possam ser relacionadas às proteínas. Tais questões podem ser mais bem estudadas utilizando-se a proteômica como abordagem. Sua proposta é a de analisar de forma global o conjunto de proteínas expressas em uma célula, tecido ou organismo (Wasinger *et al.*, 1995), permitindo avaliar quantitativa e qualitativamente as proteínas que atuam no metabolismo celular (Park, 2004; Chen e Harmon, 2006). Proteínas são determinantes na função celular e sua dinâmica de interações pode revelar muito mais sobre eventos célula/tecido específico que uma sequência genômica estática (Blakeley *et al.*, 2010).

Estudos proteômicos frequentemente utilizam técnicas de eletroforese bidimensional e de cromatografia líquida acopladas à espectrometria de massas,

permitindo a geração de dados em larga escala, como observado na comparação entre plasma seminal e sanguíneo de carpa (Dietrich *et al.*, 2014) ou no estudo do plasma humano (Jiang *et al.*, 2013). A proteômica permite estudar a estrutura, função e o controle dos sistemas biológicos pelas propriedades das proteínas, incluindo a sequência (identidade), abundância, atividade e estrutura das proteínas expressas por uma célula, assim como as modificações, interações e translocações sofridas por elas (Wilkins *et al.*, 1997). Logo o proteoma pode se modificar em resposta às condições e estímulos a que o organismo está exposto, além de características intrínsecas como presença ou ausência de componentes celulares.

Nesse projeto foi analisado pela primeira vez o perfil de abundância das proteínas dos indivíduos com presença e ausência do cromossomo B, visando identificar possíveis proteínas relacionadas ao cromossomo B em tecidos específicos de machos e fêmeas do ciclídeo *Astatotilapia latifasciata*. Em seguida, os dados foram integrados com os estudos de transcriptoma.

1.6 Integração de dados

O estudo genômico funcional se concentra nos aspectos dinâmicos, desde a transcrição de um gene, tradução e interações entre proteínas. Portanto, com dados de RNA-Seq (sequenciamento) e proteoma é possível integrar redes biológicas de interação entre proteínas, perfil de expressão de RNAs (incluindo miRNAs) e vias metabólicas em um contexto de biologia sistêmica (Rhodes *et al.*, 2005; Junker, 2008). Estas análises integrativas, a partir da construção de grafos de co-expressão dos dados do transcriptoma (dados brutos, Marques, 2016) e dos dados gerados desse trabalho, permitirão avanços na elucidação do papel biológico dos cromossomos B.

1.6.1 Redes biológicas

Conceitualmente, qualquer sistema complexo é constituído por um grande número de entidades (partículas ativas ou agentes) que podem interagir mutuamente e com o ambiente exterior (Bellomo e Carbonaro, 2011). Essa maquinaria complexa é realizada pelas interações entre as várias unidades que compõem o sistema vivo, tais como DNA, RNA, proteínas e pequenas moléculas, em conjugação com as informações vindas do meio externo. Essas interações podem ser representadas pela teoria de grafos ("redes").

São essas redes, portanto, que representam a maquinaria complexa que executa as instruções contidas nos genomas transformando-as em atividades biológicas pelas quais a vida se expressa. As redes são representadas por dois conjuntos: $G = \{V, A\}$, onde V é um conjunto de vértices e A um conjunto de pares formados por arestas que conectam cada dois vértices. Se os pares forem ordenados, a rede é classificada como direcionada, caso contrário é denominado de não-direcionada (Figura 4). As interações entre proteínas, por exemplo, por não importar a direção da interação entre elas são geralmente reportadas como grafos não-direcionados (Junker, 2008). As redes de regulação como o controle da expressão gênica, depende dos fatores que modulam a expressão do gene de interesse, no qual a ordem das interações é relevante para o resultado final. Nesses modelos, a rede é demonstrada por um grafo direcionado. As redes de transdução de sinais são representadas por grafos direcionados multi-arestas que representam as interações uni ou bi-direcionais entre proteínas, outras moléculas orgânicas ou inorgânicas (Junker, 2008; Pavlopoulos *et al.*, 2011).

Para obtermos respostas dessas redes complexas, como por exemplo, quais são os genes mais “importantes”, determinamos medidas que nos permitem caracterizar os vértices dessas redes, que são as topologias das redes complexas. As principais topologias de uma rede são: i) Grau (“*Degree*”) (o número de vizinhos do vértice v no caso de grafos direcionados, medindo o número de ligações que chegam e saem do vértice respectivamente), ii) Comprimento médio do caminho (“*Average path length*”)(número médio de conjuntos de nodos capazes de interligar dois nodos distintos v_i e v_j utilizando o menor número de nodos possível), iii) Coeficiente de agrupamento (“*Clustering coefficient*”) (mede o grau de conexões próxima a um determinado vértice), iv) Assortatividade (“*Assortativity*”) (representa a extensão dos nós de uma rede e como se associam a outros nós na rede, redes biológicas tendem a ser disassortativas, valores negativos para assortividade), v) Grau de intermediação (“*Betweenness centrality*”) (mede quanto um vértice está nos menores caminhos entre outros vértices na rede) (Albert e Barabási, 2002; Foster *et al.*, 2010).

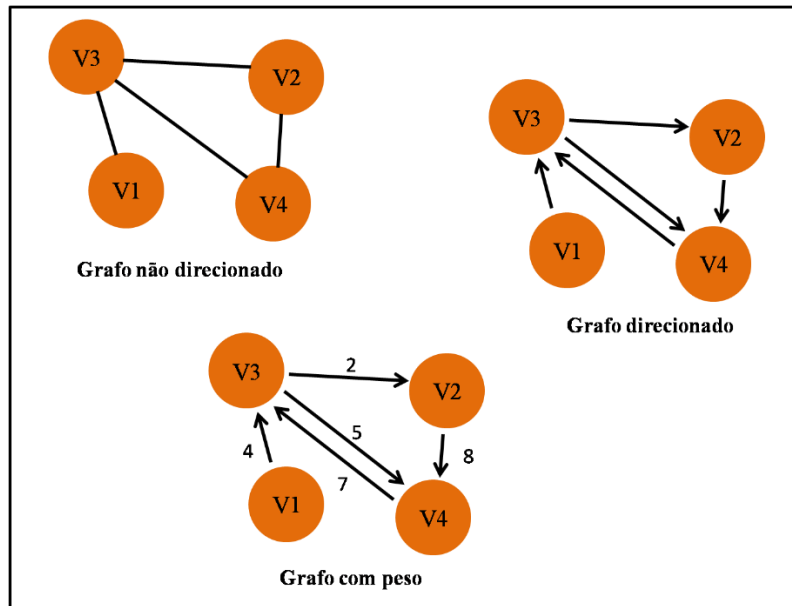


Figura 4. Exemplo de grafos direcionados, não direcionados e com peso. Figura baseada no trabalho de Pavlopoulos *et al.*, (2011).

1.7 Ciclo celular

O ciclo celular é um processo fundamental dos seres vivos multicelulares. Apresenta grande importância desde o desenvolvimento até a vida adulta, contribuem para a geração de células especializadas, que são necessárias para formar tecidos e desempenhar funções específicas. O controle do ciclo celular precisa ser intimamente preciso para não acarretar erros no desenvolvimento e surgimento de tumores. Nos eucariotos, esse processo está basicamente dividido em duas fases: Interfase e Mitose, com a presença de dois processos, entre elas G1 e G2. Nessas duas sub fases as células obtêm massa, integram sinais de crescimento, organizam a replicação gênica e preparam a segregação dos cromossomos (Barnun e O’Connell, 2014).

Durante o ciclo celular ocorrem vários processos regulatórios e de *checkpoint* para garantir o sucesso da divisão celular. Existem vários mecanismos de *checkpoint* já descritos como o crescimento celular adequado, resposta ao dano no DNA, monitoramento da replicação do DNA, dependência da fase S-M, e o *checkpoint* do fuso mitótico (SAC). Além disso, o ciclo celular é intimamente regulado, principalmente, por ciclinas dependentes de quinase (CDKs) e outras proteínas com grande importância como por exemplo: as proteínas MPS, a qual é uma das principais quinases envolvidas na localização do cinetocoro e no ponto de verificação do conjunto do fuso (SAC), que inibe

a transição da metáfase para anáfase até que as fibras do fuso estejam corretamente ligadas a todos os cromossomos, a aurora quinase A, a qual regula a ligação dos fusos mitóticos, mantendo a bipolaridade e também está relacionada com a maturação do centrômero; a proteína APC/C atua no *checkpoint* mitótico realizando a ubiquitinação da Securina liberando a Separase para separação das cromátides irmãs; o complexo NDC80 formado por um conjunto de quatro proteínas (Ndc80, Nuf2, Spc25 e Spc24), as quais são responsáveis pela ligação do cinetócoro ao microtubulos durante a divisão celular; e as proteínas da família 2.4.1.1, que atuam como estacionadores, ou seja, param o ciclo em diversos pontos da interfase e metáfase (Kaistha *et al.*, 2014; Ducat e Zheng, 2004; Hermeking e Benzinger, 2006; Ciferri *et al.*, 2007; Bolanos-Garcia e Blundell, 2011; Ruijtenberg e Heuvel, 2016).

Vários estudos referentes ao cromossomo B vem demonstrando a presença de genes envolvidos no ciclo celular (Valente *et al.* 2014; Helleu *et al.*, 2016; Clark *et al.*, 2017; Navarro-Domínguez *et al.* 2017a,b; Kichigin *et al.* 2019; Makunin *et al.* 2018; Jones, 2018). Makunin e colaboradores (2018) identificaram o gene *centromere protein N* na raposa *Vulpes vulpes L.*, Navarro-Domingues e colaboradores (2017b) encontraram genes ativos relacionados a divisão celular para os genes *CIP2A protein(cip2a)*, *cytoskeleton-associated protein 2(ckap2)*, *condensin I complex subunit G(cap-g)*, *kinesin-like protein kif20a (kif20a)* e *E3 ubiquitin-protein ligase mycbp2 (mycb2)* no gafanhoto *Eyprepocnemis plorans*, Kichigin e colaboradores (2019) encontraram genes relacionados a segregação cromossômica e divisão celular (*inner centromere protein (incep)* e *protein spire homolog 2 (spire2)*, respectivamente) no lagarto *Anolis carolinensis*. Portanto, há indícios que a manutenção do cromossomo B, em diferentes espécies, pode estar relacionadas com a presença de genes reguladores do ciclo celular.

2 HIPÓTESE

A presença do cromossomo B influencia a regulação gênica e interfere no perfil proteico no ciclídeo africano *Astatotilapia latifasciata*.

Em relação às redes biológicas, as sub redes representaram ser redes biológicas reais de acordo com as topologias de distribuição de *degree*, disassortividade, e alta conectividade. As redes demonstraram novamente a importância de genes relacionados ao ciclo celular, evidenciando ainda mais a relação entre o controle do ciclo celular e a presença do cromossomo B.

De modo geral, os resultados obtidos a partir das análises de bioinformática mostraram que o cromossomo B além de apresentar elementos repetitivos, pseudogenes e genes em sua composição, podem ainda alterar a expressão gênica e abundância proteica, apresentando evidências funcionais relacionadas ao cromossomo B na espécie de ciclídeo *A. latifasciata*.

Análises futuras relacionadas aos genes encontrados nesse estudo como os genes *daaurora kinase*, *kinetochoro nuf2 protein*, *G2/mitotic-specific cyclin-B1*, *dual-specificity protein kinase ttk* e *anaphase-promoting complex subunit 1* que participam do processo de formação do *spindle assembly* e *spindle assembly checkpoint* (SAC) durante a segregação cromossômica, a enzima 14-3-3 que se ligam a proteínas reguladoras e modulam sua função na transição das fases G1/S e G2/M e a *tubulin beta 1* que é um dos maiores constituintes dos microtubulos que atuam na segregação cromossômica, abrem novos caminhos a serem estudadas sobre a funcionalidade do cromossomo B.

9 MATERIAIS SUPLEMENTARES

Para acessar os materiais suplementares segue o link [Materiais Suplementares](#).

10 REFERÊNCIAS

Acquaviva CJ, Pines J (2006) The anaphase-promoting complex/cyclosome: APC/C. *J Cell Sci* 2006 119: 2401-2404.

Albert R, Barabási AL (2012) Statistical mechanics of complex networks. *Rev Mod Phys* 74, 47.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 403-410.

Anders S, Paul Theodor Pyl, Wolfgang Huber (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31 (2): 166-169.

Banaei-Moghaddam AM, Meier K, Karimi-Ashtiyani R, Houben A (2013) Formation and Expression of Pseudogenes on the B chromosome of Rye. *Plant Cell*, 25: 2536-2544.

Barnum KJ, O'Connell MJ (2014) Cell cycle regulation by checkpoints. *Methods Mol Biol*, 1170:29-40.

Bellomo N, Carbonaro B (2011) Toward a mathematical theory of living systems focusing on developmental biology and evolution: a review and perspectives. *Phys Life Rev*, 8(1):1-18.

Beukeboom LW (1994) Bewildering B: An impression of the 1st B–chromosome conference. *J Hered* 73:328–336.

Blakeley P, Siepen JA, Lawless C, Hubbard SJ (2010) Investigating protein isoforms via proteomics: a feasibility study. *Proteomics*, 10(6):1127–1140.

Bolanos-Garcia VM & Blundell TL (2011) BUB1 and BUBR1: multifaceted kinases of the cell cycle. *Trends in Biochemical Sci*, 36(3), 141–150.

Bolger AM, Lohse M, Usadel, B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, *et al.*, (2014) The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513: 375–381.

Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, Lee TJ, Leigh ND, Kuo TH, Davis FG, Bateman J, Bryant S, Guzikowski AR, Tsai SL, Coyne S, Ye WW, Freeman RM Jr, Peshkin L, Tabin CJ, Regev A, Haas BJ, Whited JL (2017) A Tissue-30 Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors. *Cell Rep*; 18(3):762-776.

Burrows F, Zhang H, Kamal A (2004) Hsp90 Activation and Cell Cycle Regulation, *Cell Cycle*, 3:12, 1530-1536.

Cann KL, Hicks GG (2007) Regulation of the cellular DNA double-strand break response. *Biochem. Cell Biol.* 85: 663-674.

Carmello BO, Coan RLB, Cardoso AL, Ramos E, Fantinatti BEA, Marques DF, Oliveira RA, Valente GT, Martins C. (2017) The hnRNP Q-like gene is retroinserted into the B chromosomes of the cichlid fish *Astatotilapia latifasciata*. *Chromosome Res* 25(3-4):277-290.

Chang X, Kobayashi T, Senthilkumaran B, Kobayashi-Kajura H, Sudhakumari CC, Nagahama Y (2005) Two types of aromatase with different encoding genes, tissue distribution and developmental expression in Nile tilapia (*Oreochromis niloticus*). *Gen. Comp Endocrinol* 141: 101–115.

Chen S, Harmon AC (2006) Advances in plant proteomics. *Proteomics*, 6(20):5504-16.

Ciferri C, Musacchio A, and Petrovic A. (2007) The Ndc80 complex: hub of kinetochore activity. *FEBS Lett*, 581, 2862–2869.

Clark FE, Conte MA, Ferreira-Bravo IA, Poletto AB, Martins C, Kocher TD. (2017) Dynamic Sequence Evolution of a Sex-Associated B Chromosome in Lake Malawi Cichlid Fish. *J Hered*, Volume 108, Issue 1, Pages 53–62.

Cnaani A, Lee B-Y, Zilberman N, Ozouf-Costaz C, Hulata G, Ron M, D'Hont A, Baroiller J-F, D'Cotta H, Penman DJ, Tomasino E, Coutanceau J-P, Pepey E, Shirak A, Kocher TD (2008) Genetics of Sex Determination in Tilapiine Species. *Sexual Dev* 2: 43-54.

Craig KL, Tyers M. (1999) The F-box: a new motif for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction. *Prog Biophys Mol Biol*, Volume 72, Issue 3.

D'Ambrosio U, Alonso-Lifante MP, Barros K, Kovarik A, Mas de Xaxars G, Garcia S. (2017) B-chrom: A database on B-chromosomes of plants, animals and fungi. *New Phytol*, 216, 635–642.

Dietrich MA, Adamek M, Hejmej A, Ciereszko A (2014) Characterization, expression and antibacterial properties of apolipoproteins A from carp (*Cyprinus carpio* L.) seminal plasma. *Fish Shellfish Immunol*, 41(2):389-401.

Ducat D, Zheng Y. (2004) Aurora kinases in spindle assembly and chromosome segregation. *Exp Cell Res*, Volume 301, Issue 1:60-67.

Fantinatti BE, Martins C (2016) Development of chromosomal markers based on next-generation sequencing: the B chromosome of the cichlid fish *Astatotilapia latifasciata* as a model. *BMC Genet*, 17(1):119.

Fantinatti BE, Mazzuchelli J, Valente GT, Cabral-de-Mello DC, Martins C (2011) Genomic content and new insights on the origin of the B chromosome of the cichlid fish *Astatotilapia latifasciata*. *Genetica*, 139(10):1273-82.

Farshori PQ, Shah BH, Arora KK, Martinez-Fuentes A, Catt KJ (2003) Activation and nuclear translocation of PKCdelta, Pyk2 and ERK1/2 by gonadotropin releasing hormone in HEK293. cells. *J Steroid Biochem Mol Biol* 85:337-47.

Feldberg E, Porto JIR, Alves-Brinn MN, Mendonça MNC, Benzaquem DC (2004) B Chromosomes in Amazonian cichlid species. *Cytogenet Genome Res*, 106: 195-198.

Feldberg E, Bertollo LAC (1984). Discordance in chromosome number among somatic and gonadal tissue cells of *Gymnogeophagus balzanii* (Pisces, Cichlidae). *Rev. Brasil. Genet.* 7: 639–645.

Ferreira IA, Martins C (2008) Physical chromosome mapping of repetitive DNA sequences in Nile tilapia *Oreochromis niloticus*: evidences for a differential distribution of repetitive elements in the sex chromosomes. *Micron* 39: 411-418.

Finn RD, Clements J, Eddy SR (2011) HMMER Web Server: interactive sequence similarity searching. *Nucleic Acids Res* 39: W29–W37.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) The Pfam protein families database. *Nucleic Acids Res* 42: D222–D230.

Firat-Karalar EN, Stearns T (2014) The centriole duplication cycle. *Philos Trans R Soc Lond B Biol Sci*, 369(1650):20130460.

Foster J, Foster DV, Grassberger P, Paczuski M (2010) "Edge direction and the structure of networks". *PNAS*, **107** (24): 2.

Graphodatsky AS, Kukekova AV, Yudkin DV et al (2005) The proto-oncogene CKIT maps to canid B-chromosomes. *Chromosome Res*, 13: 113-122.

Gutknecht J, Sperlich D, Bachmann L (1995) A species specific satellite DNA family of *Drosophila subsilvestris* appearing predominantly in B chromosomes. *Chromosoma* 103: 539-544.

He B, Zhao S, Chen Y, Cao Q, Wei C, Cheng X, Zhang Y (2015) Optimal assembly strategies of transcriptome related to ploidies of eukaryotic organisms. *BMC Genomics*, 16:65.

Hermeking H, Benzinger A (2006) 14-3-3 proteins in cell cycle regulation. *Semin Cancer Biol*, Volume 16, Issue 3:183-192, ISSN 1044-579X.

Hotton J, Agopiantz M, Leroux A, Charra-Brunaud C, Marie B, Busby-Venner H, Morel O, Guéant JL, Vignaud JM, Battaglia-Hsu SF, Gauchotte G (2018) Minichromosome maintenance complex component 6 (MCM6) expression correlates with histological grade and survival in endometrioid endometrial adenocarcinoma. *Virchows Arch*, 472:623-633.

Houben A (2017) B Chromosomes - A Matter of Chromosome Drive. *Front Plant Sci*, 8:210.

Huang X, Madan A (1999) CAP3: A DNA Sequence Assembly Program. *Genome Research*. 9:868–877.

Huang H, Gao Q, Peng X, et al. (2011) piRNA-associated germline nuage formation and spermatogenesis require MitoPLD profusogenic mitochondrial-surface lipid signaling. *Dev Cell*, 20(3):376-87.

Ijiri S, Kaneko H, Kobayashi T, Wang D-D, Sakai F, Paul-Prasanth B, Nakamura M, Nagahama Y (2007) Sexual Dimorphic Expression of Gene in Gonads during early differentiation of a Teleost Fish, the Nile Tilapia *Oreochromis niloticus*. *Bio Reprod*, 78: 333–341.

Jamilena M, Ruiz-Rejón C, Ruiz-Rejón M (1994) A molecular análisis of the origin of the *Crepis Capillaris* B chromosome. *J Cell Sci*, 107: 703-708.

Jiang L, et al. (2013). LC-MS/MS identification of doublecortin as abundant beta cell-selective protein discharged by damaged beta cells in vitro. *J Proteomics*, 80:268-80.

Jones RN, Rees H (1982) B chromosome. New York: Academic Press.

Jones RN, Viegas W, Houben A (2008) A century of B chromosomes in plants: so what?. *Ann Bot*, 101(6):767–775.

Jones RN (2018) Transmission and Drive Involving Parasitic B Chromosomes. *Genes (Basel)*. 9(8):388.

Junker B.H. (2008). Networks in Biology. In: Analysis of Biological Networks. Junker B.H., Schreiber F. (eds). John Wiley & Sons. p. 3-12.

Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J (2005) Repbase update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res*, 110:462–467.

Kaistha BP, Honstein T, Müller V, Bielak S, Sauer M, Kreider, R, ... Buchholz M (2014) Key role of dual specificity kinase TTK in proliferation and survival of pancreatic cancer cells. *Br J Cancer*, 111(9), 1780–1787.

Keam SP, Hutvagner G (2015) tRNA-Derived Fragments (tRFs): Emerging New Roles for an Ancient RNA in the Regulation of Gene Expression. *Life*, 5(4), 1638–1651.

Kichigin IG, Lisachov AP, Giovannotti M. *et al.* (2019) First report on B chromosome content in a reptilian species: the case of *Anolis carolinensis*. *Mol Genet Genomics* 294: 13.

Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nat Rev Genet*, 5: 288-98.

Kuroiwa A, Terai Y, Kobayashi N, Yoshida K, Suzuki M, Nakanishi A, Matsuda Y, Watanabe M, Okada N (2014) Construction of chromosome markers from the Lake Victoria cichlid *Paralabidochromis chilotis* and their application to comparative mapping. *Cytogenet Genome Res*, 142(2):112-20.

Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9:357-359.

López-León MD, Cabrero J, Dzyubenko VV, Bugrov AG, Karamysheva TV, Rubtsov NB, Camacho JPM (2008) Differences in ribosomal DNA distribution on A and B chromosomes between eastern and western populations of the grasshopper *Eyprepocnemis plorans plorans*. *Cytogenet Genome Res*, 121: 260-265.

Loreto V, Cabrero J, López-León MD, Camacho JPM, Souza MJ (2008) Possible autosomal origin of macro B chromosomes in two grasshopper species. *Chromosome Res*, 16: 233-241.

Love, M. I., Huber, W., & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 15(12), 550.

Luo W, Brouwer C (2013) Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics (Oxford, England)*; 29:1830–1.

Makunin AI, Rajičić M, Karamysheva TV, *et al.* (2018) Low-pass single-chromosome sequencing of human small supernumerary marker chromosomes (sSMCs) and Apodemus B chromosomes. *Chromosoma* 127: 301.

Marques, D.F. Functional analysis of B chromosome presence using cichlid *Astatotilapia latifasciata* as model. Available online: <http://hdl.handle.net/11449/141953> (accessed on March 2016).

Martins C, et al. (2011). Animal genomes under the focus of cytogenetics. 1. ed. Hauppauge: Nova Science Publisher, 2011. v. 1. 160.

Martins-Santos IC, Portela-Castro ALB, Julio HF Jr (1995) Chromosome analysis of 5 species of the Cichlidae family (Pisces, Perciformes) from the Paraná River. *Cytologia*, 60:223–231.

Martis MM , Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Simková H, Novák P, Neumann P, Kubaláková M, Bauer E, Haseneyer G, Fuchs J, Dolezel J, Stein N, Mayer KFX, Houben A (2012) Selfish supernumerary chromosome reveals its origins a mosaic of host genome and organellar sequences. *PNAS*. 109:13343–13346.

Nam BH, Kim YO, Kong HJ, Kim WJ, Lee SJ, Choi TJ (2009) Identification and characterization of the prepro-vasoactive intestinal peptide gene from the teleost *Paralichthys olivaceus*, 127(3-4):249-58.

Navarro-Domínguez B, Ruiz-Ruano FJ, Camacho JPM, Cabrero J, López-León MD (2017a) Transcription of a B chromosome CAP-G pseudogene does not influence normal Condensin Complex genes in a grasshopper. *Sci Rep*, 7:17650.

Navarro-Domínguez B, Ruiz-Ruano FJ, Cabrero J, Corral JM, López-León MD, Sharbel TF, Camacho JPM (2017b) Protein-coding genes in B chromosomes of the grasshopper *Eyprepocnemis plorans*. *Sci. Rep*, 7, 45200.

Nelson JS: Fishes of the world, 4th ed (John Wiley and Sons, Inc. New York 2006).

Nur U, Werren JH, Eickbush DG, Burke WD, Eickbush TH (1988) A "selfish" B chromosome that enhances its transmission by eliminating the paternal genome. *Science* 240: 512-514

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic acids research*; 27:29–34.

Pane A, Wehr K, Schüpbach T (2007) zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. *Dev cell*, 12(6), 851-62.

Park OK (2004). *Proteomic Studies in Plants*. J. Bioch. Mol. Biol, v. 37, p. 133-138.

Pavlopoulos GA, *et al.* (2011) Using graph theory to analyze biological networks. *Biodata Mining*, 4: 10.

Pawson AJ, McNeilly AS (2005) The pituitary effects of GnRH. *Anim Reprod Sci* 88(1–2):75–94.

Pines, J. (2011) Cubism and the cell cycle: the many faces of the APC/C. *Nat. Rev. Mol. Cell Biol.* 12, 427-438.

Pinhal D, Bovolenta LA, Moxon S, *et al.* (2018) Genome-wide microRNA screening in Nile tilapia reveals pervasive isomiRs' transcription, sex-biased arm switching and increasing complexity of expression throughout development. *Sci Rep*, 8(1):8248.

Poletto AB, Ferreira IA, Martins C (2010a). The B chromosome of the cichlid fish *Haplochromis obliquidens* harbors 18S rRNA genes. *BMC Genetics* 11: 1.

Poletto AB, *et al.* (2010b). Chromosome differentiation patterns during cichlid fish evolution. *BMC Genet* 11: 5-10.

Puertas MJ (2002). Nature and evolution of B chromosomes in plants: A noncoding but information-rich part of plant genomes. *Cytogenet Genome Res* 96: 198-205.

Quinlan, Aaron R., and Ira M. Hall (2010) “BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features.” *Bioinformatics* 26.6: 841–842. PMC. Web. 23 Oct. 2017.

Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, Vilo J (2016) g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic acids research*, 44(W1), W83-9.

Rice P, Longden I, Bleasby A (2000) EMBOSS: The European molecular biology open software suite. *Trends Genet* 16, 276–277.

Sansal I, Dupont E, Toru D, Rouget P (2000) NPDC-1, a regulator of neural cell proliferation and differentiation, interacts with E2F-1, reduces its binding to DNA and modulates its transcriptional activity. *Oncogene* 19, 5000-5009.

Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics Advance*.

Schulz RW, De Franca LR, Lareyre J-J, Legac F, Chiarini-Garcia H, Nobrega RH, Miura T (2010) Spermatogenesis in fish. *Gen Comp Endocrinol*, v. 165, n. 3, p. 390-411. ISSN 0016-6480.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13(11):2498-504

Smit AFA, Hubley R, Green P (2013) *RepeatMasker Open-4.0*. 2013-2015 <http://www.repeatmasker.org>.

Sun Y, Senger K, Baginski TK, *et al.* (2012) Evolutionarily conserved paired immunoglobulin-like receptor α (PILR α) domain mediates its interaction with diverse sialylated ligands. *J Biol Chem*, 287(19):15837–15850.

Supek F, Bosnjak M, Skunca N, Smuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*, 2011;6(7):e21800.

Teruel M, *et al.* (2010) B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma* 119: 217-225.

The International Cichlid Genome Consortium (2006). Genetic basis of vertebrate diversity: the cichlid fish model.

Trewavas E (1983). Tilapiine fishes of the genera *Sarotherodon*, *Oreochromis* and *Danakilia*. British Museum, Natural History, London.

Valente GT, et al. (2009). Comparative cytogenetics of cichlid fishes through genomic in-situ hybridization (GISH) with emphasis in *Oreochromis niloticus*. *Chromosome Res*, 17(6):791-799.

Valente GT, Acencio ML, Martins C (2013) The Development of a Universal *In silico* Predictor of Protein-Protein Interactions. *Plos One*, 8: e65587.

Valente GT, Conte MA, Fantinatti BEA, Cabral-de-Mello DC, Carvalho RF, Vicari MR, Kocher TD, Martins C (2014) Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Mol Biol Evol*, 31:2061–2072.

Vogel C, Marcotte EM (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics* 13, 227-232.

Wang D-S, et al. (2007). Foxl2 up-regulates aromatase gene transcription in a female-specific manner by binding to the promoter as well as interacting with ad4 binding protein/steroidogenic factor 1. *Mol Endocrinol*, 21: 712–725.

Wasinger VC, et al. (1995). Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* 16, 1090-1094.

Wilkins MR, et al. (1997). Detailed peptide characterization using PEPTIDEMASS--a World-Wide-Web accessible tool. *Electrophoresis*. 1997 Mar-Apr;18(3-4):403-8.

Wilson E (1907) The supernumerary chromosome of Hemiptera. *Science* 26:870-871.

Yoshida K, et al. (2011). B Chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. *Plos Genetics* 7(8): e1002203.

Zerbino DR, Birney E (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*, 18:821–829.

Zhou Q, et al. (2012). Deciphering neo-sex and B chromosome evolution by the draft genome of *Drosophila albomicans*. *BMC Genomics* 2012, 13:109 doi:10.1186/1471-2164-13-109.