

IGOR HENRIQUE RODRIGUES OLIVEIRA

***Psalidodon rivularis* (CHARACIFORMES: CHARACIDAE): UM COMPLEXO DE
ESPÉCIES BASEADO EM DADOS CITOGENÉTICOS, MORFOMÉTRICOS E
GENÔMICOS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós- Graduação em Manejo e Conservação de Ecossistemas Naturais e Agrários, para obtenção do título de *Magister Scientiae*.

Orientadora: Karine Frehner Kavalco

Coorientador: Rubens Pasa

Ficha catalográfica elaborada pela Biblioteca da Universidade Federal de Viçosa - Campus Florestal

T

O48p
2021

Oliveira, Igor Henrique Rodrigues, 1997-

Psalidodon rivularis (Characiformes: Characidae): : um complexo de espécies baseado em dados citogenéticos, morfométricos e genômicos / Igor Henrique Rodrigues Oliveira. – Florestal, MG, 2021.

112 f. : il. (algumas color.) ; 29 cm.

Inclui anexos.

Orientador: Karine Frehner Kavalco.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.53-61.

1. Cariomorfos. 2. Morfometria geométrica. 3. Mitogenoma. 4. Satelitoma. 5. Bacia do Alto Rio São Francisco. 6. Peixes. 7. *Psalidodon rivularis*. I. Universidade Federal de Viçosa. Instituto de Ciências Biológicas e da Saúde. Mestrado em Manejo e Conservação de Ecossistemas Naturais e Agrários. II. Título.

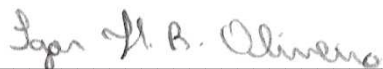
IGOR HENRIQUE RODRIGUES OLIVEIRA

***Psalidodon rivularis* (CHARACIFORMES: CHARACIDAE): UM COMPLEXO DE
ESPÉCIES BASEADO EM DADOS CITOGENÉTICOS, MORFOMÉTRICOS E
GENÔMICOS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós- Graduação em Manejo e Conservação de Ecossistemas Naturais e Agrários, para obtenção do título de *Magister Scientiae*.

APROVADA: 06 de maio de 2021.

Assentimento:



Igor Henrique Rodrigues Oliveira
Autor



Karine Frehner Kavalco
Orientadora

AGRADECIMENTOS

Em primeiro lugar à minha amada avó e madrinha Terezinha Aparecida Rodrigues. De todas as pessoas, sei que aonde você estiver, você é aquela que mais sente orgulho de mim, como sempre foi. Mesmo não estando presente fisicamente a mais tempo do que gostaria, as doces lembranças que tenho de ti são uma das minhas maiores forças.

À minha família, em especial meus pais Mauro dos Santos Oliveira e Patricia Maria Rodrigues Mathilde Oliveira, e meus irmãos Ian Henrique Rodrigues Oliveira e Iuri Henrique Rodrigues Oliveira. Mesmo estando a muitos anos morando em Rio Paranaíba, eu sempre tive um lar cheio de carinho e apoio me esperando em Embu-Guaçu e lhes agradeço muito por isso.

Aos meus orientadores Karine Frehner Kavalco e Rubens Pasa. Além de marcarem minha trajetória acadêmica e serem fundamentais na construção do cientista que hoje sou, se tornaram grandes amigos e me ajudaram em momentos críticos aos quais nunca esquecerei.

Ao professor doutor Fabiano de Bezerra Menegidio, por toda a ajuda, auxílio e orientação na minha nova caminhada no campo da genômica, ao qual devo dizer, ser muito empolgante.

Aos meus colegas de trabalho e melhores amigos, Iuri Batista da Silva e Rafael Augusto Silva Soares. Além de serem inestimáveis amigos estando ao meu lado em diversos momentos pessoais, vocês me ajudaram em pontos chave de meus trabalhos. Não consigo olhar para essa dissertação e minha trajetória acadêmica sem reconhecer a participação de vocês.

A minha namorada Priscila Martins de Assis. Eu acredito que nosso pessoal interfere muito no nosso profissional e é inegável que eu fiquei muito mais feliz desde que você chegou. Tudo ficou mais fácil quando pude contar com seu carinho, amparo e amor nesses dois anos.

As minhas melhores amigas Natália Reis de Almeida, Joyce Anna de Souza e Thamara da Silva Ramos por sempre me ampararem quando precisei. Sei que vocês sempre estarão lá por mim, por isso não posso pedir amizades melhores que vocês.

Aos meus veteranos de laboratório Matheus Lewi Cruz Bonnacorsi de Campos, Snaydia Resende Viegas, Rosana de Mesquita Alves, Renan Rodrigues Rocha, Marcos Aurélio da Silva, Francisco de Menezes Cavalcante Sassi e Dinaíza Abadia Rocha Reis. Ter uma referência ajuda muito na hora de se realizar um trabalho e quando essa referência vem de pessoas tão solícitas tudo fica ainda mais fácil.

Aos professores Rodrigo Fornel e Helder Canto Resende por terem aceitado o convite para serem banca da minha defesa de mestrado. Agradeço imensamente por suas correções que, com certeza, irão me auxiliar a aprimorar meu trabalho.

Dentre várias outras pessoas, me sinto na obrigação de citar mais alguns nomes e agradecer especialmente aos meus amigos e amigas Giovanna, Winston, Mayara e Adrielle, minha avó Nairda e meus tios Osias, Maria Alice e Márcia. Apesar de não escrever outros nomes, existem outras muitas pessoas que marcaram positivamente minha trajetória, e saibam que eu não me esqueci e também sou grato a todos vocês.

Por fim, a Universidade Federal de Viçosa e ao programa de pós-graduação em Manejo e Conservação de Ecossistemas Naturais e Agrários por fornecer toda a infraestrutura e meios para a realização deste trabalho, a Universidade de Leicester e aos professores Dr. Pat Heslop-Harrison e Dra. Trude Schwarzacher pela parceria que possibilitou a realização do capítulo de genômica deste trabalho, e as agências de fomento à pesquisa CNPq e CAPES pelas bolsas de estudo concedidas.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

A todos, o meu mais sincero e profundo muito obrigado!!!

“Aliada minha é a força, e poderosa aliada ela é, a vida a cria, crescer ela faz, é a energia que cerca-nos, e liga-nos, luminosos seres somos nós e não essa rude matéria. Você precisa a força sentir ao redor seu, sinta entre você e a árvore, a pedra, em todo lugar...”

Mestre Yoda - *Star Wars*

RESUMO

OLIVEIRA, Igor Henrique Rodrigues, M.Sc., Universidade Federal de Viçosa, maio de 2021. ***Psalidodon rivularis* (Characiformes: Characidae): Um complexo de espécies baseado em dados citogenéticos, morfométricos e genômicos.** Orientadora: Karine Frehner Kavalco. Coorientador: Rubens Pasa.

A espécie *Psalidodon rivularis* é endêmica da bacia do Alto rio São Francisco dotada de grande diversidade cariotípica, com números diplóides variando de $2n=46$ a $2n=50$ cromossomos. Trabalhos anteriores envolvendo morfometria e marcadores moleculares mitocondriais têm questionado a validade de *P. rivularis* como uma única espécie. O objetivo geral deste trabalho foi estudar a diversidade inter-populacional e inter-cariotípica de *P. rivularis* utilizando como ferramentas a citogenética, a morfometria geométrica e a genômica. Os objetivos específicos do trabalho foram avaliar a diversidade cariotípica de *P. rivularis*, avaliar a modulação da forma em diferentes populações de *P. rivularis* e uma população de *P. aff. rivularis* do Médio rio São Francisco, e comparar o mitogenoma e o satelitoma de dois citótipos ($2n=46$ e $2n=50$) de *P. rivularis*. Nós encontramos dois cariomorfos distintos de $2n=46$ cromossomos, com ocorrência de cromossomos Bs em certas populações, um cariomorfo de $2n=48$ e um de $2n=50$. Nós também encontramos diferenças nos padrões de heterocomatina constitutiva dos cariomorfos $2n=46I$, $2n=46II$ e $2n=50$, além de diferenças no enriquecimento de pares de bases AT/CG entre os cromossomos Bs dos cariomorfos de $2n=46$ cromossomos. Nós também encontramos variações na forma das populações analisadas associadas à altura corporal e à região cefálica, sendo que o recuo da placa sub-orbital marca a principal diferença entre *P. rivularis* do Alto rio São Francisco e *P. aff. rivularis* do Médio rio São Francisco. O mitogenoma dos citótipos $2n=46$ e $2n=50$ apresentou grande distância genética e o D-loop de ambos apresentou uma sequência repetitiva de 35bp que variou em 19 sítios entre eles. Além disso, o número de sequências repetitivas recuperadas no satelitoma do citótipo $2n=50$ foi bem maior que para o citótipo $2n=46$ cromossomos e ambos apresentaram sequências exclusivas. Todas as abordagens deste trabalho apontam que *P. rivularis* corresponde a um complexo de espécies crípticas com diferentes números diplóides, distribuição limitada à bacia do Alto rio São Francisco e adaptação da forma de diferentes Unidades Evolutivamente Significativas em diferentes sub-bacias hidrográficas. Além disso, os citótipos de $2n=46$ e $2n=50$ cromossomos amostrados neste trabalho representam pelo menos duas diferentes espécies crípticas encontradas em

alopatria e simpatria na sub-bacia hidrográfica do rio Abaeté, evidenciadas por dados citogenéticos e genômicos que sugerem a existência de um possível isolamento reprodutivo entre elas.

Palavras-chave: Bioinformática. Evolução alopátrica e simpátrica. Unidades Evolutivamente Significativas.

ABSTRACT

OLIVEIRA, Igor Henrique Rodrigues, M.Sc., Universidade Federal de Viçosa, May, 2021. ***Psalidodon rivularis* (Characiformes: Characidae): A species complex based on cytogenetic, morphometric and genomic data.** Adviser: Karine Frehner Kavalco. Co-adviser: Rubens Pasa.

The *Psalidodon rivularis* species is endemic to the Upper São Francisco river basin endowed with great karyotype diversity, with karyotypes ranging from $2n=46$ to $2n=50$ chromosomes. Previous work involving morphometry and mitochondrial molecular markers has questioned the validity of *P. rivularis* as a single species. The general objective of this work was to study the inter-population and inter-karyotype diversity of *P. rivularis* using cytogenetics, geometric morphometry and genomics. The specific objectives of the study were to evaluate the karyotype diversity of *P. rivularis*, to assess the shape modulation in different populations of *P. rivularis* and a population of *P. aff. rivularis* from the Middle São Francisco river, and compare the mitogenome and satellitome of two cytotypes ($2n=46$ and $2n=50$) of *P. rivularis*. We found two distinct karyomorphs of $2n=46$ chromosomes, with some cases of Bs chromosomes in certain populations, a karyomorph of $2n=48$ and one of $2n=50$. We also found differences in the constitutive heterochromatin patterns of karyomorphs $2n=46I$, $2n=46II$ and $2n=50$, in addition to differences in AT/CG base pair enrichment between Bs chromosomes of $2n=46$ chromosomes. We also found variations in the shape of the analyzed populations associated with body height and the cephalic region, with the retreat of the suborbital plate marking the main difference between *P. rivularis* of the Upper São Francisco river and *P. aff. rivularis* of the Middle São Francisco river. The mitogenome of $2n=46$ and $2n=50$ cytotypes showed great genetic distance and the D-loop of both showed a repetitive 35bp sequence that varied in 19 sites between them. Besides, the number of repetitive sequences recovered in the satellitome of $2n=50$ was much higher than for $2n=46$ and both cytotypes showed unique sequences. All the tools in this work point out that *P. rivularis* corresponds to a complex of cryptic species with different diploid numbers, limited distribution to the Upper São Francisco river basin and adaptation of the shape of different Evolutionary Significant Units in different hydrographic sub-basins. Also, the cytotypes of $2n=46$ and $2n=50$ chromosomes sampled in this work represent at least two different cryptic species, found in allopathy and sympathy in the hydrographic sub-basin of the Abaeté River,

evidenced by cytogenetic and genomic data that suggest the existence of a possible reproductive isolation between them.

Keywords: Bioinformatics. Allopatric and sympatric evolution. Evolutionary Significant Units.

SUMÁRIO

1 INTRODUÇÃO GERAL.....	12
1.1 Delimitação de Unidade Evolutivamente Significativas (ESUs).....	12
1.2 Taxonomia, sistemática e biologia de <i>Psalidodon rivularis</i>.....	12
1.3 Bacia hidrográfica do Alto rio São Francisco.....	14
1.4 Citogenética.....	16
1.5 Morfometria Geométrica.....	17
1.6 Mitogenoma, Satelitoma e o Sequenciamento de Nova Geração (NGS).....	17
1.7 Objetivos.....	19
2 CAPÍTULO 1 - POLIMORFISMOS CROMOSSÔMICOS NO COMPLEXO	
<i>Psalidodon rivularis</i> (CHARACIFORMES: CHARACIDAE).....	21
2.1 Introdução.....	22
2.2 Materiais e Métodos.....	23
2.3 Resultados.....	24
2.4 Discussão e Conclusão.....	27
3 CAPÍTULO 2 - VARIAÇÃO DA FORMA DO CORPO NO CARACÍDEO <i>Psalidodon</i>	
<i>rivularis</i> DO RIO SÃO FRANCISCO, SUDESTE BRASILEIRO.....	31
3.1 Introdução.....	32
3.2 Materiais e Métodos.....	33
3.3 Resultados.....	38
3.4 Discussão e Conclusão.....	41

4 CAPÍTULO 3 - O MITOGENOMA E O SATELITOMA DE DOIS CITÓTIPOS (2n=46 E 2n=50) DE <i>Psalidodon rivularis</i> (CHARACIFORMES: CHARACIDAE).....	46
4.1 Introdução.....	47
4.2 Materiais e Métodos.....	48
4.3 Resultados.....	51
4.4 Discussão e Conclusão.....	55
4.5 Anexos.....	58
5 CONCLUSÕES GERAIS.....	62
6 REFERÊNCIAS.....	64
7 ANEXOS.....	73
7.1 ANEXO A - Lista de exemplares usados nas análises citogenéticas.....	73
7.2 ANEXO B - Lista de exemplares usados nas análises morfométricas.....	73
7.3 ANEXO C - Descrição dos genomas mitocondriais de <i>P. rivularis</i> 2n=46 e 2n=50.....	74
7.4 ANEXO D - Descrição e homologia das sequências obtidas em <i>P. rivularis</i> 2n=46 e 2n=50.....	76
7.5 ANEXO E - Sequências obtidas por prospecção <i>In Silico</i> de <i>P. rivularis</i> 2n=46 e 2n=50.....	90

1 INTRODUÇÃO GERAL

1.1 Delimitação de Unidade Evolutivamente Significativas (ESUs)

Delimitar espécies é um processo importante para se estabelecer estratégias de manejo e conservação apropriadas. Contudo, delimitar unidades específicas em alguns grupos é um grande desafio biológico devido a existência dos chamados complexos de espécies crípticas, que seriam conjuntos de duas ou mais espécies diferentes tradicionalmente catalogadas como uma única espécie devido seus aspectos morfológicos similares (JANZEN *et al.*, 2017).

A exemplo, temos o gênero de lambaris *Psalidodon*, cujos caracteres diagnósticos utilizados na delimitação de espécies têm levado ao estabelecimento de complexos de espécies crípticas, como *P. scabripinnis* e *P. fasciatus*, evidenciados por trabalhos envolvendo citogenética, biologia molecular e morfometria (MOREIRA-FILHO, BERTOLLO, 1991; PAZZA, KAVALCO, BERTOLLO, 2006; PAZZA, CRUVINEL, KAVALCO, 2017).

Uma alternativa a este problema é delimitar, dentro das unidades específicas, populações que contém atributos genéticos significativos e que devem receber atenção especial nas estratégias de conservação, ou em outras palavras, as Unidades Evolutivamente Significativas (ESUs) (RYDER, 1986). A identificação de ESUs pode, no entanto, ser uma tarefa complicada, exigindo abordagens integradas de biologia molecular, morfometria, distribuição geográfica e temporal, citogenética e etc (RYDER, 1986). Em algumas situações, casos de distribuição geográfica disjunta em conjunto com análises moleculares podem ser o suficiente para identificar essas unidades, mesmo nos complexos do gênero *Psalidodon* (PAZZA, CRUVINEL, KAVALCO, 2017; RODRIGUES ROCHA *et al.*, 2019). Contudo, o desafio é maior em *P. rivularis*, uma vez que já foram relatados casos de diferentes citótipos (números diplóides) e linhagens mitocondriais em simpatria (ROCHA *et al.*, 2019). Dessa forma, para delimitar diferentes ESUs em *P. rivularis* se faz necessário o uso de diferentes abordagens em conjunto, como citogenética, morfometria geométrica e genômica.

1.2 Taxonomia, sistemática e biologia de *Psalidodon rivularis*

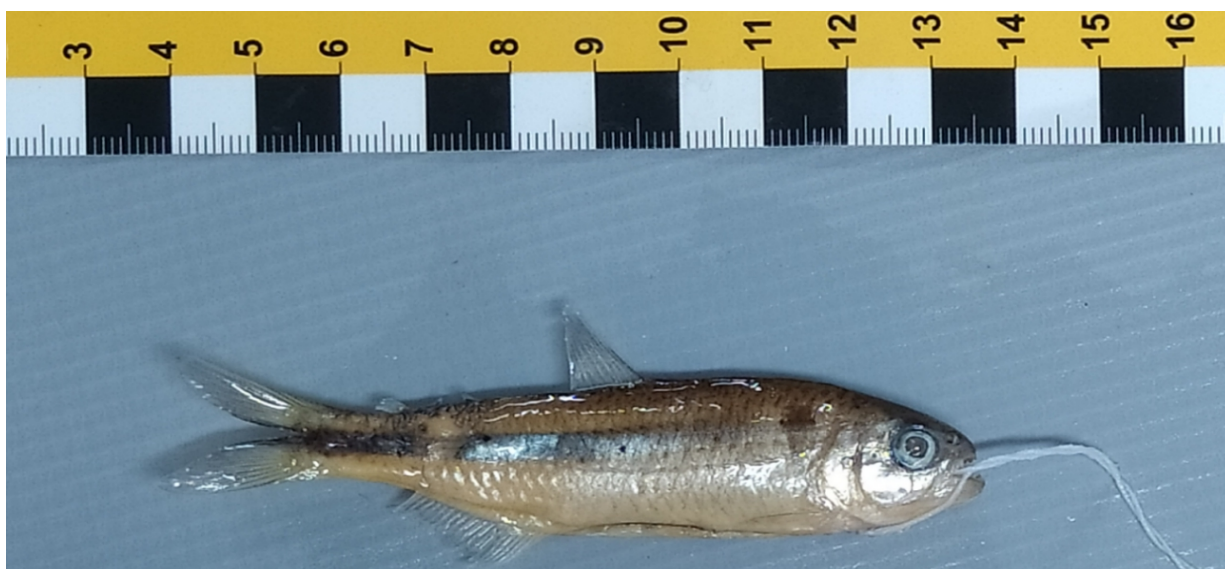
Dentre as ordens de peixes mais representativas da região Neotropical se encontra a ordem Characiformes, que possui 24 famílias, sendo a maior delas Characidae. Dos gêneros

descritos na família Characidae, *Psalidodon* se destaca por, após uma revisão recente de Terán, Benitez e Mirande (2020), passar a abranger complexos de espécies históricos antes associados ao gênero *Astyanax*, como *A. scabripinnis* e *A. fasciatus*. Entretanto, ambos os gêneros são estreitamente relacionados e filogeneticamente próximos dentro da tribo Gymnocharacini (Stethaprioninae, Characiformes). Dessa forma, novas revisões em espécies historicamente atribuídas ao gênero *Astyanax* se fazem necessárias.

A partir das características cariotípicas e morfológicas observadas em diversas populações de *P. scabripinnis*, Moreira-Filho e Bertollo (1991) chegaram à conclusão de que o grupo corresponderia a um complexo de espécies. Em uma revisão mais recente, Ingenito e Duboc (2014) apontaram a existência de pelo menos 29 espécies crípticas no complexo, habitando diferentes bacias hidrográficas, como a bacia do Alto rio São Francisco e a bacia do rio Paraná.

A única das espécies do complexo “*scabripinnis*” descrita para a bacia hidrográfica do Alto rio São Francisco é *P. rivularis* Lütken, 1875 (Figura 1), tendo como localidade tipo e área de distribuição o rio das Velhas e seus afluentes (TRIQUES, 2006). Posteriormente, sua distribuição geográfica foi ampliada para o no Parque Nacional Serra do Cipó (TRIQUES, 2006), para a região de Três Marias e para as sub-bacias do Alto rio São Francisco adjacentes ao Arco do Paranaíba (PASA *et al.*, 2019).

Figura 1: Exemplar de *Psalidodon rivularis* (sexo não identificado) coletado no rio Funchal por Igor Henrique Rodrigues Oliveira, Iuri Batista da Silva, Marcos Aurélio da Silva, Snaydia Viegas Resende, Vitor Augusto dos Santos, Francisco de Bezerra Cavalcante Sassi e Rubens Pasa.



Fonte: Igor Henrique Rodrigues Oliveira.

Assim como as demais espécies do complexo, *P. rivularis* é conhecida por habitar regiões de cabeceiras de rios e riachos. Devido a esse hábito, o isolamento populacional na espécie é intensificado, já que a própria calha do rio pode servir como uma barreira abiótica ao fluxo gênico entre populações, favorecendo processos como a evolução cromossômica e morfológica (MOREIRA-FILHO, BERTOLLO, 1991).

Embora *P. rivularis* seja uma espécie relativamente comum na bacia do Alto rio São Francisco, existem poucos estudos genéticos ou populacionais na literatura. Rocha *et al.* (2019) observou uma alta diversidade genética em uma população coletada no Parque Nacional Serra do Cipó, estimada através de marcadores moleculares mitocondriais. A distância genética superior a 2% para o gene da citocromo oxidase b (Cyt b) apresentada por um desses grupos sugere uma nova espécie do grupo *P. rivularis*. Esse trabalho evidencia o potencial das Unidades de Conservação em estudos envolvendo a diversidade e conservação de peixes neotropicais, e levanta a questão sobre *P. rivularis* constituir mais um complexo de espécies do gênero *Psalidodon*.

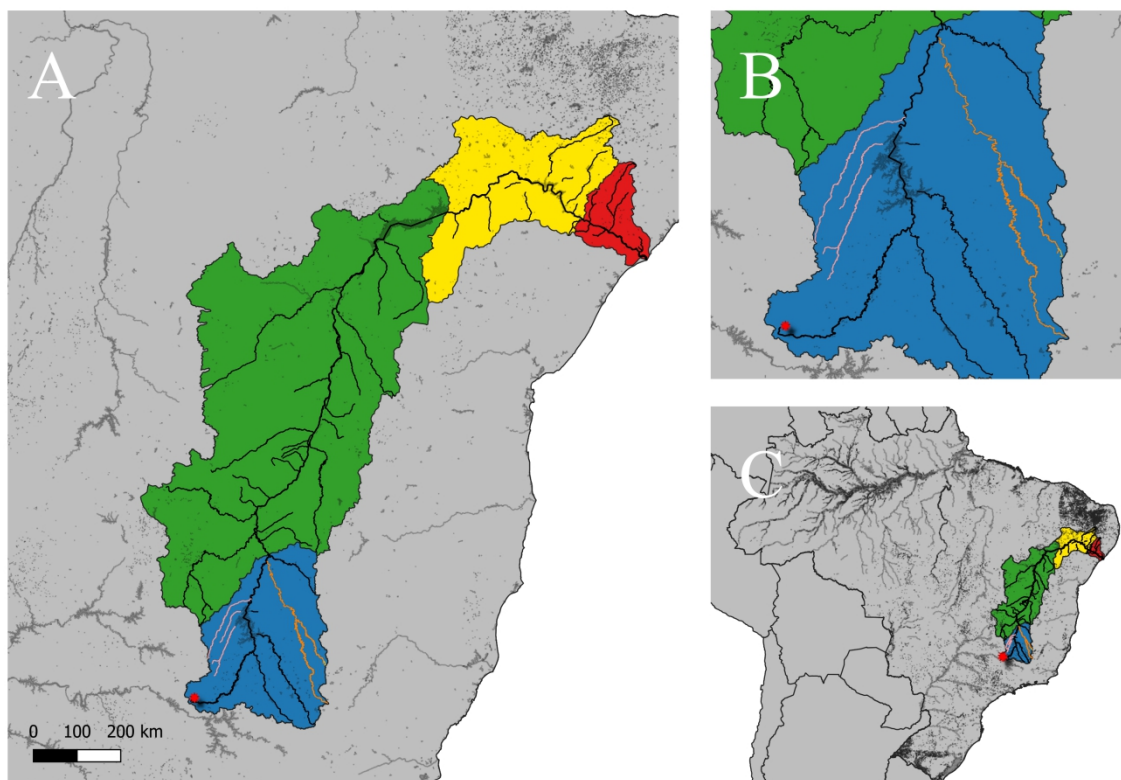
1.3 Bacia hidrográfica do Alto rio São Francisco

A bacia hidrográfica do rio São Francisco é a terceira maior bacia hidrográfica brasileira e a maior localizada integralmente em território nacional. O rio principal é o rio São Francisco, que tem sua nascente no estado de Minas Gerais e sua foz entre os estados de Sergipe e Alagoas (GODINHO, GODINHO, 2003). A bacia é subdividida em 4 porções: Alto rio São Francisco, Médio rio São Francisco, Submédio rio São Francisco e Baixo rio São Francisco (PAIVA, MELQUÍADES PINTO, BASTOS, 1982) (Figura 2).

Dentre essas 4 porções o Alto rio São Francisco se estende a partir da nascente do rio São Francisco no Parque Nacional Serra da Canastra em Minas gerais até a foz do rio das Velhas após a cidade de Pirapora (MG), tendo como suas limitações à direita e à esquerda a serra do Espinhaço e a Mata da Corda respectivamente (MMA, 2004).

Ainda à esquerda da bacia do Alto rio São Francisco se encontra a bacia do Alto rio Paraná. Ambas as bacias se isolaram através do soerguimento do Alto Paranaíba que começou seu processo de elevação a aproximadamente 117 milhões de anos atrás (CAMPOS, DARDENNE, 1997). Este evento foi responsável por separar as espécies irmãs *P. rivularis* (Alto rio São Francisco) e *P. paranae* (Alto rio Paraná) (ROCHA *et al.*, 2019).

Figura 2: (A) - Mapa das regiões hidrográficas do rio São Francisco, em azul: Alto rio São Francisco, verde: Médio rio São Francisco, amarelo: Submédio rio São Francisco, Vermelho: Baixo rio São Francisco. (B) - Enfoque da bacia do Alto rio São Francisco, ponto vermelho: nascente do rio São Francisco, em rosa: rios das sub-bacias adjacentes ao Arco do Paranaíba, em laranja: sub-bacia do rio das Velhas. (C) - Mapa demonstrando a localização da bacia do rio São Francisco no continente.



Fonte: *shapefiles* utilizados no mapa disponíveis nos endereços online: SOS GIS BR (<https://sosgisbr.com/2011/07/11/shapes-disponibilizados-pelo-comite-da-bh-do-rio-sao-francisco/>) e Google Earth na Sala de Aula (<http://www.mapasnasaladeaula.org>).

A bacia do alto rio São Francisco vem sofrendo ao longo dos anos uma série de impactos antrópicos que ameaçam sua biodiversidade. A região é a porção da bacia do rio São Francisco com maior incidência de enchentes naturais, que são responsáveis pelo carreamento de nutrientes, sedimentos, matéria orgânica, fauna e flora (JUNK, BAYLEY, SPARKS, 1989), porém a construção de represas vem gradativamente diminuindo a incidência de enchentes (E SANTOS, DOS SANTOS POMPEU, KENJI, 2012). Além disso, dois dos principais afluentes do rio São Francisco na região, o Rio das Velhas e o rio Paraopeba, sofrem constantemente com poluição por resíduos de esgotos de origem doméstica, industriais e minerários (ALVES, POMPEU, 2001; GODINHO, GODINHO, 2003). Em um episódio mais recente, no ano de 2019, uma barragem de rejeitos da mineradora VALE S.A. rompeu no município de Brumadinho (MG) contaminando recursos hídricos locais e o próprio rio Paraopeba (PEREIRA, DE BARROS CRUZ, GUIMARÃES, 2019).

No extremo oeste da bacia do Alto rio São Francisco, onde se localiza o Arco do Alto Paranaíba, os principais impactos ambientais são gerados pela agricultura (GODINHO, GODINHO, 2003). Dentro os rios da região podemos citar os rios Abaeté, Borrachudo e Indaiá que atravessam os municípios de Rio Paranaíba, São Gotardo, Arapuá, Tiros e Matutina entre outros.

1.4 Citogenética

A citogenética pode ser descrita como uma ciência interdisciplinar da Genética e da Citologia que se dedica ao estudo dos cromossomos, podendo estar isolados ou em conjuntos, condensados ou distendidos, e considerando sua morfologia, função e organização, tanto nas fases de replicação quanto nas suas variações e processos evolutivos (GUERRA, 1988).

Em geral, assume-se que a maior parte dos organismos vivos possuam um número diplóide, ou $2n$, característico ao nível de espécie. Mas o que tem se encontrado em diversos grupos de peixes são variações numéricas de cromossomos devido a rearranjos cromossômicos ou presença de cromossomos supranumerários, também chamados cromossomos Bs (PAZZA, KAVALCO, BERTOLLO, 2006).

Dentro do gênero *Psalidodon* são diversos os casos de espécies com polimorfismos numéricos e estruturais de cromossomos, levando a formação de complexos de espécies (MOREIRA-FILHO, BERTOLLO, 1991; JUSTI, 1993; PAZZA, KAVALCO, BERTOLLO, 2006). A exemplo disso, temos o complexo *P. fasciatus* como sendo o complexo com a maior diversidade de número diplóide, com $2n=45$, 46, 47, 48 e 49 cromossomos (PAZZA, KAVALCO, BERTOLLO, 2006, FERREIRA-NETO *et al.*, 2012; KAVALCO *et al.*, 2013). Outro complexo de espécies conhecido por sua diversidade cariotípica, *P. scabripinnis*, possui citótipos variando de $2n=46$ a $2n=50$ cromossomos e inclui, dentre várias outras espécies, *P. rivularis* e *P. paranae* (MOREIRA-FILHO BERTOLLO, 1991; BERTACO, LUCENA, 2006).

Além de servirem como modelo em estudos envolvendo variações numéricas e estruturais de cromossomos, as espécies do complexo *P. scabripinnis* já foram amplamente utilizadas em estudos envolvendo padrões de heterocromatina constitutiva nos cromossomos (MOREIRA-FILHO, BERTOLLO, 1991; SOUZA, MOREIRA-FILHO, BERTOLLO, 1995;

MAISTRO, OLIVEIRA, FORESTI, 2000), evolução e dinâmica dos cromossomos Bs (VICENTE, MOREIRA-FILHO, CAMACHO, 1996; MESTRINER *et al.*, 2000; MOREIRA-FILHO, GALETTI, BERTOLLO, 2004; VICARI *et al.*, 2011; DUÍLIO *et al.*, 2014) e localização de sondas de DNAs ribossomais e satélites (MANTOVANI *et al.*, 2005; ABEL, MANTOVANI, MOREIRA-FILHO; FERNANDES, MARTINS-SANTOS, 2006).

1.5 Morfometria Geométrica

A morfometria geométrica, ou análise estatística de formas, é um campo de pesquisa que surgiu entre a década de 80 e 90 através dos estudos da tese de doutorado de Colin R. Goodall em 1983 (GOODALL, 1991) e dos artigos de David G. Kendall (KENDALL, 1984) e Fred L. Bookstein (BOOKSTEIN, 1984). A técnica é baseada na análise da informação contida na co-variação espacial entre marcos anatômicos (ROHLF, MARCUS, 1993).

Comparada a morfometria tradicional, a morfometria geométrica é mais eficiente na captura de informações relacionadas a variação da forma nos organismos, além de apresentar maior robustez estatística (ROHLF, MARCUS, 1993). A técnica também permite visualizar reconstruções das formas inferidas estatisticamente e observar como, onde e em que magnitude as mudanças estão ocorrendo (FORNEL, CORDEIRO-ESTRELA, 2012).

Muitos trabalhos têm usado essa técnica para responder uma ampla gama de perguntas em organismos biológicos. Dentro do grupo dos peixes, inclusive nos gêneros *Astyanax* e *Psalidodon*, a ferramenta já foi usada em trabalhos envolvendo sistemática, ecomorfologia, evolução da forma, alometria, dimorfismo sexual, assimetria e filogeografia (AGUILAR-MEDRANO, 2013; CASTRO *et al.*, ORNELAS-GARCÍA, BASTIR, DOADRIO, 2014; CASTRO *et al.*; SANTOS, ARAÚJO, 2015; POWERS *et al.*, 2017; PERAZZO *et al.*, 2018; ROCHA *et al.*, 2019; DA SILVA *et al.*; DE MESQUITA ALVES *et al.*, 2020).

1.6 Mitogenoma, Satelitoma e o sequenciamento de nova geração (NGS)

O genoma de um organismo eucarioto está organizado em diferentes moléculas de DNA que compõem os cromossomos, mitocôndrias e, em alguns organismos, plastídios.

Grande parte do genoma cromossomal dos eucariotos é formada por elementos repetitivos, que podem ser sequências repetidas em Tandem ou dispersas pelo genoma (LÓPEZ-FLORES, GARRIDO-RAMOS, 2012).

Dentre os elementos repetitivos em Tandem temos as sequências moderadamente repetitivas (RNAs ribossômicos, regiões teloméricas curtas e famílias de genes codificadores de proteínas) e as sequências altamente repetitivas (microssatélites, minissatélites e satélites não codificantes). Já entre os elementos dispersos pelo genoma temos retrogenes, retropseudogenes, famílias de genes compostas por membros dispersos e elementos transponíveis (LÓPEZ-FLORES, GARRIDO-RAMOS, 2012).

Os DNAs satélites (satDNAs) representam uma porção não codificante do genoma e estão entre os componentes mais abundantes do genoma eucarioto (PLOHL, MEŠTROVIĆ, MRAVINAC, 2012). Esses elementos podem ser organizados em famílias de satDNAs em respeito às variações de suas sequências, o tamanho do motivo de repetição, sua abundância e localização no genoma (GARRIDO-RAMOS, 2017). Dentro de um único genoma eucarioto podem haver centenas de famílias de satDNAs, cada uma com suas próprias dinâmicas evolutivas, e o conjunto dessas bibliotecas é denominado Satelitoma (RUIZ-RUANO *et al.*, 2016). Essas diferentes famílias podem ser espécie-específicas ou ser compartilhadas por espécies próximas (GARRIDO-RAMOS, 2017).

Essas sequências possuem grande aplicabilidade na citogenética podendo ser usadas em estudos citotaxômicos de complexos de espécies (ABEL, MANTOVANI, MOREIRA-FILHO, 2006; SAITO, EDPALINA, ABE, 2007; KANTEK *et al.*, 2009; VICARI *et al.*, 2010). Contudo estudos genômicos e citogenômicos envolvendo DNAs satélites ainda são escassos quando comparados a outros elementos repetitivos, como os RNAs ribossômicos (REBORDINOS, CROSS, MERLO, 2013).

O mitogenoma por sua vez é o nome dado ao genoma mitocondrial (mtDNA) dos eucariotos. Esse genoma possui uma organização e composição gênica conservada em vertebrados, com 13 genes codificadores de proteínas (PCGs), 22 genes de tRNAs, 2 genes de rRNAs e uma região controle denominada D-loop (TAANMAN, 1999; SATOH *et al.*, 2016). Essa organização conservada facilita o estudo evolutivo entre diferentes espécies e grupos de organismos (SACCONE *et al.*, 1999). Até o momento, as únicas espécies do

gênero *Psalidodon* com genoma mitocondrial descrito são *P. paranae* (SILVA *et al.*, 2016) e *P. fasciatus* (CALEGARI *et al.*, 2019).

O sequenciamento de nova geração (NGS) surgiu nos anos 2000 com o intuito de solucionar os problemas biológicos complexos provenientes do sequenciamento do genoma humano (GOODWIN, McPHERSON, MCCOMBIE, 2016). Graças ao NGS, e a ferramentas de bioinformáticas mais acessíveis desenvolvidas nos últimos anos, a anotação, montagem e caracterização de genomas biológicos, como o satelitoma, o mitogenoma e mesmo o genoma completo, tem se tornado mais fáceis e baratas (NOVÁK *et al.*, 2013; AFGAN *et al.*, 2018; RODRIGUES *et al.*, 2019). O NGS utilizando a plataforma Illumina é um tipo de sequenciamento de leituras curtas (short reads) por síntese que usa terminadores de cadeia semelhantes aos utilizados no método de Sanger (GOODWIN, McPHERSON, MCCOMBIE, 2016).

Aliadas aos novos métodos de sequenciamento por NGS, as plataformas on-line públicas de bioinformática permitem acessibilidade, replicabilidade, reprodutibilidade e integrabilidade com bancos de dados genômicos, facilitando o estudo do genoma em diversas instâncias (GOECKS, NEKRUTENKO, TAYLOR., 2010; AFGAN *et al.*, 2018). A exemplo, podemos citar a plataforma Galaxy e seus três servidores (europeu: <http://usegalaxy.eu>, australiano: <https://usegalaxy.org.au/> e americano <https://usegalaxy.org/>) dedicada a manipulação, alinhamento, montagem e anotações genômicas (AFGAN *et al.*, 2018), a plataforma Repeat Explorer (NOVÁK *et al.*, 2013) e seu servidor público baseado em galaxy (<https://repeatexplorer-elixir.cerit-sc.cz/galaxy/>) voltada para o estudo do satelitoma e o servidor MittoFish (<http://mitofish.aori.u-tokyo.ac.jp>) que contém a pipeline MitoAnnotator (IWASAKI *et al.*, 2013) para anotação do mitogenoma de peixes.

1.7 Objetivos

Uma vez que nossa hipótese é que *P. rivularis* representa um complexo de espécies, nossa pergunta era solucionar “quantas ESUs de *P. rivularis* seria possível identificar através do uso de uma abordagem integrada com ferramentas citogenéticas, morfométricas e genômicas?”. Dessa forma, nosso objetivo geral foi avaliar a diversidade citogenética, morfológica e genômica do grupo *P. rivularis*, endêmico da bacia do Alto Rio São Francisco, com enfoque nas populações existentes nas sub-bacias adjacentes ao Arco do Paranaíba.

Além disso, buscou-se identificar possíveis ESUs e discutir como essas ferramentas suportam *P. rivularis* como um complexo de espécies.

Os objetivos específicos do trabalho foram:

1- avaliar e descrever a diversidade cariotípica do grupo *P. rivularis*, com a premissa de identificar diferentes cariomorfos em diferentes locais;

2- avaliar a modulação da forma de *P. rivularis* em diferentes pontos de coletas e sub-bacias e correlacionar com fatores biológicos e ambientais, buscando identificar possíveis aspectos que diferem a forma entre diferentes grupos amostrais;

3- descrever e comparar o mitogenoma de dois citótipos ($2n=46$ e $2n=50$) de *P. rivularis* com a premissa que, se cada citótipo corresponde a diferentes espécies crípticas, seria possível identificar aspectos nos mitogenomas que corroborem essa afirmação;

4- realizar a prospecção *in Silico* do satelitoma dos dois citótipos, buscando identificar as principais diferenças entre eles, satélites exclusivos compartilhados e satélites exclusivos não compartilhados.

2 CAPÍTULO 1 - POLIMORFISMOS CROMOSSÔMICOS NO COMPLEXO *Psalidodon rivularis* (CHARACIFORMES: CHARACIDAE)

Igor Henrique Rodrigues-Oliveira, Karine Frehner Kavalco, Rubens Pasa

RESUMO

Dentre os peixes Neotropicais dotados de grande diversidade cariotípica se encontra o complexo de espécies *Psalidodon rivularis*. O grupo possui números diplóides variando de $2n=46$ a $2n=50$ cromossomos, distribuídos por tributários desde a montante do rio São Francisco até a região de Três Marias. O objetivo deste trabalho foi caracterizar seis populações de *P. rivularis* através da citogenética, com o intuito de encontrar diferentes cariomorfos em diferentes locais como foi observado na literatura. As preparações cromossômicas foram obtidas a partir do tecido renal de 33 espécimes. Conduzimos também a técnica de identificação da heterocromatina constitutiva em três cariomorfos amostrados e a técnica de dupla coloração de CMA₃/DAPI em indivíduos portadores de cromossomos Bs de dois cariomorfos distintos. Foram observados dois cariomorfos de $2n=46$ cromossomos (46I e 46II), com e sem presença de cromossomos Bs, um cariomorfo de $2n=48$ e um de $2n=50$. A técnica de bandeamento C mostrou variações nos padrões de heterocromatina constitutiva entre todos os cariomorfos analisados. A dupla coloração com CMA₃/DAPI indicou uma composição rica em pares de bases AT apenas no cromossomo B do cariomorfo $2n=46I$, embora tenham sido identificados múltiplos sítios GC-ricos. Uma vez que o cariomorfo de $2n=48$ foi obtido numa localidade pertencente a bacia do rio das Velhas, local de descrição do holótipo de *P. rivularis*, sugerimos que este cariótipo seja o representativo para a espécie e os demais devam ser considerados diferentes ESUs.

Palavras-chave: cariomorfos, citogenética, cromossomos Bs, evolução cromossômica.

2.1 Introdução

Dentro dos diversos gêneros de peixes neotropicais, o gênero *Astyanax* é historicamente conhecido pela ocorrência de complexos de espécies baseados em dados citogenéticos (PAZZA, KAVALCO, 2007). Dentre os complexos de espécies do gênero, *A. scabripinnis* é evidenciado pela presença de populações com diferentes cariótipos, números diplóides, cromossomos Bs, padrões de heterocromatina e outros marcadores citogenéticos (MOREIRA-FILHO, BERTOLLO, 1991; SOUZA, MOREIRA-FILHO, BERTOLLO, 1995; MAISTRO, OLIVEIRA, FORESTI, 2000; MOREIRA-FILHO, GALETTI, BERTOLLO, 2004; FERNANDES, MARTINS-SANTOS, 2005; BIAVATTI, MAISTRO, 2007; CASTRO *et al.*, 2014; KLASSMANN, DOS SANTOS, 2017). Após uma ampla revisão taxonômica de Terán, Benitez e Mirande (2020) os dois maiores complexos de espécies do gênero *Astyanax*, *A. scabripinnis* (MOREIRA-FILHO, BERTOLLO, 1991) e *A. fasciatus* (JUSTI, 1993) foram movidos para o gênero *Psalidodon*.

Psalidodon rivularis pode ser descrita como a espécie característica do antigo complexo “*Astyanax scabripinnis*” (MOREIRA-FILHO, BERTOLLO, 1991) na Bacia hidrográfica do Alto rio São Francisco (BERTACO, LUCENA, 2006; PASA *et al.*, 2019). Populações pertencentes a bacia hidrográfica do Alto rio São Francisco referidas como *P. scabripinnis* foram abordadas nos trabalhos de Moreira-Filho e Bertollo (1991), Mantovani *et al.* (2005), Abel, Mantovani e Moreira-Filho (2006), Biavati e Maistro (2007) e Peres, Bertollo e Moreira-Filho (2008). Juntos esses trabalhos evidenciam a presença de dois citótipos ($2n=46$ e $2n=50$) e quatro cariomorfos ($2n=46: 6m+22sm+8st+10a$ e $2n=50: 6m+30sm+8st+6a$ (MOREIRA-FILHO, BERTOLLO, 1991), $2n=50: 8m+20sm+6st+16a$ (BIAVATI, MAISTRO, 2007) e $2n=50: 12m+24sm+8st+6a$ (PERES, BERTOLLO, MOREIRA-FILHO, 2008)) em cinco populações amostradas, com diferenças nos padrões de heterocromatina constitutiva (MOREIRA-FILHO, BERTOLLO, 1991), quantidade e localização do rDNA 18S (MANTOVANI *et al.*, 2005) e satDNA As51 (ABEL, MANTOVANI, MOREIRA-FILHO., 2006), sem evidências de cromossomos supranumerários ou Bs.

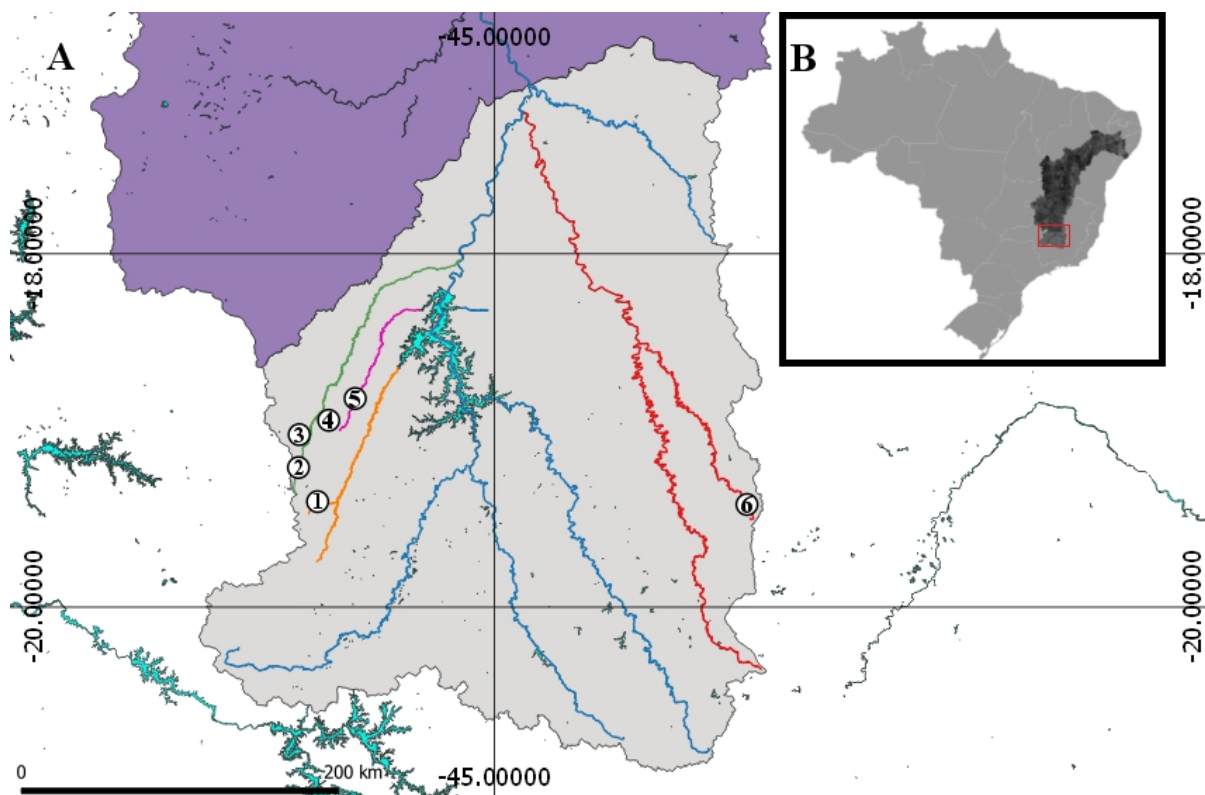
O objetivo deste trabalho foi ampliar o conhecimento acerca dos padrões citogenéticos da espécie para a região, com o intuito de se encontrar novos cariomorfos, realizando a caracterização cariotípica de seis populações de *P. rivularis* na bacia do Alto rio São Francisco, cinco delas pertencentes à sub-bacias adjacentes ao Arco do Paranaíba (rios

Abaeté, Borrachudo e Indaiá) e uma pertencente ao Parque Nacional Serra do Cipó (rio das Velhas).

2.2 Materiais e Métodos

Utilizamos neste trabalho, exemplares pertencentes a seis pontos de coleta de quatro sub-bacias hidrográficas do Alto rio São Francisco: Abaeté, Borrachudo, Indaiá e das Velhas (Figura 1, Tabela 1). Todos os espécimes foram identificados como *P. rivularis* por Pasa *et al.* (2019).

Figura 1: (A) Mapa dos pontos de coleta amostrados nas análises de citogenética. 1- rio Funchal, 2 - Usina Abaeté, 3 - córrego Lage, 4 - córrego Tiros, 5 - córrego Lambari, 6 - ribeirão Bandeirinhas. Em verde: sub-bacia do rio Abaeté, em rosa: sub-bacia do rio Borrachudo, em laranja: sub-bacia do rio Indaiá, em vermelho: sub-bacia do rio das Velhas. (B) Mapa do Brasil destacando a bacia do rio São Francisco em preto, retângulo vermelho delimitando a bacia do Alto rio São Francisco.



Todas as coletas foram realizadas conforme licenças emitidas pelo SISBIO (Sistema de Autorização e Informação em Biodiversidade, número de licença 1938128) e SISGEN (Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado, número de licença A9FE946) atribuídas a Rubens Pasa. Os procedimentos de amostragem e eutanásia (via imersão em Eugenol 1%) dos animais foram realizados de

acordo com a Resolução Normativa N°37 do Conselho Nacional de Controle de Experimentação Animal (CONCEA).

Obtivemos as suspensões celulares a partir da dissecação dos espécimes e o material se encontra depositado no Banco de Tecidos, Suspensão e DNA do Laboratório de Genética Ecológica e Evolutiva da Universidade Federal de Viçosa - Campus Rio Paranaíba. Para a obtenção dos cromossomos mitóticos realizamos o procedimento de *Air Drying* conforme protocolo adaptado de Bertollo *et al.* (1978) e posteriormente coramos as lâminas com Giemsa 10%. Também realizamos a técnica de bandamento C (SUMNER, 1978) em cada cariomorfo amostrado. Realizamos a técnica de dupla coloração com Cromomicina A₃ (CMA₃) e 4',6-diamidino-fenilindole (DAPI) (SOLA *et al.*, 1982, com modificações) em ambos os cariomorfos de 2n=46 com cromossomos Bs amostrados, visando identificar regiões ricas em pares de bases GC/AT nestes elementos supranumerários.

Capturamos as fotomicrografias com o auxílio de um fotomicroscópio Olympus BX41 (Olympus Inc., Japão) acoplado com uma câmera de 3 megapixels, usando o software QCapture Pro 6.0 (QImagine, Surrey, BC, Canadá). Caracterizamos a morfologia dos cromossomos de acordo com a razão entre os comprimentos dos braços conforme Levan, Fredga e Sandberg (1964) e montamos as pranchas cariotípicas no software GIMP 2.8.

2.3 Resultados

Os resultados obtidos evidenciaram diversidade cariotípica inter e intra-populacional no que diz respeito a número diplóide, fórmula cariotípica e presença/ausência de cromossomos Bs (Figura 2, Tabela 1). Uma vez que dois diferentes cariomorfos de 2n=46 cromossomos foram encontrados, os nomeamos de 2n=46I e 2n=46II (Tabela 1).

Figura 2: Cariótipos obtidos. Barra de escala: 5µm.

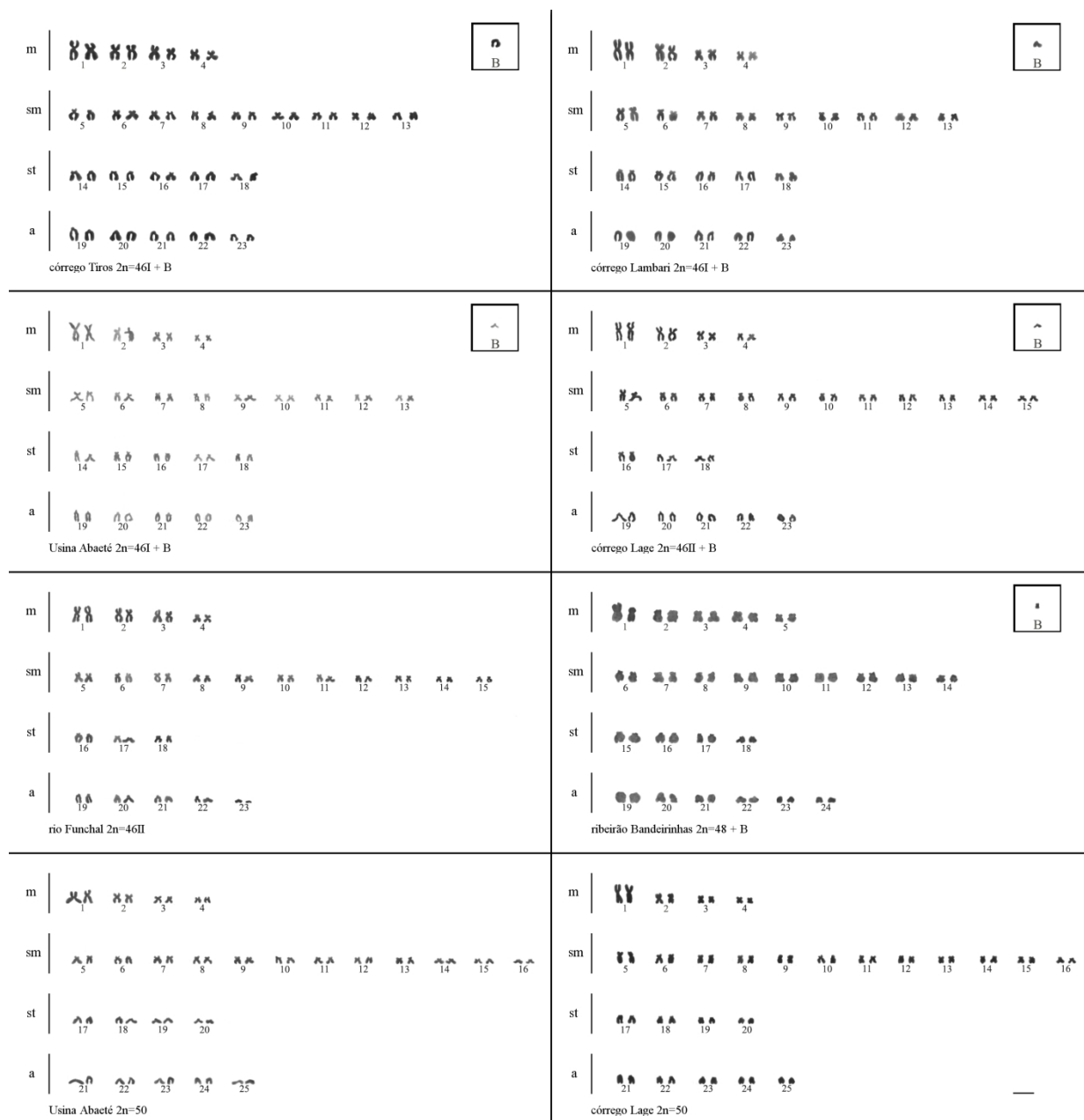
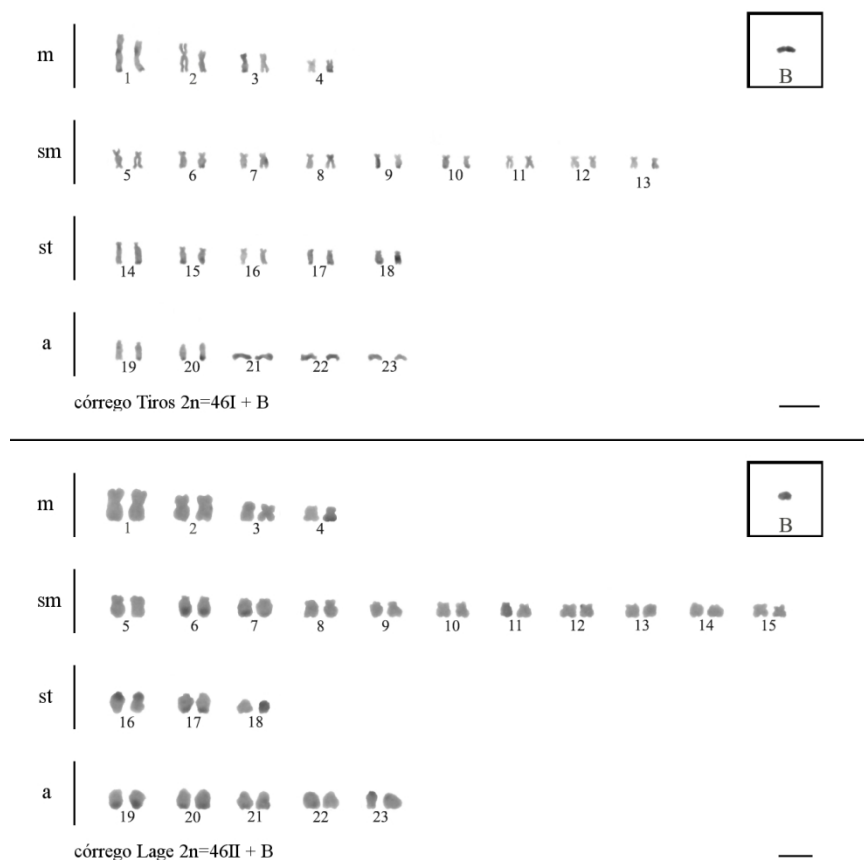


Tabela 1: Números diplóides e fórmulas cariotípicas obtidas agrupados por pontos de coleta. Classificação dos cromossomos: m - metacêntricos, sm - submetacêntricos, st - subtelocêntricos e a - acrocêntricos.

Sub-bacia Hidrográfica	Ponto de coleta	n amostral	2n	no. de Bs por célula	Tipos de cromossomos			
					m	sm	st	a
Abaeté	córrego Lage (19°1'25.18"S e 46°6'18.74"O)	11	46II	0-1	8	22	6	10
		3	50	0	8	24	8	10
	Usina Abaeté (19°12'35.69"S e 46°6'33.79"O)	1	46I	0-1	8	18	10	10
		4	50	0	8	24	8	10
	córrego Tiros (18°56'34.08"S e 45°56'18.20"O)	6	46I	0-1	8	18	10	10
Borrachudo	córrego Lambari (18°49'8.49"S e 45°47'17.72"O)	4	46I	0-1	8	18	10	10
das Velhas	ribeirão Bandeirinhas (19°25'8.33"S e 43°34'12.37"O)	2	48	0-1	10	18	8	12
Indaiá	rio Funchal (19°24'9.54"S e 46°0'4.61"O)	2	46II	0	8	22	6	10

No cariomorfo $2n=46I$ visualizamos blocos discretos de heterocromatina constitutiva na região terminal dos braços longos dos pares 1, 9, 10, 15, 17, 20 e 21; além do mesmo tipo de marcação em um único cromossomo dos pares 3, 4, 6, 14, 18 e 19. Por sua vez, no cariomorfo $2n=46II$ visualizamos grandes blocos conspícuos na região terminal dos braços longos dos pares 6, 17 e 19, na região pericentromérica e nos braços curtos do par 16, além de blocos discretos na região terminal dos braços longos dos pares 7 e 20. Ambos os cariomorfos apresentaram cromossomos B totalmente heterocromáticos (Figura 3). Não foi possível visualizar marcações positivas no bandamento C em ambos os cariótipos do cariomorfo $2n=50$, indicando ausência de blocos conspícuos de heterocromatina constitutiva. Não foi possível obter resultados para o cariomorfo $2n=48$ do ribeirão Bandeirinhas, devido à pouca quantidade de material obtido e às limitações impostas pela pandemia do Covid-19, que limitou as coletas de exemplares.

Figura 3: Bandamento C nos cariomorfos $2n=46I$ e $2n=46II$. Barra de escala: $5\mu m$.

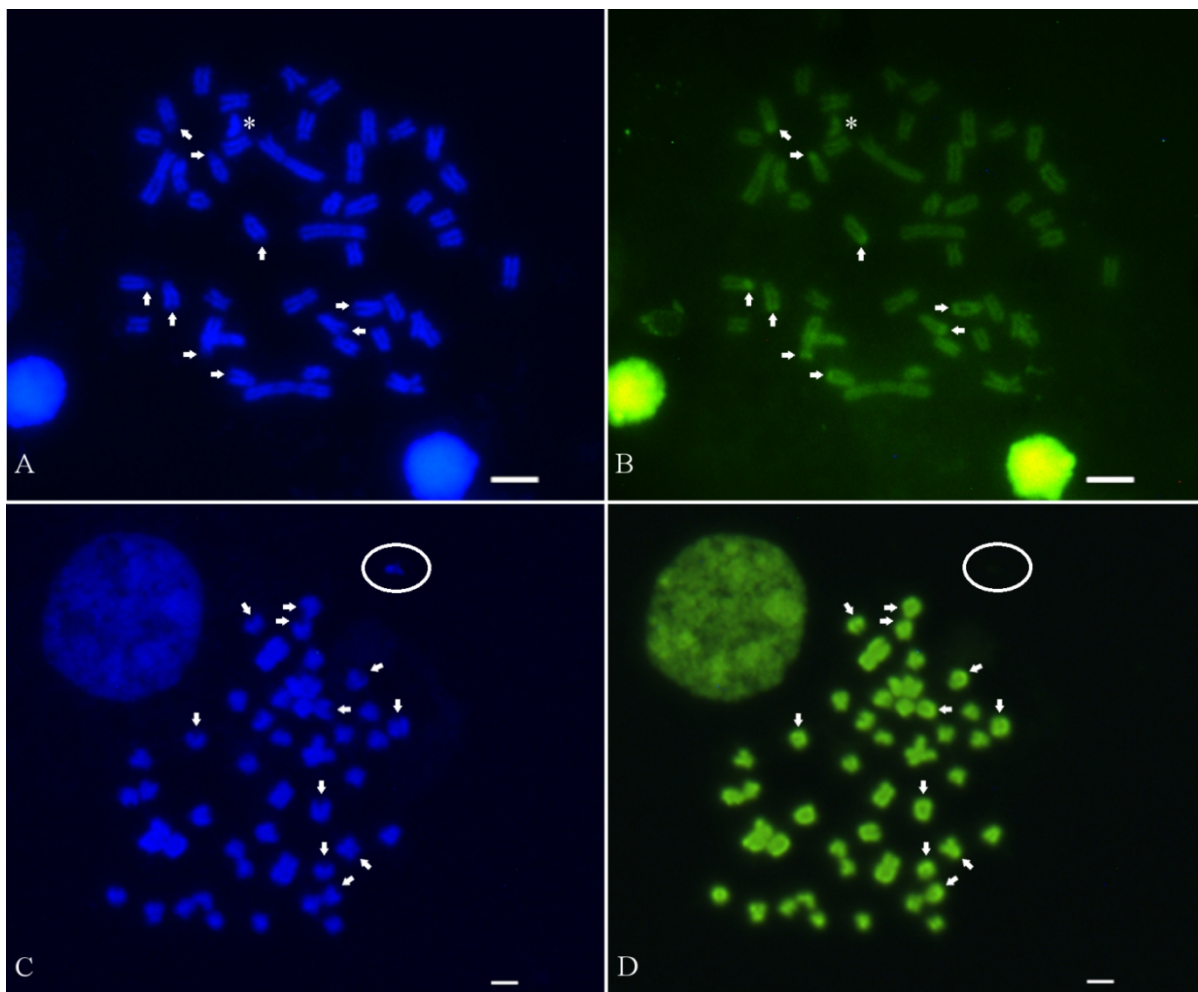


Nas análises de dupla coloração com CMA₃/DAPI, observamos nas metáfases do citótipo $2n=46II + B$ (córrego Lage) a presença de 9 sítios terminais GC-ricos, sem nenhuma região AT-rica. Em compensação, no cariomorfo $2n=46I + B$ (córrego Tiros), observamos o cromossomo B fortemente visível na DAPI e ausente na CMA₃, sugerindo um conteúdo rico em pares de bases AT em toda sua extensão, além de 11 sítios terminais GC-ricos (Figura 4).

2.4 Discussão e Conclusão

Os resultados observados demonstraram um polimorfismo cromossômico no complexo *P. rivularis* muito além do que havia sido exposto na literatura para os “*P. scabripinnis*” da bacia do Alto rio São Francisco (MOREIRA-FILHO, BERTOLLO, 1991; BIAVATI, MAISTRO, 2007; PERES, BERTOLLO, MOREIRA-FILHO, 2008). Pazza e Kavalco (2007) citam a evolução cromossômica como um fator delimitante na história do gênero *Psalidodon* e é provável que a evolução cromossômica tenha sido o marco inicial para a especiação de *P. rivularis* na bacia do Alto rio São Francisco, originando um complexo de espécies endêmicas.

Figura 4: Dupla coloração de CMA₃/DAPI em metáfase cariomorfo 2n=46II + B (A: DAPI e B: CMA₃) e do cariomorfo 2n=46I + B (C: DAPI e D: CMA₃). Setas indicando regiões GC ricas, asterisco indicando cromossomo B (córrego Lage) e círculo indicando cromossomos B AT rico (córrego Tiros). Barra de escala: 5µm.



Espécies historicamente atribuídas ao complexo *A. scabripinnis*, como *P. rivularis*, são conhecidas por habitar pequenas cabeceiras de riachos, onde a própria calha do rio pode servir como barreira biológica, o que faz com que estes grupos tendem a evoluir independentemente entre diferentes bacias hidrográficas, ou mesmo dentro de uma mesma bacia (MOREIRA-FILHO, BERTOLLO, 1991). Desta forma, eventos vicariantes podem ocasionar a evolução alopátrica em populações relativamente próximas, ocasionando endemismo em peixes de cabeceiras (CASTRO, 1999).

As sub-bacias da região adjacentes ao Arco do Paranaíba apresentaram fixação de cariomorfos diferentes em riachos relativamente próximos, o que pode ser ocasionada por eventos vicariantes como a captura de cabeceiras de um rio pelo outro (VICARI *et al.*, 2008).

Uma vez em situações de isolamento reprodutivo, eventos de inversões e translocações podem acentuar diferenças no cariótipo desses grupos (SOUZA, MOREIRA-FILHO, BERTOLLO, 1995; PAZZA e KAVALCO, 2007). Os mesmos processos vicariantes de capturas de cabeceiras podem ocasionar situações opostas, onde citótipos antes alopátricos se unem em condições de simpatria (MAISTRO, OLIVEIRA, FORESTI, 2000; VICARI *et al.*, 2008). Acreditamos que esses fatores possam explicar a distribuição dos cariomorfos na região do Arco do Paranaíba, em especial os casos de simpatria observados no córrego Lage ($2n=46II$ e $2n=50$) e na Usina Abaeté ($2n=46I$ e $2n=50$).

Enquanto tanto o cariomorfo $2n=46I$ quanto o $2n=46II$ possuem fórmulas cariotípicas semelhantes ao $2n=46$ descrito no córrego Curral das Éguas (MOREIRA-FILHO, BERTOLLO, 1991), os cariomorfos $2n=50$ amostrados diferem bastante dos $2n=50$ descritos previamente (MOREIRA-FILHO, BERTOLLO, 1991; BIAVATI, MAISTRO, 2007; PERES, BERTOLLO, MOREIRA-FILHO, 2008). Os resultados reforçam a ideia do $2n=50$ como simplesiomórfico em *Psalidodon* (PAZZA, KAVALCO, 2007). Dessa forma, é possível que este citótipo tenha acumulado mais diferenças cariotípicas ao longo do tempo na bacia do Alto rio São Francisco.

Padrões de heterocromatina constitutiva para *P. rivularis* foram descritos para uma população com $2n=46$ (MOREIRA-FILHO, BERTOLLO, 1991) e três com $2n=50$ (MOREIRA-FILHO, BERTOLLO, 1991; BIAVATI, MAISTRO, 2007). Os padrões de heterocromatina observados neste trabalho não só diferem entre cada cariomorfo amostrado como também com todos os registros literários para a espécie (MOREIRA-FILHO, BERTOLLO, 1991; BIAVATI, MAISTRO, 2007). Curiosamente, ao contrário do que foi relatado por Moreira-Filho e Bertollo (1991), onde as populações com $2n=50$ cromossomos apresentaram marcações positivas espalhadas nas regiões teloméricas de diferentes pares de cromossomos e a população com $2n=46$ apresentou apenas um par de cromossomos com marcação positiva para banda C, nosso cariomorfo de $2n=50$ não apresentou marcações positivas na banca C enquanto que os cariomorfos de $2n=46$ apresentaram marcações positivas em vários pares em regiões teloméricas (Figura 3). Essas informações evidenciam o potencial do bandamento C como um marcador populacional no grupo (MOREIRA-FILHO, BERTOLLO, 1991).

O cromossomo B acrocêntrico descrito nas populações dos córregos Lage, Lambari, Tiros e Usina Abaeté se assemelham muito ao encontrado nos *P. scabripinnis* do rio Ivaí,

bacia do Paranapanema (FERNANDES, MARTINS-SANTOS, 2005). É possível que esse cromossomo tenha se originado a partir dos menores cromossomos acrocêntricos desses cariótipos, mas sem a formação de um isocromossomo, uma vez que os cromossomos Bs de *P. rivularis* permaneceram como cromossomos acrocêntricos ao invés de duplicarem seus braços longos formando isocromossomos metacêntricos como foi proposto para *P. scabripinnis* e *P. paranae* (VICENTE, MOREIRA-FILHO, CAMACHO, 1996; MESTRINER *et al.*, 2000; MOREIRA-FILHO, GALETTI, BERTOLLO, 2004; VICARI *et al.*, 2011; DUÍLIO *et al.*, 2014).

Além de descrever a presença de cromossomos Bs, pudemos levantar informações sobre a composição desses cromossomos em *P. rivularis*. Os cromossomos Bs de ambos os cariomorfos de 46 cromossomos parecem ser heterocromáticos (Figura 3), com composição AT-rica no cariomorfo 2n=46I, embora o padrão não tenha se repetido no 2n=46II (Figura 4). É bastante incomum a presença de marcações AT-ricas em cromossomos do grupo *Psolidodon*, muito embora seja reconhecido que as regiões GC-ricas sejam associadas às heterocromatinas intercalares às Regiões Organizadoras Nucleolares em ostariófitos.

O cariomorfo 2n=48 do ribeirão Bandeirinhas, única população amostrada fora das bacias adjacentes ao arco do Paranaíba, além de representar a descrição de um novo número diplóide para *P. rivularis* é o primeiro cariótipo descrito para a bacia do rio das Velhas, local do holótipo da espécie. Rocha *et al.* (2019) observou que alguns espécimes de *P. rivularis* dessa localidade apresentam uma alta distância genética do gene mitocondrial citocromo oxidase b (Cyt b) em relação às populações adjacentes ao Arco do Paranaíba e deve corresponder a uma nova espécie. Partindo do pressuposto que este cariomorfo foi amostrado na bacia do rio das Velhas, localidade tipo da espécie, sugerimos que este é o cariótipo representativo de *P. rivularis*, e os demais cariomorfos deste trabalho e da literatura (MOREIRA-FILHO, BERTOLLO, 1991; BIAVATI, MAISTRO, 2007; PERES, BERTOLLO, MOREIRA-FILHO, 2008) correspondem a diferentes ESUs de *P. rivularis*.

3 CAPÍTULO 2 - VARIAÇÃO DA FORMA DO CORPO NO CARACÍDEO *Psalidodon rivularis* DO RIO SÃO FRANCISCO, SUDESTE BRASILEIRO

Igor Henrique Rodrigues-Oliveira, Karine Frehner Kavalco, Rubens Pasa

RESUMO

Psalidodon rivularis é uma espécie de lambaris endêmica da bacia do Alto rio São Francisco e trabalhos envolvendo citogenética e biologia molecular vem demonstrando que representa um complexo de espécies. O objetivo do presente trabalho foi identificar diferenças morfológicas na forma corporal de 6 populações de *P. rivularis* e 1 população de *P. aff. rivularis* da bacia do Médio rio São Francisco (córrego Crico, sub-bacia do rio Paracatu), através da morfometria geométrica. Ao todo foram fotografados 174 indivíduos na face direita do corpo e sobre cada imagem foram digitalizados 17 landmarks. Definimos como classificadores pontos de coleta, sub-bacias hidrográficas e sexos. Para estudar os efeitos da alometria na forma realizamos a análise de Regressão e, para estudar a modulação da forma em diferentes pontos de coleta e sub-bacias, as análises de CVA e Neighbor-Joining. Encontramos diferenças na forma entre pontos de coleta e sub-bacias hidrográficas, associadas à altura corporal relativa e recuo da placa sub-orbital, além de uma influência significativa do tamanho na forma dos espécimes (alometria) associada a robutez ventral crânio, órbitas e placa sub-orbital. Não visualizamos diferenças na forma corporal entre machos e fêmeas. Vários trabalhos com peixes relacionam a altura corporal com a velocidade da água, enquanto o recuo da placa sub-orbital mostra um potencial taxonômico ou ecológico, marcando a principal diferença das populações do Alto rio São Francisco e do Médio rio São Francisco.

Palavras-chave: alometria, altura corporal, morfometria geométrica, placa sub-orbital.

3.1 Introdução

A subfamília Stethaprioninae (Characiformes: Characidae) inclui dois gêneros conhecidos por seus complexos de espécies: *Astyanax* Baird & Girard, 1854 e *Psalidodon* Eigenmann, 1911. Até recentemente, dois dos complexos de espécies mais conhecidos de *Psalidodon*, *P. scabripinnis* (MOREIRA-FILHO, BERTOLLO, 1991) e *P. fasciatus* (JUSTI, 1993), eram formalmente parte do gênero *Astyanax* (TERÁN, BENITEZ, MIRANDE, 2020). Ambos os gêneros têm sido utilizados em diversos trabalhos envolvendo morfometria geométrica nos últimos anos (AGUILAR-MEDRANO, 2013; CASTRO *et al.*; ORNELAS-GARCÍA, BASTIR, DOADRIO, 2014; Castro *et al.*; SANTOS, ARAÚJO, 2015; POWERS *et al.*, 2017; PERAZZO *et al.*, 2018; ROCHA *et al.*, 2019; DE MESQUITA ALVES *et al.*; DA SILVA *et al.*; DELARIVA, NEVES, 2020). Trabalhos com o gênero *Astyanax* evidenciaram como a forma corporal do grupo varia em função de pressões seletivas do meio, como a velocidade da água e a existência de fluxo gênico entre rios e áreas represadas (SANTOS, ARAÚJO, 2015; DA SILVA *et al.*, 2020), e como a forma varia em função de mudanças nos hábitos tróficos associada à ocupação de diferentes nichos ecológicos (ORNELAS-GARCÍA, BASTIR, DOADRIO, 2014; DELARIVA, NEVES, 2020).

O gênero *Psalidodon*, por sua vez, tem mostrado grande potencial em estudos envolvendo a variação da forma associada a características genéticas (ROCHA *et al.*, 2019; DE MESQUITA ALVES *et al.*, 2020). Rocha *et al.* (2019) demonstraram a correspondência entre dados moleculares e variação da forma corporal em diferentes linhagens mitocondriais de *P. rivularis* e *P. paranae*. O estudo também reforçou a validade taxonômica entre as duas espécies, tanto através dos dados moleculares quanto morfométricos.

Abordagens envolvendo marcadores citogenéticos em conjunto com morfometria geométrica foram utilizadas por Castro *et al.* (2014) para inferir um processo de especiação alopátrica em duas populações de *P. scabripinnis* de diferentes altitudes do mesmo rio, isoladas por uma cachoeira e geograficamente próximas. Além de ser uma ferramenta importante no estudo de espécies crípticas, a técnica também demonstrou a existência de dimorfismo sexual associado à forma corporal em pelo menos 4 populações do complexo *P. scabripinnis* (CASTRO *et al.*, 2014; 2015).

No presente trabalho apresentamos a caracterização de 6 populações de *P. rivularis* da bacia do Alto rio São Francisco e uma população de *P. aff. rivularis* do Médio rio São Francisco. Nossos objetivos eram demonstrar possíveis diferenças na forma corporal

associadas aos diferentes ambientes dos pontos de coleta, Sub-bacias hidrográficas, sexos e a influência da alometria na forma. Também buscamos realizar inferências fenéticas com base nos resultados das análises morfométricas.

3.2 Materiais e Métodos

Amostragem

Os espécimes utilizados neste estudo encontram-se depositados na Coleção de Vertebrados da Universidade Federal de Viçosa - *Campus* de Rio Paranaíba. A amostra inclui espécimes a partir de sete populações (Figura 1, Tabela 1). Todas as coletas foram feitas de acordo com licenças emitidas pelo SISBIO (Sistema de Autorização e Informação em Biodiversidade, número de licença 1938128) e SISGEN (Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado, número de licença A9FE946) atribuídas a Rubens Pasa.

Coletamos os espécimes através de pesca ativa, com o auxílio de varas de pesca, tarrafas, peneiras e puçá. Após a coleta, transportamos os indivíduos em baldes de água limpa contendo bombinhas de ar comprimido até o Laboratório de Genética Ecológica e Evolutiva da Universidade Federal de Viçosa - *Campus* de Rio Paranaíba, onde foram aclimatados e armazenados em aquários de 100 litros com água filtrada e aeradores, e alimentados diariamente até o dia da eutanásia.

Realizamos a eutanásia através de imersão em Eugenol 1% diluído em 1L de água, conforme a Resolução Normativa N°37 do Conselho Nacional de Controle e Experimentação Animal, do Ministério da Ciência, Tecnologia, Inovações e Comunicações (CONCEA - MCTI). Após a eutanásia cada indivíduo recebeu um número de tombo, foram fixados em Formaldeído 10% por 48 horas e armazenados em Álcool 70%.

Figura 1: (A) Mapa dos pontos de coleta amostrados nas análises de morfometria. 1- rio Funchal, 2 - Usina Abaeté, 3 - córrego Lage, 4 - córrego da Espinha, 5 - rio Vereda Grande, 6 - ribeirão Bandeirinhas, 7 - córrego Crico. Em verde: sub-bacia do rio Abaeté, em rosa: sub-bacia do rio Paracatu, em laranja: sub-bacia do rio Indaiá, em vermelho: sub-bacia do rio das Velhas, em azul claro: barragem de Três Marias/rio São Francisco. (B) Mapa do Brasil destacando a bacia do rio São Francisco em preto, retângulo vermelho delimitando a bacia do Alto rio São Francisco.

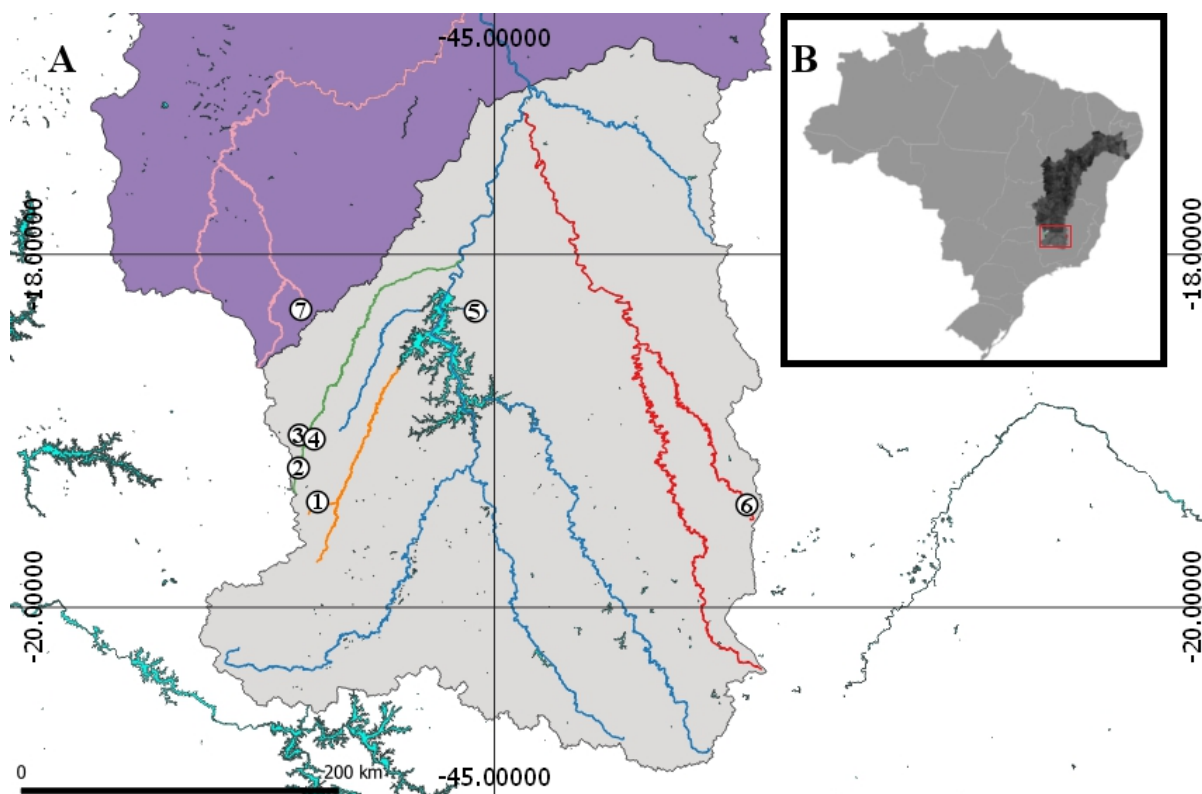


Tabela 1: Pontos de coleta usados no trabalho, sub-bacias e bacias hidrográficas a qual pertencem, coordenadas e “N” amostral.

Pontos de coleta	Sub-bacia	Bacia	Coordenadas	N amostral
Lagoa da Usina Abaeté	rio Abaeté	Alto rio São Francisco	19°12'35.69"S 46°6'33.79"O	25
córrego Lage	rio Abaeté	Alto rio São Francisco	19°1'25.18"S 46°6'18.74"O	18
córrego da Espinha	rio Abaeté	Alto rio São Francisco	19°2'47.18"S 46° 1'12.93"O	25
rio Funchal	rio Indaiá	Alto rio São Francisco	19°24'9.54"S 46°0'4.61"O	17
ribeirão Bandeirinhas	rio das Velhas	Alto rio São Francisco	19°25'8.33"S 43°34'12.37"O	28
rio Vereda Grande	rio São Francisco (Três Marias)	Alto rio São Francisco	18°19'18.62"S 45°6'32.80"O	26
córrego Crico	rio Paracatu	Médio rio São Francisco	18°18'44.36"S 46° 5'44.46"O	18

Caracterização dos pontos de coleta

Os pontos de amostragem no presente trabalho pertencem a cinco sub-bacias hidrográficas diferentes (Tabela 1). Todos os pontos de coleta pertencem à bacia do Alto rio São Francisco, com exceção do córrego Crico, que pertence à bacia do Médio rio São Francisco. Uma caracterização de cada ponto de coleta segue abaixo:

Lagoa da Usina do Abaeté (Rio Paranaíba, MG): o ponto de coleta se encontra logo abaixo da Barragem Usina do rio Abaeté localizada no município vizinho de São Gotardo. O córrego proveniente da barragem deságua em uma lagoa através de uma estreita porém alta queda d'água. No caminho até a queda d'água o córrego percorre um terreno rochoso formando várias pequenas piscinas naturais onde espécimes de *P. rivularis* podem ser coletados com extrema facilidade. Durante a época de cheia o volume de água vindo da barragem sobe drasticamente ocasionando o desaparecimento das piscinas naturais e transformando a queda d'água em uma cachoeira. O local apresenta mata ciliar abundante e o ambiente é predominantemente lântico em toda sua extensão, com exceção dos meses de cheia, onde a velocidade da água aumenta muito no trecho entre a barragem e a lagoa, mas estas continuam fornecendo um ambiente lântico para os peixes durante todo ano.

Córrego Lage (Arapuá, MG): o ponto em questão é um córrego com correnteza de vazão média a rápida, contendo várias locas e poços fundos onde a velocidade da água diminui e fundo barroso. A vegetação ripária é abundante em uma das margens do córrego enquanto que na outra margem alguns trechos estão afetados por pastagens. O local é muito usado por pescadores locais.

Córrego da Espinha (Tiros, MG): apesar de se localizar próximo a fazendas e pastagens, o córrego da Espinha possui vegetação ripária abundante em ambas as suas margens. O córrego possui a correnteza mais rápida dentre os locais amostrados, com várias pequenas quedas d'água e terreno predominantemente pedregoso.

Rio Funchal (São Gotardo, MG): o ponto de coleta se localiza abaixo de uma ponte. Em uma das margens do rio a vegetação ripária é constante enquanto que na outra ela se alterna com pastos, principalmente no trecho próximo à ponte. O ambiente é extremamente heterogêneo, com profundidade variando de alguns centímetros a alguns metros e fundo variando de cascalho, nos locais mais rasos, a arenoso, tanto nos locais rasos quanto fundos. A correnteza é predominantemente rápida, embora alguns locais fundos e próximo às

margens forneçam ambientes de correnteza mais lenta. Assim como o córrego Lage, o local é frequentado costumeiramente por pescadores.

Ribeirão Bandeirinhas (Parque Nacional Serra do Cipó - Santana do Riacho, MG): a amostragem dos espécimes foi realizada em um local conhecido como Cânion dos Bandeirantes, e apesar de estar localizado em uma Unidade de Conservação e se tratar de um ambiente conservado e pouco antropizado, o ponto de coleta em questão não apresenta Mata ciliar, uma vez que as margens do corpo d'água são as rochas do cânion, tendo apenas pouca vegetação arbustiva entre as rochas. Apesar de apresentar água corrente, a velocidade da água no cânion é lenta, e a profundidade varia bastante, chegando a vários metros em alguns trechos.

Rio Vereda Grande (Três Marias, MG): assim como o Ribeirão Bandeirinhas, o local, apesar de apresentar certa correnteza, possui uma velocidade da água bem lenta, sendo um afluente da Usina Hidrelétrica de Três Marias. Apesar de apresentar vegetação ripária, a área está sob ameaça de ações antrópicas, devido à proximidade com pastos e com uma rodovia, que inclusive passa por cima do rio próximo ao ponto coletado.

Córrego Crico (Presidente Olegário, MG): único ponto amostrado fora da bacia do Alto rio São Francisco, o local da coleta possui fundo arenoso, com lâmina d'água bem baixa (cerca de 20cm), correnteza média a baixa e mata ciliar bem preservada. O ponto se localiza no distrito de Galena que é conhecido por possuir várias cavernas calcárias, inclusive no entorno do ponto de coleta.

Digitalização e análise dos dados

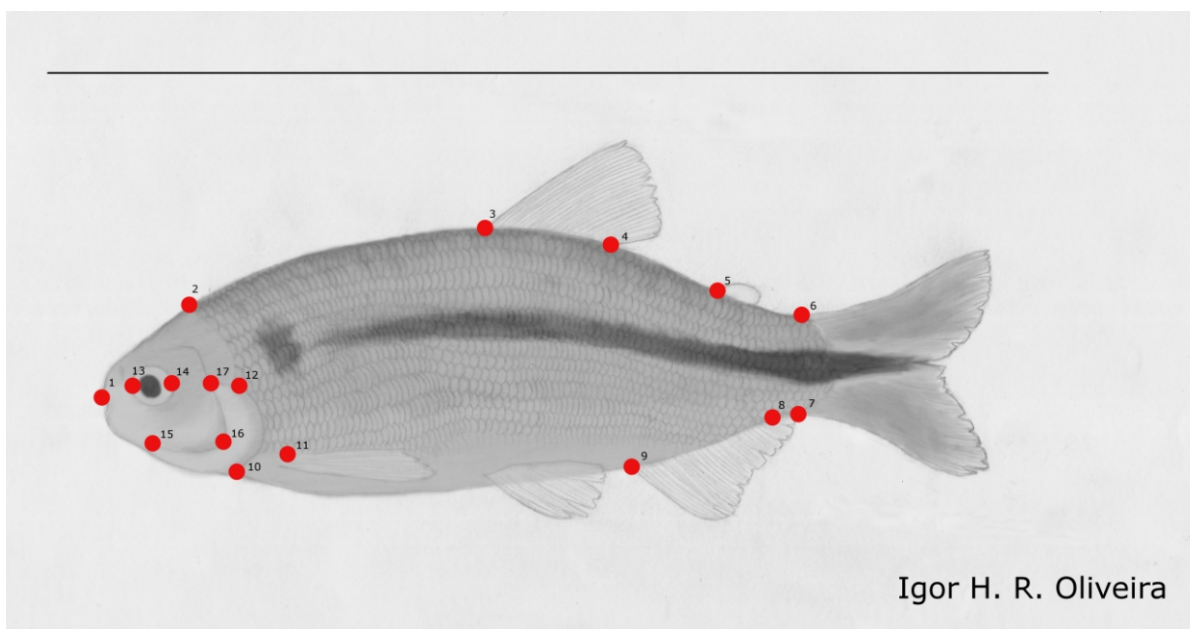
Utilizamos 174 espécimes no estudo. Depositamos os indivíduos sobre uma lona cinza e fotografamos na face direita do corpo com o auxílio de um celular Xiaomi Redmi Note 8 (resolução de 48 MP e zoom 1.8x) acoplado a um tripé e voltado em um eixo perpendicular em relação aos espécimes. A distância focal entre câmera e objeto foi padronizada para todas as fotografias. Junto aos espécimes colocamos uma barra de escala de 20 cm e uma etiqueta contendo nome da espécie, número de tombo e ponto de coleta.

Após digitalizadas, agrupamos as imagens em um arquivo no formato .tps pelo software TPSUtil 1.64 (ROHLF, 2015). Em cada espécime digitalizamos 17 landmarks para

inclusão em um sistema de coordenadas X/Y com o fator de escala calibrado em 1 cm (Figura 2). Conduzimos a digitalização dos landmarks no software TPSDig2 2.32 (ROHLF, 2016).

Conduzimos as análises morfométricas no software MorphoJ 1.06 (KLINGENBERG, 2011). Estabelecemos dois classificadores: ponto de coleta e sub-bacia hidrográfica. No classificador para sub-bacias agrupamos as populações dos córregos Lage e Espinha e a Lagoa da Usina do Abaeté como rio Abaeté, e os demais pontos de coleta foram renomeados de acordo com a sub-bacia ao qual pertencem (rios Indaiá, das Velhas, São Francisco ou Paracatu).

Figura 2: Landmarks utilizados no estudo: 1 - Ponta do focinho, 2- Crista supraoccipital, 3 - inserção anterior da nadadeira dorsal, 4 - inserção posterior da nadadeira dorsal, 5 - inserção da nadadeira adiposa, 6 - inserção dorsal da nadadeira caudal, 7 - inserção ventral da nadadeira caudal, 8 - inserção posterior da nadadeira anal, 9 - inserção anterior da nadadeira anal, 10 - margem ventral do opérculo, 11 - inserção da nadadeira peitoral, 12 - margem média do opérculo, 13 - margem anterior da órbita, 14 - margem posterior da órbita, 15 - origem ventral da placa sub-orbital, 16 - curva da placa sub-orbital, 17 - margem dorsal da placa sub-orbital. Barra de escala: 10 cm.



Fonte: Igor Henrique Rodrigues Oliveira.

Para retirada dos efeitos de posição, orientação e tamanho das amostras, e acessar as coordenadas de forma, conduzimos uma Análise Generalizada de Procrustes (GPA, ROHLF, SLICE, 1990). A fim de verificar a normalidade dos dados e identificar previamente possíveis outliers, contrastamos a curva de distribuição de distâncias com a curva de distribuição normal multivariada, através da opção *Find Outliers*. Verificamos a hipótese de nulidade dos dados para ambos os classificadores através da Análise de Variância de Procrustes (Procrustes ANOVA, KLINGENBERG, MCINTYRE, 1998).

A fim de se verificar os efeitos da alometria na amostra, e remover a influência do tamanho na forma para análises posteriores, conduzimos uma análise de regressão multivariada das coordenadas de Procrustes (variável dependente) vs tamanho do centroide transformado em log (variável independente) (KLINGENBERG, 2016). Incluímos um teste de permutação com 10 mil rodadas a fim de se testar a hipótese de independência entre forma e tamanho do centroide.

Para visualizar a variação morfológica entre pontos de coleta e sub-bacias hidrográficas, conduzimos análises discriminantes utilizando Análises de Variação Canônica (CVAs) para cada fator de grupo. Em ambas as análises foi incluído um teste de permutação para hipótese nula de média de grupos iguais com 10 mil rodadas.

Através distâncias de Mahalanobis entre pares de grupos obtidas nas CVAs, construímos matrizes de distâncias que foram importadas no software Paup 4.0 (SWOFFORD, 2003) para construção de fenogramas, utilizando o algoritmo Neighbor-Joining (NJ). Em seguida, importamos os fenogramas no MorphoJ 1.06 (KLINGENBERG, 2011) e utilizamos a opção *Map Onto Phylogeny* para mapear os dados das pontuações de variáveis canônicas e coordenadas de procrustes nos fenogramas e assim visualizar como os ramos se organizam no morfo-espço.

Criamos um segundo conjunto de dados, após o procedimento de GPA, apenas com os indivíduos com sexo identificado (através de observação das gônadas após abertura dos espécimes). Submetemos esse conjunto de dados às análises de Regressão, Procrustes ANOVA e Análise de Componentes Principais (PCA) com o intuito de se verificar a existência ou não de dimorfismo sexual relacionada à forma nos espécimes.

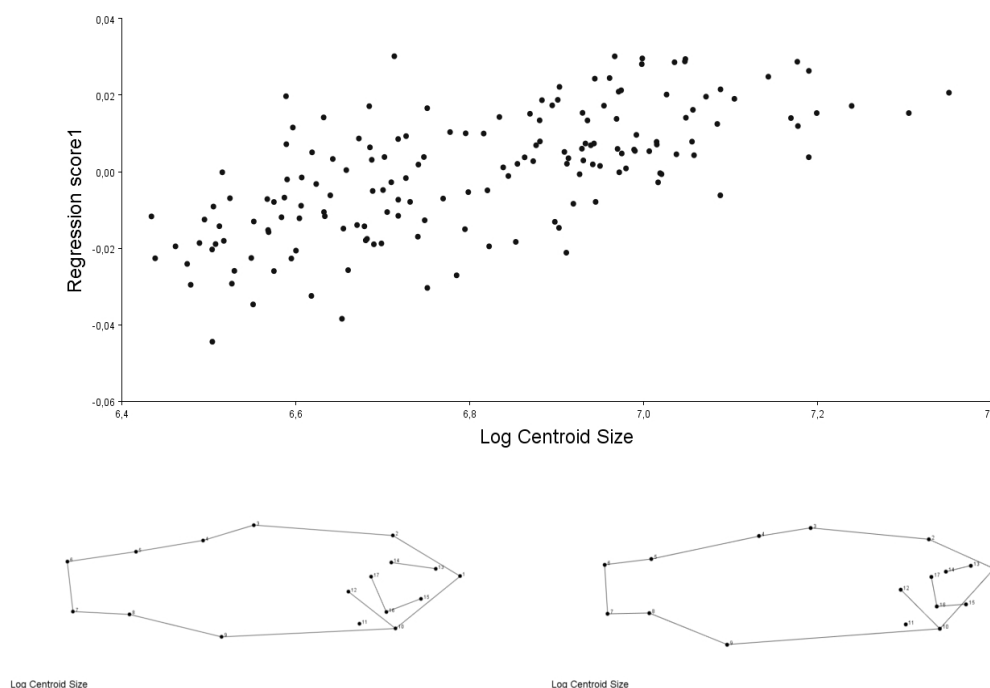
3.3 Resultados

Os testes da Procrustes ANOVA rejeitaram a hipótese de nulidade em ambas as análises, sugerindo diferença significativa na forma entre pontos de coleta ($SS = 0,11386$; $MS = 0,00063$; $df = 180$; $F = 9,66$; e $p < 0,0001$) e sub-bacias hidrográficas ($SS = 0,07912$; $MS = 0,00066$; $df = 120$; $F = 9,17$ e $p < 0,0001$).

A análise de Regressão evidenciou uma possível influência do tamanho na forma dos espécimes ($R^2 = 4,4517\%$; $p\text{-value} < 0,0001$). Através dos wireframes, foi possível notar que

essa diferença parece estar concentrada no crânio dos espécimes, com os menores indivíduos amostrados apresentando a órbita, a placa sub-orbital e o próprio crânio maior em relação ao corpo que os maiores indivíduos (Figura 3).

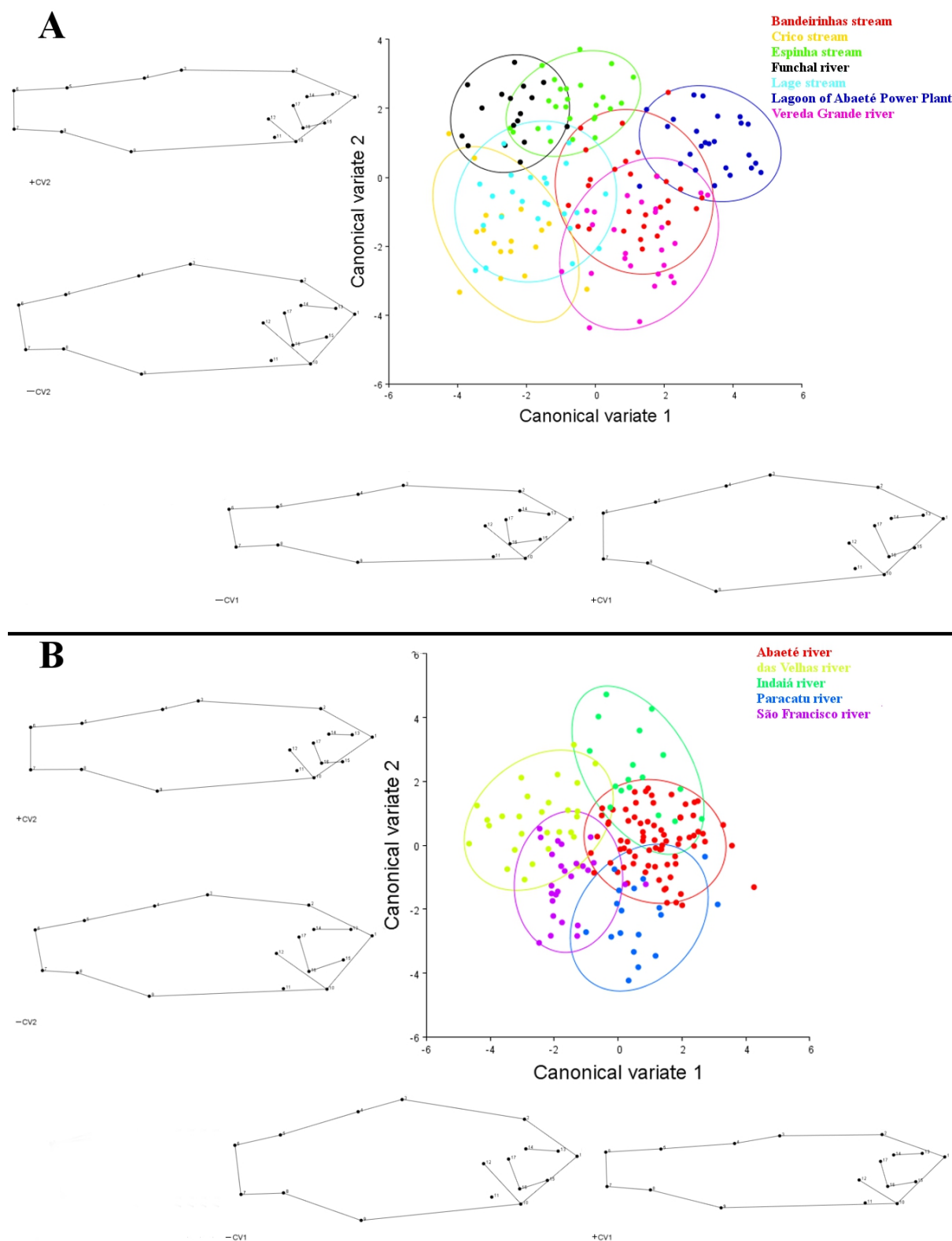
Figura 3: Gráfico de distribuição dos escores do efeito alométrico (acima). Wireframes com as formas encontradas nos extremos da variável independente (Log Centroid Size) (abaixo).



A CVA entre pontos de coleta demonstrou que cada população possui uma localização limitada no morfo-espaço, com as populações do córrego Lage e ribeirão Bandeirinhas apresentando maior distribuição dos escores. A hipótese de nulidade foi rejeitada para todos os pares de comparação ($p < 0,05$). Foi possível notar que tanto a CV1 (41,2% da variação) quanto a CV2 (22,28%), estavam relacionadas a altura do corpo e recuo da placa sub-orbital, apresentando diferentes combinações desses caracteres (Figura 4A).

A CVA entre sub-bacias hidrográficas apresentou uma clara estruturação entre cada sub-bacia ($p\text{-value} < 0,01$). A CV1 (41,36% da variação) demonstrou estar principalmente relacionada a altura do corpo, com as sub-bacias dos rios São Francisco e das Velhas (- CV1) apresentando uma maior altura do corpo em relação às sub-bacias dos rios Abaeté, Indaiá e Paracatu (+ CV1). Por outro lado, a CV2 (28,19%) parece estar relacionada a robustez na região ventral do crânio, recuo da placa sub-orbital e diâmetro da órbita, com as sub-bacias dos rios Paracatu e São Francisco apresentando uma região anterior mais robusta, placa sub-orbital mais recuada em direção ao opérculo e maior diâmetro da órbita (Figura 4B).

Figura 4: (A) Análise de variação canônica (CVA) para pontos de coleta. Wireframes contendo as formas nos extremos de cada variável canônica e gráfico de distribuição de escores dos Pontos de coleta. (B) Análise de variação canônica (CVA) para Sub-bacias Hidrográficas. Wireframes contendo as formas nos extremos de cada variável canônica e gráfico de distribuição de escores das Sub-bacias.



Distinguimos três agrupamentos no fenograma construído a partir das distâncias de Mahalanobis da CVA por pontos de coleta. Com base nos wireframes, construídos a partir das distâncias de Procrustes, foi possível observar que esses agrupamentos estão principalmente relacionados à altura no corpo dos espécimes, com um dos agrupamentos apresentando os locais com maior altura (Lagoa da Usina do Abaeté, rio Vereda Grande e ribeirão Bandeirinhas), outro com altura média (córrego Lage e córrego Crico) e outro com menor altura (córrego da Espinha e rio Funchal) (Figura 5A).

No fenograma para sub-bacias foi possível distinguir dois agrupamentos, um composto pelas populações de maior altura corporal (rio São Francisco e rio das Velhas), e outro para sub-bacias do alto Paranaíba (rio Abaeté e rio Indaiá). A sub-bacia do rio Paracatu ficou em uma posição isolada no fenograma. Através dos wireframes, foi possível visualizar que a sub-bacia do Paracatu se distingue das demais pelo alto grau de recuo placa sub-orbital, maior diâmetro da órbita e crânio mais robusto especialmente na região ventral (Figura 5B).

Para as análises de variação de forma entre os sexos a hipótese de nulidade da Procrustes ANOVA foi corroborada ($p = 0,2367$) e não foi observada separação morfo-espacial entre sexos na PCA, ao contrário do que foi relatado para outras espécies de *Psalidodon* (CASTRO *et al.*, 2014, 2015) (Figura 6).

3.4 Discussão e Conclusão

O presente estudo trouxe considerações sobre a variação da configuração corporal entre 6 pontos de coleta de *P. rivularis* e um ponto de coleta de *P. aff. rivularis* pertencentes a 5 sub-bacias hidrográficas diferentes do Médio e Alto rio São Francisco. Nós visualizamos variações na forma relacionada a alometria, altura corporal e perfil cefálico. Também demonstramos a importância da utilização da placa sub-orbital no estudo da forma em *P. rivularis*.

Vários estudos têm demonstrado como as pressões ambientais têm forte influência na modulação da forma em espécies historicamente atribuídas ao gênero *Astyanax* (ORNELAS-GARCÍA, BASTIR, DOADRIO, 2014; SANTOS, ARAÚJO, 2015; e POWERS *et al.*, 2017). Contudo a presença do fluxo gênico também representa um papel importante na modulação da forma restringindo a estruturação entre diferentes ambientes (SANTOS, ARAÚJO, 2015).

Figura 5: (A) Fenograma construído a partir do algoritmo Neighbor-Joining para Pontos de coleta no morfo-espaço da CVA, a direita pontos de coleta com água parada ou lenta onde foram encontrado espécimes com maior altura corporal proporcional ao corpo, a esquerda inferior pontos de coleta com velocidade da água intermediária onde foram encontrados espécimes de altura corporal proporcional intermediária e a esquerda superior pontos de coleta com correnteza rápida onde foram encontrados espécimes de menor altura corporal proporcional ao corpo. Wireframes construídos a partir das distâncias de procrustes de cada ponto de coleta. (B) Fenograma construído a partir do algoritmo Neighbor-Joining para Sub-bacias Hidrográficas no morfo-espaço da CVA, a direita sub-bacias onde foram amostrados espécimes com maior altura corporal proporcional ao corpo, a esquerda inferior sub-bacia do rio Paracatu do Médio rio São Francisco e a esquerda superior sub-bacias adjacentes ao Arco do Paranaíba. Wireframes construídos a partir das distâncias de procrustes de cada Sub-bacia.

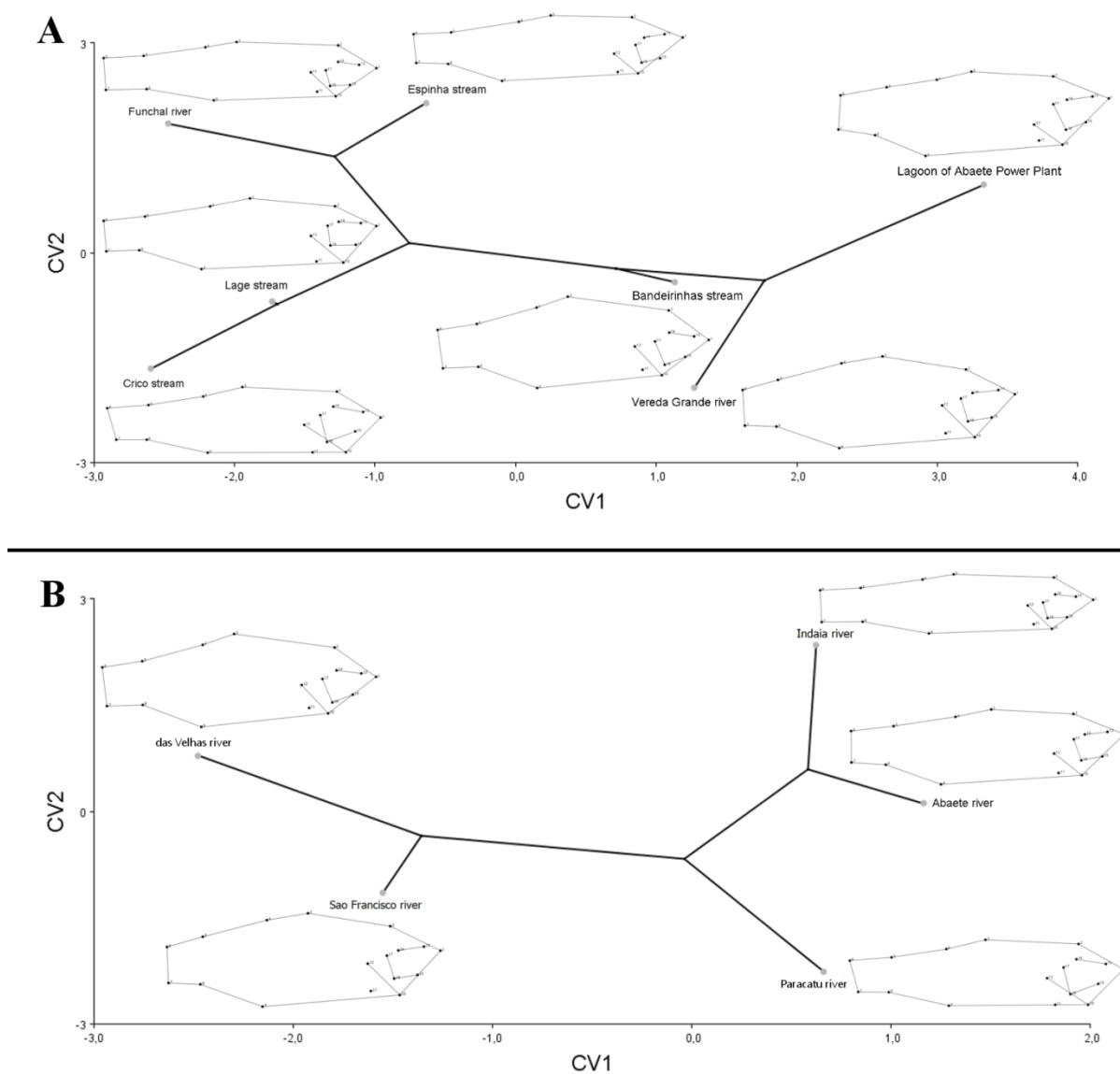
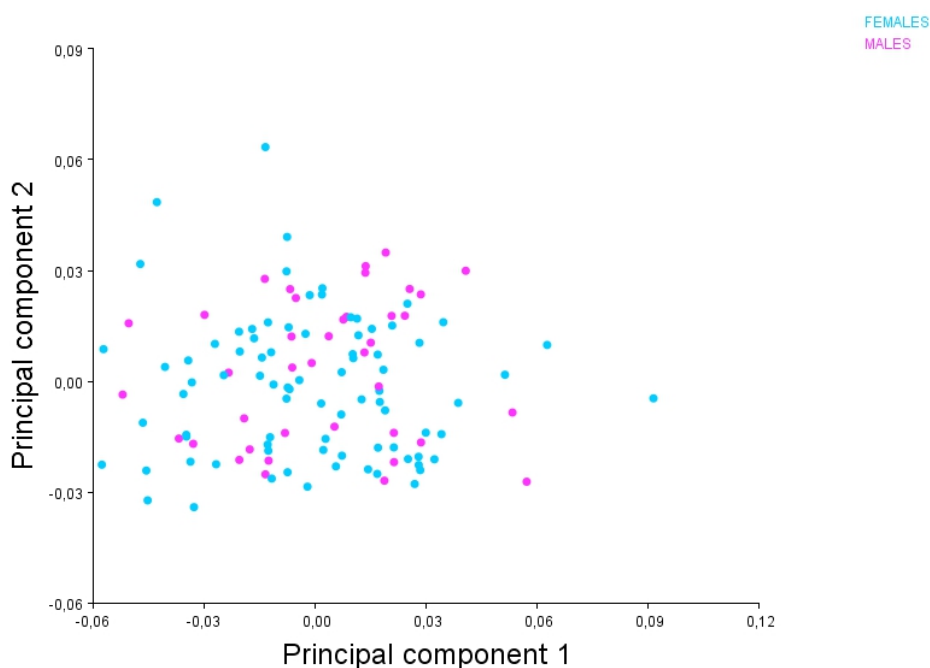


Figura 6: Análise de componentes principais (PCA) entre sexos, azul identificando fêmeas e rosa identificando machos.



Apesar da quantidade limitada de estudos publicados utilizando morfometria geométrica em *P. rivularis* (ROCHA *et al.*, 2019; DA SILVA *et al.*, 2020), é possível notar diferentes cenários para a espécie, onde fatores genéticos e ecológicos atuam juntos e em diferentes níveis na modulação da forma entre populações.

A análise de regressão demonstrou uma influência positiva da alometria na forma dos espécimes, essa influência ocorre principalmente na forma e tamanho do crânio em relação ao corpo que foi maior em espécimes menores (Figura 3). Mudanças na forma do crânio em função do estágio de crescimento já foram relatadas em populações cavernícolas de espécies próximas a *P. rivularis*, como *A. mexicanus*, com a assimetria do crânio só aparecendo na fase adulta após a ossificação do mesmo (POWERS *et al.*, 2017).

A variação da altura corporal dentro do gênero *Astyanax*, e mesmo em outros grupos de peixes, tem sido atribuída a fatores ecológicos e comportamentais (WEBB, 1984; AGUILAR-MEDRANO, 2013; SANTOS, ARAÚJO, 2015; DA SILVA *et al.*; DELARIVA, NEVES, 2020). Normalmente ambientes lóticos favorecem um corpo mais fusiforme, reduzindo o arrasto e o gasto de energia e favorecendo o nado prolongado, enquanto que ambientes lênticos selecionam um corpo mais alto, favorecendo explosões súbitas de velocidade e maior manobrabilidade (ALEXANDER, 1967; WEBB, WEIHS, 1986).

No presente trabalho encontramos uma grande variação na altura corporal nas sete localidades amostradas. Essa variação foi o fator determinante na topografia do fenograma para pontos de coleta (Figura 4A). É possível que a velocidade da água tenha sido um fator determinante para isso, já que os locais da Lagoa da Usina do Abaeté, rio Vereda Grande e ribeirão Bandeirinhas, agrupadas no fenograma, possuem as maiores alturas corporais do estudo e vem de ambientes lênticos ou de correnteza lenta, e as populações do rio Funchal e córrego da Espinha, também agrupadas no fenograma, representam as duas com a menor altura corporal e foram coletadas nos ambientes de correnteza mais rápida (SANTOS, ARAÚJO, 2015; ROCHA *et al.*, 2019; DA SILVA *et al.*, 2020).

O último agrupamento parece abranger as populações de correnteza e altura corporal intermediária entre os dois extremos observados neste estudo (Figura 5A). Como cada um dos pontos de coleta da sub-bacia do rio Abaeté ficou em um agrupamento diferente no fenograma e ao lado de populações com características ambientais similares, é muito provável que os agrupamentos observados reflitam processos de homoplasia em populações sujeitas a pressões ambientais semelhantes (DA SILVA *et al.*, 2020).

Modificações na forma da região anterior do corpo, principalmente atreladas ao crânio, têm sido relatadas em vários estudos envolvendo os gêneros *Astyanax* e *Psalidodon* (CASTRO *et al.*; ORNELAS-GARCÍA, BASTIR, DOADRIO, 2014; CASTRO *et al.*, 2015; POWERS *et al.*, 2017). Diferenças na forma do crânio foram atribuídas a diferenças nos hábitos tróficos em diferentes morfotipos de *A. aeneus* (ORNELAS-GARCÍA, BASTIR, DOADRIO, 2014) e espécies simpátricas de *Herichthys* (PERES-MIRANDA *et al.*, 2020). Da mesma forma, variações na posição do focinho e olho, e tamanho do perfil cefálico, foram relatadas no gênero de peixe marinho *Chromis*, e parecem estar relacionadas à posição que esses peixes ocupam na coluna da água (AGUILAR-MEDRANO, 2013).

Neste trabalho observamos que, depois da altura corporal, modificações na região cefálica foram as maiores responsáveis pela variação da forma entre diferentes pontos de coleta e até sub-bacias (Figura 4). Essa variação parece estar principalmente concentrada no recuo da placa sub-orbital (landmarks 16 e 17) e em modificações na região ventral do crânio (landmarks 10 e 15) (Figura 4). Curiosamente, os trabalhos feitos com morfometria geométrica da morfologia externa de *Astyanax/Psalidodon* não exploraram a fundo a placa sub-orbital na delimitação de landmarks (CASTRO *et al.*; ORNELAS-GARCÍA, BASTIR,

DOADRIO, 2014; CASTRO *et al.*; SANTOS, ARAÚJO, 2015; ROCHA *et al.*, 2019; DE MESQUITA ALVES *et al.*; DA SILVA *et al.*, 2020).

Essas modificações são o principal aspecto que diferencia a população do córrego Crico, pertencente a sub-bacia do rio Paracatu, que apresentam órbita maior, região ventral do crânio mais robusta e placa sub-orbital mais recuada em direção ao opérculo (Figura 5B). No fenograma de sub-bacias hidrográficas, a sub-bacia do rio Paracatu ocupou uma posição isolada em relação às demais. Uma vez que essa amostra é a única que não pertence a bacia do Alto rio São Francisco no estudo, suas particularidades podem então significar que as populações do Médio São Francisco constituam uma unidade taxonômica diferente da existente na bacia adjacente, com uma possível característica diagnóstica relacionada a placa sub-orbital, embora uma maior amostragem na região seja necessária para confirmar tal suposição.

Através das considerações já apresentadas, concluímos que altura corporal e modificações no crânio são os principais responsáveis pela variação da forma em *P. rivularis*, com o primeiro fator sendo provavelmente modulado por pressões ambientais como a velocidade da água. Os landmarks da placa sub-orbital (15, 16 e 17) revelaram-se relevantes ao se estudar a variação da forma, sobretudo entre populações de diferentes bacias, e podem ter implicações filogenéticas e até taxonômicas no grupo. Dessa forma, é aconselhável explorar diferentes landmarks nessa região em futuros trabalhos envolvendo *Astyanax* e *Psalidodon*.

4 CAPÍTULO 3 - O MITOGENOMA E O SATELITOMA DE DOIS CITÓTIPOS (2n=46 E 2n=50) DE *Psalidodon rivularis* (CHARACIFORMES: CHARACIDAE)

Igor Henrique Rodrigues-Oliveira, Rubens Pasa, Fabiano Bezerra Menegídio, John Seymour (Pat) Heslop-Harrison, Trude Schwarzacher, Karine Frehner Kavalco

RESUMO

A descrição do genoma mitocondrial tem grande aplicabilidade na montagem de filogenias e no conhecimento da diferenciação molecular das espécies. Devido suas rápidas dinâmicas evolutivas, o estudo de DNAs satélites pode ter grande aplicabilidade no rastreamento de rearranjos cromossômicos. Dessa forma, tanto o mitogenoma quanto o satelitoma são ferramentas eficientes no estudo da evolução molecular. Neste trabalho nós apresentamos a montagem do genoma mitocondrial completo e realizamos a prospecção *In Silico* do satelitoma de dois citótipos de *Psalidodon rivularis*. Enviamos o DNA de dois exemplares (2n=46 e 2n=50 cromossomos) para sequenciamento em um Novaseq 6000, plataforma Illumina. Ambos os citótipos apresentaram mitogenomas com composição semelhantes, e a variação de tamanho entre ambos se deu principalmente pelo número de repetições de uma sequência de 35pb encontrada na região controle ou D-loop. O motivo da repetição variou em 19 sítios entre ambos os citótipos e a distância genética para o genoma mitocondrial completo foi de 1,76%. O número de sequências repetitivas encontradas no satelitoma do citótipo de 2n=50 foi bem maior que no 2n=46 cromossomos (66 contra 41). O citótipo 2n=46 apresentou 19 sequências não encontradas no 2n=50, seis sem similaridade alguma no banco de dados, já o 2n=50 apresentou 43 sequências não encontradas no 2n=46 cromossomos, 12 sem similaridade no banco de dados. As diferenças tanto no satelitoma quanto no mitogenoma entre os dois citótipos sugerem que ambos, na verdade, correspondem a diferentes espécies crípticas do grupo *P. rivularis*.

Palavras-chave: Bioinformática, especiação, genômica, satDNA, mtDNA

4.1 Introdução

A descrição do genoma mitocondrial de uma espécie é fundamental para a montagem de filogenias e para reconstruir a evolução molecular em um grupo. O sequenciamento de nova geração e montagem de genomas completos, em especial genomas mitocondriais, têm se tornado cada vez mais comuns nos últimos anos (SMITH, 2015) e provavelmente seja a forma mais rápida na atualidade de resgatar genomas mitocondriais completos.

Psalidodon rivularis é uma espécie endêmica da bacia hidrográfica do Alto rio São Francisco historicamente atribuída ao complexo “*Astyanax scabripinnis*”, conhecido por sua diversidade cariotípica, com citótipos variando de $2n=46$ a $2n=50$ cromossomos (MOREIRA-FILHO, BERTOLLO, 1991). Após uma ampla revisão taxônomica de Terán, Benitez e Mirande (2020), a maior parte das espécies do complexo *A. scabripinnis*, incluindo *A. rivularis*, foram movidas para o gênero *Psalidodon*. As espécies mais próximas de *P. rivularis* com genoma mitocondrial descrito são *P. paranae* (SILVA *et al.*, 2016) e *P. fasciatus* (CALEGARI *et al.*, 2019). Além destas, outras espécies próximas, também incluídas na tribo Gymnocharacini (Characiformes: Characidae: Stethaprioninae), com genoma mitocondrial descrito são: *Astyanax mexicanus* (Nakatani *et al.*, 2011), *A. altiparanae* (CALEGARI *et al.*, 2019; DE CAMPOS *et al.*, 2020) e *Deuterodon giton* (BARRETO *et al.*, 2017).

Assim como o DNA mitocondrial, o estudo do DNA satélite também pode fornecer informações importantes sobre a evolução molecular, tanto no que diz respeito a micro-evolução (UGARKOVIC, PLOHL, 2002) quanto a macro-evolução (PONS, PETITPIERRE, JUAN, 2002; PONS, GILLESPIE, 2003; KANTEK *et al.*, 2009). Esse tipo de DNA constitui a maior parte da heterocromatina constitutiva em genomas eucariotos (PLOHL, MEŠTROVIĆ, MRAVINAC, 2012), e é caracterizado por sequências únicas repetidas em tandem várias vezes, onde a unidade de repetição não apresenta outros tipos de sequência, como ocorre com genes ribossomais (RUIZ-RUANO *et al.*, 2016). Graças às suas rápidas dinâmicas evolutivas, esses elementos têm grande potencial como marcador citogenético (SAITO, EDPALINA, ABE, 2007; KANTEK *et al.*, 2009; VICARI *et al.*, 2010; PAZZA *et al.*, 2018).

O termo DNA satélite foi cunhado devido ao padrão que esses DNAs apresentavam após centrifugação em gradiente de alta velocidade do DNA, já que por serem abundantes e terem composição de bases diferentes do resto do DNA, formavam bandas satélites (GALL,

PARDUE; JOHN, BIRNSTIEL, JONES, 1969; BISCOTTI, OLMO, HESLOP-HARRISON., 2015). Esses elementos estão entre os grupos de DNAs repetitivos de evolução mais rápida no genoma (RUIZ-RUANO *et al.*, 2016; DUÍLIO *et al.*, 2017) e podem ser organizadas em diferentes famílias de acordo com as variações em suas sequências de bases, tamanho do motivo de repetição, cobertura e localização no genoma (GARRIDO-RAMOS, 2017). O conjunto, ou biblioteca, de famílias de DNA satélites pode ou não ser compartilhado entre espécies próximas, o que é conhecido como “hipótese da biblioteca” (RUIZ-RUANO *et al.*, 2016; GARRIDO-RAMOS, 2017).

Até o momento, o único satDNA de que temos conhecimento para *P. rivularis* é o satDNA As51 (MESTRINER *et al.*, 2000), um DNA satélite rico em pares de bases AT, presente em populações de $2n=50$ de *P. rivularis* e ausente em populações de $2n=46$ (ABEL, MANTOVANI, MOREIRA-FILHO, 2006). Além disso, a única espécie do gênero *Psalidodon* com o satelitoma descrito é *P. paranae* (DUÍLIO *et al.*, 2017). Alguns dos satélites descritos para esta espécie também foram encontrados em *P. fasciatus* e *P. bockmanni* (DUÍLIO *et al.*, 2017). Entre os demais representantes da tribo Gymnocharacini, *A. altiparanae* também teve seu satelitoma descrito (DE CAMPOS *et al.*, 2020).

Neste trabalho nós apresentamos a montagem do genoma mitocondrial completo e estudo *in Silico* do satelitoma de dois citótipos ($2n=46$ e $2n=50$) de *P. rivularis*, amostrados na mesma sub-bacia hidrográfica (Sub-bacia do rio Abaeté). Realizamos a prospecção de diferentes satDNAs compartilhados, exclusivas com similaridade e exclusivas sem similaridade em ambos os citótipos, que requerem validação futura via hibridização fluorescente *In Situ* (FISH).

4.2 Materiais e Métodos

Extração e sequenciamento do DNA

Selecionamos para sequenciamento o DNA extraído a partir do fígado de dois exemplares machos adultos de *P. rivularis* previamente cariotipados ($2n=46$: $8m+18sm+10st+10a$ e $2n=50$: $8m+24sm+8st+10a$) armazenados no Banco de Tecidos, DNA e Suspensão do Laboratório de Genética Ecológica e Evolutiva da Universidade Federal de Viçosa *Campus* Rio Paranaíba, sob os números de voucher: LAGEEVO-2030 e LAGEEVO-

2614 respectivamente. O DNA foi extraído utilizando-se kit apropriado e seguindo as instruções do fabricante (Invitrogen).

Verificamos a qualidade do DNA extraído através de eletroforese em gel de Agarose 1% e a pureza em nanodrop. Em seguida, enviamos as amostras para sequenciamento em um Novaseq 6000 (Illumina, São Diego, CA) na empresa Novogene, Reino Unido. Foram geradas duas bibliotecas de 2x150bp para cada citótipo.

Montagem do genoma mitocondrial dos dois citótipos

Montamos o mitogenoma para ambos os citótipos a partir de leituras brutas no Novoplasty v3.7 (DIERCKXSENS, MARDULYN, SMITS, 2017) em um cluster paralelo de computadores (64 Gb de RAM), usando o mito-genoma de *Psalidodon paranae* disponível no GenBank (SRR5461470) como seed (suplemento). Então, anotamos as sequências no MitoAnnotator (IWASAKI *et al.*, 2013) no servidor MitoFish (<http://mitofish.aori.u-tokyo.ac.jp>).

Importamos os dados brutos 2x150 das sequências de ambos os citótipos na plataforma on-line Galaxy (AFGAN *et al.*, 2018) através do servidor público *usegalaxy.eu* (<http://usegalaxy.eu>), onde verificamos a qualidade das leituras através do FastQC (Babraham Bioinformatics) e realizamos a filtragem com a ferramenta Fastp (CHEN *et al.*, 2018).

Acessamos a região repetitiva do mitogenoma, presente no D-loop (região controle), com o Tandem Repeats Finder (<https://tandem.bu.edu/trf>) (BENSON, 1999). Alinhamos as sequências FASTA com o ClustalW (THOMPSON, HIGGINS, GIBSON, 1994) e calculamos a distância p no software MEGA X (KUMAR *et al.*, 2018). O mitogenoma para ambos os citótipos estão disponíveis no GenBank com os códigos de acesso MT428069 para o 2n=46 e MT428070 para o 2n=50.

Montagem do satelitoma dos dois citótipos

A fim de acessarmos o satelitoma, importamos os dados brutos obtidos pelo sequenciamento dos genomas completos de *P. rivularis* 2n=46 e *P. rivularis* 2n=50 na plataforma online Galaxy Europe (AFGAN *et al.*, 2018) no servidor público *usegalaxy.eu*

(<http://usegalaxy.eu>), onde verificamos a qualidade das leituras através do FastQC (Babraham Bioinformatics) e filtramos com a ferramenta Fastp (CHEN *et al.*, 2018) com todos os parâmetros no *default*.

Após os procedimentos de corte de qualidade baixamos os arquivos gerados e importamos os dados resultantes da filtragem na plataforma online Repeat Explorer (NOVÁK, NEUMANN, 2013), um servidor público baseado em Galaxy próprio para análises de elementos repetitivos, como DNAs satélites (<https://repeatexplorer-elixir.cerit-sc.cz/galaxy/>).

Iniciamos o pré-processamento das leituras Fastq emparelhadas utilizando as ferramentas do RepeatExplorer (*RepeatExplorer Utilities: Preprocessing of FASTQ paired-end reads*) com todos os parâmetros no *default*. Assim foram gerados arquivos entrelaçados após corte, filtragem de qualidade, filtragem da adaptação de corte e entrelaçamento, e descarte dos pares quebrados. Em seguida, utilizamos a pipeline Tarean (*Tandem Repeat Analysis*) (NOVÁK *et al.*, 2017) com base nos arquivos entrelaçados para identificação de repetições genômicas em tandem.

Comparamos os DNAs satélites de alta confiança, baixa confiança, LTRs (*Long Terminal Repeats*) e rDNAs obtidos pela Tarean com os DNAs presentes nos bancos de dados Blast (*Basic Local Alignment Search Tool*) do NCBI, com *Word Size* = 16, e dFAM (HUBLEY *et al.*, 2016). Consideramos bons alinhamentos sequências com e-value menor que $1e-5$ (DE CAMPOS *et al.*, 2020). Avaliamos a similaridade entre algumas sequências dos dois citótipos com tamanhos iguais/semelhantes e que não apresentaram similaridade alguma nos bancos de dados, alinhando ambas no Blast e anotando seus valores de cobertura, identidade e e-value.

Nós nomeamos as sequências de satélites obtidos da seguinte forma: Pri (abreviação de *Psalidodon rivularis*) 46 ou 50 (identificação correspondente ao número diplóide de cada citótipo) Sat (caso a sequência não teve alta similaridade com outros tipos de DNA repetitivo, como rDNA) XX (número de cluster em ordem decrescente de frequência no genoma) - xx (número de pares de bases da sequência) (ex: Pri46SatXX-xx).

4.3 Resultados

Dados brutos

O sequenciamento do genoma de *P. rivularis* 2n=46 resultou em 121.066.528 Megabases de reads com comprimento médio de 150pb, 18.159.979 Gigabases de bases nitrogenadas e 39,89% de conteúdo GC. Após as filtragens de qualidade esses números caíram para 119.264.232 Megabases de reads com comprimento médio de 148pb, 17.761.347 Gigabases de bases nitrogenadas e 39,65% de conteúdo GC.

O sequenciamento do genoma de *P. rivularis* 2n=50 resultou em 75.515.128 Megabases de reads com comprimento médio de 150pb, 11.327.269 Gigabases de bases nitrogenadas e 39,3% de conteúdo GC. Após as filtragens de qualidade esses números caíram para 74.549.690 Megabases de reads com comprimento médio de 149pb, 11.123.436 Gigabases de bases nitrogenadas e 39,14% de conteúdo GC.

Genoma mitocondrial

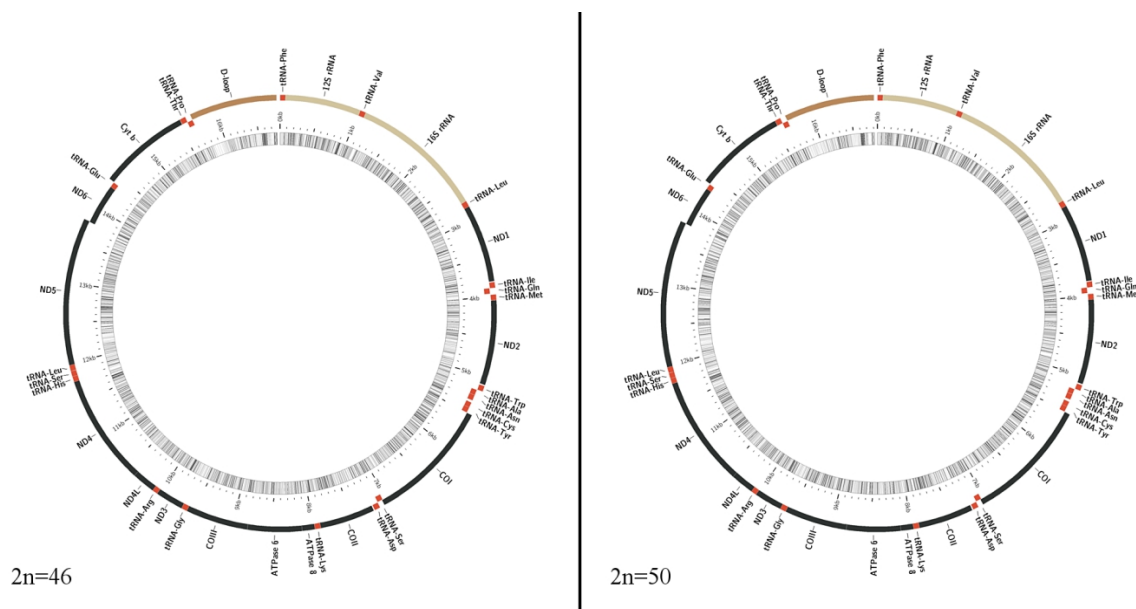
Ambos os genomas foram idênticos em conteúdo e ordem dos genes, apresentando 13 genes codificadores de proteína (PCGs), 22 genes de tRNA, 2 genes de rRNA e uma região controle (D-loop) (Figura 1). A organização encontrada nos mito-genomas está de acordo com a descrita previamente nas espécies relacionadas *P. paranae* (SILVA *et al.*, 2016) e *P. fasciatus* (CALEGARI *et al.*, 2019).

A composição de bases nitrogenadas nos mitogenomas foi: 42,9% GC; 29,5% A; 27,6% T; 27% C e 15,9% G no 2n=46 cromossomos e 42,8% GC; 29,6% A; 27,6% T; 27% C; e 15,8% G no 2n=50 cromossomos.

Ambos os citótipos apresentaram as PCGs com o mesmo tamanho, porém todas as PCGs variaram na posição de início e fim por um nucleotídeo. A razão disso, é que o gene rRNA 12S, o segundo no mitogenoma atrás do gene tRNA-Phe, é um nucleotídeo maior em *P. rivularis* com 2n=46 cromossomos (949pb contra 948pb no 2n=50).

Todos os genes de RNAs, com exceção do já citado rRNA 12S, possuíram o mesmo tamanho em ambos citótipos, e assim como ocorreu com as PCGs, todos, exceto o tRNA-Phe, variaram na posição de início e fim por um nucleotídeo em razão da variação de tamanho no rRNA 12S.

Figura 1: Representação gráfica da reconstrução do mito-genoma de ambos os citótipos.



Nós encontramos codificação sobreposta entre os genes tRNA-Gln e tRNA-Ile, tRNA-Cys e tRNA-Tyr, COI e tRNA-Ser (primeira cópia), ATP6 e ATP8, ND4L e ND4, ND5 e ND6, e tRNA-Thr e tRNA-Pro.

Visualizamos os códons atípicos UAA no tRNA-Leu, UGA no tRNA-Ser (primeira cópia), GCU no tRNA-Ser (segunda cópia) e UAG no tRNA-Leu (segunda cópia). Também visualizamos códons de parada incompletos no COII, ATP6, COIII, ND3 e ND4.

Encontramos uma repetição de 35pb no D-loop de ambos os citótipos. O D-loop foi o principal responsável pela variação de tamanho entre os genomas, apresentando 1.102pb no 2n=46 e 1.136pb no 2n=50 cromossomos. Os motivos da repetição, encontrados através do alinhamento, variaram em 19 sítios (em negrito e sublinhados) entre os citótipos:

2n=46: TAT**GCATAATTATACATATATGTATTAGTACATAT**

2n=50: TA**ATTATACATATATGTATTAGTACATATTATGCA**

Todos os PCGs, exceto o gene ND6, foram de cadeia pesada. Todos os tRNA, exceto oito (tRNA-Gln, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, tRNA-Ser (primeira cópia), tRNA-Glu e tRNA-Pro), também foram de cadeia pesada. O tamanho do mito-genoma foi 16.779pb no 2n=46 e 16.812pb no 2n=50. A distância genética calculada entre os dois citótipos foi 1,76%.

Satelitoma

Visão geral do satelitoma dos citótipos de $2n=46$ e $2n=50$ cromossomos

Ao comparar os dois satelitomas encontramos diversas similaridades, o que era esperado para grupos relacionados. Houve 19 casos em que duas sequências alinharam com uma mesma sequência no banco de dados, sendo em sete ocasiões com satélites de *P. paranae* (DUÍLIO *et al.*, 2017) (Tabela 1).

Localizamos dois pares de sequências nos dois citótipos que não apresentaram nenhuma similaridade no banco de dados mas que possuíam similaridade o suficiente para serem consideradas da mesma família de satDNA: Pri46Sat138-27 e Pri50Sat256-27, Pri46Sat116-22 e Pri50Sat110-22. Também localizamos dois pares de sequências nos dois citótipos que não apresentaram nenhuma similaridade no banco de dados mas que possuíam similaridade o suficiente para serem consideradas a mesma variante de satDNA: Pri46Sat198-43 e Pri50Sat147-43, Pri46Sat237-190 e Pri50Sat198-191 (Tabela 2).

Apesar das semelhanças já mencionadas, algumas diferenças entre os satelitomas foram notáveis. O número de DNAs repetitivos recuperados para o citótipo de $2n=50$ foi consideravelmente maior que o citótipo $2n=46$ (66 contra 41 sequências). Além disso, se tomarmos como base o $2n=46$, mesmo com menos sequências recuperadas este citótipo apresentou 19 sequências não encontradas no $2n=50$, seis destas sem similaridade alguma no banco de dados. Já o $2n=50$ apresentou 43 sequências não encontradas no $2n=46$, 12 sem similaridade no banco de dados (Tabela 3).

Satelitoma do citótipo $2n=46$ cromossomos

O tamanho dos DNAs repetitivos em *P. rivularis* com $2n=46$ variou de 22pb a 8.910pb. Nós obtivemos 17 satélites putativos de alta confiança, 22 satélites putativos de baixa confiança, 1 LTR e 1 rDNA. Das 17 sequências de alta confiança, quatro não apresentaram similaridade no banco de dados, e das 22 sequências de baixa confiança, seis não apresentaram similaridades no banco de dados.

No Blast, nós obtivemos alinhamentos com 26,67% dos DNAs satélites descritas no satelitoma de *P. paranae* (DUÍLIO *et al.*, 2017), uma sequência de rDNA 5S de *P. fasciatus* (HASHIMOTO *et al.*, 2011), o satDNA MS7 descrito em *Moenkhausia sanctaefilomenae*

(UTSUNOMIA *et al.*, 2016), o satDNA As51 descrito em *P. scabripinnis* (MESTRINER *et al.*, 2000), o satDNA MmaSat049-1709 de *Megaleporinus macrocephalus* (UTSUNOMIA *et al.*, 2019), dois lncRNAs de *A. mexicanus* (PASQUIER *et al.*, 2016) e uma região contendo os rDNAs 18S/5.8S/28S e inter-espaçadores de *Cyprinus carpio* (VERA *et al.*, 1997). Outras 20 sequências alinharam com sequências de mRNAs, ncRNAs e porções do genoma de diversas espécies.

Nós também identificamos dois satélites putativos de lncRNAs no genoma do *P. rivularis* $2n=46$. O Pri46Sat_ncRNA132-51 alinhou com cinco porções de um lncRNA de 690-858pb descrito no locus LOC111193191 do cromossomo 9 de *A. mexicanus* (PASQUIER *et al.*, 2016). Já o Pri46Sat_ncRNA258-96 alinhou com uma porção de um lncRNA de 762-764pb descrito no locus LOC111189946 e com uma porção dos lncRNAs de 848pb e 727pb descritos nos loci LOC111197525 e LOC111196590 de *A. mexicanus*, respectivamente (PASQUIER *et al.*, 2016).

No dFAM, seis das nossas sequências de satélites putativos de alta confiança apresentaram alinhamentos com 12 tipos tRNAs (tRNA-Gly, tRNA-Ile, tRNA-Ala, tRNA-Val, tRNA-Asp, tRNA-Arg, tRNA-Met, tRNA-His, tRNA-Thr, tRNA-Ser, tRNA-Tyr, tRNA-Glu), uma sequência de satélite putativo de baixa confiança alinhou com uma sequência de rRNA 5S, a sequência de rDNA putativo alinhou com rRNAs diversos e outras quatro sequências alinharam com diferentes famílias de elementos transponíveis.

Satelitoma do citótipo $2n=50$ cromossomos

No genoma de *P. rivularis* $2n=50$ o tamanho dos DNAs repetitivos variou de 21pb a 8.707pb. Encontramos 27 satélites putativos de alta confiança, 35 satélites putativos de baixa confiança, 3 LTRs e 1 rDNA. Das 27 sequências de alta confiança, 10 não apresentaram similaridade no banco de dados, e das 35 sequências de baixa confiança, seis não apresentaram similaridades no banco de dados.

No Blast, nós obtivemos alinhamentos com 31,11% dos DNAs satélites descritas no satelitoma de *P. paranae* (DUÍLIO *et al.*, 2017), o satDNA MS7 descrito em *Mo. sanctaefilomenae* (UTSUNOMIA *et al.*, 2016), quatro satDNAs descritos em *Me. macrocephalus* (UTSUNOMIA *et al.*, 2019), dois lncRNAs de *A. mexicanus* (PASQUIER *et al.*, 2016), o snRNA U2 de *Mo. sanctaefilomenae* (UTSUNOMIA, dados não publicados), o

snRNA U1 de *P. paranae* (UTSUNOMIA, dados não publicados) e uma região contendo os rDNAs 18S/5.8S/28S e inter-espaçadores de *Cyprinus carpio* (VERA *et al.*, 1997). Outras 35 seqüências alinharam com seqüências de mRNAs, ncRNAs e porções do genoma de espécies variadas.

Nós também identificamos três satélites putativos de lncRNAs no genoma de *P. rivularis* com $2n=50$. O Pri50Sat220_ncRNA-70 alinhou com cinco porções de um lncRNA de 3.786pb descrito no locus LOC111191672 de *A. mexicanus* (PASQUIER *et al.*, 2016). O Pri50Sat129_ncRNA-51 apresentou 15 alinhamentos com um lncRNA de 690-858pb descrito no locus LOC111193191 de *A. mexicanus* (PASQUIER *et al.*, 2016). Já o Pri50Sat186_ncRNA-236 alinhou com quatro porções de um lncRNA de 1.780-1.800pb além de duas de um lncRN de 2.597-3.200bp de *A. mexicanus* (PASQUIER *et al.*, 2016).

No dFAM, sete satélites putativos e um LTR putativo alinharam com 11 tipos de tRNAs (tRNA-Gly, tRNA-Ala, tRNA-Arg, tRNA-Met, tRNA-His, tRNA-Thr, tRNA-Asn, tRNA-Lys, tRNA-Ser, tRNA-Leu, tRNA-Pro), dois satélites putativos alinharam com os snRNAs U2/U5 e U1/U6, um satélite putativo de baixa confiança alinhou com um rRNA 5S, nosso rDNA putativo alinhou com seqüências de rRNAs diversos e outras 5 seqüências alinharam com diferentes famílias de elementos transponíveis.

4.4 Discussão e Conclusão

P. rivularis tem se mostrado um complexo de espécies sustentado com dados citogenéticos, moleculares e morfométricos (ROCHA; PASA *et al.*, 2019; RODRIGUES-OLIVEIRA *et al.*, em preparação). Mantendo essa problemática em mente, utilizamos dois diferentes citótipos para a montagem dos mitogenomas e satelitomas, e assim obtivemos importantes informações sobre os genomas do grupo em diferentes conjuntos diplóides. Além disso, ampliamos o conhecimento do genoma mitocondrial da tribo Gymnocharacini na bacia do Alto rio São Francisco, uma vez que dos mitogenomas descritos na tribo apenas *P. fasciatus* habita a bacia em questão (CALEGARI *et al.*, 2019).

Como já era esperado, a organização do genoma mitocondrial entre os citótipos é muito similar entre si e com mitogenomas de espécies próximas já descritos (NAKATANI *et al.*, 2011; SILVA *et al.*, 2016; BARRETO *et al.*, 2017; CALEGARI *et al.*, 2019; DE

CAMPOS *et al.*, 2020). Com exceção da região controle (D-loop), ambos os citótipos variaram em tamanho no mitogenoma em apenas 1 nucleotídeo.

O D-loop se demonstrou a região do mitogenoma mais variável entre os dois citótipos, inclusive no tamanho, com 34pb de diferença. A variação de tamanho no D-loop se deve principalmente a presença de uma sequência repetitiva com 35pb, que embora esteja presente em ambos os citótipos, varia em 19 sítios entre eles. Variações no tamanho do mitogenoma em decorrência de repetições no D-loop já foram relatados em outros complexo de espécies de peixes, como *Hypostomus ancistroides* (ROCHA REIS *et al.*, 2020). Essa variabilidade característica do D-loop ocorre devido essa região ser a principal região não codificadora do genoma mitocondrial, possuindo a função de regular a replicação e transcrição, a tornando passível de evolução mais rápida e maior variabilidade (CLAYTON, 1991; SHADEL, CLAYTON, 1997). Dessa forma, o D-loop representa uma região de potencial interesse em estudos com filogenias, comparação de genomas mitocondriais e evolução molecular de mitogenomas (ROCHA REIS *et al.*, 2020).

Uma vez que os DNAs satélites são apenas uma das várias classes de DNAs repetitivos, é natural que nossa análise tenha encontrados outros grupos de DNAs repetitivos, como rDNAs, algumas famílias de genes codificadores de proteínas e elementos transponíveis, como os LTRs (LÓPEZ-FLORES, GARRIDO-RAMOS, 2012). A detecção das sequências contendo o rDNA 18S em ambos os bancos de dados utilizados e com valores de e-value = 0 é congruente com o fato de que os organismos da subfamília Stethaproninae possuem múltiplos sítios desse marcador no genoma (PAZZA, DERGAM, KAVALCO, 2018).

Nós encontramos LTRs putativos em ambos os citótipos, assim como obtivemos alinhamentos com outros grupos de TEs no banco de dados dFAM (HUBLEY *et al.*, 2016). O conteúdo de elementos transponíveis (TEs) no genoma de diferentes grupos de peixes varia significativamente, mas estudos têm demonstrado uma correlação positiva entre certos grupos de TEs, em especial LTRs, com o tamanho do genoma em peixes (SHAO, HAN, PENG, 2019). Vários estudos têm correlacionado a atividade de TEs com formação de espécies e irradiação adaptativa (GRAHN *et al.*, 2005; OLIVER, GREENE, 2011; PLATT *et al.*, 2014; SHAO, HAN, PENG, 2019).

Algo notável nas nossas sequências obtidas são o número de sequências para ncRNAs, em especial lncRNAs, que são sequências de genes que expressam RNAs que não são

codificadas em proteínas. Ambos os citótipos apresentaram lncRNAs, sendo dois compartilhados entre ambos e um deles classificado como LTR pela análise do Tarean (NOVÁK *et al.*, 2017) (Tabela 1). Em peixes teleósteos, lncRNAs foram correlacionados com diversas funções biológicas, como desenvolvimento, diferenciação celular, resposta imunológica e regulação da expressão gênica (BASU *et al.*; BOLTAÑA *et al.*; WANG *et al.*, 2016; WANG *et al.*, 2018).

Algumas sequências que merecem destaques são os satélites putativos de lncRNAs Pri46Sat132_ncRNA-51, Pri46Sat258_ncRNA-96, Pri50Sat186_ncRNA-236, Pri50Sat119_ncRNA-51 e Pri50Sat220_ncRNA-70. Todas essas sequências curtas alinharam com diferentes segmentos de lncRNAs de *A. mexicanus* (PASQUIER *et al.*, 2016). Apesar de lncRNAs serem sequências grandes (mais de 200pb), muitos desses elementos possuem sequências repetitivas funcionais em seus domínios (MATTICK, 2003). Dessa forma, essas sequências podem ser satélites putativos de lncRNAs e serem um indicativo da presença desses elementos nos genomas de *P. rivularis*, embora isso ainda requeira validação.

As sequências Pri46Sat132_ncRNA-51 e Pri50Sat129_ncRNA-51, além da similaridade alta já mencionada com segmentos de lncRNAs de *A. mexicanus* (PASQUIER *et al.*, 2016), também demonstraram similaridade superior a 50% com os satélites AalSat24-51 de *A. altiparanae* (DE CAMPOS *et al.*, 2020) e ApaSat15-51 de *P. paranae* (DUÍLIO *et al.*, 2017). De Campos *et al.* (2020) propôs que tanto AalSat24-51 quanto ApaSat15-51 fossem possíveis sequências da família do As51 de *P. scabripinnis* (MESTRINER *et al.*, 2000). Entretanto, baseado nos limites de similaridade para classificação de sequências como mesma variante (95%), família (80%) ou superfamília (50%) utilizadas por Ruiz-Ruano *et al.* (2016) e Duílio *et al.* (2017), consideramos que essas sequências pertençam a uma superfamília distinta do As51, uma vez que possuem menos de 50% de similaridade com o As51 original (MESTRINER *et al.*, 2000), o ApaSat01-51 (DUÍLIO *et al.*, 2017) e Pri46Sat03-51. Desta forma, propomos a criação de uma nova superfamília de satélites putativos que inclui Pri46Sat132_ncRNA-51, Pri50Sat129_ncRNA-51, ApaSat15-51 e AalSat24-51, que nomearemos PriSat_ncRNA-51, e que possivelmente correspondem a satélites de lncRNAs nos grupos *Astyanax* e *Psalidodon*.

Embora o tamanho e organização dos genes mitocondriais tenha se mantido estável entre os dois citótipos, o mesmo não pode ser dito de seu conteúdo. A distância genética encontrada (1,76%) foi significativamente maior que a distância genética do gene

mitocondrial Citocromo Oxidase I (0,6%) e próxima da distância genética do gene ATPase 6/8 encontrada entre populações de *P. fasciatus* das bacias hidrográficas dos rios Paraná e São Francisco (PAZZA, CRUVINEL, KAVALCO, 2017), sendo consideravelmente alta considerando que ambos os citótipos de *P. rivularis* coexistem em uma pequena área geográfica. Aliada à diferença cariotípica, a distância genética se torna um forte indicativo que ambos os citótipos estão em isolamento reprodutivo, e correspondem a diferentes espécies crípticas com distribuição limitada a bacia do Alto rio São Francisco. É possível que a especiação e evolução em *P. rivularis* tenha tido como marco inicial a diversificação cromossômica, com a evolução molecular acontecendo posteriormente, tal como foi relatado para *P. fasciatus* (KAVALCO *et al.*, 2016).

Baseado no nosso objetivo de prospecção de satélites para a espécie *P. rivularis* e nos resultados obtidos, propomos a construção de primers para os satélites exclusivos de cada citótipo, assim como os exclusivos para *P. rivularis* compartilhado por ambos os citótipos, para validação futura via reação de cadeia em polimerase em tempo real (qPCR) e Hibridização Fluorescente *In Situ* (FISH). Através desses mecanismos de validação será possível compreender melhor a evolução molecular nos dois citótipos e no grupo *P. rivularis*.

4.5 Anexos.

Tabela 1: Sequências de satDNAs nos dois citótipos com correspondências na mesma sequência no banco de dados. Dados do Dfam possuem apenas e-value

satDNA 2n=46 (Cobertura/Identidade/e-value Blast) (e-values Dfam)	satDNA 2n=50 (Cobertura/Identidade/e-value Blast) (e-values Dfam)	Homologia
Pri46Sat04-184 (100%/94,23%/9e-57)	Pri50Sat59-177 (100%/93,88%/9e-32)	<i>Astyanax paranae</i> satellite ApaSat27-178 sequence (Blast)
Pri46Sat08-236 (100% / 94,63% / 1e-56)	Pri50Sat06-236 (100% / 84,17 % / 4e-17)	<i>Astyanax paranae</i> satellite ApaSat02-236 sequence (Blast)
Pri46Sat08-236 (100% / 94,63% / 1e-56)	Pri50Sat29-236 (99% / 89,89 % / 2e-59)	<i>Astyanax paranae</i> satellite ApaSat02-236 sequence (Blast)
Pri46Sat24_lncRNA-631 (44% / 77,86% / 3e-37)	Pri50Sat61_lncRNA-631 (44% / 77,64% / 1e-45)	<i>Astyanax mexicanus</i> uncharacterized LOC111190794 (LOC111190794), ncRNA (Blast)

Pri46Sat74-35 (88% / 100% / 4e-06)	Pri50Sat113-35 (94% / 100% / 3e-07)	<i>Astyanax paranae</i> satellite ApaSat08-35 sequence (Blast)
Pri46tRNA_Arg/Met/His-1251 (27% / 97,62% / 3e-29) (3,4e-25; 2,4e-24; 3,5e-14)	Pri50tRNA_Arg/Met/His-1251 (38% / 91,30% / 3e-32) (3,4e-25; 2,4e-24 e 2,6e-13)	<i>Danio rerio</i> genome assembly, chromosome: 4 (Blast) Vertebrata tRNA_Arg (CGY), tRNA_Met_v, tRNA_His (CAY) (Dfam)
Pri46Sat193_tRNA_Ala/Thr-548 (28% / 95,18% / 1e-25) (2,2e-23 e 3,5e-21)	Pri50Sat244_tRNA_Ala/Thr-552 (15% / 195% / 1e-25) (2,2e-23; 2,5e-06)	<i>Megaleporinus</i> <i>macrocephalus</i> satellite MmaSat049-1709 sequence (Blast) Vertebrata tRNA_Ala(GCG), tRNA_Thr(ACG) (Dfam)
Pri46lncRNA_tRNA_Ser/Ser-350 (61% / 96% / 1e-42) (1,5e-27 e 5,2e-05)	Pri50tRNA_Leu/Ser/Ser-563 (40% / 87,31% / 1e-54) (9,9e-19; 3e-05)	<i>Astyanax mexicanus</i> uncharacterized LOC111188592 (LOC111188592), ncRNA (Blast) Vertebrata tRNA_Ser(TCY), tRNA_Ser(TCG) (Dfam)
Pri46rDNA5S-195 (9,2e-35)	Pri50rDNA5S-195 (9,2e-35)	Vertebrata 5S rDNA gene / pseudogene (Dfam)
Pri46Sat83-107 (67% / 98,61% / 5e-26)	Pri50Sat135-107 (66% / 97,18% / 8e-24)	<i>Astyanax paranae</i> satellite ApaSat38-107 sequence (Blast)
Pri46Sat104-69 (97% / 98,51% / 1e-23)	Pri50Sat98-69 (97% / 98,51% / 1e-23)	<i>Astyanax paranae</i> satellite ApaSat12-69 sequence (Blast)
Pri46Sat121-103 (98% / 97,03% / 1e-37)	Pri50Sat109-106 92% / 93,88% / 2e-30	<i>Moenkhausia</i> <i>sanctaefilomenae</i> voucher LBP19830 69723 clone 18 satellite MS7 sequence (Blast)
Pri46Sat132_ncRNA-51 (100% / 96,08% / 4e-13) (56% / 100% / 3e-04)	Pri50Sat129_ncRNA-51 (100% / 96,09% / 4e-13) (92% / 100% / 3e-14)	<i>Astyanax mexicanus</i> uncharacterized LOC111193191 (LOC111193191), transcript variant X4, ncRNA (Blast) <i>Astyanax paranae</i> satellite ApaSat15-51_a sequence (Blast)
Pri46Sat147-42 (100% / 90,48% / 4e-05)	Pri50Sat82-42 (97% / 90,24% / 2e-04)	<i>Astyanax mexicanus</i> nectin-3-like protein (LOC111193783), mRNA (Blast)

Pri46Sat151-80 (85% / 88% / 3e-12)	Pri50Sat138-80 (85% / 88% / 3e-12)	<i>Astyanax mexicanus</i> mucin-2-like (LOC103026144), mRNA (Blast)
Pri46Sat172-52 (98% / 92,16% / 8e-10)	Pri50Sat276-52 (100% / 94,23% / 5e-12)	<i>Astyanax mexicanus</i> uncharacterized protein C1orf194 homolog (LOC103046376), mRNA (Blast)
Pri46Sat243-27 (100% / 100% / 5e-04)	Pri50Sat157-27 (100% / 100% / 5e-04)	<i>Astyanax paranae</i> satellite ApaSat25-27 sequence (Blast)
Pri46LTR_IncRNA- 721 (60% / 90,32% / 2e-157)	Pri50LTR_IncRNA-1998 (48% / 86,75% / 0)	<i>Astyanax mexicanus</i> (LOC111189805), transcript variant X2, ncRNA (Blast)
(9,7e-13; 3,1e-05)	(2,4e-39; 2,8e-30)	<i>Danio rerio</i> Helitron- 1N4_DR / DNA2-7_DR (Dfam)
Pri46rDNA18S/5.8S/28S01-8910 (22% / 99,02% / 0)	Pri50rDNA18S/5.8S/28S13-8708 (22% / 99,24% / 0)	<i>Cyprinus carpio</i> intergenic spacer, partial sequence; 18S ribosomal RNA gene, internal transcribed spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2, and 28S ribosomal RNA gene, complete sequence; and intergenic spacer, partial sequence (Blast)
(0; 0; 0; 9,9e-140; 0,0074)	(0; 0; 0; 5,9e-140)	Metazoa LSU-rRNA_Cel / SSU-rRNA_Hsa / LSU- rRNA_Hsa gene / pseudogene (Dfam)

Tabela 2: sequências exclusivas de *P. rivularis* compartilhadas entre os dois citótipos.

Sequências consideradas como variantes da mesma família ou como a mesma variante (*)				
2n=46	2n=50	cobertura	identidade	e-value
Pri46Sat138-27	Pri50Sat256-27	92%	96%	1e-10
Pri46Sat116-22	Pri50Sat110-22	77%	100%	2e-08
Pri46Sat198-43*	PriSat147-43*	100%	100%	8e-22
Pri46Sat237-190*	Pri50Sat198-191*	99%	98,94%	3e-98

Tabela 3: Sequências citótipo-exclusivas.

2n=46	2n=50
Pri46Sat190-53	Pri50Sat75-152
Pri46Sat262-31	Pri50Sat160-47
Pri46Sat214-52	Pri50Sat230-187
Pri46Sat245-143	Pri50Sat236-68
Pri46Sat256-59	Pri50Sat264-40
Pri46Sat260-602	Pri50Sat272-33
	Pri50Sat278-21
	Pri50Sat282-41
	Pri50Sat19-32
	Pri50Sat136-24
	Pri50Sat212-43
	Pri50Sat235-21

5 CONCLUSÕES GERAIS

A diversidade cariotípica observada neste trabalho é maior que a registrada previamente para a espécie *P. rivularis*, com a descrição de novos cariomorfos, um novo número diplóide ($2n=48$) e presença de cromossomos Bs. A maior parte das populações de *P. rivularis* da sub-bacia do rio Abaeté e rios vizinhos (Borrachudo e Indaiá) parece ter $2n=46$ ou $2n=50$ cromossomos. Além das diferenças cariotípicas, as diferenças no satelitoma e mitogenoma entre ambos os citótipos corroboram a existência de pelo menos duas ESUs nesta região. Os espécimes desta região também apresentam baixa altura corporal relativa, provavelmente em função da pressão ambiental exercida pelos ambientes lóticos em que os espécimes foram coletados.

A população do ribeirão Bandeirinhas (sub-bacia do rio das Velhas), coletada no Parque Nacional Serra do Cipó, apresenta não só um cariomorfo divergente a das populações adjacentes ao Arco do Paranaíba, com $2n=48$ cromossomos, como diferenças notáveis na configuração corporal, apresentando um corpo mais alto, que pode ser em função do ambiente com velocidade da água mais lenta em que os espécimes foram coletados, e modificações no recuo da placa suborbital. Como esta localidade está inserida na sub-bacia em que o holótipo da espécie foi descrita, sugerimos que essa população, além de representar uma ESU diferente das demais, seja a população representativa da espécie *P. rivularis*.

Nas análises morfométricas, a população do rio Vereda Grande se mostrou a de maior altura corporal, apresentando uma configuração corporal semelhante a do ribeirão Bandeirinhas. Apesar de não ter sido possível obter cromossomos mitóticos para esta população, ela é geograficamente próxima das populações analisadas por Moreira-Filho e Bertollo (1991) na região de Três Marias, que possuem $2n=50$ cromossomos e fórmula cariotípica bem distinta dos $2n=50$ deste trabalho. As diferenças cariotípicas, além das diferenças na forma corporal, são bem sugestivas da existência de outra ESU nos afluentes da barragem de Três Marias no rio São Francisco.

Também não foi possível obter cromossomos mitóticos para a população do córrego Crico (sub-bacia do rio Paracatu, bacia do Médio rio São Francisco), mas as análises morfométricas evidenciaram uma diferença na configuração corporal em relação às demais populações, associada principalmente ao recuo da placa sub-orbital em direção ao opérculo. Acreditamos que essa característica possa ter uma importância taxonômica em *P. rivularis*, marcando a principal diferença diagnóstica desta localidade em relação às demais.

Todas as ferramentas aqui utilizadas indicam uma diversidade genética e morfológica maior do que seria esperado para uma única espécie. Diferenças cariotípicas no número e morfologia dos cromossomos tem grande potencial de acarretar isolamento reprodutivo, o que pode impulsionar diferenças não só morfológicas, como genéticas inter e intra-populacionais. Dessa forma, sugerimos que *P. rivularis* corresponde a um complexo de espécies crípticas da bacia do Alto rio São Francisco, sendo necessários trabalhos que permitam a identificação e descrição das diferentes espécies do grupo.

6 REFERÊNCIAS

- ABEL, Luciano Douglas dos Santos; MANTOVANI, Monique; MOREIRA-FILHO, Orlando. **Chromosomal distribution of the As51 satellite DNA in two species complexes of the genus *Astyanax* (Pisces, Characidae)**. *Genetics and Molecular Biology*, v. 29, n. 3, p. 448-452, 2006.
- AFGAN, Enis *et al.* **The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update**. *Nucleic acids research*, v. 46, n. W1, p. W537-W544, 2018.
- AGUILAR-MEDRANO, Rosalía. **Body shape evolution of *Chromis* and *Azurina* species (Perciformes, Pomacentridae) of the eastern Pacific**. *Animal Biology*, v. 63, n. 2, p. 217-232, 2013.
- ALEXANDER, R. M. **Functional design in fishes**. Hutchinson Univ. Library, London, 1967.
- ALVES, Carlos Bernardo Mascarenhas; POMPEU, P. dos S. **A fauna de peixes da bacia do rio das Velhas no final do século XX. Peixes do Rio das Velhas: passado e presente**. Cap. v. 3, p. 165-187, 2001.
- BARRETO, Cynthia Aparecida Valiati *et al.* **Complete mitochondrial genome sequence of neotropical fish *Astyanax giton* Eigenmann 1908 (Ostariophysi; Characidae)**. *Mitochondrial DNA Part B*, v. 2, n. 2, p. 839-840, 2017.
- BASU, Swaraj *et al.* **The *Tetraodon nigroviridis* reference transcriptome: developmental transition, length retention and microsynteny of long non-coding RNAs in a compact vertebrate genome**. *Scientific reports*, v. 6, n. 1, p. 1-14, 2016.
- BENSON, Gary. **Tandem repeats finder: a program to analyze DNA sequences**. *Nucleic acids research*, v. 27, n. 2, p. 573-580, 1999.
- BERTACO, Vinicius A.; DE LUCENA, Carlos AS. **Two new species of *Astyanax* (Ostariophysi: Characiformes: Characidae) from eastern Brazil, with a synopsis of the *Astyanax scabripinnis* species complex**. *Neotropical Ichthyology*, v. 4, n. 1, p. 53-60, 2006.
- BERTOLLO, Luiz Antonio Carlos. **Cytotaxonomic considerations on *Hoplias lacerdae* (Pisces, Erythrinidae)**. *Brazil. J. Genet.*, v. 1, p. 103-120, 1978.
- BIAVATI, Sandra Helena; MAISTRO, Edson Luis. **Cytogenetical Analyses in Three *Astyanax scabripinnis* Populations (Pisces, Characidae) from Minas Gerais State, Brazil**. *Cytologia*, v. 72, n. 3, p. 357-361, 2007.
- BISCOTTI, M. A.; OLMO, E.; HESLOP-HARRISON, J. P. **Repetitive DNA in eukaryotic genomes**. *Chromosome Research*, v. 23, p. 415-420, 2015.
- BOLTAÑA, Sebastian *et al.* **Long noncoding RNAs (lncRNAs) dynamics evidence immunomodulation during ISAV-Infected Atlantic salmon (*Salmo salar*)**. *Scientific reports*, v. 6, n. 1, p. 1-13, 2016.
- BOOKSTEIN, Fred L. **A statistical method for biological shape comparisons**. *Journal of Theoretical Biology*, v. 107, n. 3, p. 475-520, 1984.

CALEGARI, Rodrigo Milan *et al.* **The complete mitochondrial genome sequences of five *Otophysi* species (Vertebrata, Teleostei).** Mitochondrial DNA Part B, v. 4, n. 2, p. 4198-4199, 2019.

CASTRO, R. M. C. **Evolução da ictiofauna de riachos sul-americanos: padrões gerais e possíveis processos causais.** [Evolution of the ichthyofauna of South American streams: General patterns and possible causal processes.]. 1999.

CASTRO, Jonathan P. *et al.* **Evidence of incipient speciation in *Astyanax scabripinnis* species complex (Teleostei: Characidae).** Neotropical Ichthyology, v. 12, n. 2, p. 429-438, 2014.

CASTRO, Jonathan Pena *et al.* **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, v. 25, n. 1, p. 231-244, 2015.

CAMPOS, José Eloi Guimarães; DARDENNE, Marcel Auguste. **Origem e evolução tectônica da Bacia Sanfranciscana.** Revista Brasileira de Geociências, v. 27, n. 3, p. 283-294, 2017.

CHEN, Shifu *et al.* **fastp: an ultra-fast all-in-one FASTQ preprocessor.** Bioinformatics, v. 34, n. 17, p. i884-i890, 2018.

CLAYTON, David A. **Nuclear gadgets in mitochondrial DNA replication and transcription.** Trends in biochemical sciences, v. 16, p. 107-111, 1991.

DA SILVA, Marcos Aurélio *et al.* **Shape diversity of the fish genus *Astyanax* Baird & Girard, 1854 (Teleostei, Characidae) in adjacent basins.** Biologia, p. 1-9, 2020.

DE CAMPOS, Matheus Lewi Cruz Bonaccorsi *et al.* **O mitogenoma e o satelitoma de *Astyanax altiparanae* Garutti & Bristiski, 2000: considerações iniciais.** 2020.

DE MESQUITA ALVES, Rosana *et al.* **A new species of genus *Psalidodon* Eigenmann, 1911 related to the *P. paranae* complex (Characiformes Characidae) from Upper Paranaíba river basin, Brazil, supported by genetic and morphometric data.**

DELARIVA, Rosilene Luciana; NEVES, Mayara Pereira. **Morphological traits correlated with resource partitioning among small characin fish species coexisting in a Neotropical river.** Ecology of Freshwater Fish, v. 29, n. 4, p. 640-653, 2020.

DIERCKXSENS, Nicolas; MARDULYN, Patrick; SMITS, Guillaume. **NOVOPlasty: de novo assembly of organelle genomes from whole genome data.** Nucleic acids research, v. 45, n. 4, p. e18-e18, 2017.

DUILIO, MZ de A. *et al.* **Delimiting the origin of a B chromosome by FISH mapping, chromosome painting and DNA sequence analysis in *Astyanax paranae* (Teleostei, Characiformes).** PLoS One, v. 9, n. 4, p. e94896, 2014.

DUÍLIO, MZ de A. *et al.* **High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*.** Scientific reports, v. 7, n. 1, p. 1-12, 2017.

E SANTOS, Hersília de Andrade; DOS SANTOS POMPEU, Paulo; KENJI, Danilo Okuma Lessa. **Changes in the flood regime of São Francisco River (Brazil) from 1940 to 2006.** Regional Environmental Change, v. 12, n. 1, p. 123-132, 2012.

FERREIRA-NETO, Maressa *et al.* **Three sympatric karyomorphs in the fish *Astyanax fasciatus* (Teleostei, Characidae) do not seem to hybridize in natural populations.** Comparative Cytogenetics, v. 6, n. 1, p. 29, 2012.

FERNANDES, Carlos Alexandre; MARTINS-SANTOS, Isabel Cristina. **Sympatric occurrence of three cytotypes and four morphological types of B chromosomes of *Astyanax scabripinnis* (Pisces, Characiformes) in the River Ivaí Basin, state of Paraná, Brazil.** Genetica, v. 124, n. 2, p. 301-306, 2005.

FERNANDES, Carlos Alexandre; MARTINS-SANTOS, Isabel Cristina. **Mapping of the 18S and 5S ribosomal RNA genes in *Astyanax altiparanae* Garutti & Britski, 2000 (Teleostei, Characidae) from the upper Paraná river basin, Brazil.** Genetics and Molecular Biology, v. 29, n. 3, p. 464-468, 2006.

FORNEL, Rodrigo; CORDEIRO-ESTRELA, Pedro. **Morfometria geométrica e a quantificação da forma dos organismos.** Temas em Biologia: Edição comemorativa aos, v. 20, p. 101-120, 2012.

GALL, J. G.; PARDUE, M. L. **Formation and detection of RNA-DNA hybrid molecules in cytological preparations.** Proceedings of the National Academy of Sciences of the United States of America, v. 63, p. 378–383, 1969.

GARRIDO-RAMOS, Manuel A. **Satellite DNA: an evolving topic.** Genes, v. 8, n. 9, p. 230, 2017.

GODINHO, Alexandre Lima; GODINHO, Hugo Pereira. **Breve visão do São francisco.** Águas, peixes e pescadores do São Francisco das Minas Gerais. Belo Horizonte: PUC Minas, v. 468, p. 15-23, 2003.

GOECKS, Jeremy; NEKRUTENKO, Anton; TAYLOR, James. **Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences.** Genome biology, v. 11, n. 8, p. 1-13, 2010.

GOODALL, Colin. **Procrustes methods in the statistical analysis of shape.** Journal of the Royal Statistical Society: Series B (Methodological), v. 53, n. 2, p. 285-321, 1991.

GOODWIN, S.; McPHERSON, J. D.; MCCOMBIE, W. R. **Coming of age: ten years of next generation sequencing technologies.** Nature Reviews Genetics, v. 17, p. 333-351, 2016.

GRAHN, Robert A. *et al.* **Extinction of LINE-1 activity coincident with a major mammalian radiation in rodents.** Cytogenetic and genome research, v. 110, n. 1-4, p. 407-415, 2005.

GUERRA, M. dos S. **Introdução à citogenética geral.** Guanabara koogan, 1988.

HASHIMOTO, D. T. *et al.* **Chromosome mapping of H1 histone and 5S rRNA gene clusters in three species of *Astyanax* (Teleostei, Characiformes).** Cytogenetic and Genome Research, v. 134, n. 1, p. 64-71, 2011.

HUBLEY, Robert *et al.* **The Dfam database of repetitive DNA families.** Nucleic acids research, v. 44, n. D1, p. D81-D89, 2016.

INGENITO, Leonardo FS; DUBOC, Luiz F. **A new species of *Astyanax* (Ostariophysi: Characiformes: Characidae) from the upper rio Iguaçú basin, southern Brazil.** Neotropical Ichthyology, v. 12, n. 2, p. 281-290, 2014.

IWASAKI, Wataru *et al.* **MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline.** Molecular biology and evolution, v. 30, n. 11, p. 2531-2540, 2013.

JANZEN, Daniel H. *et al.* **Nuclear genomes distinguish cryptic species suggested by their DNA barcodes and ecology.** Proceedings of the National Academy of Sciences, v. 114, n. 31, p. 8313-8318, 2017.

JOHN, H. A.; BIRNSTIEL, M. L.; JONES, K. W. **RNA-DNA hybrids at the cytological level.** Nature, v. 223, p. 582-587, 1969.

JUNK, Wolfgang J. *et al.* **The flood pulse concept in river-floodplain systems.** Canadian special publication of fisheries and aquatic sciences, v. 106, n. 1, p. 110-127, 1989.

JUSTI, A. J. **Caracterizacão cariotípica de populações de *Astyanax fasciatus* (Suvier, 1819) Pisces, Characidae, em três bacias hidrográficas.** Master's thesis. Universidade Federal de São Carlos, Brazil., 1993.

KANTEK, D. L. Z. *et al.* **Chromosomal location and distribution of As51 satellite DNA in five species of the genus *Astyanax* (Teleostei, Characidae, Incertae sedis).** Journal of Fish Biology, v. 75, n. 2, p. 408-421, 2009.

KAVALCO, Karine Frehner *et al.* **Biogeographic patterns in the chromosomal distribution of a satellite DNA in the banded tetra *Astyanax fasciatus* (Teleostei: Characiformes).** Organisms Diversity & Evolution, v. 13, n. 1, p. 67-76, 2013.

KAVALCO, Karine Frehner *et al.* **Chromosomal diversification higher than molecular variation in *Astyanax aff. fasciatus* (Teleostei, Characidae).** Zebrafish, v. 13, n. 4, p. 345-353, 2016.

KENDALL, David G. **Shape manifolds, procrustean metrics, and complex projective spaces.** Bulletin of the London mathematical society, v. 16, n. 2, p. 81-121, 1984.

KLASSMANN, Josiele Paula; DOS SANTOS, Isabel Cristina Martins. **Karyotype diversity in *Astyanax sacripinnis* (Pisces, Characiformes) from the São Francisco river basin, Brazil.** Cytologia, v. 82, n. 2, p. 199-203, 2017.

KLINGENBERG, Christian Peter; MCINTYRE, Grant S. **Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods.** Evolution, v. 52, n. 5, p. 1363-1375, 1998.

KLINGENBERG, Christian Peter. **MorphoJ: an integrated software package for geometric morphometrics.** Molecular ecology resources, v. 11, n. 2, p. 353-357, 2011.

KLINGENBERG, Christian Peter. **Size, shape, and form: concepts of allometry in geometric morphometrics.** Development genes and evolution, v. 226, n. 3, p. 113-137, 2016.

- KUMAR, Sudhir *et al.* **MEGA X: molecular evolutionary genetics analysis across computing platforms.** *Molecular biology and evolution*, v. 35, n. 6, p. 1547-1549, 2018.
- LEVAN, Albert; FREDGA, Karl; SANDBERG, Avery A. **Nomenclature for centromeric position on chromosomes.** *Hereditas*, v. 52, n. 2, p. 201-220, 1964.
- LÓPEZ-FLORES, I.; GARRIDO-RAMOS, M. A. **The repetitive DNA content of eukaryotic genomes.** *Repetitive DNA*, v. 7, p. 1-28, 2012.
- MAISTRO, Edson Luis; OLIVEIRA, Claudio; FORESTI, Fausto. **Sympatric occurrence of two cytotypes of *Astyanax scabripinnis* (Characiformes, Characidae).** *Genetics and Molecular Biology*, v. 23, n. 2, p. 365-369, 2000.
- MANTOVANI, Monique *et al.* **Conserved 5S and variable 45S rDNA chromosomal localisation revealed by FISH in *Astyanax scabripinnis* (Pisces, Characidae).** *Genetica*, v. 123, n. 3, p. 211-216, 2005.
- MATTICK, John S. **Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms.** *Bioessays*, v. 25, n. 10, p. 930-939, 2003.
- MESTRINER, Carlos A. *et al.* **Structural and functional evidence that a B chromosome in the characid fish *Astyanax scabripinnis* is an isochromosome.** *Heredity*, v. 85, n. 1, p. 1-9, 2000.
- MMA - Ministério do Meio Ambiente. **Programa de revitalização da Bacia Hidrográfica do Rio São Francisco.** Caderno 3 - Diagnóstica da bacia, 2004.
- MOREIRA-FILHO, Orlando; BERTOLLO, L. C. A. ***Astyanax scabripinnis* (Pisces, Characidae): a species complex.** *Brazil J. Genet.*, v. 14, p. 331-357, 1991.
- MOREIRA-FILHO, O.; GALETTI JR, P. M.; BERTOLLO, L. A. C. **B chromosomes in the fish *Astyanax scabripinnis* (Characidae, Tetragonopterinae): An overview in natural populations.** *Cytogenetic and Genome Research*, v. 106, n. 2-4, p. 230-234, 2004.
- NAKATANI, Masanori *et al.* **Evolutionary history of Otophysi (Teleostei), a major clade of the modern freshwater fishes: Pangaeen origin and Mesozoic radiation.** *BMC evolutionary biology*, v. 11, n. 1, p. 1-25, 2011.
- NOVAK, Petr *et al.* **RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads.** *Bioinformatics*, v. 29, n. 6, p. 792-793, 2013.
- NOVÁK, Petr *et al.* **TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads.** *Nucleic acids research*, v. 45, n. 12, p. e111-e111, 2017.
- OLIVER, Keith R.; GREENE, Wayne K. **Mobile DNA and the TE-Thrust hypothesis: supporting evidence from the primates.** *Mobile DNA*, v. 2, n. 1, p. 1-17, 2011.
- ORNELAS-GARCÍA, Claudia P.; BASTIR, Markus; DOADRIO, Ignacio. **Morphometric variation between two morphotypes within the *Astyanax* Baird and Girard, 1854 (Actinopterygii: Characidae) genus, From a Mexican tropical lake.** *Journal of morphology*, v. 275, n. 7, p. 721-731, 2014.

- PAIVA, MELQUIADES PINTO; BASTOS, S. A. **Marcações de peixes nas regiões do alto e médio São Francisco (Brasil)**. *Ciência e Cultura*, v. 34, p. 1362-1365, 1982.
- PASA, Rubens *et al.* **Distribution and morphological diversity of *Astyanax rivularis* Lütken, 1874 (Teleostei Characiformes) in the upper São Francisco River basin, Brazil**.
- PASQUIER, Jeremy *et al.* **Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database**. *BMC genomics*, v. 17, n. 1, p. 1-10, 2016.
- PAZZA, R.; KAVALCO, K. F.; BERTOLLO, L. A. C. **Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 1. Karyotype analysis, Ag-NORs and mapping of the 18S and 5S ribosomal genes in sympatric karyotypes and their possible hybrid forms**. *Cytogenetic and Genome Research*, v. 112, n. 3-4, p. 313-319, 2006.
- PAZZA, RUBENS; KAVALCO, KARINE FREHNER. **Chromosomal evolution in the neotropical characin *Astyanax* (Teleostei, Characidae)**. *NUCLEUS-CALCUTTA-*, v. 50, n. 3, p. 523, 2007.
- PAZZA, Rubens; CRUVINEL, Leticia Aparecida; KAVALCO, Karine Frehner. **Parallel evolution evidenced by molecular data in the banded-tetra (*Astyanax fasciatus*)**. *Biochemical Systematics and Ecology*, v. 70, p. 141-146, 2017.
- PAZZA, Rubens; DERGAM, Jorge A.; KAVALCO, Karine F. **Trends in karyotype evolution in *Astyanax* (Teleostei, Characiformes, Characidae): insights from molecular data**. *Frontiers in genetics*, v. 9, p. 131, 2018.
- PERAZZO, Giselle Xavier *et al.* **Shape and size variation of *Jenynsia lineata* (Jenyns 1842)(Cyprinodontiformes: Anablepidae) from different coastal environments**. *Hydrobiologia*, v. 828, n. 1, p. 21-39, 2019.
- PEREIRA, Luís Flávio; DE BARROS CRUZ, Gabriela; GUIMARÃES, Ricardo Morato Fiúza. **Impactos do rompimento da barragem de rejeitos de Brumadinho, Brasil: uma análise baseada nas mudanças de cobertura da terra**. *Journal of Environmental Analysis and Progress*, p. 122-129, 2019.
- PERES, Wellington Adriano Moreira; BERTOLLO, Luiz Antonio Carlos; MOREIRA FILHO, Orlando. **Physical mapping of the 18S and 5S ribosomal genes in nine Characidae species (Teleostei, Characiformes)**. *Genetics and Molecular Biology*, v. 31, n. 1, p. 222-226, 2008.
- PÉREZ-MIRANDA, Fabian *et al.* **The role of head shape and trophic variation in the diversification of the genus *Herichthys* in sympatry and allopatry**. *Journal of fish biology*, v. 96, n. 6, p. 1370-1378, 2020.
- PLATT, Roy N. *et al.* **Large numbers of novel miRNAs originate from DNA transposons and are coincident with a large species radiation in bats**. *Molecular biology and evolution*, v. 31, n. 6, p. 1536-1545, 2014.
- PLOHL, Miroslav; MEŠTROVIĆ, Nevenka; MRAVINAC, Brankica. **Satellite DNA evolution**. *Repetitive DNA*, v. 7, p. 126-152, 2012.
- PONS, Joan; PETITPIERRE, Eduard; JUAN, Carlos. **Evolutionary dynamics of satellite DNA family PIM357 in species of the genus *Pimelia* (Tenebrionidae, Coleoptera)**. *Molecular biology and evolution*, v. 19, n. 8, p. 1329-1340, 2002.

- PONS, Joan; GILLESPIE, Rosemary G. **Common origin of the satellite DNAs of the Hawaiian spiders of the genus *Tetragnatha*: evolutionary constraints on the length and nucleotide composition of the repeats.** *Gene*, v. 313, p. 169-177, 2003.
- POWERS, Amanda K. *et al.* **Cranial asymmetry arises later in the life history of the blind Mexican cavefish, *Astyanax mexicanus*.** *PLoS One*, v. 12, n. 5, 2017.
- REBORDINOS, L.; CROSS, I.; MERLO, A. **High evolutionary dynamism in 5S rDNA of fish: state of the art.** *Cytogenetic and Genome Research*, v. 141, n. 2-3, p. 103-113, 2013.
- ROCHA, Renan Rodrigues *et al.* ***Astyanax rivularis*: ferramentas para a avaliação da diversidade genética e conservação da espécie.** 2019.
- ROCHA REIS, Dinaíza Abadia *et al.* **The Complete Mitochondrial Genome of Two Armored Catfish Populations of the Genus *Hypostomus* (Siluriformes, Loricariidae, Hypostominae).** *Frontiers in Ecology and Evolution*, v. 8, p. 421, 2020.
- RODRIGUES, Pedro Henrique de Mira *et al.* **Chromosomal and genomic dynamics of satellite DNAs in Characidae (Characiformes, Teleostei) Species.** *Zebrafish*, v. 16, n. 4, p. 408-414, 2019.
- RODRIGUES ROCHA, Renan *et al.* **Morphological and Genetic Structure of Two Equivalent *Astyanax* Species (Characiformes: Characidae) in the Region of Paranaíba Arc.** *The Scientific World Journal*, v. 2019, 2019.
- ROHLF, F. James; SLICE, Dennis. **Extensions of the Procrustes method for the optimal superimposition of landmarks.** *Systematic Biology*, v. 39, n. 1, p. 40-59, 1990.
- ROHLF, F. James; MARCUS, Leslie F. **A revolution morphometrics.** *Trends in ecology & evolution*, v. 8, n. 4, p. 129-132, 1993.
- ROHLF, F. J. **tpsUtil. Version 1.64.** Department of Ecology and Evolution, State University of New York, Stony Brook (program). 2015.
- ROHLF, F. J. **tpsDig2 ver. 2.32.** Ecology & Evolution, SUNY. 2016.
- RUIZ-RUANO, Francisco J. *et al.* **High-throughput analysis of the satellitome illuminates satellite DNA evolution.** *Scientific reports*, v. 6, n. 1, p. 1-14, 2016.
- RYDER, Oliver A. **Species conservation and systematics: the dilemma of subspecies.** *Trends Ecol. Evol.*, v. 1, p. 9-10, 1986.
- SACCONI, Cecilia *et al.* **Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system.** *Gene*, v. 238, n. 1, p. 195-209, 1999.
- SAITO, Yasuhiro; EDPALINA, Rizalita R.; ABE, Syuiti. **Isolation and characterization of salmonid telomeric and centromeric satellite DNA sequences.** *Genetica*, v. 131, n. 2, p. 157, 2007.
- SANTOS, Alex Braz Iacone; ARAÚJO, Francisco Gerson. **Evidence of morphological differences between *Astyanax bimaculatus* (Actinopterygii: Characidae) from reaches above and below dams on a tropical river.** *Environmental Biology of Fishes*, v. 98, n. 1, p. 183-191, 2015.

- SATOH, Takashi P. *et al.* **Round and pointed-head grenadier fishes (Actinopterygii: Gadiformes) represent a single sister group: evidence from the complete mitochondrial genome sequences.** *Molecular Phylogenetics and Evolution*, v. 40, n. 1, p. 129-138, 2006.
- SHADEL, Gerald S.; CLAYTON, David A. **Mitochondrial DNA maintenance in vertebrates.** *Annual review of biochemistry*, v. 66, n. 1, p. 409-435, 1997.
- SHAO, Feng; HAN, Minjin; PENG, Zuogang. **Evolution and diversity of transposable elements in fish genomes.** *Scientific reports*, v. 9, n. 1, p. 1-8, 2019.
- SILVA, Duilio MZ de A. *et al.* **The complete mitochondrial genome sequence of *Astyanax paranae* (Teleostei: characiformes).** *Mitochondrial DNA Part B*, v. 1, n. 1, p. 586-587, 2016.
- SMITH, David Roy. **The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs?.** *Briefings in functional genomics*, v. 15, n. 1, p. 47-54, 2016.
- SOLA, L. *et al.* **Cytogenetics of bisexual/unisexual species of *Poecilia*.** *Cytogenetic and Genome Research*, v. 60, n. 3-4, p. 229-235, 1992.
- SOUZA, I. L.; MOREIRA-FILHO, O.; BERTOLLO, L. A. C. **Cytogenetic Diversity in the *Astyanax scabripinnis* (Pisces, Characidae) Complex II. Different Cytotypes Living in Sympatry.** *Cytologia*, v. 60, n. 3, p. 273-281, 1995.
- SUMNER, A. T. **Banding as a level of chromosome organization.** *Current chromosome research*, p. 17-22, 1976.
- SWOFFORD, D. L. PAUP* ver 4.0. b10. **Phylogenetic Analysis Using Parsimony and Other Methods.** Sunderland, MA: Sinauer Associates, Sunderland, 2003.
- TAANMAN, Jan-Willem. **The mitochondrial genome: structure, transcription, translation and replication.** *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, v. 1410, n. 2, p. 103-123, 1999.
- TERÁN, Guillermo E.; BENITEZ, Mauricio F.; MIRANDE, J. Marcos. **Opening the Trojan horse: phylogeny of *Astyanax*, two new genera and resurrection of *Psalidodon* (Teleostei: Characidae).** *Zoological Journal of the Linnean Society*, v. 190, n. 4, p. 1217-1234, 2020.
- THOMPSON, Julie D.; HIGGINS, Desmond G.; GIBSON, Toby J. **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic acids research*, v. 22, n. 22, p. 4673-4680, 1994.
- TRIQUES, Mauro L. **Adições à Ictiofauna do Parque Nacional da Serra do Cipó, Minas Gerais, Brasil (Teleostei, Characidae).** *Lundiana: International Journal of Biodiversity*, v. 7, n. 1, p. 73-75, 2006.
- UGARKOVIĆ, Đurđica; PLOHL, Miroslav. **Variation in satellite DNA profiles—causes and effects.** *The EMBO journal*, v. 21, n. 22, p. 5955-5959, 2002.
- UTSUNOMIA, Ricardo *et al.* **Uncovering the ancestry of B chromosomes in *Moenkhausia sanctaefilomenae* (Teleostei, Characidae).** *PLoS One*, v. 11, n. 3, p. e0150573, 2016.

UTSUNOMIA, Ricardo *et al.* **Satellitome landscape analysis of *Megaleporinus macrocephalus* (Teleostei, Anostomidae) reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome.** Scientific reports, v. 9, n. 1, p. 1-10, 2019.

VERA, María Inés *et al.* **Seasonal acclimatization of the carp involves differential expression of 5.8 S ribosomal RNA in pituitary cells.** Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, v. 118, n. 4, p. 777-781, 1997.

VICARI, M. R. *et al.* **Satellite DNA and chromosomes in Neotropical fishes: methods, applications and perspectives.** Journal of Fish Biology, v. 76, n. 5, p. 1094-1116, 2010.

VICARI, Marcelo Ricardo *et al.* **New insights on the origin of B chromosomes in *Astyanax scabripinnis* obtained by chromosome painting and FISH.** Genetica, v. 139, n. 8, p. 1073-1081, 2011.

VICENTE, V. E.; MOREIRA-FILHO, O.; CAMACHO, J. P. M. **Sex-ratio distortion associated with the presence of a B chromosome in *Astyanax scabripinnis* (Teleostei, Characidae).** Cytogenetic and Genome Research, v. 74, n. 1-2, p. 70-75, 1996.

WANG, Jian *et al.* **Identification and functional prediction of large intergenic noncoding RNAs (lincRNAs) in rainbow trout (*Oncorhynchus mykiss*).** Marine Biotechnology, v. 18, n. 2, p. 271-282, 2016.

WANG, Man *et al.* **Non-coding RNAs function as immune regulators in teleost fish.** Frontiers in immunology, v. 9, p. 2801, 2018.

WEBB, Paul W. **Form and function in fish swimming.** Scientific American, v. 251, n. 1, p. 72-83, 1984.

WEBB, Paul W.; WEIHS, Daniel. **Functional locomotor morphology of early life history stages of fishes.** Transactions of the American Fisheries Society, v. 115, n. 1, p. 115-127, 1986.

7 ANEXOS

7.1 ANEXO A - Lista de exemplares usados nas análises citogenéticas.

Ponto de coleta	2n	Vouchers
córrego Lage	46II	1973, 1974, 1975, 1976, 1977, 1981, 1982, 1984, 1985, 1986, 1997
	50	2008, 2011, 2025
Usina Abaeté	46I	2014
	50	2614, 2615, 2616, 2634
córrego Tiros	46I	2028, 2030, 2033, 2034, 2039, 2045
córrego Lambari	46I	2050, 2067, 2070, 2071
ribeirão Bandeirinhas	48	4042, 4057
rio Funchal	46II	4335, 4336

7.2 ANEXO B - Lista de exemplares usados nas análises morfométricas.

Ponto de coleta	Vouchers
córrego Lage	1973, 1974, 1977, 1978, 1981, 1982, 1984, 1985, 1986, 1987, 1988, 1989, 1994, 1998, 2000, 2008, 2011, 2019
Usina Abaeté	2614, 2618, 2620, 2629, 2631, 2632, 2634, 2643, 3696, 3697, 3699, 3700, 3701, 3705, 3710, 3711, 3715, 3717, 3992, 3994, 3997, 3998, 3999, 4329, 4340
ribeirão Bandeirinhas	4039, 4040, 4041, 4042, 4043, 4044, 4045, 4046, 4047, 4048, 4049, 4050, 4051, 4052, 4053, 4054, 4055, 4057, 4058, 4059, 4062, 4063, 4064, 4065, 4066, 4067, 4068, 4069
rio Funchal	3968, 3969, 3970, 3971, 3972, 3973, 3974, 3975, 3976, 3977, 4248, 4232, 4234, 4335, 4336, 4338, 4339
córrego da Espinha	3851, 3853, 3855, 3858, 3860, 3867, 3868, 3869, 3870, 3871, 3872, 3873, 3874, 3877, 3878, 3885, 3886, 3887, 3888, 3889, 3890, 3891, 3892, 3893, 3894

rio Vereda Grande	2146, 2147, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 3936, 3941, 3944, 3948, 3951
córrego Crico	4323, 4325, 4326, 4333, 4334, 4341, 4343, 4344, 4346, 4348, 4349, 4350, 4351, 4352, 4353, 4354, 4355, 4356

7.3 ANEXO C - Descrição dos genomas mitocondriais de *P. rivularis* 2n=46 e 2n=50.

Gene	Tamanho (bp) 2n=46	Tamanho (bp) 2n=50	Início-Parada 2n=46	Início-Parada 2n=50
tRNA-Phe	67	67	1-68	1-68
rRNA 12S	949	948	69-1018	69-1017
tRNA-Val	71	71	1019-1090	1018-1089
rRNA 16S	1669	1669	1091-2760	1090-2759
tRNA-Leu (1)	74	74	2761-2835	2760-2834
ND1	971	971	2836-3807	2835-3806
tRNA-Ile	71	71	3017-3888	3016-3887
tRNA-Gln	70	70	3957-3887	3956-3886
tRNA-Met	69	69	3972-4041	3971-4040
ND2	1055	1055	4043-5098	4042-5097
tRNA-Trp	70	70	5115-5184	5114-5183
tRNA-Ala	68	68	5260-5192	5259-5191
tRNA-Asn	72	72	5334-5262	5333-5361
tRNA-Cys	67	67	5432-5365	5431-5364
tRNA-Tyr	70	70	5502-5432	5501-5431
COI	1559	1559	5504-7063	5503-7062
tRNA-Ser (1)	71	71	7122-7051	7121-7050
tRNA-Asp	71	71	7127-7198	7126-7197
COII	690	690	7213-7903	7212-7902
tRNA-Lys	72	72	7904-7976	7903-7975
ATPase 8	167	167	7978-8145	7977-8144
ATPase 6	681	681	8136-8817	8135-8816

COIII	783	783	8818-9601	8817-9600
tRNA-Gly	72	72	9602-9674	9601-9673
ND3	438	438	9675-10023	9674-10022
tRNA-Arg	68	68	10024-10092	10023-10091
ND4L	296	296	10093-10389	10092-10388
ND4	1380	1380	10383-11763	10382-11762
tRNA-His	68	68	11764-11832	11763-11831
tRNA-Ser (2)	67	67	11833-11900	11831-11899
tRNA-Leu (2)	72	72	11902-11974	11901-11973
ND5	1838	1838	11975-13813	11974-13812
ND6	515	515	14325-13810	14324-13809
tRNA-Glu	67	67	14393-14326	14392-14325
Cyt b	1136	1136	14399-15535	14398-15534
tRNA-Thr	71	71	15537-15608	15536-15607
tRNA-Pro	69	69	15676-15607	15675-15606
D-loop	1102	1136	15677-16779	15676-16812

7.4 ANEXO D - Descrição e homologia das sequências obtidas em *P. rivularis* 2n=46 e 2n=50.

Nome	Classificação TAREAN	Homologia BLAST	Cobertura / Identidade e-value	Homologia Dfam	e-value
2n=46					
Pri46Sat03-51	satélite putativo alta confiança	<i>Astyanax scabripinnis</i> KpnI satellite family member sequence	76% / 97,44% 3e-08	-	-
		<i>Astyanax paranae</i> satellite ApaSat01-51_c sequence	49% / 100% 0,047		
Pri46Sat04-184	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat27-178 sequence	100% / 94,23% 9e-57	-	-
Pri46Sat08-236	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat02-236 sequence	100% / 94,63% 1e-56	-	-
Pri46Sat24_lncRNA-631	satélite putativo alta confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111190794 (LOC111190794), ncRNA	44% / 77,86% 3e-37	-	-
Pri46Sat26-91	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat03-91_e sequence	89% / 95,06% 4e-26	-	-
Pri46Sat66-2069	satélite putativo alta confiança	<i>Astyanax mexicanus</i> uncharacterized LOC103021673 (LOC103021673), transcript variant X3, mRNA	17% / 75,34% 2e-28	<i>Danio rerio</i> Helitron- 1N4_DR, Rolling circle transposon	3,2e-11 / 1,1e-06
Pri46Sat74-35	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat08-35_a sequence	100% / 97,14% 1e-06	-	-

Pri46Sat138-27	satélite putativo alta confiança	-	-	-	-
Pri46tRNA_Gly/Ile/Ala/Val/Asp-1270	satélite putativo alta confiança	<i>Danio rerio</i> genome assembly, chromosome: 4	35% / 98,75% 3e-29	Vertebrata tRNA-Gly-GGA gene / pseudogene, tRNA-Ile-ATT gene / pseudogene, tRNA-Ala-GCY gene / pseudogene, tRNA-Val-GTY gene / pseudogene, tRNA-Asp-GAY gene / pseudogene	4e-25 4.9e-25 2.8e-24 8e-23 6.7e-20
Pri46tRNA_Arg/Met/His-1251	satélite putativo alta confiança	<i>Danio rerio</i> genome assembly, chromosome: 4	28% / 97,62% 3e-29	Vertebrata tRNA-Arg-CGY gene / pseudogene, tRNA-Met_v gene / pseudogene, tRNA-His-CAY gene / pseudogene	3.4e-25 2.4e-24 2.6e-13
Pri46Sat190-53	satélite putativo alta confiança	-	-	-	-
Pri46Sat193_tRNA_Al Thr-548	satélite putativo alta confiança	<i>Megaleporinus macrocephalus</i> satellite MmaSat049-1709 sequence	28% / 95,18% 1e-25	Vertebrata tRNA-Ala-GCG gene / pseudogene, tRNA-Thr-ACG gene / pseudogene	2.2e-23 3.5e-21
Pri46Sat198-43	satélite putativo alta confiança	-	-	-	-
Pri46lncRNA_tRNA_Ser/Ser-350	satélite putativo alta confiança	<i>Astyanax mexicanus</i> uncharacterized	61% / 96,43% 1e-42	Vertebrata tRNA-Ser-TCY	1.5e-27

		LOC111188592 (LOC111188592), ncRNA		pseudogene / pseudogene, tRNA-Ser-TCG gene / pseudogene	5.2e-05
Pri46tRNA_Val/Val-192	satélite putativo alta confiança	<i>Apteryx australis mantelli</i> genome assembly AptMant0, scaffold scaffold1350	47% / 93,62% 8e-08	Vertebrata tRNA-Val-GTG gene / pseudogene, tRNA-Val-GTY gene / pseudogene	8.8e-10 6.8e-08
Pri46tRNA_Tyr/Glu-658	satélite putativo alta confiança	<i>Echeneis naucrates</i> genome assembly, chromosome: 2	33% / 97,53% 2e-28	Vertebrata tRNA-Tyr-TAC gene / pseudogene, tRNA-Glu-GAG_v gene / pseudogene	2.5e-23 1.1e-20
Pri46Sat262-31	satélite putativo alta confiança	-	-	-	-
Pri46lncRNA-431	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111191906 (LOC111191906), ncRNA <i>Astyanax paranae</i> satellite ApaSat06-86 sequence	35% 96,98% 7e-62 11% / 98% 5e-13	<i>Danio rerio</i> DIRS-N4_DR, DIRS- type LTR retrotransposon	8.1e-13
Pri46rDNA5S-195	satélite putativo baixa confiança	<i>Astyanax fasciatus</i> voucher LAGENPE2771 5S ribosomal RNA gene region	98% / 99,18% 2e-53	Vertebrata 5S rDNA gene / pseudogene	9.2e-35
Pri46Sat83-107	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat38-107 sequence	67% / 98,61% 1e-23	-	-
Pri46Sat104-69	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat12-69 sequence	97% / 98,51% 1e-23	-	-

Pri46Sat116-22	satélite putativo baixa confiança	-	-	-	-
Pri46Sat121-103	satélite putativo baixa confiança	<i>Moenkhausia sanctaefilomenae</i> voucher LBP19830 69723 clone 18 satellite MS7 sequence	98% / 97,03% 1e-37	<i>Danio rerio</i> Tc1-4_DR, autonomous Tc1/Mariner type DNA transposon	1,2e-4
Pri46Sat132_ncRNA-51	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111193191 (LOC111193191), transcript variant X4, ncRNA <i>Astyanax paranae</i> satellite ApaSat15-51_a sequence	100% / 96,08% 4e-13 56% / 100% 3e-04	-	-
Pri46Sat147-42	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> nectin-3-like protein (LOC111193783), mRNA	100% / 90,48% 4e-05	-	-
Pri46Sat151-80	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> mucin-2-like (LOC103026144), mRNA	100% / 90,12% 7e-19	-	-
Pri46Sat165-250	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> zinc finger protein 239-like (LOC107196801), transcript variant X5, misc_RNA	64% / 78,05% 2e-15	-	-
Pri46Sat172-52	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized protein C1orf194 homolog (LOC103046376), mRNA	98% / 92,16% 8e-10	-	-
Pri46Sat178-165	satélite putativo	<i>Astyanax paranae</i> satellite	94% / 99,36%	-	-

	baixa confiança	ApaSat31-165 sequence	7e-12		
Pri46Sat185-75	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> chondroitin sulfate N- acetylgalactosaminyltransfe rase 1-like (LOC103045746), mRNA	100% / 91,94% 2e-13	-	-
Pri46Sat214-52	satélite putativo baixa confiança	-	-	-	-
Pri46Sat222-112	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat33-112 sequence	89% / 99% 1e-41	-	-
Pri46Sat237-190	satélite putativo baixa confiança	-	-	-	-
Pri46Sat243-27	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat25-27 sequence	100% / 100% 5e-04	-	-
Pri46Sat245-143	satélite putativo baixa confiança	-	-	-	-
Pri46Sat249-915	satélite putativo baixa confiança	Molluscum contagiosum virus subtype 1 isolate MCV1_MC505, complete genome	14% / 76,64% 4e-07	-	-
Pri46Sat256-59	satélite putativo baixa confiança	-	-	-	-
Pri46Sat258_ncRNA-96	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111189946 (LOC111189946), transcript variant X2, ncRNA	73% / 94,37% 2e-20	-	-
Pri46Sat260-602	satélite putativo	-	-	-	-

	baixa confiança				
Pri46LTR_incRNA-721	LTR element	<i>Astyanax mexicanus</i> uncharacterized LOC111189805 (LOC111189805), transcript variant X2, ncRNA	60% / 90,32% 2e-157	<i>Danio rerio</i> Helitron-1N4_DR, Rolling circle transposon DNA2-7_DR, DNA transposon	9.7e-13 3.1e-05
Pri46rDNA18S/5.8S/28S-8910	rDNA	<i>Cyprinus carpio</i> intergenic spacer, partial sequence; 18S ribosomal RNA gene, internal transcribed spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2, and 28S ribosomal RNA gene, complete sequence; and intergenic spacer, partial sequence	66% / 90,46% 0	Metazoa LSU-rRNA_Cel gene / pseudogene SSU-rRNA_Hsa gene / pseudogene LSU-rRNA_Hsa gene / pseudogene	0 0 0
2n=50					
Pri50Sat06-236	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat02-236 sequence	100% / 84,17% 4e-17	-	-
Pri50Sat29-236	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat02-236 sequence	99% / 89,89% 2e-59	-	-
Pri50Sat49_incRNA-1614	satélite putativo alta confiança	<i>Colossoma macropomum</i> uncharacterized LOC118808736 (LOC118808736), ncRNA	23% / 83,26% 4e-49	<i>Danio rerio</i> Penelope-2_DR Penelope-like retrotransposon	4,9e-51 / 1,3e-16
Pri50Sat59-177	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat27-178 sequence	100% / 93,88% 9e-32	-	-
Pri50Sat61_incRNA-631	satélite putativo alta confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111190794	44% / 79,64% 1e-45	-	-

		(LOC111190794), ncRNA			
Pri50Sat75-152	satélite putativo alta confiança	-	-	-	-
Pri50Sat113-35	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat08-35_b sequence	100% / 94,29% 7e-05	-	-
Pri50Sat122_snRNA_U2/U5- 1036	satélite putativo alta confiança	<i>Moenkhausia sanctaefilomenae</i> U2 snRNA gene, complete sequence	18% /94,48% 6e-91	Vertebrata U2 snRNA (spliceosomal RNA) gene / pseudogene U5 snRNA (spliceosomal RNA) gene / pseudogene	4,3e-55 1,4e-17
Pri50Sat147-43	satélite putativo alta confiança	-	-	-	-
Pri50Sat160-47	satélite putativo alta confiança	-	-	-	-
Pri50Sat172-578	satélite putativo alta confiança	<i>Astyanax mexicanus</i> FYN- binding protein-like (LOC103040774), transcript variant X2, mRNA	13% / 88,61% 1e-15	-	-
Pri50tRNA_Arg/Met/His-1251	satélite putativo alta confiança	<i>Danio rerio</i> genome assembly, chromosome: 4	28% / 94,62% 1e-29	Vertebrata tRNA-Arg-CGY gene / pseudogene, tRNA-Met_v gene / pseudogene, tRNA-His-CAY gene / pseudogene	3.4e-25 2.4e-24 3.5e-14
Pri50Sat177-178	satélite putativo alta confiança	<i>Astyanax paranae</i> satellite ApaSat10-179 sequence	100% / 97,73% 6e-33	-	-

Pri50Sat181-41	satélite putativo alta confiança	<i>Colossoma macropomum</i> uncharacterized LOC118808059 (LOC118808059), transcript variant X2, mRNA	68% / 96,43% 0,023	-	-
Pri50Sat186_lncRNA-236	satélite putativo alta confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111192019 (LOC111192019), transcript variant X2, ncRNA <i>Astyanax paranae</i> satellite ApaSat04-233 sequence	100% / 876,75% 6e-25 94% / 78,15% 5e-19	-	-
Pri50Sat199-78	satélite putativo alta confiança	<i>Astyanax mexicanus</i> cell wall protein DAN4-like (LOC111196889), partial mRNA	97% / 98,68% 2e-28	-	-
Pri50tRNA_Tyr/Arg-548	satélite putativo alta confiança	<i>Danio rerio</i> genome assembly, chromosome: 4	40% / 91,30% 3e-32	Vertebrata tRNA-Thr-ACY gene / pseudogene, tRNA-Arg-AGA gene / pseudogene	5,8e-25 3,1e-23
Pri50Sat230-187	satélite putativo alta confiança	-	-	-	-
Pri50Sat236-68	satélite putativo alta confiança	-	-	-	-
Pri50Sat247-88	satélite putativo alta confiança	<i>Plectropomus leopardus</i> DNA, chromosome 3, nearly complete sequence	82% / 87,84% 3e-13	-	-

Pri50Sat256-27	satélite putativo alta confiança	-	-	-	-
Pri50tRNA_Asn/Lys- 419	satélite putativo alta confiança	<i>Epinephelus fuscoguttatus</i> DNA, LG23, complete sequence	40% / 96,70% 7e-32	Vertebrata tRNA-Asn-AAC gene / pseudogene, tRNA-Lys-AAG gene / pseudogene	8,7e-26 1,4e-15
Pri50Sat264-40	satélite putativo alta confiança	-	-	-	-
Pri50Sat270_tRNA_Arg/Gly-358	satélite putativo alta confiança	<i>Megaleporinus macrocephalus</i> satellite MmaSat104-752 sequence	36% / 98,65% 2e-26	Vertebrata tRNA-Arg-AGG gene / pseudogene, tRNA-Gly-GGG gene / pseudogene	9,5e-18 6,5e-07
Pri50Sat272-33	satélite putativo alta confiança	-	-	-	-
Pri50Sat278-21	satélite putativo alta confiança	-	-	-	-
Pri50Sat282-41	satélite putativo alta confiança	-	-	-	-
Pri50Sat19-32	satélite putativo baixa confiança	-	-	-	-
Pri50rDNA5S-195	satélite putativo baixa confiança	<i>Colossoma macropomum</i> 5S ribosomal RNA (LOC118808737), rRNA	61% / 97,48% 2e-48	Vertebrata 5S rDNA gene / pseudogene	9,2e-35
Pri50TE_RC-1656	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> zinc finger BED domain-containing protein	47% / 81,98% 2e-136	<i>Danio rerio</i> Helitron- 1N4_DR, Rolling circle transposon	1,9e-28 / 1,5e-14

		1-like (LOC111189369), mRNA			
Pri50Sat82-42	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> nectin-3-like protein (LOC111193783), mRNA	97% / 90,24% 2e-04	-	-
Pri50Sat86-54	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase TPTE2-like (LOC111196018), transcript variant X9, misc_RNA <i>Astyanax paranae</i> satellite ApaSat18-58_a sequence	94% / 96,08% 4e-13 57% / 100% 2e-05	-	-
Pri50Sat98-69	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat12-69 sequence	97% / 98,51% 2e-23	-	-
Pri50Sat109-106	satélite putativo baixa confiança	<i>Moenkhausia sanctaefilomenae</i> voucher LBP19830 69723 clone 18 satellite MS7 sequence	92% / 93,88% 2e-30	<i>Danio rerio</i> Tc1-4_DR, autonomous Tc1/Mariner type DNA transposon	2,6e-05
Pri50Sat110-22	satélite putativo baixa confiança	-	-	-	-
Pri50TE_hAT-61	satélite putativo baixa confiança	<i>Danio aesculapii</i> genome assembly, chromosome: 19	100% / 93,65% 9e-16	<i>Danio rerio</i> hAT-N48_DR, Putative non-autonomous hAT DNA transposon	4e-11
Pri50Sat129_ncRNA-51	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111193191	100% / 96,09% 4e-13	-	-

		(LOC111193191), transcript variant X4, ncRNA <i>Astyanax paranae</i> satellite ApaSat15-51_a sequence	92% / 100% 2e-14		
Pri50Sat132-48	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111197609 (LOC111197609), mRNA	95% / 97,83% 4e-12	-	-
Pri50Sat135-107	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat38-107 sequence	66% / 97,18% 8e-24	-	-
Pri50Sat136-24	satélite putativo baixa confiança	-	-	-	-
Pri50Sat138-80	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> mucin-2-like (LOC103026144), mRNA	100% / 90,12% 7e-19	-	-
Pri50Sat143-48	satélite putativo baixa confiança	<i>Sparus aurata</i> genome assembly, chromosome: 12	95% 93,62% 8e-09	-	-
Pri50Sat150-29	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> NIM1 serine/threonine protein kinase (nim1k), mRNA	100% / 100% 4e-05	-	-
Pri50Sat157-27	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat25-27 sequence	100% / 100% 5e-04	-	-
Pri50Sat163-1345	satélite putativo baixa confiança	<i>Megaleporinus macrocephalus</i> satellite MmaSat093-953 sequence	25% / 82,2% 2e-16	-	-
Pri50Sat165-52	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat16-54_b sequence	63% / 100% 2e-06	-	-

Pri50Sat193_snRNA_U2/U5- 1114	satélite putativo baixa confiança	<i>Astyanax paranae</i> clone 29 U1-II snRNA gene, complete sequence	13% / 99,33% 1e-67	Vertebrata U2 snRNA (spliceosomal RNA) gene / pseudogene U5 snRNA (spliceosomal RNA) gene / pseudogene	1e-49 3,7e-29
Pri50Sat198-191	satélite putativo baixa confiança	-	-	-	-
Pri50tRNA_Leu/Ser/Ser-563	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111188592 (LOC111188592), ncRNA	40% / 87,31% 1e-54	Vertebrata tRNA-Leu-CTG gene / pseudogene, tRNA-Ser-TCY gene / pseudogene, tRNA-Ser-TCG gene / pseudogene	2,9e-27 9,9e-19 3e-05
Pri50Sat210-187	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> zinc fingers and homeoboxes protein 1-like (LOC103044808), transcript variant X4, mRNA	100% / 84,94% 9e-37	-	-
Pri50Sat212-43	satélite putativo baixa confiança	-	-	-	-
Pri50ncRNA220-70	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized LOC111191672 (LOC111191672), transcript variant X2, ncRNA	100% / 92,65% 2e-17	-	-
Pri50Sat225-189	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat40-189 sequence	89% / 98,24% 7e-77	-	-
Pri50Sat235-21	satélite putativo	-	-	-	-

	baixa confiança				
Pri50Sat240-142	satélite putativo baixa confiança	<i>Gadus morhua</i> genome assembly, chromosome: 11	40% / 84,46% 0,005	-	-
Pri50Sat242-87	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> sulfiredoxin 1 (srxn1), transcript variant X2, mRNA	98% / 91,01% 9e-23	-	-
Pri50Sat244_tRNA_Ala/Thr-552	satélite putativo baixa confiança	<i>Megaleporinus</i> <i>macrocephalus</i> satellite MmaSat049-1709 sequence	22% / 98,18% 1e-25	Vertebrata tRNA-Ala-GCG gene / pseudogene, tRNA-Thr-ACG gene / pseudogene	2,2e-23 2,5e-06
Pri50Sat255-76	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> C-type natriuretic peptide 2-like (LOC103038817), mRNA	98% / 81,33% 1e-05	-	-
Pri50Sat260-91	satélite putativo baixa confiança	<i>Astyanax paranae</i> satellite ApaSat03-91_a sequence	94% / 100% 3e-08	-	-
Pri50tRNA_Pro-153	satélite putativo baixa confiança	<i>Echeneis naucrates</i> genome assembly, chromosome: 2	61% / 97,83% 1e-34	Vertebrata tRNA-Pro-CCA gene / pseudogene	8,4e-26
Pri50Sat271-52	satélite putativo baixa confiança	<i>Megaleporinus</i> <i>macrocephalus</i> satellite MmaSat085-52 sequence <i>Astyanax paranae</i> satellite ApaSat29-52 sequence	100% / 100% 0,004 88% / 97,83% 4e-12	-	-
Pri50Sat276-52	satélite putativo baixa confiança	<i>Astyanax mexicanus</i> uncharacterized protein C1orf194 homolog (LOC103046376), mRNA	100% / 94,23% 5e-12	-	-

Pri50LTR_lncRNA-1998	LTR element	<i>Astyanax mexicanus</i> uncharacterized LOC111189805 (LOC111189805), transcript variant X2, ncRNA	48% / 86,75% 0	<i>Danio rerio</i> Helitron-1N4_DR, rolling simple transposon DNA2-7_DR, DNA transposon	2,4e-39 2,8e-30
Pri50LTR-999	LTR element	<i>Astyanax mexicanus</i> leucyl-tRNA synthetase (lars), mRNA	40% / 76,69% 8e-45	-	-
Pri50LTR_tRNA_Ala/Thr-458	LTR element	<i>Acomys russatus</i> genome assembly, chromosome: 25	39% / 98,75% 1e-29	Vertebrata tRNA-Ala-GCG gene / pseudogene, tRNA-Thr-ACG gene / pseudogene	6,4e-26 5,7e-26
Pri50rDNA18S/5.8S/28S-8708	rDNA	<i>Cyprinus carpio</i> intergenic spacer, partial sequence; 18S ribosomal RNA gene, internal transcribed spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2, and 28S ribosomal RNA gene, complete sequence; and intergenic spacer, partial sequence	66% / 90,60% 0	Metazoa LSU-rRNA_Cel gene / pseudogene SSU-rRNA_Hsa gene / pseudogene LSU-rRNA_Hsa gene / pseudogene	0 0 0

7.5 ANEXO E - Sequências obtidas por bioprospecção *In Silico* de *P. rivularis* 2n=46 e 2n=50.

Nome	Sequência consenso
	2n=46
Pri46Sat03-51	CACTTTTTGACCGTACCCATGGTAGACCAAGTGGGACTTAGTCATTTTTTC
Pri46Sat04-184	TCTTCTCAGCCCAAATAAGGAGAGGACCTTGAGGGTCCCCAAGAAGACCGCCCGGACCCGGTCTCGAAC CGGCGACCTTCGGCACTCCGAAGCCTCCGAGGGCCAAAGGCCGGAGGACTCTACCGCTGAGCCACCGAGGC CCACATGAGTTCTGGGGAGCTTTTGCATCCTTTGTAATTCGT
Pri46Sat08-236	TGAAATACAACTTTGCTGCATTCAACTTACAAATGAACCAAAACACTTAAGAAATGTTTAGAAGAAGATAT TAGTGCATTTTACACTGTAATAATGTGTTTCAGGTATTAATAACTCATTTTTGAGCGAAAGAGAAGTTTTAG TTCAAACAACACTTGATGCAGTTTTTCGCTGTAAGGAGTGAGAGAGGAGCTCAAACTATCCCCAAAACGT TTTTCTCAGCCTGAAACGC
Pri46Sat24_IncRNA-631	GAGTTTTTGGAGCATTCCCGTAATGTCAAATATGCATTGAAACCTAGGTGGCGCTATAGAGCCAATGAAATAC GTATATATGAAATTTCAATTCAGATCAAAGGACCGCGTAAATCAAGCAGGCCTGTAAAGTTTTGTGAGTTT TCATCCATTTTAACTCCTCAAACACCTACCAACACCAGAATGTATTCCTCCTGTTTTAATGGGGCATCTC AAGCAATCAACTGTCCGCCATTTTCCTCGTTAAGATATCGACTATTCCTCACAATTTATCATCAGGTA TGTTAGTAGTTTATTATTGCCAAATATCAAGAGAATTCAACAAATTCTCTAGGAGGAGTTTGACAAAGTATG CAAAAAATGTGTGTAATAATGCACATATTTCAAATGGCCGACTTCCTGTTGGGCGGAGCTAATACAAGCCAA TTGCAATATGTGCAGAGTCACAGTATCTACATTCCTACCAAAATTTGAGAAGATTAGACAAATTTTGACAC ATCCACAGCTAATTTATTAACCGAACAAATTAGGGGGCGCTGTAGAGTCCGCTAGCTACATATGAAAAAATT ATCACAATATTTGTAAAGTGTTTTTAGGTCCTATGGAATCTGACAATTTTCAA
Pri46Sat26-91	TCCGCGGGTGCCTTTAAACCCTCGGCGATCTAAATTAAGGACTGTGGTATGCATTCCAACAGCGTGCATC TGAATTAACGCGCGGCCG
Pri46Sat66-2069	GGCTCCGTCCACATCCACACTGCCATATTTAGCAAAATCTATAAAACACAAATTTTTATACTACTTCTGGAT TTTTGAAAATATTCCAACAATATAGGCATCAAAATGTTTCAGCAGAGTCTCGTGATCAAGATTGATTAACAA ATATTCATATTTATTAACAAGTGTGCTGATATATGCAATTAGTCATCCTGGGGTGGAGCTTAGTTCACATAA ATGGCTAAAAATCAAGATCACTAGACTTTATCCAATGTAAGTACTAGCAGTTTTGTTAAAGACATAATTCTTTG GAAGGCATCCAAGTCTCACTCCAATTAAGCGATAGGTGGCGCTATAAAATGCAAAAATCTGTTTCTCAACAA CTACTTCTCAATTAAGTTGAATCTTTGCATGCTGCAATCTTTGGCCATTTGATGAGTGATTTATTAAGGAC AAATCAATAATTAAAAAAATACGGCTGCAGTTTATTAACATTGAAAAATACTAAAATGTAAGTACTGTCCTG

GGTTTTCATCAAATGTTTGCAAATCAGGTCTCACAACAATCTTTGGAATGTCTAAGTATGAAATGATTA
GAACTTGAGTTTTATTGAATCTAATGGCTGCAGTAGATAAAATCAATAAGTGCAGGGTGGTATGTGTTAGCTT
GATAACTGTAAGTCTATAACTTTACTCTGGATTATGAAATTTGTGACCAAACTGACAACCTTATGTCCACA
AGATCACCATGACACAAAATCTTAAATATTTTTCAATAATATTTAATAATAATGCTAATAATTTAAAGTTTG
TAAATGCTTGCAAAAAAAGGTGCTGGGTTCCCTTTTTACTCTAAACAATATACTGCCTGAAAACGTGCTGTGG
AGTCAAGGATCCTGGGAGGCATCTCGCTTAAGCACTCGCCTCTGAACAGGAAGGTCGTGGGTTTCGATCCAG
CCTCTGACAGCTCTGGGTTTTCTACGTATCACAACATAGCTGGCATGTGCCAATTAATCCTACTGGAGTTCA
GCTTAGTTCACATAAATCTATGAATATTGTTAACATGGAATCCGATGAAACTTTGAAATTTGATTTGCCATGA
TTCTTTGGAAGGTATCTAATTTTCAAAGTATCAATCATCTGTTTCTCAACAACCACTACTGAGTCCAATTGAG
GTGAAACTTTGCCTGCTGGATCTGTGGCCCATTTGATAAGGGTCTTATTAAGGACAACCTCAATAATTA
AATATTGCTGCAGCTTATTAAGTAAAAAATAATTTCAAACCTACTGTTTTCATCAGATTTTACAAAATCAA
GTCTCAAAAACAATATTTGGAACCTCAAAGTAATTAATTTAAAAAATAAGTGTGGTGGTCTTTTTGCT
AAAGTACATAAATAAGCAGAGCATGCTGGATTGTGTGTATAGCCGGATCTATAAATTTGTTGTAGACTGTCA
AATGGTGAACAAATTGACAAATTTTCAAGGTCATCTTTAATATTTTTATTTTATTTTATTTAATAAAA
ACCAATAATACTGCAACGCTTGTAATTTTCAAAGAAGATGCTGGGTTCCCTTTGCATTCTTTTGAATAAAG
TGCCTGAAAAATCTCAATTTTTCTTCTGTGACGAGCGGCATCTCTGTAAAGCACATGTCTATGAAGTGCAA
GGTTGTGGAACCTCTCTATCTTCACTGCTTCTTATTATTTTCACTCGTGTGCTTGGACCCCGTCATTGCTGCTGC
AGCTATATTTAGGGGTCCAAGCACCGTAGGTGCTGGAACCTATTGAAATTGCTCGGATTTTCTTATTATTA
TTATTTTTCTCCACTAAAACGGCTGTAGCACAGTAACCGTAAGTCATCCAACTCCAAATTTGGAATATAG
GTACAGTAACCTACCCACTACTCAGGCAAAAAAATGACCTCAGTTGGCCTGATGGTGGCGCTATAGAGCA
GTGGTTTTCGTTTGGGTCAATATCTCAGCAACCGTGACATCTAGAGTGAAAATTCTTTTTGAATCTTATCCTG
GAAAGGTTTCTGCATCAATGATTCACCTTGAAGTTTAA

Pri46Sat74-35

ATGACTTTTTTCTGTTTTAAAGGACGGTAGACTCT

Pri46Sat138-27

ACCTCAGAGTGTAAGAGGAAAAGTTTTT

Pri46tRNA_Gly/Ile/Ala/Val/Asp-1270

TGAGGGTCTTTCAGCAAAACAATGGGGTGCTCGTTTGTGGCGTTGGTGGTATAGTGGTGAGCATAGCTGCCT
TCCAAGCAGTTGACCCGGGTTTCGATTCCCGGCCAACGCATTGCCTTTAGCGCGTGGCACCTTTCTGCTCCAAA
ACCTTATCAGCTTTTTGGGACACAGGTTTCAGAAAAAATAATTTTCTCTCCCGTCGGGGAATCGAACCCCGGT
CTCCGCGTGACAGGCGGAGATACTGTCCACTATACTAACGAGGAAGCTGTGAGCAGCTGAATCTCTGCAAC
ATACAACAGGGAATACTGGCCCCTTCTCAGTCTTACAGAAGGTATATGTGTTCTTGTGCCTTTTTATTAGG
GAATGCATGGTGAGTGAGCTTGTGAGGGATGTGAAGTC
CCAGGAACCTGAAGCTGTGCATAGTTCTCCCTGTGCTACTGCTGAAGTGAGTCACCTGTGATCAACATGCTT
CAGCCTTCTAAATGGTATGCTTTTGTGATGAGAGGCTAATAGGTCTTCTTTATTCAAGGGGAATTAGCTCAA
TGGTAGAGCGCTCGCTTAGCATGCGAGAGGTAGCGGGATCGATGCCTGCATTCTCCAACGTAGGTTTTGACT
GTGGGAATGCCACTGAGAACAGCTTTAAAACCTTTACGTTTGTGAGTGTATATCTTCAACGTACGTTATTCACCG

TTTGCTTCTGCTCGACAGATGACCGTCTGAAAGCATGTTATAGCAAAACAAAGAGAAGAATATTTGGCCCCGT
ACGGGGATCGAACCCGCGACCTTGGCGTTATTAGCACACGCTCTAACCAACTGAGCTAACCGGCCACAG
ATGAAAGCTGCACCCTCAGGCATGGCGGATGAATGCGCTCTGAATCGAGGCCTTTGCTTTGTGCGTGATG
CCACCAGGTATAACTGTTTTATTCTGTTGTTTTCCACATCCGAAGCCAAACACAGAAGGCACAACAACAGCA
CAAGCACGGCATCATTGCTTCTGTGCTCGCAATTGTCTTAATAAGGCCTAACACAAAGTACAGATTGTTCC
CATCTGCTACTGAGGTGTGATGACAGGCTTGCCCATAGTGTGACACAAAAGCCAAGAAGTAAAGGCATGTTT
CTGCCTGGTTTCGAACCAGGGACCTTTCGCGTGTAAGGCGAACGTGATGACCACTACACTACAGAAACCACA
GCATTCCACCGAGAGTGCATGTGTCCCGCAGTGCAAATGACAGAGTTTGGTCATCTGTCCCAGTGGAAAATT
TGCAC

Pri46tRNA_Arg/Met/His-1251

ATCTACATTAAGACTTATGTTCTCCTTGCCTTGCACACACATGTGTAGTTAATTACAACATCTTGGAACGG
GGAAAATGAATCTGATCCCAGCATTACTCAGCCTTGCAAGCATATCAGAAAAACGCGTACAGTGGATCATG
TAACAGGAGGCTCTTGTTATGATCCAGGAGGTGTTGTTCCACCTCTGTGTCGGATGTAGAAATCTAGTCATG
TGCCGTACAATCAGTTGACAAGCCTATTGCCTGGGAAGACAGCAAGACCTCTGCCAATGGATGGTGGGGTG
CCGTGATCGTATAGCGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCCCGGTTTGAATCCGGGTCACGGCAA
GTTCTTTAAGCTTACAGTAGGCTTTTCTTGCTTGTCTTTGGAATCGTTTTGAGTAAAACAGGTGCATGGCTTG
TTTTTCAGTGGTACTGCTCCAGCAAAGCGTTGTGCAACATGTCCTAAGGATAAGGCATGTGTCTGCTTTACAG
ATAATTCTACCTATTGTTTCTGGATGCTTGACAAAATTCTTACGGATTCCAACACTTGAAAAATAGTCCTGTA
GGAGAGCTTATTTCCAAACTCACAAAGTCTAATGACACAATTTTCTCGTGCCTTTTTTGTCTCAGAGTCCCAT
ATGGTCTAGCGGTTAGGATTCCTGGTTTTACCCAGGCGGCCCGGGTTCGACTCCCGGTATGGGAAACTTCTT
GCTCGCTGAGAAAAGCAGAGTGATGCCTCAACAATTTGATACGTTCCAATGGCGATCCTGTGCGAACTTTTA
ATAGTGCAGCTTGGTTGTTGGGCTCTGGCTTTACTTTGAAGCTAAGAGGCCCATTTTGGTGTGACGTAGGCCA
GAAATGTTCTTATCCACTTGCCTGACAGGTTCTTGCGGGTGGGCCAGTGGCGCAATGGATAACGCGTCT
GACTACGGATCAGAAGATTCTAGGTTGACTCCTGGCTGGCTCGGCTCCTCTGTTTTTGTGCTCCCAACAT
TCTCAAAGATTACTTTGGTTGTTGCCATCCCCTGTGGTGCTGGCAGGTGACCAGATGCATGCCTGAGCCTCGT
TAGCGCAGTAGGTAGCGCGTCAGTCTCATAATCTGAAGGTCGTGAGTTCGATCCTCACACGGGGCAATAGTC
ATTTTTAAATGAGAAACATGCCCCAATATCTTAAAAGACTCACCTTGTATGTAAGGAGCATGGTTAATGC
AAGTGTTTAAGCGCCTGA

Pri46Sat190-53

GTGTGACTAGAATAGCACTTCTGAAGTAGCTAGCTAACCTAGATTGGCTAAAT

Pri46Sat193_tRNA_Ala/Tyr-548

CCTGTCTCTTAGTAGATACAACATAGAACCCGCAATCAATCAATGCACTCAATGCATGCTGTAGCATGCAAT
TATGAAGCCTATAAGAAATCAGATTTATCAAGCTACTTCATCAGCAAATGCCTTGTATAAATGTGCATTTTTG
GCTAATTCTATAGATGTGTCCTTGATGTTAATTTCAAGGTGCAAACAGCAGAAATGGAGATGCTGGGGAT
TGAACCCAGGACCTCATAACATGCAAAGCATGCGCTCTACCACTGAGCTACATCCCCTACTAAGCACATCTG
TTTGTGTAAGTGTGTCATCCTGGTAACAAAACCTCTGCAGACTTGCAAAGACAGGATCCTTAAACTGCTGCATTGTT
TCCATGCTTTTAGGACAACCTACAGTGTGAGAGATAGGTTTGAAGAGCATGCTAAAGCAAGCCAGAAATTT
TCAACAGAAAGACACATCAGCCCTCAGGCACGGACAGGGATCGAACCCATGATCTTCGGTTTTACGAGACC

GACGCCTTACCACTTGGCCACCATGCCAGACAACAACCAACGT

Pri46Sat198-43 CCGTCTCACATTTACCAACTAACATCAAAAACAACACAACCAC

Pri46tRNA_Ser/Ser-350 GGTAAGAGGCTGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGTCTCCCCGCGCA
GGCTCAAATCCTGCCGACTACGTTTCACTGTTTATTTTTGCGTAGCTTTGGTTGTTTTGCTTGCATTTGCAT
TCTTACAGAGAATGGTCAGCTGACCAATAAAAATATTTTACGCTCGACAGCCAGTTACACACTCGAAAACAT
TTGCGGTCGCCCTGCTCAGGTTCAAATCCTTTTAACTGCGTTGTTCTTTTTGTCAGTAGGGCATGCCATAGTTT
GCTAGGCCCATTTTTGACAACAGTCACATGAAAATCCTGAAAATTTATAGGGAAGGTTAC

Pri46tRNA_Val/Val-192 CGCGAAAGGTCCCTGGTTCGAAACCAGGCAGAAACAGGCTTTTTTCCATCTTGATTCTGTCCCATCCTATTG
TACGTTTTCTACAAGTTCTCTTACTAAGACAAGGTGCATCATGTGCTTCCATGTGGATCTTATGCAAAGCCTT
AAAGCAGTTGTTTCTGTAGTGTAGTGGTTATCACGTTCCGCTCACA

Pri46tRNA_Tyr/Glu-658 GAGCCTGATCGTAAGTGAGGTGTTTGGCACACTGCTTTGCCACAACGGGTTTTAAAAAAAAGCAGTCCTTC
GAGCCGGATTTGAACCAGCGACCTAAGGATGACAATTTAACACCTACAGTCCTCCGCTCTACCAACTGAGC
TATCGAAGGAGCTCGGACACCTGTGTTTATTTGAAATGTAATTGGTACATGTATGATGTAGTTTTTTGCCAT
CAAGCAGCTTAAACATTTAGAAACCTGGTTTTAATCCAAATTCTGAAGAATGCACACAATTTTGTCCGTC
ACAACTATCAACTGAGGTATCAGTGGCAGTAGAGTGCCTTGGGCAGCTCATTTTTTTCACAGCTCCACATTA
GTTACCTCCAGAAACATGGGCCACATACTGTGGTCTAGAACAGTGTTGAAACAAAAAGAAATTGAGGGCTTT
TTGCTTGTGTGCAATTAGGTGTTGGACCTTCCTTTTACTCCATGCCTGCTTGCATGTACGCAGCATGGGTGG
GTCAGCCATGCTGAGCTTCAAAAAGTTTTTTGACAGGCAGTTCGTCAGCTTGTTTCCCTGGTGGTCTAGTGGC
TAGGATTCGGCGCTCTACCCGCCGCGCCCGGGTTCGATTCCCGGTCAGGGAAGTCCTGTTTTAGCTGACGA
CGAACAC

Pri46Sat262-31 CTATTCTCACTGTAATATACACATTTATTCA

Pri46Sat05_lncRNA-431 CTGCCGCAGTTTTTCGAGATATCGACTTCGTTCAAAGAGAAAATGTGCGTCCATATCGGGAATGGTGGGCTT
GTATACAGCTTTCCGATCGGACTTACGTTTTTCGTAAAAACTTCGTAAATGTGCGCTTTTTTGCCATTGAAA
ATGAATGGGGCCATCGAAGGACGGCAGCTAGAACTCCAGTATAGTGAGTGTATGGAGCAGCTATGGGATAA
AGCAGCTATTGCTCAAAAAGTTTTGACACCCCCGCCCCCAGCCAGGTTTTAGCTACCACTTATCAACACCT
CAGCAACCACCTAGCAACACCTTAGCAACCATTTTAACTTTTCAACGTTATTAGCGTACTTTTCCAAGCCAAC
TTAAAGTTCGTTACGAACTTTGCCTTTTCTAGTTATTATTATCCACCTGTTTTTTGTCCGGCTTTCAA

Pri46rDNA5S-195 CCTTGAAGGATAATTTAGCATAAGTAGTCTTGATGAGAGTTCTGCAATCGCTTACGGTCACACTACCCTGAG
AACGCCGATCTCGTCTGATCTCGGAAGCTAAGCAGGGTCTGGCCTGGTTAGTACTTGGATGGGAGACCACC
TGGGAATACCAGGTGCTGTAAGCTTTCCCTCTAGCTGCTATTACTATGG

Pri46Sat83-107 CCCCACCACCCAGAGGTCTGAAATCGCAGTATGTTGTCCAAACGGCGCCCCTGGGGCTGCCTGCAGAGTTT
CGTGCCGCAGAAGTCAACGTAGGCCAAACTCCTTC

Pri46Sat104-69 TGATTCTGTATCTACTGTAAGTGTAGATTAATTACAGAGCAGTATGGGTACAACAGATTACGGTGCTGG

Pri46Sat116-22 AGTGTCCTTTTGGGCATCTCCA

Pri46Sat121-103 GGTCATGGTGAACATAAGGCTTGTGTTGTGTACAGTAAAGTTGTAAGTGGTATTAGTGAATATAGTAACTCT
GTATTGTAGTAGAGAAGTGGGCGGAGCTTCT

Pri46Sat132ncRNA-51 CTACTAATCTAACCAGCACAGCACCATCCATCTGTTATCAAACCCCAGAAC

Pri46Sat147-42 CAGACCTCACTACAACCTCCATAATTCACCTGCACCCACTAT

Pri46Sat151-80 CAGTGATCAGTGAGAGTGTCTACCTGTAATTAAGTCTATATAACATTAGAATACTAAACCCAAATAAACATT
AATAAAGT

Pri46Sat165-250 TCGGACAAAGGCTGAGCTGAGCGATTCTCACTAGAGCGGGGCGTGGCAAACCCCCTCATTACACATTTTGT
GTCTGCATACCTCAGGGGCGTCACTCACGGCGTTTAAAAGACAGTCTGGTGGAGCTGATTTCCATAAATAAGA
GCGGGCACACTACGAGGGGCGAACAGGAAGTGAAGTACACCAGACTGCGCTTTTACTCCACTTAACTGC
AGCTACGCTGTAAGAGCAACAGCAAACGCGCTTTT

Pri46Sat172-52 GTGTAAGTAGTTAGCTGAATTGGTAATATTTGGGTGAAATAAACATTAATAAAA

Pri46Sat178-165 AAAAGTTACTGGGTGAAATAGTAATAGACACACAGTTTATCAACATGATTAATGTACCCATAATCCTCTCCA
GCTGTAATAAACAGAGGGGGAACCACTTTTCTTAGGTTCTGGTTAAATATGTTCCCCCTGTGTTCTCAGAGA
TAAAGAGGTATAGAGAGATT

Pri46Sat185-75 CTCATTCTGACAGACTGTAATACACTACAGGAAGCGTAGGGGGCGCTCTATAATGTTTCATATGATCCACAAC
ACG

Pri46Sat214-52 GCTAGGTCAGCTGGTTGACCAGTTAGAGCAGCTGGTCGAGGCTGGAAGACCA

Pri46Sat222-112 GTGTTACACACACTGACCACTCTACACTGAGACAAATTTAGGACGTTACACACATCTGTTTCTAATGCTACA
ATGGTAAAGTTAAATTAACCTAAGGTAACATCAGGACA

Pri46Sat237-190 GGTGCATTATGGCTCGGTATGCCATGTTGCCAGTTCATTTTAAAGTCTAAAATCAGCTCGGGCCAGGCTAATT
CTAGGTTTTCTTAAAGTAGAAGACTAGAAGTCTTCATCTGATGTGAATTTGTTCTGGCCCAAGATGAAAAA

GTCCTTCAAAATCACACTGGAACATCAGCTTTATATTCAATTTTT

Pri46Sat243-27

ACCCAATATTCTCTCCCTTATCCTG

Pri46Sat245-143

GTGAATTAATTTCAATAGCTTTTGAACCGTTGAGGGTATCACAACAGAACAAGTTGTATTAGATCAAGTCCC
ACCCACATGACATACTACAACCTGATATTGGTCAAATTAACCCTTGCTGATCTGTGGCAACAGGTTTTTT

Pri46Sat249-915

GCCACCAAAGATGGACGTAGCGGGTAAACCCTGGACGAGCAGCTCCAGCCGCTTCTGCGGTACCTGCTCTAC
CTGCCACGGTCTGCCTGGATCAGCAGCTCCAGCCGCGTCAGCGGTACCTGCTGTTCCCGCCCTGGCTCTCGTC
TGCTCCTGCAGATACCGCCATGCCTGCACCTGCAGGGACCGCCGCGACCTGCATGGACTGCACCCGCCGCGC
CTGCGCCTACGGATACCGCCGCACCTGCACGGATTGCACCTGCCGCGACCTGTAAGGACCGCAACCGCTGCG
CCTGCGCCTGCAGACACCGCCGCGCCTGCACGGACTGCATCCACCGCGACCTGCACGGGCTTCGTTTGCTGC
GCCTGCGCCCGCAGAGACTGCCGCGACCTGCACTGGCTGCGCCTGCTGCGCCTGCAGATACCGCCGGCCCTG
CAAAGGACTGCGCTCGCCGCGAGCTCCCTGGACTTCTCCCGCTCCATGTTTTCTGTTTCGCCCCCTCTGCC
AGTCCCTGTTTTGTTTTCGCGCCCCGGGAGTTGCGCGTTTTCGGAGGGGGCTTATGTCATGTCTCAGCCCTGT
GTTTTGTTTACGTCTCGTTTTAGTCTTTTATTTAGCTCCTGTATTGTTTTCGTCACGTGTTCTTTGTTCTGTGT
GCCGCTCAGTCTGTGTCTCCACGTGTTTGTAGCTCCGCCCCTACCTTCCCAGGTGTTGCCACTTACCTGTCC
CTCCATGTGTGTATAAATACCCCGTGTCTACCTGTCTACCCTGTGGCTTGTGTACGTGTTATACGTCTGTTAG
GTCTACTGCAGGTTTTTACCTGTTTCTGTGTGTTTCTGGCAATAAATGTTTACCTGCGTTTACGTCCGACTCC
TCGTGCTTCCCTGCAGTATGCCCTCCGCGTGACAGAACA

Pri46Sat256-59

ACCAGGTACAGACCAGCTACAAGCTGGTTTTAGATGTTCTGACCAGCTATACACCAGTT

Pri46Sat258ncRNA-96

GCAGAAATAAAAGTTTCAGATTTATCTCAGATTCAGTGAATATGAACGGGAGCAGAGCGCGGGGTCCTGCA
GCACGGAGCGTTTTCACTCTAAGGA

Pri46Sat260-602

CTTAAGTGGAGTATTCTCAGATTTCTCACTGCATGCAGGCACTGTACCACTAATGGAAGGTCCTTTCTCAGGA
ATAGCCTAACAGTTTCAGAAACACGTATTTACATGTATCATTGCTTGGAAACCAATGCTATGTGGTTCTGTGT
TGTCATGTAATGAAATGGTGTAAAGAATAATGTAAGAGTGTGTCTAGAACATGTTCTAGAGCCTAAGAAA
AAGAAGTCCCTCGAGACCACAATTTAGTCATATACTACACACACAGGCTGATCTTAATATGCAGGTAACACC
CTCTGTAGAGTCCCTGTAAAGTGTAGTGTCTTATAAACCAGCTATGGACCTTACACTCCACAAAACACACA
CAATCCCCTAAACAATGCAGTCCCAGTAAAGTACTGAAGTCCTTGTATAGAAGGTTTCTGTACACATGTTG
ATACACACAGACCAACCTTACACAGAAAACCTTCACTCAAACCTTAAAATTATCAGTAAATCCAGTGCATGT
CCTTGTAAAACACACTGTACTACAATACATGCTTGGACATACACAGCTTAAATACACATACTGTGAGAAAT
CTGAGAATACTCCAGTATATATACT

Pri46LTR_incRNA-721

ATAGAAGGCCCTCATATACATGTAGGGTATAAGTATCGTCCTGATTGGTCATTGTTTAGCCTGTCAAAGCTT

GCTGGAAACTGATTGGCTAATAGCGATCGCAAAAATTTACATATCAAAAATTCCTTCAGTGTAACATTAGGA
TATAGGCCCAAGATGATAACTGCCAAAACCGAAGCTTGATAGCTTGTATGGTTCCTGAGAAACAGATTTCTA
GCTTTGGCTCCGCCCCCTGGGGCGTATGTCTACATAAACTGACAGGGTACCTCAGAATCATGTTGTGATCA
AAGGTCTAAAGTGGCATTAGTGTCAAGTTAAGCATTCAAAAAGTTACAGCTGTTAGAGTAAATTTGAGTGTGC
CACGGTAAGACTAATTTGAATATGGCGGCCATATTGTTACAAAATTTCAAGATTTTTTTAATAATTATTGAGG
AGTAAGCTCTGCTGAGTTGTTTGACACCAAACATGTCATGATTGGTCAATGAGCCAAGGACAAGTTCTCAA
AGTAGGTTTTGCATATTATGCAAATTTGCAAAAAACTTAAGTGGGTGGAGCATAAACAAGAATGACCCAGG
TGGCTGACCAGCATGACCAAGAAGGATAAAGATGTTCAACCAGCAAGACCAGCAAGATTAATAAATTCAG
CAGAGCTTTAATTTAGAATTATTCACCAGTAGCTGTTTACCAAATGACCACCAGAGCGCAGCAAGAGACGT
G

Pri46rDNA18S/5.8S/28S-8910

AGAGTCGCGGCCGCTGCTCCGCGGCCCGCCGCCCGCCGTACCCCGCCCTACCGGCCCGGAGCGTAGAGCC
GCAACCCGGACCCCCCGCCGCCCGGAGCCACCCGACCGCCGAAACGGCCGGGCACCCGGACGACGAGA
GAGGACGGGCGCGACGCGCTCCGAGGCTTCCGGACGAGGGGCGGAACGACGGGGCGCGGGGCAGCCGC
TCCCCAGTCGCGGCTCGAGCCCAGCCCCACTTCGCTCCCTGGCCCCGACCGACCCAGCCCTTAGAGCCAATC
CTTATCCCGAAGTTACGGATCTGACTTGCCGACTTCCCTTACCTACATTGTTCCAACATGCCAGAGGCTGTT
ACCTTGGAGACCTGCTGCGGATATGGGTACGGCCCCGCGCGAGACTTACACCCTCTCCCCGGATTTTCAAG
GGCCAGCGAGAGCTTACCGGACGCCGCGAGAGCCGCGGCGCTTTCCAGGGCGTGGGCCCTATCTCGGGGT
GAACCCATTCCAGGGAGCCCTGCCCTTACAAAAGAAAAGAGAAGCTTCCCGGTGCCCCGCTGGCGTCTCC
AGGTTGTTGCGTTACCGCACTGGGCGCCTCGCGGCGCCTATCTCCGCCACTCCGGGTTCGGGGATCTGAA
CCCGACTCCCTTTCGATCGGCCGGGGCGACGGAGGCCATCGCCCCACGCTTCCGAACGGCGTTCCGCCATC
CCTTAGGACCGACTGACCCATGTTCAACTGCTGTTACATGGAACCCCTTCTCCACTTCGGCCTTCAAAGCTCT
CGTTTGAATATTTGCTACTACCACCAAGATCTGCACCCGCGGCGGCTCCACCCGGGCCCGCCCTAGGCTT
CCGCGCCACCGCGGCGGCCCTCTACTCGTCGCGGCGTAAGCCCCGAAAGGACTTGCCCTCTCTGCCAG
CGACGGCCGGGTATGGGCCCGACGCTCCAGCGCCATCCATTTTCAGGGCTAGTTGATTTCGGCAGGTGAGTTT
TTACACACTCCTTAGCGGATTCCAACCTCCATGGCCACCGTCTGCTGTCTATATCAACCAACACCTTTCATG
GGGTCTGATGAGCGTCGGCGTCGGGCGCCTTAACCCGGCGTTCGGTTCATCCCGCAGCGCCAGTTCTGCTTA
CCAAAAGTGGCCACTGGGCACTCGCATTCCATGCCCGGCTCCAAGCCAGCGAGCCGGGCTTCTTACCCATT
TAAAGTTTGAGAATAGGTTGAGATCGTTTCGGCCCCAATGCCTCTAGTCATTAGCTTTACCGGATAAACTG
CTGAGGCGAGTGCCAGCTATCCTGAGGGAACTTCGGAGGGAACCAGCTACTAGATGGTTTCGATTAGTCTTT
CGCCCCATACCCAGGTCGGACGACCGATTTGCACGTCAGGACCGCTGCGGGCCTCCACCAGAGTTTCCCT
GGCTTCGCCCTGCCAGGCATAGTTCACCATCTTTCGGGTACCATCGCGTACGCTCTAGCTCCACCTCCCCGA
CAGAGCGGGAGAGACGGGCCGGTGGTGCGCCCCCGACCCGATAGGGGGGGATCCCACCTGGCCCGCGT
GCGCCGCTTACCTTCAATGCGCCGTGGGTTTCTTTCAGCCCTTTGACTCGCGCACGCGCTTAGACTCCTT
GGTCCGTGTTTCAAGACGGGTGCGGTGGGTAGCCGGCATCGCCGCTGACCCCGAGCGCCCTTTTTTTCGCGAG
GCCGGTCCCCGCCCTGGCTGCGGACGCGGTTGGACGCGGACTGAGGACAGTCCGGCCCGGTCGACGGCAC
GCGCAGGGGACGGAGGGGCCCGTCCCTCCCCCGGGGGAAGGAAGGCGCGGAGGTACTCACACGGCC
CCGGGGTAAGCGGCGAAGTCGGGGCGGGAGGGCGCTGTAAAGCTCGCGGGCCGAAGCTCGCGAGCCACC

TTCGCCCCCTGACCCTTCCAAGCCGAACCGAGGCCGGTCGCGGGCGCACACCAGCGGAGGAAATGCGCCCGA
CGGCGGCCGTTAGCCCCGAGCCACAGGCGGCCTCCCTCGCGGGAGGAGCTCGCCGTGCGGACCGGGACGAC
ACCGCCGCGGGTTGAATCCCCGGGGCGGACTGTGCGGACCCACCCGTTTACCTCTTAACGGTTTACGCC
CTGTTGAACTCTCTCTTCAAAGTTCTTTTCAACTTTCCTTACGGTACTTGTTCGCTATCGGTCTCGTGCCAGT
ATTTAGCCTTAGATGGAGTTTACCACCCACTTTAGGCTGCATTCCCAAGCAACCCGACTCCGAGAAGAACC
TGACCTCAGCGCGGGCGGGCCGTTACCGGCCTCACACCGTCCCTGGGCTGAGCCTCTATCAGAAGGACTCA
GGCCCCCTGGCCCGCGCTGAGAGATACGGACTTCTATACGCCACATGTCCCAGCGCCCGTGAAGGACGGGG
GATTCGGCGCTGGGCTCTCCCTCTTCGCTCGCCGCTACTAGGGGAATCCTTGTTAGTTTCTTTTCTCCGCTT
AGTAATATGCTTAAATTCAGCGGGTTCGTCTCGTCTGATCTGAGGTCGCGCTCGAAGGGCTGTCGTCGCCAGG
CATACGCCGCTGCAGTGTGGGGGAGACCCCCACACGCAGTCTGGCTGGAACCATCCGGACCAGGGAACCTT
TCCAGACGAGGCTCGATTCCGCCGAGGCCGGGATGAAAGCTGGAGGGAACAGAGTCCCCCTCTTCGCAT
TCTACCGGAGCAATCTTTACCAACCTCTTTTTGCCGCGCACTCGGCCGACTGAAGGAGCGGGAGGACGGG
GGTCCGCTTAGCGGAGAGACGGCGCGCTTCGCGGGCAGCCAGAGAGGGTAAACCCACCGGCAGCCGCGGCC
ACGACGTCTCTCCACGGACACCCGCCACCACTCCCTCGCCGGAACGAGGAACGGTAGAGGACGGCGTGAA
AGCCGGCCTCCGAGACCCATCGCTGGGCATCGGAGACGAAAGCGAGAGGGTAAACCCGAGTTTTACCCC
AGCAATCAAGCCGTCGAGCGGCTTTTATCGAGTCCCGAACAAAGAAGGGAACCTGACCGAGTCTGCATTT
AGGAAGACGAAGGTGCCTCGGGTGGGCACCTGCGAACCTCCAGCCGCGGTCCCCTCGCACTCTCATCTCTTA
CCCCGGAGGGAAGAGGGAGAGAAGAACGAAAAGGTCCGATAGATTGGAATAGCGACCCTCAGACAGGCGT
GGCCCCGGGATTGACCCGGGGCCGCAATGTGCGTTCAAAGTGTGATGATCAATGTGTCCTGCAATTCACAT
TAGTTCTCGCAGCTGGCTGCGTTCTTCATCGACGCACGAGCCAAGTGATCCACCGCTAAGAGTTGTACTCTTT
TTGTTTGGTTTTTCAAAGCCATGTAACCAACAGACACAGTGTGACTGGTTGAAAAAGCGTTTTACCGCAATC
AAGGGTACGGGGCCCCCGGAACGCCGAAGCGAACTCGGGGTCAAAACCCGCGCCGGCCCGCCTCCGAGGA
ACCGCAGACGGAAGTGGGCCAGGTACCCGTCTCCCTTTGGCGGTAACCGTGGGAGTTTTCTTAGGACACCT
TGCGTTTTTGTGCTTCTGGAGGAGAGGAGGTGCGCCGCCCATGCCGACCCCGTCACTCTGGAAGTCTTTCTC
CTTTCGATCAAGTCATTGAGGCCGACGGGAATCCCTGTAGTTCCCGCCATCGGGCCGTGGCTTGTAAGGGGG
TGGTTTTTTTGGAAAAGGGAGTTCGCGCCGTGAGGCACTACCAACCCAGTCCCGGTAATGATCCTTCCGCAGG
TTCACCTACGGAAACCTTGTTACGACTTTTACTTCTCTAGATGATCAAGTTCGATCGACTTTTTCAGCACAGC
ACCAAGGCCCCGGAAGGAGCCCCGGCGTGGCCGATCCGAGGACCTCACATAACCATCCAATCGGTAGTAG
CGACGGGCGGTGTGTACAAAGGGCAGGGACTTAATCAGCGCGAGCTGATGACCCGCACTTACTGGGAATTC
CTCGTTCATGGGGAATAATTGCAATCCCCAATCCCCGTCACGAACGGGTTTAAGCGGATTACCCGCGCCTCT
CGGCGTAGGGTAGGCACACGCTGGCCCGCCATTGTGGCGCGCGTGCAGCCCCGGACATCTAAGGGCATCA
CAGACCTGTTATTGCTCCATCTCGCGTGGCTATACGCCACTGGTCCCTCTAAGAAGTTAACCAGCACAAATCACTCCACCA
ACTAAGAACGGCCATGCACCACCACCCACAGAATCGAGAAAGAGCCATCAGTCTGTCAATCCTTTCCGTGTC
CGGGCCGGGTGAGGTTCCCGTGTGAGTCAAATTAAGCCGAGGCTCCACTCCTGGTGGTGCCTTCCGCTC
AATTCCTTTAAGTTTTCAGTTTTGCAACCATACTCCCCCGGAACCCAAAGACTCTGGTTTTCCCGCACGCTGC
CCGGCGGGTTCATGGGAATAACGCCCGGATCGCGGGTCCGCATAGTTTACGGTCCGAACTACGACGGTAT
CTGATCGTCTTCGAACCTCCGACTTTCGTTCTTGATTAATGAAAACATTCTTGCAAATGCTTTCGTTTCGTC

CGTCTTGCGCCGGTCCAAGAATTTACCTCTAGCGGCGCAATACGAATGCCCCCGGCCGTCCCTCTTAATCAT
GGCCCTGGTTCAGGAAACCCACAAAATAGAACCGGAGTCTATTCCATTATTCTAGCTCAGGCATTACGGC
GAGTTTGGTGCCCGCTTTGAACACTCTGATTTTTTCAAAGTAAACGCTCCGGACCCCTGAGGACCGGACACC
CAACCAAGGGCATCCGGGGGGCGCCGGGAGGCAGGGTCTGGGACAGGCGGTAGCTCGCCTCGCGGGGAC
CGCCAGCCCACTCCCGAAGTCCAACACTACGAGCTTTTTAACTGCAGCAAATTTAATATACGCTATTGGAGCTG
GAATTACCGCGGCTGCTGGCACCAGACTTGCCTCCAATGGGTCCCTACCCATGTGTTTAGGATACGTTTATT
CCGATTACAGGGCCTCGAAAGAGACCTGCATCGTTATTTTTCGTCACTACCTCACCGTGTGCGTAATGGGTA
ATTTGCGCGCCTGCTGCCTTCCTTGGATGTGGTAGCCATTTCTCAGGCTCCCTCTCCGGAATCGAACCCCTGAT
TCCCCGTTACCGTGGTCACCATGGTAGCCCCCTATGCTACCATCGACAGCTGATAGGGCAGACATTCGAAT
GAGACGTCGCCGCCGCGGTGGGCGCGGATCGACCCGAGGTTATCTAGAGTCACCAAAGGGCCGGGGGAAA
CCCTCGCGGGTCCCCCGGGTGGTTTTTGGGTCTGATAAATGCACGCATCCCGGCCCCCCCGAGAGAGGACC
GGTCAGCGCTCGTTTGCATGTATTAGCTCTGGAATTACCACAGTTATCCGAGTAACGTGTGGAGCGATCAAA
GGAACCATAACTGATTTAATGAGCCATTCGCAGTTTACTGTACCTTCCGTGTGCACTTAGACCTGCATGGCT
TAATCTTTGAGACAAGCATATGTTACTGGCAGGATCAACCAGGTTGCCCGAGTGAAGGTTTGACACACACAA
ACCCAGAGGTTTATGAGGGGAGAGGGATCGGCCCGTTCGGCCCGACCCCGCCGGCATGGGGTAACGGGCG
GAGAAGCTCGTCCAACCTACCCCTCAATTATCCCGGTGGATATAGAGTGGCAAACCTCCGCTGCGGAAGTTTT
CCTTTTGCTTCGGGGATAACAAGGAAATGATCGTTTTCAAACCGGGGCCTGGAGCTTTTTGCCGGGGGAGC
TCGGGTGCTAGGAGCCCCGAAAGGACTCGCCAGAGGCATCCGTCCGACGGCTCCAGGCCGACCGGCTTTC
GTTTAGGGCAGATTGACAGGTAGACGGGCCGTCTACGACTCCCTTGGATAGGGTCGAAAGGTTCTCCCCCT
GTCGGAGAACCTCCCTGGAGGCTGGGCCGCTTGGTTCGGGTAGGATCACCGCGCATAGTGGCCCCCCCCGGT
CGAGGGATGGAGGCCGCGGATTGCCTACCCAGGCGGGCTTGGTAATCCTTCAATCAAGTAAAAATAAAA
AAAAAAAAAAAAAAAAATTTGACCTCCATCTTACCCGGGGGCGAGGGGCGTGATGGAGGCCGACCTGAGGA
GCCCCAAATCGGGGTCGAAATCGGCCTTAAAACGGTCCGAATAGCCCCATTGGAACCATATTTGAGAATTTT
TCTAAGTTCTGCCACACTTAGAAAATATTCGGGAATCTCATTCTGTGCTTGGACCCCTAGTTATTGCACCTCT
CCTCATATAACCCGACGGCACCTCTCTGCAAAAACCCCCAACGAACGAACCTCCAGAGCAAGGGAAAATCCG
CTCCGAGGCCGACCAGAGTGCACCTTCGTGCGATACGAGAGCCGGGCGAAGGGGGTTCAGAAATCCTGGAGC
TCGCCACCCCTCTAACTTGGGTGTGAAATGGTAGACAAAAGTGGGGACTTAGTCATTTTTTCCACTTTTCG
GACTTGGCACTGGTAGACCGAGTGGGACATAGTCTTTTTTTTTCCACTTTTTCGGCCAATACAGCCCGAATTAGG
GTTAGGGAATTTCCATTGAAATTGCGGTGAAAGTGCAAGCCTTCAAGTTGCAAAAATGGCAATCCCTCCAAA
TGCAATTTTTCCGGCTTGCATTTTCGCAGACGGCCAATGTGCAATCTGCAATTTCCAGTTCTGGGCCTCTCTA
CAAGGCCTGTAACCGGGGGGATCCAGCCCTGGAGGTCTCTCCAGTCTGGGGCCTCCCTCCAAAGCCAGTCC
CCTGGCTCTGGAGGTCTACCCAGCCCTGGGCCTCCCTCCCAAGCCAGTCCCCTGGCCCTGGAGGTACACC
CAGTCCCCTGGATGCCTCTCCAGTCCCCTGGCACTATCCTGGGCATGGCTCCTCCACACATCCGGGGGGA
ACCAGCCCTGGAGGTCCCAGCCGACCAAGGTGACCCCTGGTAGCCAAAGGGCCACTGGTAGACCGGGTGGAC
CACATCAGCGCGGCCCTCAGAGCTCCGGGCCACTGGTAGACAAAGGGCCACTGGTAGACCGGGTGGAC
TTAGAGGAATTTTCTATTTTTGGTTTTCTGTCTCCAGGGGAGGGAGACCGCGAGGGCGCGAGGCCCTACCCAC
GGCCGTGCCTTACCTACGGGCGGACTCGTGAGCTTGGTCGCCCCCTCCCCTCCGGGGCCACTGGTAGACCG
AGGACTTACCAAAAATAGGTGCTTTTGACGCGGAACGCCGCGCCGCTCCAGACCCCGAGAGGGTAGCAC

CATTTGGAGACAAAAGTTAGAGTACAGGGATGATTCTTAATGGATCGCAGCGCGGTGCTGCTCTGCCACTTA
CGAAACCCTGACACTGAATCAGGTCGTCTACGAGTCATTTACGCCGTGAACCACAAACATGCGGTAAACGC
GTTTTCGGAGATGGGGTACGGCATTCTTCGGCCGACCCCGGGACCGTTAGCGAGCGGCTCTGCCACCGGG
GAGTGGACCCCCCGGCTACGCTGACCAACCGTAGGTCCTCGGCACTGTGGATATCGTTCGGTCTAGGCGG
GATTCTGACTTAGAGGCGTTCAGTCATAATCCCACAGATGGTAGCTTCGCACCATTGGCTCCTCAGCCAAGC
ACACGCACCAAATGTCTGAACCTGCAGTTCCTCTCGTACTGAGCAGGATTACTATTGCAACAACACATTATC
AGTAGGGTAAAATAACCTGTCTCACGACGGTCTAAACCCAGCTCACGTTCCCTGTTAGTGGGTGAACAATC
CAACGCTTTGTGAATTCTGCTTCACAATGATAGGAAGAGCCGACATCGAAGGATCAAAAAGCGACGTGCT
ATGAACGCTTGGCCGCCACAAGCCAGTTATCCCTGTGGTAACTTTTCTGACACCTCCTGCTTAAAACCCAAA
AAGCCAGAAGGATCGTGAGGCCCGCTTTCGCGGTCCGTATTCATACTGAAAATCAAGATCAAGCGAGCTTT
TGCCCTTCTGCTCCACGGGAGGTTTCTGTCTCCCTGAGCTCGCTTAGGACACCTGCGTTACGGTTTGACAG
GTGTACCGCCCCAGTCAAACCTCCCACCTGCCACTGTCTCGGAGCGGGTTCGCGGCCCGAGACCCTCGCGG
GCCCCGAGCGCTTGACACCAGAAGCGAGAGCCGCGGGGGCCCGCCTCCCCGCCTACCGAGTTAGTGA
GGAAACGATAAGAGTAGTGGTATTTCACTGGCGACACCCGACGAGCGGGGCTCCCACTTATTCTACACCCC
TCATGTCTCTTACAGTGCCAGACTAGAGTCAAGCTCAACAGGGTCTTCTTTCCCCGCTGATTCTGCCAAGCC
CGTTCCTTGGCTGTGGTTTCGCTAGATAGTAGGTAGGGACAGTGGGAATCTCGTTCATCCATTATGCGCGT
CACTAATTAGATGACGAGGCATTTGGCTACCTTAAGAGAGTCAATAGTACTCCCGCCGTTTACCCGCGCTTC
ATTGAATTTCTTCACTTTGACATTCAGAGCACTGGGCAGAAATCACATCGTGTGAGCAGCCGCTTAGAGCCC
TCACGATGCTTTGTTTTAATTAACAGTCGGATTCCCCTGGTCCGCACCAGTTCTAAGCCAGCTGTTAGGCGC
CGGCCGAGGCACCGCGCCGGGAGGGTTCCCCCGCGCGCCGCGACCCCTCCCGGCGCGAACCAGGGAGAGGA
CGGACGCGGAGGTCCCGACGCAGCACCGTAGCCGGGGAGATCCGCGAGAAGGGCCCGGCGCACGCC

2n=50

Pri50Sat06-236	ACTACATTCAACACACAAATGAACCTAAACGCTTCAGACATGTTGAAAACAAAATGTTAGTGCATTTTACAC AGTAACATTATGTTTTATAGCTTAAAAACTCATTTCTGAACGAAAGAGAAATTTTGAGTTCAAAAACAACGCT TGATACAGTTTTTCTGTGTAAGGAGTGAGAGAGGAGCTCAGAACTATCCGCAAAACGCTTTTCTCCGCTGA AAGCTTCAAATACAGCCTTT
Pri50Sat29-236	GAGCTCCTCTCTCACTCCTTACAGCGAAAAACTGCTTCAAGTGTTGTTTTGAACTGAAAATTCTCTTTTCGAT CAGAAATGAGTTTTTAAGACCTAAAACACATTATTACAGTGTAATAATGCACTAATATCTTCTTCTCAACATTT CTTAAGTGTTTGGGTTCAATTGTAAGTTAAATGTAGTAAAGTGTTTTTTCAGACGATTTAAGCTGAGAAAAG CGTTTTGCGTATAGTTTT
Pri50Sat49_lncRNA-1614	TAAAATTTAAAACACCAAATCCTAAAACCTAATTAGATCCTCACCCAGGGTATCCAACCTGGACCCCTCAA ATTCAAAAAAACCCCTAAGGCCAGCAATAGACCGTTGTGGGCCCTTTAAGAAAATATGTAAATCTATAACAC ACAGAATTATGGACCATGAGTTCTTTTTAACTTTGAACATTAATAACTTAAAACTATAGTAGATAAAAAAC

	<p>TCACGGTTTTTTGCGTCAAAAAGACCACAATCTCCTCTACAAGATAGCGTAGAAAACCCATCTGAAGTCTGG TCCACCCTGATGATGAACTGTGACACTATAGAAGGCTTTACTGCCCTCAGCACGGTGAAATCTAGGCGGGGG GAGCGCTGGTGGTAGAAGCTTGGCGGCATCCAACAGGTGTGGATGCTGTCCGCCATTAGTGTACACCCCCAG GGGATCCAAAGCAGTAGAAGCTTGGTATCTTCCCAGAAAGTATTCTAGTGTCTGGGGAATACTCGACTAGAAT GGTAGGGAAGCTATCCGCTGCCTTGGATTTCCACCCAAGACTGTCTGGGAGGTGACGTATTATACGGGACCT ACCTTCCCAGGTCTTGGTATCAGCCACCTGGCAACGGGTGGCTGGGGCATGGGTTCTGGAGCCGGGGGGCCT GGGAGGCAGCGATTTCGTTTCGCGGCTGGCCTTGGTCCTCTTGCCTCTGGAGTCTTAGCTCAGGCTTCGGCCAG AGTAGGCCCTACCAGCACGAAGCTACCTGGGATACCTTGCGGTGTTAACGAACAGGTGGCTTGGGCAGGA GGTTCTGGAAGTCCTGTGGTTGGGAGGCAGCGGTCCGTCCGCGGCTGGCTCCATCCTCGGTGGCTCTGGAA ACTCAGCTCTCACCTTCTCCTCTCAGGGAGATGTTGGAGTATGCCCAAGATGAGACCACTTCGCAGGCTCT GCCTGGATGACCTGGGAAACCTAGCGGTGTTAACAAAGGGTCATTGGGGTAGGGGTAAGTTTCCAGCCTCT CTGGGTCCACTAGGACAGCACGCAGAGAAGGTGGTCTGGGGCATAACAGACCTTAAATGGGTGCCAGTCTT CTATCGGTGACTAGCACCAACCACGCAACAGCACACCTCACACACATTTTAAAATTTAATATCAGTTCTA GCCGACCTTAAACAGACCAATCAATTTATAGTGGTTAGTAACACTTTATCATGAAGGTACCAACAGCTAATT AGATTGGAGGCTGCAGCTTTAACCTACTAACCATCCATAATATATAATGGCCATTATTAGTAAATATTTTTAA GCAATAAGAACACCAAAAATATAAAAAATTGACCTCAGGACACAAGCCTAGGACTGATCAACCTAAGGCGGG CCACCTCAGAAGAGGGTGTATAGTGTAAAAGGCCGAAACACCTGATGACTCTGAGGCACACTACTGATG ATGTGTCCTTTATTAGAATTTCAACTCAAATAACAGATTCACTTGGACTTCTGTAGAGTTATACTAATCAC TCTAGTCATCCTGCATAGGACCTATTCCATATAACTCCACACCTATTCACTCCCTGGTTGGACAGCACATTA GGGTCACCCAGGATATCCAACCTGGACCCCT</p>
Pri50Sat59-177	<p>GTCTTCTGGGGACCCTCAAGGTCTTCTTTGGTTTGGGCTGAGAAGAACGAAGTTCAAAGGATGCAAAA TCTCCCAGAACCCATGTGGGCCTCGGTGGCCAGCGGTAGAGTCCACCGGCCCTTGGGGTCTTCGGAGTGC CGAAGGTCGCCGGTTCGAGACCGGGTCCGGGCG</p>
Pri50Sat61_lncRNA-631	<p>CAATTCAACTGTCCGCCATTTTCGTCCTCGTTTAAAGATATCGACTATTCCTTCGCAATTTATCATCAAATATGTC AGATGTTTATTAACAGCTAATTTCAAGAGAATTCAACAAATTCCTTAGGAGCAGTTTGACAAAGTACGCAAA AAATGTGTGAAAATTGCACAAATTTCAAAATGGCCGACTTCTGTGGGCGGAGCTAATACAAGCCAATTGC AAATATGTGCAGGGTCACAGTATCTACATTCTACCAAAATTTGAGAAGATTAGACAAATTTTGACACATCC ACAGCTAATTTATTAACCGAACAAATTAGGGGGCGCTGTAGAGTCCGCTAGCTACATATGAAAAAATTATCA CAATATTTGTAAAGTGTTTTGGTCTATGGAATCTGACAATTTCAAGAGTTTTTGAGCATTCCCCTAATG TCAAATATGCATTGAAACCTAGGTGGCGCTATAGAGCCAATGAAATACGTATATATGAAATTTCAATTCAG ATCAAAGGACCGCGTAAATCAAGCAGGCCTGTAAAGTTTTGTGAGTTTTTCATCCATTGTAACCTCCTCAAAC ACCTACCAACACCAAGATGTATTCACTTCTGTTTAAATGGGGCATCTCAAG</p>
Pri50Sat75-152	<p>TTTTTATTTATATTAATAATGCTTTAAATTCATTTTCTCACGATTGGCCCATAAATACCACAACGGTCTACTA CCATTCGAAAGAGACTGTCCTTGGCCATCAGCGGTGCGAGTTTGGGCAAATTCACCCAATGGTTGATTTTA TATGAA</p>

Pri50Sat113-35	CGTCCTTTAAAACAGAAAAAGTCATAGAGTCTAC
Pri50Sat122_snRNA_U2/U5-1036	CAATTTTTCTTTTTATTTGATATGTCGTTTAGCTTGTCTCTTTGTGTTCTTTTGTGTCCCTTTCTGCTCTTAAGG CAACTACGACTGTACCCGCCAGGTGTGACGTCATCAGGACTCGGGCGTCTTAACTGGAGGAAAAGACAAA AAAAGTTGCTGATTTGCCTTCCCTCTCTATCTTCCTTCGCTCCCTGCCTTCCCTCCAGCTGCCTCGGCCCTTGCT GGACCTCGGAGGGGTCTGCGGACTTACGGGTATAGCTCGGGTTTCTCTTCATTACAGTATAAATCTTTCGCCT TTACTAAAGATTTCCGTGGAGGGGAACGTTGACGAGTATTGTTTATTTTTGAAGGCTTTGCTTTTAGCGAA GCTGCATATTTGGTAAATAAAAGAAAGAAATAATTAATAAGTAAATAAAAAAAGGACAGGACAGGAA AGGACAGGACAGACCTCTTTGAACAATGGCTAGAGTAGGGTGACATCAAATGGGTGGGAAAGCTGTTGGCT TCGCCCCCTTAGTAGCTACATAGTAAATGAAGGGCCGCGCAACCAGGTGGCGATATTCGGGTTATCGCTTCT CGGCCTTTTGGCTAAGATCAAGTGTAGTATCTGTTCTTATCAGTTAATATCTGATACGCTCTCTATACGAGG GCTACATATTAATGGATTTTTAGAACAGGGAGCCAGAACAGGAGCTTGCTCCTTCTGCTCCACGCATCGAC CTGGTATTGCAGTACTTCCAGGAGCGGTGCACCTTCCCTACAGCGGTGCAAAGAAAGAAGAAACGACCGGG TCGTTGTCTCAAGTACTCAAGAGGACGTGTGGTGGACGGCTGAGGGGGAGGGTTCGTTTCTTGTGATGTGTCT TGTTTTTTTATTCCTCTGCTTGTGTTTTGAAGTGAATGAATGCATGTTTGTGTTGTTTTCTTGTCTATTTAT TTATTTATTTACGTATTTTCTGTCAGGGGGAGGGTCTTTTTTCGAGCGTGTCCGTTTTCCCTCCGTTTTTGTCT TTTTTTTTTTTTTTTTTTA
Pri50Sat147-43	CCGTCTCACATTTACCAACTAACATCAAAAACAACACAACCAC
Pri50Sat160-47	CTAACATTAGCTACCCAGTTAAATAAAGCACTGGTAACTTTAGCTAG
Pri50Sat172-578	GAAAAAAATACTGAAAATTGATCAGTGAAAAAGCAACTTTCTGCTAATTTTACACTCTTCCCATTAGAT TTCCCTTGAATTTTTGACATCACATACTTTAGACAAGCACCAGTTACCAGTGTTTCAGTAATGACTCCAGAGC CAGTACAAGAGAGCC TGGGAGACGGGGGCCTTTTAAAGCACGGTCCCGGCACATTTTACTTCTTCATGGTCACTGCTCAGAATTTTCA AAAAACACAAAAAAACGCAAAAACAACCCCTAAAACATCTATTTACCACAAAATAACACCACAATACCT CAGAAAAAATCCTCCATACACAATTCCTTAATAATTAATCAAATTTAAAAATCAAGTGCACCCCAATGCAT CATGGGAAAGCTCTATGTCACAGTGTACATGACACTGTGTGGGTGAGTGTGTTGTTTATATTTATATATA AATGTACACAGTAAATACAAAGCCAGTGCCACAAAGGAACTAACTCACTAAAGAGCTCATGACAAA GATGTGATGAAAAATAGGACTGATATTGTAACACAAGTAAAGATTTGTAAAAA
Pri50tRNA_Arg/Met/His-1251	CAAATTGTTGAGGCATCACTCTGCTTTTCTCAGCGAGCAAGAAGTTTCCCATACCGGGAGTCGAACCCGGGC CGCCTGGGTGAAAACCAGGAATCCTAACCGCTAGACCATATGGGACTCTGAGACAAAAAAGGCACGAGGAA AATTGTGTCATTAGACTTGTGAGTTTGGAAATAAGCTCTCTACAGGACTATTTTTCAAGTGTGGAATCCGT AAGAATTTTGTCAAGCATCCAGAAACAATAGGTAGAATTATCTGTAAAGCAGACACATGCCTTATCCTTAGG ACATGTTGCACAACGCTTTGCTGGAGCAGTACCCTGAAAAACAAGCCATGCACCTGTTTTACTCAAAACGA TTCCAAAGACAAGCAAGAAAAGCCTACTGTAAGCTTAAAGAAGTGGCGTGACCCGGATTTCGAACCCGGGT

	TGCTGCGGCCACAACGCAGAGTACTAACCCTATACGATCACGGCACCCCACCATCCATTGGCAGAGGTCTT GCTGTCTTCCCAGGCAATAGGCTTGTCAACTGATTGTACGGCACATGACTAGATTTCTACATCGGACGACAG AGGTGGAACAACACCTCCTGGATCATAACAAGAGCCTCCTGTTACATGATCCACTGTACGCGTTTTTCTGAT ATACTTGCAAGGCTGAGTAAATGCTGGGATCAGATTCATTTTCCCCGTTCCAAGATGTTGTAATTAACACAC ATGTATGTGCAAGGACAAGGAGAACATAAAGTCTTTAATGTAGATTCAGGCGCTTAAACACTTGCATTAACCA TGTCCTTACATAACAAGGTGAGTCTTTTAAAGATATTAGGGGCATGTTTCTCATTTAAAAATGACTATTGCC CGTGTGAGGATCGAACTCACGACCTTCAGATTATGAGACTGACGCGCTACCTACTGCGCTAACGAGGCTCAG GCATGCATCTGGTCACCTGCCAGCACCACAGGGATGGCAACAACCAAAGTAATCTTTGAGAATGTTGGGG AGCGACAAAAACAGAGGAGCCGAGCCAGCCAGGAGTGAACCTAGAATCTTCTGATCCGTAGTCAGACGCG TTATCCATTGCGCCACTGGCCACCCGCAAGAACCTGTCAGGCAAAGTGGATAAGAGAACATTTCTGGCCTA CGTCACACCAAATGGGCCTCTTAGCTTCAAAGTAAAGCCAGAGCCCAACAACCAAGCTGCACTATTA GTTCCGACAGGATCGCCATTGGAACGTAT
Pri50Sat177-178	CAAAATTTGGTCCAATTTACCATGCCATGCATTTTCCAACCTCGTTTGAGTGATTTTTTCATGGAAAAACAAG GTTTCATGAATGGTCAAAATTTTACCAAATACATACCGCGGGTCAAAGCTTTTTTCTGATTTGGGTCAA AATTGAAAAGTGCTTTTTGTCCCAATTCTGAG
Pri50Sat181-41	CTGTGATGTCACAAACACAGTCCATTTTGAATGCACTGTTC
Pri50Sat186_ncRNA-236	CATATTAGTACTGTATAACAATGCACTAACATCTTGTTTTCAACATTTATCAAGTGTTTGGGTTCATTTGTGTGT TGAATATAGTAAACGCGTTTTTCCGTCGTAACAGGCTGTGAGAATCGTTTTGAGCATAGTTTTGAGCTCCTCT CTCACTCCTTTGAGATGAAAACCTGCATTTAGTGTTGTTTTGAACTGAAAACGTTTCTTTTCGTACAGAAATGAG TTTTTACGACCGAAAA
Pri50Sat199-78	TCAGTTGACTTGTGGAGGAGCTGTAGTACTTGTAGACATTGTAGTTGAAGTTGGAGCAGCTGTTGTGGAG GTCATT
Pri50tRNA_Tyr/Arg-548	GGCACTGCAGGCTCTGTGGCGCAATGGATAGCGCATTGGACTTCTAGATGATGAGCTGGAGCAATTCAAAG GTTGTGGGTTTCGAGTCCCACCAGAGTCGAAGGTTGTGCTGTTTATTTGCTTTGTTGCAGAACAAACCAAGCTCA AAATTTAGAAATATTCCACTTGTCTTCATAGAATTGCAATGACTTTGTGGTGCCATAGAGAGTGCATTCACT TGTAGTTGAAAGACTGAAGCAATTGAAACATGTTGTTGAGTTGATTTTTGAATTTTTTGGAGGAACTGTGG GTCTTATTTAGAAAAGTTCTGTTTGTGTTGTCACAAGGTTTTAAACCAATTAGAAAAATTAGCACCAGTTTC CCTTAAACAACATTGTAATGATCTGCACCTTGACCATGGAGCGCGTGGCGCCGTGGCTTAGTTGGTTAAAGT GCCTGTCTAGTAAACAGGAGATCCTGGGTTGCAATCCCAGCGGTGCCTTTGCTTTGATCAATATGCATTGATC TTTTAGTTAGGACTGCGTGTCTTAGACATGCTACAGTTGTTTTCTCGAAAAATTGGTCAGTCTTGGTCAAG CGA

Pri50Sat230-187	CTATAGGCGATACATCCCATTAATGATATTATACAGAATGGGCATAGTTTTATCAGTCTACTGTCTAAACTA TACATTGTGTCAACATAAACCCATAGTCAAATTAATTCATCAGACCCTCCTCTGTCTAGAAAACATGGAGATT ACTATCAAAAAGCGTAATTGTACATGTAGGAATCACTTATA
Pri50Sat236-68	CCCCTCACTGCCTCCACCTCTTACAGTTAATAAACACTAATACTGTACTCTGACTGTAATCCAACA
Pri50Sat247-88	TTCTGACACTCGCAGTGTGTTTTACGACACTCCAATGTGTATTTTTAACACTAAATTGTGTTATTCTTTTTCA CTTTTAGTGTTAAT
Pri50Sat256-27	CTCAGAGTATAAGAGGAAAGTTTTTAC
Pri50tRNA_Asn/Lys-419	CTACCGACTGAGCTAGCCGGGCTGATGCATTTAGCCTCCATGTCGTAGGTCCTACACTGCAACTATCGGGGG CACTGTTAAATGTTTACAAGCCATGTGTGAAAGCGTTTGCCAGCACCGTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTGGTGGTTCGAGCCACCCAGGGACGCTGAGTTTTGTGCTATCTTGCAC ACTTCTGTTCCAACCTCTGACTTACTGCACCTTTTCAGTTTGACTAGTGCTTTTAAAAGCGCTTGATTGGCTGTG ACATGTTTGGCAGATGTGAGTTAGGCTTGGAGGTGAAACCTGCAGGAAGTTCAGCCTTCTGTTAAAAAGTGT TCTATTCGCCAACGTGGGGCTCGAACCCACGACCCTGAGATTAAGAGTCTCATGCT
Pri50Sat264-40	TACACTCTCTATAATATCAGATCCTGATGCTAAAGTACAG
Pri50Sat270_tRNA_Arg/Gly-358	GTTTTAAATCAATGTCATAAATGTAAGAAATAAATGTAAGCAAGCCTTAAAGATGCGCTAACCGGGAATCG AACCCGGGTCGCAAGAATGGGAATCTTGCATGATACCACTACACCATTAGCGCTGATAAAGACGACTGCTA GTTTCAAATGTAAGGATTGGACTTTCTCCAGAGCATTCAAATGAAGATCTCTGATTTACACAAAGGCCT TTTTGCAAAAGAAGCATGACATGATCTTTGCCTGCACGCATATCGTGACCCAGATGGGACTCGAACCCACAA TCCTCAGCTCCGAAGGCTGATGCCTTATCCATTAGGCCACTGGGCCTGCTAATCTTCAAATTTTCTTTTTAT
Pri50Sat272-33	GAGGAGTAATGGAGGTGATTTAATGTATTAGAG
Pri50Sat278-21	CAGTATGTGTCCAGTATAATG
Pri50Sat282-41	TGTAAAGTCGTGTATTTACTGTGATTATTGAAGGAGAAGTG
Pri50Sat19-32	AAGGTGTTGCTAACGTTTCCAAGGTGGTTGCT
Pri50rDNA5S-195	GGGAAAAGCTTACAGCACCTGGTATTCCCAGGTGGTCTCCCATCCAAGTACTAACCCAGGCCAGACCCTGCTT AGCTTCCGAGATCAGACGAGATCGGGCGTTCTCAGGGTAGTGTGACCGTAAGCGATTGCAGAACTCTCATCA AGACTACTTATGCTAAATTATCCTTCAAGGCCATAGTAATAGCAGCTAGAG

Pri50TE_RC-1656	GGGTTCAAGCACTGAAAGTGCGTAGAACCCCTATTGTTTTTGCTCAGATTTTTCTTCTTCTTCTTATTATTC TTTTTCCCTCCGAAAAGCAGTTGTGCAGCCTAAACCGTAACTCGTAGAGATATGAAACTTGGTAGGTAGATG TAGGATCAGTGCATCTCGGTGGACAACAAAAATGGCACCATTGATCAATGGGTGGCGCTATAACAAGGA AAGTTCATTTTTCACTTTTTCTCAATAACTCCAGAACCATAAGACCTACATTCAAAATCTTTTTTTGATGGA TTCCTTGGGTCAATACCTACAACCTTTCCAATTTGGACCATGCACCTCCGTCTACTTAGATTTTTTGTAAATTTG CATAATATGCAAAACATACTTTTGCGAAGTAGTTCTAGGAATTTTGACCAATTCTGACATGTTTGGTATCAA ATACTCGGTAGAGCATGCGCTTCAATATTCATTAAGAAAAAGTTGAAATTTGTACACAATATGGCCGCCATA TGCAAATTAGTCCTTCCGGATATATGCCCATTCCTCAACAGTTAAATGTGGAGCACTGTTTCTCAGCAAC CGTGCAACCTACCAAGCTGCAACTTTGCATGCAGAATCTATCATAACAGTCTAAAGGATGTTCAAAGGGCA ACTTAATTAATCAACATGGCTGAACACCATCAGCCAATCAGCATTGAGCAGACATTTTGGCAGGCTGAATGT TGCACAATCTGGATGAAACTTGGCAGTTATGTTTCAAGGTGGAGACATTTTAGTGACCTGCAAAGTGCTGAAAC AATCCGGCCACTGGGGGCGCTGTTTACAAAAAAACTAGTATATGTCAATCATGCTGTAATATTTGACATC TAATTGGTTTTGCCATATTAAGTACAGTGGCTCTGACAAATTTGTCAATACAACCTGTGTTTAAAAAATGCTTT GTTTATTCATAATTGCTAATTGTTTAAAAATAGCTAATTTGAAACTAGTCTCTGGATATTTGACCTATCACCT CCATTTCACTGTTGTTGCACACTGTAGAGTCTGCAGGTAAATAGTCATCAAAAACAATTGTGAAATTGTCAC TTACAATAGTGTCCAGGCATCAAGCAATCTGTGCGTCCTTGGAAATGTCTAACCTGAATTGCTTTAACTCTGCA GTATTTGCTCTAATCATCTCACAATGTTAAACACCTATATACAAAATCACAGTGATGAAGCGCCACTTAATTC AGAAGTAGATTTAGAACAACTTTGTAATTACATGTCATGTCAGAACATGCACCTCCTTACTTGTGTTTGTCAAAC ACTTCAGCTTATTTGAAAAGGTCAAAAGGTCAATCTGAACACCACTCTGTGAATATTCTGGTTAAAGTCTCC TGATCAAAAACCGTAACATGTGTTTACTAATTGTAATTAATTAATGAATATTCCACAAAGTTCAACAGAT TTATGTGAATTTAAGCACTGATCATGTCGAATAGCCTTTGCCCATTAATAACAGGTGAAATACTTTTGAT AATCTATAAATGTGACAAAGTGTCTCAAATTTGATTAGACCTTCTTACACATCACCATCCTATGCTCTAATTG GCGCCATTGTGCCACTTGGTGCTTGAACCCGATGGTTGCCGCTTGGCGCTATATTTA
Pri50Sat82-42	CACTATCAGACCTCACTACAACCTCCATAATTCACCTGCACC
Pri50Sat86-54	TTGATCAGGGCTGGCGCTCTCGTTCACTCAGTGATTCGTTCACTCACTGAACTG
Pri50Sat98-69	TGATTCTGTATCTACTGTAAGTGTAGATTAATTACAGAGCAGTATGGGTACAACAGATTACGGTGCTGG
Pri50Sat109-106	TCCGCCCACTTCTCTACTACAATACAGAGTTACTATATTCATAATACCACTTACAACCTTACTACTGTACAC AACACAAGCCTTATGTTCAACATGACCAGAAGC
Pri50Sat110-22	CCTTTTGGGCATCTCCAAGTGT
Pri50TE_hAT-61	GTGAAGTAAAAATAATACTCCAGTAGATACAGATACAGCATTTTAGTACTTAAGTACAGTA
Pri50Sat129_ncRNA-51	GTTATCAAACCCAGAACCTACTAATCTAACCAGCACAGCACCATCCATCT

Pri50Sat132-48	ACATTCCACACTGAAATAAAGTAGTTAACATTAGCTAACCTAGCTAAA
Pri50Sat135-107	CCCCGACCACCCAGAGGTCTGAAATCGCAGTATGTTGTCCAAACGGCACCCCTGGGGTTGCCTGCAGAGTT TCGTGCCGCAGAAGTCAACGTAGGCCAAACTCCTT
Pri50Sat136-24	CTAGCTAACCTAGCTTTAATTTAA
Pri50Sat138-80	CAGTGATCAGTGAGAGTGTCTACCTGTAATTAECTCTATATAACATTAGAATACTAAACCCAAATAAACATT AATAAAGT
Pri50Sat143-48	AACGTTACTAACGTTACCTCCATCACCTCAAAGTCCCTGTGTTATCCT
Pri50Sat150-29	GGGGCTAAACTGCTGTAGGCTGAATGGGT
Pri50Sat157-27	TCCCTTATCCTGACCCAACTATTCCTC
Pri50Sat163-1345	TCCAACAAATAGCACTAACTGAATTACTATGGAAAGGTCAGTGCCGTTGTTAGAGTACCAATTGTGACATTG ATCCAGTAGACACTGTTTTTAATCTAAATGTCATTTAAAATATTAAGATTACATTTTCTTTGTATTTACCTT CTTGATGTCCTGCAGATTCTAGAAAAGGTGATCGTCCTGCAGCAGATCTTGTTTGTGCTCACCTTGTCTAAA GTCTTACTATTTATGTATTCTGAAATCTTTCCATAAATTGAGGTCATTCACAGCTCCTCCCTGTA AAAATCACTC TCTGACATCACAATAGGCCATCCTGATTTCTGTAAAGATGGATACTTCACACTTGTAAAAGCATTGAAAAT TCATATTCGAATATCAGTTCATAACTACTAATTGTAAAACAAAAAAGGATTTGTTAAGAGTTTATTTGTAA TAGGTTGGTAAAAAAAATTA AACTTTACCAAAATGAGTCTAGCAAGCTTTCTATTCTACTACAAAATATTA GCCAATTTTGACAGAATCTTTTGTACATAATAAAGGAAGAAGTTACTGATGACGGAGACTAATTTAGTCACA TTACATCAGGTTCTGTGTAGAAATATAAGACAAACACCCCTTGTTGCTTGTGCCACTTAA AACTACCCCAAC AAATAGCACTTACCAACTTTTACTAAAAGGTCAGCACTATTGTTAAAGTAGCAATTGTGACATTTATTTTAA CGTTTTCTTCGATACAACGATCCATAAGA ACTAATAAATCTTTAATATTCTTTCAATATATAAGAGTGCTTTA CACCAAGAAACCCCATGTTAAGAAAGTATTTTTATATGCTTACTTCTATTGGTATATAATATAATGAAATCA TG TAGTTCATATTCAATTTAAATTGAATAGTAACTTTAATTGCTGAAAATGAACAGAATATGGGTA ACTTTC ATTTAGTTGCTCAATAGTACCATTTTGACTACCCCAAAGAGAAGCACTTACTCCTTCTATTAAATAATCAATG CTTCACTTTGTGATGGTCACCATAGATATTCAAACTATAATTGTA CTGATAAAAATTTACCATGCAAGTTAACAC TGTTCAAAA ACTTACACCAAATTGATAGAACC TTTTTAACACTGTGTTTGTAAATAAATGCTAAAATTTGTACCA GATGGACTTGGTCTCACTGTTGATGGTGTACTTCAAGTGATGCCCAACTTGTTAGCACTTACATTTCAATT TAGGTAAGATGGCTACAGTTAGAGTGACACTCAGAGATAGAAGAGATAGAAATGTGCACAGGTTTATCAGC AAAATGTATTTGGCTGTTAGAGTCACTTCAA ACTAC
Pri50Sat165-52	CGGAGTCTGAGAGAGAGAGTACAGTAGGTCATGCTGCTTCTCCATCAGCAGC

Pri50Sat193_snRNA_U2/U5-1114	<p>GCGCGCTCTCCCCTGACGTGCTTGTGTGGACAATGTAGAGTTGTGTTTCTGCCGCTTTCTTTTTGGCTGAGCA AAGTATATTGAAAGCAGCTGTGGGCGTTCCTTGCCGTAGAAAATTGTGGTCCCTTGTGTCCCACGCTTTGGA GATCTGCTCCAGTGTCAAGCTCAGGTCAGACTTGCTGATGTTTTCCAGTAAGCATATTCCTATGAAAAAAA ATAAGTGGAAAGGTAGCTCCAAAGGCTTGAATTTCATAGATCTGAGCAGCACCTTCCTATGACAATCTTCAA TGCAGCTCAGTCATTGCTTCTCAGCTCATGCTTGATTTGATTTAGTCACCTGGCTTCCAAAGCACCTTCTTCTG CATTGGAAGGTAGAAAATGTGCTATGCATACGACATTTCTAAGTACTGCCGCGTGTGAAACAGAAAGCGTCA CTGGAGTCCAAAGTGGGGAGGGCTTGTCACTCCCAGCCTGTGAGGTGTCAACAGGACACCTGATAAGCCTCT GAGGCTGCCCTCTTGCAGAAACCCTGTTGAGGCGCGGTTGGTATATGGGAGAGACGTGGGCAACGAGGA GCGTGCTTGCTACGGCAGCACATATACTAAAATTGGATCGATACAGAGAAGATTAGCATGGCCCTGCGAA AGGATGACACGCAAATCCGTGAGGCGCTCCTTGTTTTGTGCCGTCTAACGAAAAGTAAGGCCCATGTCAGCC TCTGCCCCGTAAAGTCTTGAGATTGCTGTGCAAAGGCATGAGAAATGCCCGCAAGGCTGCTAAGTAAATGC CCTGTCAGAGCCTAAGAGGTCGCTACATGCAAGGACTCCACTAGAGCCTTTAGTTCAGGTTGAGAGGCGGG GCTATGATCTCATCAAGCCCCTCCCCAAAAGAATTACACCATCACCTCTTCCCTGCATTAGCTGCGGTCAGG ATGAGCCTTCCCCAAAGACAACAACCTCTCATACTTACCTGGCAGGGGTGATACCATGATCAGAAAGGTGGTT CGCCCAGAGCGAGGCTCAGCCATTGCACTTGGGCTGTGCTGACCTTTGCGAATTCCCCAAATGCGGGAATCT CGACTGCATAATTTCTGGTAGTGGGGGACTGCGT</p>
Pri50Sat198-191	<p>GTGCATTATGGCTCGGTATGCCATGTTGCCAGTTCATTTTAAAGTCTAAAATCAGCTCGGGCCAGGCTAATTC TAGGTTTTCTTAAAGTAGAAGACTAGAGCTTTCATCTGATGTGAATTTGTTCTGGCCCAAGATGAAAAAG TCCTTCAAATCACACTGGAACATCAGCTTTATATTCATTTTTTTG</p>
Pri50tRNA_Leu/Ser/Ser-563	<p>TAACCACTCGGCCACGACTACAGCCTCTTACCGTGACCTTCCCTATAAAATTTTCAGGATTTTCATGTGACTGT TGTCAAAATGGGCTAGCAAATATGGCATGCCCTACTGACAAAATGAACAACACAGTTGAAAGGATTTGA ACCTGAGCAGAGCGACCCCATATGTTTTAGGGTGTGTAACCTTTAACAACCTCAGAAATGACTACAAGAGTTGG GAATTGCCATACTATATTTCTTGAATTGAACAACACCGCCTTTCATTCCAGTAATACATTTAAGTTTCCTT GCATAAAGATTGAAGTAAACCTGTGCAAGTGGGATTCGAACCCACGCTCCAGAGGAGACTGCGACCTGA ACGCAGCGCCTTAGACCGCTCGGCCATCCTGACACACTGGCTGTGGAGCGTAGAACATTTTAATTGGTCATC TGACCATTCTGTGAAGAATGCAAAATGCAAAGCAAAAACAACCAAAGCTACGCAAAAATAAACAGTGAAC GTAGTCGGCAGGATTTGAACCTGCGCGGGGAGACCCCAATGGATTTCTAGTCCATCGCCT</p>
Pri50Sat210-187	<p>GGGTCATAATAAAAAAACTGTACAGTAATATTTTTAGTTATTAAGGCTATTCACAGTCTAATAGGTATATTTT CATAATGTGGTGTCAAATTCAACAGACTATTTGACAATCTGGAGTTAGGGCAACTGGAAATATGCAGAACA GGCAATTTATTCTCCTTAATCTTCTCAAATGCCAAGTTCAGA</p>
Pri50Sat212-43	<p>AATAAAGCGCACTGTGATAGGACTGGAGTGAGGGTACTGTAAT</p>
Pri50Sat220_ncRNA-70	<p>GCGCTGTAATAATCAGGTGAGTTATTGTGATATTGAGTGTGGGTACAGTACAGTTAGAGGACAGTAGAGG</p>

Pri50Sat225-189	TGTCCTTCTTCAGGATTTTGACAGTGTGGGGTATTGGGTATTAGGTCTGATAGTCATTTTTGTGAGACGATTT AAGACTGAAGAGATAATGGTATTGGAATTAGGCGAAAAGTGC AATTTGACCATTCAATTCTATTATATTACAT GCTATGTTACAGGGTCATTATGGAAAAATTCCACAGTAAAATT
Pri50Sat235-21	CTTTTAAACTGAACCGCGGAT
Pri50Sat240-142	TCTAAAGACTACAAGTGATTTTATATCAATATTTTCATATTAGTACATTACAGTAGCAATTTTTAGTATTA AAA AATTAATAAAAAATCAACCGAATCATATTTTTTCAA AACTAAAAGTTGTAGTATGTTTATCAGGTGTGTC
Pri50Sat242-87	CTTCGTCTCCGAAAAATTTTAAAAAAAAGGTGTCGTGCCAAATTTGGCCCACTTGTCAAAAACTGCTACCCA TGGGGACCGAAAAA
Pri50Sat244_tRNA_Ala/Thr-552	TTGAAGACCATGGGTTTCAATCCCAGTCCCTGCCTCATGGGCTGAAGTGTCTTTCTGTAAAAATTTCTGGCT TGCTTTAGCATGCTCTTTCAAACATATCTCTGACACTGTAGGTTGTCCTAAAAGCATGGAAACAATGCAGCA GTTTAAGGATGCTGTTTTGCAAGTCTGCAGAGTTTTGTTACCAGGATGACAGTATACAAACACATGTACTTA GTGAGGGGATGTAGCTCAGTGGTAGAGCGCATGCTTTGCATGTATGAGGTCCTGGGTTCAATCCCCAGCATC TCCATTTCTGCTGTTTTGCGACTTGAAAGTAAACATAAAGCGCATGTCTCTGGAATTTGCAAAAAAAAATCG GCATGTGTACAAGGCATTTGCTGATGAAGTAGCTTGATAAATCTGATTTCTTATAGGCTTCATAATTGCATGC TACAGCATGCATTGAGTGCATTGATTGATTGCGGGTGCTATGTTGTATCTACTAAGAGACAGGACATTGCTT GTTGTCTGGCATGGTGGCCAAGTGGTAAGGCGTTGGTCTTGTA A
Pri50Sat255-76	TGTATATCACAATACCTACATTTAGAGTGTTATTAATATATATTCACCCTAATGAGAGACTGTAGCTCCACCT CAT
Pri50Sat260-91	AGTCGTTTAATTTAGATCGCCAAGGGTTTAAACGCACCCGCGGACGGCCGCGCGTTTAATTCAGATGGCACG CTGTTGGAATGCATACCAC
Pri50tRNA_Pro-153	AGTCTTTTTTTTTTAAAGCACAACTAAAGCATTAGCTCAGACTACTTTCACGAACTGGGCTCGTCCGGGAT TTGAACCCGGGACCTCTCGCACCCAAAGCGAGAATCATACCCTAGACCAACGAGCCACGTGAACTGCAGC CCTCCGGAC
Pri50Sat271-52	GTTTATAGACTAATATCTCCTATTCTAGTGTATAGGAGTGTTTATTGAGGGT
Pri50Sat276-52	TGGGTGAAATAAGCCTTTAAATCTGTAAGTAGTTAGCTAAAGTGGTAATATT
Pri50LTR_lncRNA-1998	ATATGCGCTTTTCAACTATTTTCAGTCACACTGGTCATCTAGCAGGGTCATTCTGGTCATGCTGGTCATCCAGA TTGATCTTTATGGTAATTCGGTCATCCAGCTTTTTCATGTTGGTCATCCAGCTTGGTCTTACTGGTCATCCAG

CTATTGGTCATGCCAGTTAACCAGCTTGGTCACGCTGGTCATGCTGACCAACCAGTTTGGTCATACTGGTCAT
CCAGCTTGGTCTTTTTGGTATTGCTGGTCATCCAGCTTGGTCACATTGGTAATCCAGCTTGGTGAGGCTGGCC
AACCGGCAATGTGTTGTAAGCCTAACTGGCCTTAAGGGGCGCTCTTTGCCTTTGATGCTCGGACCCCGTAAAT
CGCCGCTTGCGGCTATATTGGTATTTATAATCGCTTCCACTTCTTTTTCTGACCTAGAAGTGATTGCACAGAA
CAAACCGTAAGGCCTACAGGGCTGAGACTTGGTCATATGGTAGTACTTCTCACCGCTACTCAGATTCAAAG
ATGAGCCCGATCGGCCTCAAGGGGCGCTATGGCGAAGGTCAACGCGTTAGGCCTCGTAACTCCCACGCCCT
GATAGCTAGACAAAAAATTCTTTCGTTATATGATTCCCTTGGTTTATGGCAAATCAAAAAGGTCAATAGAACC
ACTAAGCTCCGCCCACTTAGATTTTTTGGCTATTTAGCATAATATGCAAAACCTACTTTTGAGAACTTGTTCTA
AAGTCATTGACCAATCCTGTCAATTTGGTGTCAAACAACCTCAGCAGAGTCCCATCCTCAATAATTATTA
AAAATCTTGAAATTTGTGAACAATATGGCCGCCATATTTAAATTAGTCTTACCGTGGCACACCCAAATTTACT
CTAACAGCTGTAACCTTTTGAATGCTTAATCTGACACTAATGCCACTTTAGATCTTTGATCCCAACATGATTCT
GAGGTAGCATGTCAATCCGTGTAGACATACGCCCCAGGGGGCGGAGCCAAAGCTAGAAATCTGTATCTCA
GGAACCATAACAAGCTATCAAGTTCTGGATTGGTAGTTATCATCTTGGGCCTATGCCCTCGTGGTATATTGAA
GGAAAATTTGATATGTAAATTTTTGCGATCGCTATTAGCCAATCAGTTTCCAGCAAGCTTTTGACAGGCTAA
ACAATGACCAATCAGGACGATACTTATACCCTACATGTATATGAGGGCCTTCTATCATCCATTAATAATATGG
CGTTGATTGGCCTCTAGGGGGCGCTATAGCAGAAAATCAGTTATATCTAAATGTCTTTGTGGCTTAGGCATG
TTTTATTTTTGATTTGTATACTTTGCTGTAGCCTTTACAACATTGTAATTACATGTCTTTACCAAAAATGCAA
GGATTTTCATCAGTGGCCAAAAGAGTAAAAATGGCTCTTTGAGAACTTGTCTTTAAGTATTCAATCAATCAA
AACAAATGTGGTCTTAATGCAATCTTGAGACTCTCTAGGTAAATAATGTTTGAAAAAATCTTGACATTTTAA
CTATTTGTGGTCCATAAACCTTGATCAAACATGTGCGTGAAAGCCTTTTCAGAGTTATTTGGCCAAAGCTCTG
TCAAAGTTTTAAACATGTCAAAACTGTTGAAGCTACTAATTAACAATGACATTTGAAGCACATCTGCCAAGT
ATGAGCCAAATAAGTCTTCAGTAGGCTCTACAGTGACAAATTAACCTATTTTCAATTTAATCTATTTTTCGCTA
TTTTCAAACCTAAATGGACAGAAGTCAGGAACCTTTGACCCTATCAACACACAAATTTACATACACATATAT
ACAGAACACTTGATCTTGACTACCAATTTTGAGCCTAATCGGACTTTTCTCTCTCTGCAATAAATACAAAGG
CAGAGATAGGAACAAAACCTTTTTGCTGATAGAGTGATAATTTTTCAGCAGGTTATATGTGGTCTCTTGCTGC
GCTCTGGTGGTCATTTGGTGAAACAGCTTTAGCTCAATTG

Pri50LTR-999

TAATACTGAGCCACAGCTGAACTCAATGCCTCGCGGGAGAAAGAGAGGGAAGCCCGCCCCGGCTACTCAC
CGCTCAGCTGCGATGACGCGGGGGATTTCCTTCTCCTTCTCCTCCTGAGCTACAGAGAGCCGCTGCTCC
GCCTCCAGCCCCGGCGCTACAGCGAGCCCGCGCTCCACCCGTTCCGGGGCCACGAGCCGCTGCTGCGCCTGT
CACGGCGCCCAAAGCCCGCGCTCCTCCTGCCGCGGCGAGAGCCTCTGGCACAGAGGCTGGGAGGTGGGCTG
CCATCTCCCCGGCCTCTCTCTGGAGAGGGGGGCCCTAGCGAACGAACCTGCATCCCCGCTCTCCTCACCT
GCCTCTCTTCGTTTCGCTCCAGTCACCAGGTCACGCCTCACCTTCGACCTTGGTGGCTGGAAGCTTAGGGGGG
AGGACTGTGGCGTGAGGAGGCGGGGAGTGGTCCGCCCACGCGCACAGCTCTGCCGCTCCCCAG
CTGATCTGCGATCCGGACTGATTAGCAGATCAGCTGGGAGACCATAAAAACGGAGCTTACGCCCTCATTGGA
GAAGGACTGCACTGGTGAGCAGAGGCTGAGGCTCCACGGACCTTTAAACACTGGAAAAAGACTGAGAACT
GCTGAACTCTAGAGCCCCATTGGGCTGAGCATTATTATGCTCATTAAAGTTTGTGGTGGTGGAGGGC
ATTTACAAATAAAGGTATGATTTGAGTTCATTTACTCCCCGTGCTGTGTGTCTCTGCACTTCCCTCCTCCG

	GGTCACAGTGGTAACGCCCCATCCCCAGGGGCGCGCCACATTGGTGGAGAATGCGGGCACCGCCCCTTTTG GGCGGGGTCCAGTGACCCGGTGGGAAGGAAGTCGCGGAGTGGAGCACAGCGCGGGGTGGATCGAGCTCA GGTCATACCTTTGTTTGGCCGCCGAAGCCACCCACAGTTTGTTTTGCCTTTTTATAAAAAAA
Pri50LTR_tRNA_Ala/Thr-458	CAAGTGCCCTGCAAGTGCCTCACAAAACCTGTGTAACTGAGGTAATTTTAAAGCTGTGGACAACTGGTTA GACTCAATGACTAGATTAAGCAACACTTGACCCGACGTGATTTGAACACGCAACCTTCTGATCTGGAGT CAGACGCGCTACCGTTGCGCCACGAGGTCCATCATGCTGCTTGCCTCATGGTGCCTGGTGCCAGAAGATAGT CAGAGGCCAAGGCTAAAAAGTACTCCTCATTGCTACTTGACCACCATGCATCAGTTGAACAATGTCAAGT CGCTATAAGAAGGCACAGGCAAGGGCCCGGATAGCTCAGTCGGTAGAGCATCAGACTTTTAACTGAGGGT CCAGGGTTCAAGTCCCTGTTTCGGGCGGCCAAAGCCTGTGCTTCTTTGTTACCAATTAACCTTTAAACTCCT CAACTGAGCATGAACTAAAGACGTTA
Pri50rDNA18S/5.8S/28S-8708	GGGTTCCCCGCGCGCCGCGACCCCTCCCGGCGCGAACC GGGAGAGGACGGACGCGGAGGTCCCGACGCA GCACCGTAGCCGGGAGATCCGCGAGAAGGGCCCGGCGCACGCCAGAGTCGCGGCCGCTGCTCCGCGGCC CGCCGCCCGCCGTACCCCGCCCTACCGGCCCGGAGCGTAGAGCCGCACCCGGATCCCCCGCCGCCGCC GAGCCACCCGACCGCCGAAACGGCCGGGACCCGGACGACGAGAGAGGACGGGCGCGACGCGCTCCGAGG CTCCGGACGAGGGGCGGAACGACGGGGCGCGCGGGGACGCCGCTCCCCAGTCGCGGCTCGAGCCAGCC CCACTTCGCTCCCTGGCCGACCGACCCAGCCCTTAGAGCCAATCCTTATCCCGAAGTTACGGATCTGACTTG CCGACTTCCCTTACCTACATTGTTCCAACATGCCAGAGGCTGTTACCTTGGAGACCTGCTGCGGATATGGGT ACGGCCCGGCGCGAGACTTACACCCTCTCCCCGGATTTCAAGGGCCAGCGAGAGCTTACCGGACGCCGCC AGAGCCGCGGCGCTTTCAGGGCGTGGGCCCTATCTCGGGGTGAACCCATTCCAGGGAGCCCTGCCCTTCA CAAAGAAAAGAGAACTCTTCCCGGTGCCCCCGCTGGCGTCTCCAGGTTTCGTTTGCCTTACCGACTGGGCGC CTCGCGGCGCCTATCTCCGCCACTCCGGGTTCCGGGATCTGAACCCGACTCCCTTTCGATCGGCCGGGGCG ACGGAGGCCATCGCCCCACGCTTCCGAACGGCGTTCGCCCATCCCTTAGGACCGACTGACCCATGTTCAACT GCTGTTACATGGAACCTTCTCCACTTCGGCCTTCAAAGCTCTCGTTTGAATAATTTGCTACTACCACCAAGA TCTGCACCCGCGGCGGCTCCACCCGGGCCCGCGCCCTAGGCTTCCGCGCCACCGCGGCGGCCCTCCTACTCG TCGCGGCGTAAGCCCGAAAGGACTCGCCCCTCTCTGCCAGCGACGGCCGGGTATGGGCCCGACGCTCCA GCGCCATCCATTTTCAGGGCTAGTTGATTCGGCAGGTGAGTTTTTACACACTCCTTAGCGGATTCCAACCTCC ATGGCCACCGTCCTGCTGTCTATATCAACCAACACCTTTCATGGGGTCTGATGAGCGTCGGCGTCGGGCGCC TTAACCCGGCGTTCGGTTCATCCCGCAGCGCCAGTTCTGCTTACCAAAGTTTGAATAAGTTGAGATTGATCGTTT CGCCCCAATGCCTCTAGTCATTAGCTTTACCGGATAAAACTGCTGAGGCGAGCGCCAGCTATCCTGAGGGA AACTTCGGAGGGAACCACTACTAGATGTTTCGATTGCTTTTCGCCCTATACCCAGGTCGGACGACGCGAT TTGCACGTCAGGACCGCTGCGGCCCTCCACCAGAGTTTCCCCTGGCTTCGCCCTGCCAGGCATAGTTACC ATCTTTCGGGTACCATCGCGTACGCTCTAGCTCCACCTCCCCGACAGAGCGGGAGAGACGGGCCGGTGGTGC GCCCCGCGACCCGTAGGGGGGGGATCCACCTCGGCCGGCGCGCGCCGGCCTTACCTTCAATTGCGCCGTG GGGTTTCTTTCAGCCCTTTGACTCGCGCACGCGTTAGACTCCTTGGTCCGTGTTTCAAGACGGGTGCGGTGGG TAGCCGGCATCGCCGCTGACCCCGAGCGCCCTTTTGTACGGAGGCGGTCCCCGCCCTGGCTGCGCGACGCG

GTTGGACGCGGACTGAGGACAGTCCGGCCCCGGTTCGACGGCACGCGCAGGGGGACGGAGGGGCCCGTCCCT
CCCCGCGGGGAAGGAAGGCGCGGAGGTACTCACCGCGGGCCCCGGGGTAAGCGGCGAAGTCGGGGCGG
GAGGGCGCTGTAAAGCTCGCGGGCCGGAGCTCGCGAGCCACCTTCGCCCCGTGACCCTTCCAAGCCGAACC
GAGGCCGGTTCGCGGCGCACCGCCGCGGAGGAAATGCGCCCCGACGGCGGCCGTTAGCCCCGAGCCCCAGGGC
GCCTCCCTCGCGGGAGGAGCTCGCCGTGCGGACCCGGACACCGCCGCCGGTTGAATCCCCCGGGCGG
ACTGTGCGGACCCACCCGTTTACCTCTTAACGGTTTCACGCCCTGTTGAACTCTCTCTCAAAGTTCTTTTCA
ACTTCCCTTACGGTACTTGTTCGCTATCGGTCTCGTGCCAGTATTTAGCCTTAGATGGAGTTTACCACCCGC
TTAGGCTGCATTCCCAAGCAACCCGACTCCGAGAAGAACCGCGACCTCGGCGCGGGGGGCCGTTACCG
GCCTCACACCGTCCCTGGGCTGAGCCTCTATCAGAAGGACTCAGCCCCCTGGCCCCGCGCGAGAGATACGG
ACTTCTATACGCCACATGTCCCGACGCCCCGTAAGGACGGGGGATTTCGGCGCTGGGCTCTTCCCTCTTCGCT
CGCCGCTACTAGGGGAATCCTTGTAGTTTCTTTTCTCCGCTTAGTAATATGCTTAAATTCAGCGGGTTCGTC
TCGCTGATCTGAGGTCGCGCTCGAAGGGCTGTCGTCGCCAGGCATATGCCGCTGCAGGGTGGGGGAGACCC
CCAACACGCAGTCTGGCTGGAACCGTCCGGACCAGGGAACCTTCCAGACGAGGCTCGATTTCGGCCGAGGC
CGGGATGAAAGCTGGAGGGAACCGAGCCCCACCCTCTTCGCATTCTACCGGAGCGATCTTTTACCAACCTCT
CTTTTGGCGCGCACTCGGCCCGACTAAGGGAGCGGGAGGACGGGGGTCCGCTTAGCGGAGAGACGGCGCGC
TTCGCGGGCAGCCAGAGAGGGTAAAACCCACCGGCAGCCGCGGCCACGACGTCTCTCCACGGACACCCGCC
ACCACTCCCTCGCCGGAACGAGGAACGGTAGAGGACGGCGTGGAAAGCCGGCCTCCGAGACCCATCGCTGG
GCAACGGAGACGAAAGCGAGAGGGGTAAACTGAGTTTACCCCCGGAATCAAGCCCGTCGAGCGGCTTTT
ATCGAGTCCCGAACAAAGAAGGGAACCTGACCGAGTCTGCATTTAGGAAGACGAAGGTGCCTCGGGTGGGC
ACCTGCGAACCTCCAGCCGCGGTCCCCTCGCACTCTCATCTCTTAACCCCGGAGGGAAGAGGGGAGAGAAGA
ACGAAAAGGTCCGATAGATTGGAATAGCGACCCCTCAGACAGGCGTGGCCCCGGGATTGACCCGGGGCCGCA
ATGTGCGTTCAAAGTTCGATGATCAATGTGTCCTGCAATTCACATTAGTTCTCGCAGCTGGCTGCGTTCTTC
ATCGACGCACGAGCCAAGTGATCCACCGCTAAGAGTTGTAATCTTTTGTGTTTGGTTTTCAAAGCCATGTAA
CCAACAGACAGTGTGATGGTTGAAAAAGCGTTTTACGGCAGCCGAGGGTACGGGCGCCCCGGAACGCCGA
AGCGAACTCGGGGTCATCAAACCGCGCCGGCCCCGCTCCGGAGAACCGCAGACGGACTTGGGCCAGGTACC
CGTCTCCCTTTGGCCGTAACCGTGGGGGTTTTTCTTAGGACACCTTTGCTTTGTGCTTCTGGAGGAGAGGA
GGTCGGCCGCCCCGTGCCGACCCCGTCACTCTGGAAGTCTTCTCCTTTTCGATCAAGTCATTGAGGCCGACG
GGAATCCCTGTAGTTCCCGCCATCGGGCCGTGGCTTGTAAAGGGGGGTTTTTTGGAAAAGGGAGTCGCACCG
TGAGGCACTACCAACCCAGTCCCGGTAATGATCCTTCCGACAGGTTACCTACGGAAACCTTGTACGACTTTT
ACTTCTCTAGATGATCAAGTTCGATCGACTTTTTCAGCACAGCACCAAGGCCCGCGAAGGAGCCCCGGCGT
GGCCGATCCGAGGACCTCACATAACCATCCAATCGGTAGTAGCGACGGGCGGTGTGTACAAAGGGCAGGGA
CTAATCAGCGCGAGCTGATGACCCGCACTTACTGGGAATTCCTCGTTCATGGGGAATAATTGCAATCCCCA
ATCCCCGTCACGAACGGGTTTAAAGCGGATTACCCGCGCCTCTCGGCGTAGGGTAGGCACACGCTGGCCCCGC
CATTGTGGCGCGCGTGCAGCCCCGACATCTAAGGGCATCACAGACCTGTTATTGCTCCATCTCGCGTGGCT
ATACGCCACTGGTCCCTCTAAGAAGTTAACCAGCAAACTACTCCACCAACTAAGAACGGCCATGCACCACCCACA
TCTCGTTCGTTATCGGAATTAACCAGACAACTACTCCACCAACTAAGAACGGCCATGCACCACCCACA
GAATCGAGAAAGAGCCATCAGTCTGTCAATCCTTTCCGTGTCCGGGCCGGGTGAGGTTCCCCGTGTTGAGTC
AAATTAAGCCGACGGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTCCTTTAAGTTTCAGTTTTGCAACCATA

CTCCCCCGGAACCCAAAGACTCTGGTTTCCCCGCACGCTGCCCGGGGGTTCATGGGAATAACGCCGCCGGA
TCGCGGGTCGGCATAAGTTTACGGTCGGAACACTACGACGGTATCTGATCGTCTTTCGAACCTCCGACTTTCGTTCT
TGATTAATGAAAACATTCTTGGCAAATGCTTTCGCTTTCGTCCGTCTTGCGCCGGTCCAAGAATTTACCTCT
AGCGGCGCAATACGAATGCCCCCGGCCGTCCCTCTTAATCATGGCCCTGGTTCAGGAAACCCACAAAATAGA
ACCGGAGTCCTATTCCATTATTCTAGCTCAGGCATTCAGGCGAGTTTGGTGCCCGCTTGAACACTCTGATT
TTTCAAAGTAAACGCTCCGGACCCCTGAGGACCCGGACACCCAACCAAGGGCATCCGGGGGGCGCCGGGAG
GCAGGGTCTGGGACAGGCGGTAGCTCGCCTCGCGGCGGACCGCCAGCCACTCCCGAAGTCCAACACTACGAG
CTTTTAACTGCAGCAAATTAATATACGCTATTGGAGCTGGAATTACCGCGGCTGCTGGCACCAGACTTGC
CCTCCAATGGGTCTCACCCATGTGTTTAGGATACGTTTATTCCGATTACAGGGCCTCGAAAGAGACTGCA
TCGTTATTTTTCGTCACTACCTCACCGTGTGCGTAATGGGTAATTTGCGCGCCTGCTGCCTTCTTGGATGTG
GTAGCCATTTCTCAGGCTCCCTCTCCGGAATCGAACCTGATTCCCCGTTACCCGTGGTCACCATGGTAGCCC
CCTATGCTACCATCGACAGCTGATAGGGCAGACATTCGAATGAGACGTCGCCGCCGCGGTGGGCGCGCGAT
CGACCCGAGGTTATCTAGAGTCACCAAAGGGCCGGGGAAACCCCTCGCGGGTCCCCCGGGTGGGTTTTGG
GTCTGATAAATGCACGCATCCCGGCCCCCCGAGAGAGGACCGGTACGCGCTCGTTTGCATGTATTAGCTCT
GGAATTACCACAGTTATCCGAGTAACGTGTGGAGCGATCAAAGGAACCATAACTGATTTAATGAGCCATTTCG
CAGTTTCACTGTACCTTCCGTGTGCACTTAGACCTGCATGGCTTAATCTTTGAGACAAGCATATGTTACTGGC
AGGATCAACCAGGTTGCCCCGAGAGAAGGTTTTGACCACACAAACCCAGAGGTTTCATGAGGGGAGAGGGATCG
GCCCCGTACCGTCCGAACCCCCCGCAGGGGGACGGGCGGAGAAGCTCGTCCGAACCTACCCCTCAATTAT
CCCGGTGGATAGAGTGGGCAAACCTCCGCTCCGGGGAGTTTTCCCTTTGCTTCGGGGATAACAAGGAAATGATC
GTTTCAAACCGGGCCTGGAGCTTTTTGCCGGGGAGCTCGGGTTCGCTAGGAGCCCCAGAGGACTCGCCCCGA
GGCATCCGTCCGACGGCTCCAGGCCGACCGGCTTTTGTGTAGGGCAGATTGACAGGTAGACGGGCCGTCTAC
GACTCCCTTGGATTGGGTCCGAAAGGTTCTCCCCCTGTCCGAGAACCTCCCTGGAGGCTGGGCCGCTTGGTT
CGGGTAGGATCACCGCGCATAGTGGCCCCCGGGGGGGCCGCGCGATTGCCTACCCAGGCGGGCTTGGTAAT
CCTTTCAGTCAAGTGAATAATAAAAAAAAAAAAAAAAAAAAAAAAAATTTTTGGACCTCCATCTTCCCCGGGGGGC
GGGGGGGGTGATGGAGGCCGACCTGAGGAGCTCGAAATCGGGGTAGAATTCCGCCTTAAAACGGTCCGTAT
ATCCCCTTTGGAACCATATTTGAGAATTTTTCTAAGTCCTGCCACACTTAGAAAATATTCAGGAATAATATTC
AGTGCTTGAACCTAGTCATTGCACCTCTCCTCATATACCCGACGGCACCTCTCTGCAAAACCCCCGACG
AACAAACCTCCAGAGCGCGGGAAAACCCGCTCCGAGGCCGACCAGAGTGCACCTTCGTCCGATACGAGAGC
CGGGCGAAGGGGGTACAGAAATCCTGGAGCTCGCCACCCTCTCTAACTTTAGTGTGAAATGGTAGACAAAA
GTGGGGACTTAGTCATTTTTTCCACTTTTCGGACTTAGCACTGGTAGACCAAGTTGGACTCAACCAGGCCTGG
GCCTCCCTCCGAAGCCAGTCCCCTGGGGGATTCAGCCCTGGAGGTCTACCCAGTCCCTGGCCTCCCTCCCA
AGCCAGTCCCCTGGCTCTGGAGGTCACACCCAGTCCCCTGGATGCCTCTCCAGTCCCCTGACACTACCCTG
GGCATGGCTCCTCCACACATCCGGGGGGAACCCAGCCCCGGAGGTCCCGGCGCACCCAGGTCCACTTGGTACC
CCAGACTTAAGCACCCCTCCTCGGACTACACACATCAGCCGCGCCCTCGGAGCTCCGGGCGACTGGTAGAC
CAACGTGCCACTGGTAGACCCGGCGGTCCACTGGTAGACCCGGTGGGACTTAGAGGAATTTTTTTTTTTTATTT
TGGTTTCTGTCTCCAGGGGAGGGAGACCCGAGGGCGGAGGCCTCACCCACGGCCGCGCCTTACCTACGG
GCGCGACTCGTGACTTGGTCGCCACCTCCCCTCCGGGGGGCCACTGGTAGACCCGAGGACGTACCAAAAAA
AGGTGCTTTTGACGCGGGACGCCGCGTGCCTCCAGACCCCCGAGAGGGTAGCACCATTTGGAGACAAAAG

TTAGAGTACAGGGATGATTCTTAATGGATCGCAGCGCGGTGCTGCTCTGCCACTTACGAAACCCTGACACTG
AATCAGGTCGTCTACGAGTCATTTACGCCGTGAACCACAAACATGCGGTAAACGCGTTTTTCGGAGATGGGG
TACGGCATTCTTCGGCCGCACCCCGAGACCGTTAGCGAGCGGCTCTGCCACCCGAGGGGGGACCCCCCCC
CGGCTACGCCTGACCAACCGTAGGTCCTCGGCACTGTGGATATCGTTCCGTCTAGGCGGGATTCTGACTTAG
AGGCGTTCAGTCATAATCCCACAGATGGTAGCTTCGCACCATTGGCTCCTCAGCCAAGCACACGCACCAAAT
GTCTGAACCTGCAGTTCCTCTCGTACTGAGCAGGATTACTATTGCAACAACACATTATCAGTAGGGTAAAAC
TAACCTGTCTCACGACGGTCTAAACCCAGCTCACGTTCCCTGTTAGTGGGTGAACAATCCAACGCTTTGTGA
ATTCTGCTTCACAATGATAGGAAGAGCCGACATCGAAGGATCAAAAAGCGACGTCGCTATGAACGCTTGGC
CGCCACAAGCCAGTTATCCCTGTGGTAACTTTTCTGACACCTCCTGCTTAAAACCCAAAAAGCCAGAAGGAT
CGTGAGGCCCCGCTTTCGCGGTCCGTATTCACTGAAAATCAAGATCAAGCGAGCTTTTGCCCTTCTGCTCC
ACGGGAGGTTTCTGTCTCCCTGAGCTCGCCTTAGGACACCTGCGTTACGGTTTGACAGGTGTACCGCCCA
GTCAAACCTCCCCACCTGCCACTGTCTCGGAGCGGGTCGCGGCCCGAGACCTCGCGGGCCCCGGAGCGCT
TGACACCAGAAGCGAGAGCCCGCGGGGGCCGCCTCCCCGCCTCACCGAGTTAGTGAGGAAACGATAAGA
GTAGTGGTATTTCACTGGCGACACCCGACGAGCGGGGCCTCCCACTTATTCTACACCCCTCATGTCTTTCAC
AGTGCCAGACTAGAGTCAAGCTCAACAGGGTCTTCTTTCCCGCTGATTCTGCCAAGCCCGTTCCCTGGCTG
TGGTTTCGCTAGATAGTAGGTAGGGACAGTGGGAATCTCGTTCATCCATTCATGCGCGTCACTAATTAGATG
ACGAGGCATTTGGCTACCTAAGAGAGTCATAGTTACTCCCGCGTTTACCCGCGCTTCATTGAATTTCTTCA
CTTTGACATTCAGAGCACTGGGCAGAAATCACATCGTGTCAACACCCGCTTAGAGCCCTCACGATGCTTTGT
TTAATTAACAGTCGGATTCCCCTGGTCCGCACCAGTTCCTAAGCCAGCTGTTAGGCGCCGGCCGAGGCGCC
GCGCCGGA
