



UNIVERSIDADE FEDERAL DE GOIÁS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA  
MOLECULAR

**ANDREZZA ARANTES CASTRO**

**DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA  
*Ololygon centralis* (ANURA: HYLIDAE) POR SEQUENCIAMENTO DE  
SEGUNDA GERAÇÃO COM BAIXA COBERTURA**

GOIÂNIA, GO – BRASIL

2018

**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR  
VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES  
NA BIBLIOTECA DIGITAL DA UFG**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

**1. Identificação do material bibliográfico:**    ☒ **Dissertação**    ☐ **Tese**

**2. Identificação da Tese ou Dissertação:**

Nome completo do autor: Andrezza Arantes Castro

Título do trabalho: Desenvolvimento de marcadores microsatélites para *Ololygon centralis* (anura: Hylidae) por sequenciamento de segunda geração com baixa cobertura

**3. Informações de acesso ao documento:**

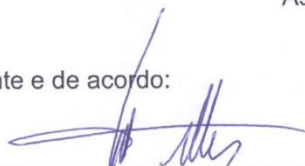
Concorda com a liberação total do documento ☒ **SIM**    ☐ **NÃO**<sup>1</sup>

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Andrezza Arantes Castro

Assinatura do(a) autor(a)<sup>2</sup>

Ciente e de acordo:



Assinatura do(a) orientador(a)<sup>2</sup>

Data: 12 / 04 / 2018

<sup>1</sup> Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Versão atualizada em setembro de 2017.

<sup>2</sup> A assinatura deve ser escaneada.

**ANDREZZA ARANTES CASTRO**

**DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA  
*Ololygon centralis* (ANURA: HYLIDAE) POR SEQUENCIAMENTO DE  
SEGUNDA GERAÇÃO COM BAIXA COBERTURA**

Dissertação de Mestrado apresentada ao  
Programa de Pós-Graduação em Genética  
e Biologia Molecular como requisito  
parcial à obtenção do título de Mestre em  
Genética e Biologia Molecular.

Orientadora: Dra. Mariana Pires de Campos Telles  
Co-orientadora: Dra. Cíntia Pelegrineti Targueta de Azevedo Brito

GOIÂNIA, GO – BRASIL

2018

Ficha de identificação da obra elaborada pelo autor, através do  
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Castro, Andreza Arantes  
DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES  
PARA *Oligodon centralis* (ANURA: HYLIDAE) POR  
SEQUENCIAMENTO DE SEGUNDA GERAÇÃO COM BAIXA  
COBERTURA [manuscrito] / Andreza Arantes Castro. - 2018.  
XCVII, 97 f.

Orientador: Profa. Dra. Mariana Pires de Campos Telles; co  
orientadora Dra. Cintia Pelegrineti Targueta de Azevedo Brito.  
Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto  
de Ciências Biológicas (ICB), Programa de Pós-Graduação em Genética  
e Biologia Molecular, Goiânia, 2018.

Bibliografia. Anexos.

Inclui lista de figuras, lista de tabelas.

1. anura. 2. microsatélite. 3. montagem de novo. 4. NGS. 5.  
transposons. I. Telles, Mariana Pires de Campos, orient. II. Título.

CDU 575



SERVIÇO PÚBLICO FEDERAL  
UNIVERSIDADE FEDERAL DE GOIÁS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

**ATA DA SESSÃO PÚBLICA DE DEFESA DE DISSERTAÇÃO DE Nº 063**

Aos doze dias do mês de março do ano de dois mil e dezoito (12/03/2018), às 14hs00min, no (a) Sala de Reuniões do INCT – ICB I, reuniram-se os componentes da banca examinadora: **Profa. Dra. Mariana Pires de Campos Telles, Profa. Dra. Cíntia Pelegrineti Targueta de Azevedo Brito, Dr. Willian Vaz Silva e Profa. Dra. Daniela de Melo e Silva** para, em sessão pública presidida pelo primeiro examinador citado, procederem à avaliação da defesa de dissertação intitulada: **"DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA *Ololygon centralis* (ANURA: HYLIDAE) POR SEQUENCIAMENTO DE SEGUNDA GERAÇÃO COM BAIXA COBERTURA"**, em nível de mestrado, área de concentração em **Genética e Biologia Molecular**, de autoria de **Andreza Arantes Castro**, discente do Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal de Goiás. A sessão foi aberta pela presidente, que fez a apresentação formal dos membros da banca. A palavra, a seguir, foi concedida ao autor da dissertação que, em cerca de 30 minutos, procedeu à apresentação de seu trabalho. Terminada a apresentação, cada membro da banca arguiu o examinado, tendo-se adotado o sistema de diálogo sequencial. Terminada a fase de arguição, procedeu-se à avaliação da dissertação. Tendo-se em vista o que consta na Resolução nº 1294/2014 do Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC), que regulamenta o Programa de Pós-Graduação em Genética e Biologia Molecular, a dissertação foi aprovada, considerando-se integralmente cumprido este requisito para fins de obtenção do título de Mestre em Genética e Biologia Molecular pela Universidade Federal de Goiás. A conclusão do curso dar-se-á quando da entrega da versão definitiva da dissertação na secretaria do programa, com as devidas correções sugeridas pela banca examinadora, no prazo de trinta dias a contar da data da defesa.



SERVIÇO PÚBLICO FEDERAL  
UNIVERSIDADE FEDERAL DE GOIÁS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

Cumpridas as formalidades de pauta, às 16 horas e 00 minutos,  
encerrou-se a sessão de defesa e, para constar, eu, Clayton L. Borges,  
Coordenador do PGBM do Instituto de Ciências Biológicas, da Universidade  
Federal de Goiás, lavrei a presente ata que, após lida e aprovada, será assinada  
pelos membros da banca examinadora em três vias de igual teor.

**Profa. Dra. Mariana Pires de Campos Telles**  
**Presidente da Banca**  
**UFG/GO**

**Profa. Dra. Cíntia Pelegrineti Targueta de Azevedo Brito**  
**UFG/GO**

**Dr. Willian Vaz Silva**  
**PUC/GO**

**Profa. Dra. Daniela de Melo e Silva**  
**UFG/GO**

## AGRADECIMENTOS

Agradeço aos órgãos de fomento que permitiram a realização deste trabalho, especialmente à FAPEG (DOC-FIX) no projeto aprovado na Chamada 04/2014 (Processo: 201410267001235), à CAPES pela bolsa de pós-graduação que me permitiu dedicar integralmente ao mestrado, bem como agradeço ao INCT em Ecologia, Evolução e Conservação da Biodiversidade (EECBio) pela possibilidade de gerar os dados desse trabalho utilizando a estrutura de equipamentos e bolsistas DTI, disponíveis no Laboratório de Genética & Biodiversidade da UFG.

Agradeço ao IBAMA pela autorização concedida para a coleta dos indivíduos (SISBIO/ICMBIO 46522-3) e também ao Comitê de Ética em Uso de Animais da Universidade Federal de Goiás (número 109/14) por permitir a realização dos procedimentos da pesquisa. Gostaria de agradecer aos que participaram da coleta dos animais: Vinícius B. Guerra, especialmente por passar horas em busca dos animais e Rogério P. Bastos, por além de coletar, nos mostrar o caminho para a localidade-tipo da espécie.

Agradeço a minha orientadora, professora Dra. Mariana Pires de Campos Telles pela oportunidade de realizar uma pós-graduação e acreditar na minha capacidade, por todo aprendizado adquirido, pelo exemplo de profissionalismo e empenho nas multitarefas que exerce. À minha co-orientadora, Dra. Cíntia Pelegrineti Targueta de Azevedo Brito, meus sinceros agradecimentos por me auxiliar em cada etapa do trabalho, cada correção, por me incentivar e me motivar academicamente e por ser um exemplo de superação e competência.

Agradeço as amigas que o convívio no LGBio me presenteou, obrigada pelo relacionamento e por compartilharem experiências e conhecimentos comigo. Em especial agradeço à Adriana Taquary, Anita Pessoa, Ramilla Braga, Vanessa Bernardes, Lucas Donizetti e Rhewter Nunes, por me auxiliarem diretamente na execução de algumas etapas essenciais para a conclusão deste trabalho e a todas as pessoas que indiretamente cooperaram para a realização desse estudo.

Agradeço aos profissionais do Programa de Pós-Graduação em Genética e Biologia Molecular (PGBM) – UFG pelo ensino de qualidade que contribuíram significativamente para o meu desenvolvimento científico e profissional.

Agradeço aos professores e amigos da graduação que me incentivaram a prosseguir na carreira acadêmica.

Agradeço aos meus irmãos (Daniella Castro e Dilson Neto) parentes e amigos que mesmo longe torceram para que esse momento se consolidasse. Agradeço a Nattane da Costa por me acolher em “sua” casa no início do mestrado mesmo quando eu não podia pagar pelas despesas, pela sua irmandade e solidariedade. Agradeço aos meus pais Neime F. Castro e Josana Arantes por sempre me incentivarem a prosseguir na vida acadêmica e me desejarem um futuro do qual não puderam obter. Ao meu esposo Ramon Queiroz por compreender minha ausência, por nunca medir esforços para me auxiliar e incentivar, obrigada por seu amor e companheirismo.

Grata sou a Deus por ter me conduzido em mais uma etapa da minha vida, por me mostrar o Seu agir extraordinário em todas as circunstâncias.

## RESUMO

A facilidade para o acesso às metodologias de Sequenciamento de Segunda Geração (*Next Generation Sequence*) tem impulsionado o desenvolvimento de diversos projetos, uma vez que possibilita a montagem *de novo* de sequências, gerando uma grande quantidade de dados e permitindo o acesso a genomas de diversos organismos ainda pouco conhecidos. A espécie de anuro *Ololygon centralis* (Pombal e Bastos 1996) pertence à família Hylidae, considerada endêmica do Bioma Cerrado e que ainda não dispõe de nenhuma informação genômica. O objetivo desse trabalho foi realizar uma montagem parcial do genoma de *Ololygon centralis* a fim de identificar, caracterizar, quantificar elementos repetitivos e prever genes nas sequências genômicas, além de desenvolver marcadores microssatélites para a espécie. O genoma parcial montado resultou em 13.989 *scaffolds*, que correspondem a 4.926.069 pb com N50 de 336. Os elementos repetitivos encontrados no genoma de *O. centralis* (1.795 elementos transponíveis, 957 microssatélites, um DNA satélite e 25 pequenos RNAs nucleares) totalizaram 531.337 pb. Entre os retrotransposons, o elemento LINE (49,83%) foi o mais abundante, seguido do LTR (48,66%) e SINE (1,56%). Foram identificados os genes que codificam as subunidades 5S, 18S e 28S de ribossomos, 18 tipos de tRNAs no genoma e 26 genes. A partir das sequências foi possível desenhar 87 pares de *primers* flanqueando regiões microssatélites. Dentre elas, 47 foram regiões compostas de tetranucleotídeos, 39 dinucleotídeos, e um trinucleotídeo. Deste total, 31 pares de *primers* foram selecionados para padronização em gel de poliacrilamida (6%) e 18 deles apresentaram produto de amplificação adequado. Durante a padronização, as temperaturas de anelamento variaram entre 50° e 60° C e sete marcadores apresentaram polimorfismo. Dos sete, apenas cinco marcadores (OCE05, OCE11, OCE20, OCE21, OCE26) foram padronizados e utilizados para caracterização em 30 indivíduos de *O. centralis* no analisador automático ABI3500. O número médio de alelos foi igual a 10, a amplitude alélica variou entre 148pb (OCE 20) a 318 pb (OCE 21). A Heterozigosidade média esperada foi 0,7 e a observada igual a 0,5. A probabilidade de exclusão de paternidade (Q) foi igual a 0,993 e a probabilidade combinada de identidade (I) igual a  $1,13 \times 10^{-6}$ . O valor global para o índice de fixação ( $F_{ST}$ ) foi significativo e igual a 0,2 ( $p < 0,05$ ). Os dados obtidos a partir do NGS plataforma Illumina representam um importante passo para o conhecimento genômico dessa espécie e dos anfíbios em geral, o desenvolvimento de *primers* contribuirá para o estudos genéticos-populacionais de *Ololygon centralis* e espécies correlacionadas por meio da transferibilidade.

Palavras-chave: Anura, microssatélite, montagem *de novo*, NGS, transposons.

## ABSTRACT

The ease access to Next Generation Sequence methodologies has improved the development of several projects, once it makes possible the de novo assembly of DNA sequences, creating a great amount of data and giving access to unknown genomes of different organisms. The anuran species *Oloolygon centralis* belong to Hylidae family. It is considered as endemic of Cerrado Biome and little is known about its genomics. The aim of this work was to realize a draft of the genome of *O. centralis* in order to identify, to characterize, to quantify repetitive elements and to predict possible genes. Also, we aimed to develop microsatellite markers for the species. The genome draft assembled in 13989 scaffolds, which correspond to 4926069 bp with N50 = 336. The repetitive elements found in the *O. centralis* genome (1795 transposons, 957, microsatellite, one satellite DNA and 25 small nuclear RNAs) comprised 531337 bp. Between the retrotransposons, the LINE element (49,83%) was the most abundant, followed by LTR (48,66%) and SINE (1,56%). It was possible to identify coding genes for the 5S, 18S and 28S subunits of ribosomes. Also, 18 types of coding regions for tRNA and 26 putative genes were found in the *O. centralis* genome. From the microsatellite regions, it was possible to design 87 pairs of primers for the flanking sites. There were found 47 regions composed of tetranucleotides, 39 dinucleotides and on trinucleotide on the microsatellite sites. From them, 31 pairs of primers were selected for amplification and 18 showed good results. The annealing temperatures varied from 50° e 60° C. Seven markers showed polymorphism. From them, only 5 markers (PCE05, OCE11, PCE20, OCE21, OCE26) were used to characterize the genotype of 30 individuals. The medium number of alleles was 10, the allelic amplitude varied from 148 bp (OCE20) to 318 bp (OCE21). The Expected Heterozygosity was 0,7 and the observed was 0,5. The probability of parentage exclusion (Q) was 0,993 and the probability of combined identity (I) was  $1,13 \times 10^{-6}$ . The global value of the fixation index ( $F_{ST}$ ) was significant and equal to 0,2 ( $p < 0.05$ ). The data obtained from the NGS Illumina platform represent one important step for the knowledge of genomics of this species and of the anuran in general. The development of primers will contribute to genetic populations studies of *O. centralis* and correlated species through primer transferability

Keywords: Anura, draft *de novo*, microsatellite, NGS, transposons.

## SUMÁRIO

|                                                                                                                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| LISTA DE TABELAS .....                                                                                                                                                                                           | 12 |
| LISTA DE FIGURAS .....                                                                                                                                                                                           | 13 |
| <b>1. INTRODUÇÃO GERAL</b> .....                                                                                                                                                                                 | 15 |
| <b>2. OBJETIVOS</b> .....                                                                                                                                                                                        | 16 |
| 2.1 OBJETIVO GERAL .....                                                                                                                                                                                         | 16 |
| 2.2 OBJETIVOS ESPECÍFICOS .....                                                                                                                                                                                  | 16 |
| <b>3. REVISÃO BIBLIOGRÁFICA</b> .....                                                                                                                                                                            | 17 |
| 3.1 Composição do genoma eucarioto .....                                                                                                                                                                         | 17 |
| 3.2 Sequenciamento, montagem e anotação de genomas .....                                                                                                                                                         | 19 |
| 3.3 Marcadores Moleculares .....                                                                                                                                                                                 | 22 |
| 3.4 Desenvolvimento de marcadores microsatélites .....                                                                                                                                                           | 24 |
| 3.5 Diversidade e conservação de anfíbios .....                                                                                                                                                                  | 26 |
| 3.6 A espécie <i>Oloolygon centralis</i> .....                                                                                                                                                                   | 28 |
| <b>4. REFERÊNCIAS</b> .....                                                                                                                                                                                      | 31 |
| <b>5. CAPÍTULO 1: MONTAGEM E ANOTAÇÃO DE SEQUÊNCIAS DO GENOMA DE <i>Oloolygon centralis</i> (Pombal e Bastos, 1996) (ANURA: HYLIDAE) OBTIDAS POR SEQUENCIAMENTO DE SEGUNDA GERAÇÃO COM BAIXA COBERTURA</b> ..... | 44 |
| RESUMO .....                                                                                                                                                                                                     | 44 |
| INTRODUÇÃO .....                                                                                                                                                                                                 | 45 |
| MATERIAL E MÉTODOS .....                                                                                                                                                                                         | 46 |
| RESULTADOS E DISCUSSÃO .....                                                                                                                                                                                     | 48 |
| CONCLUSÃO .....                                                                                                                                                                                                  | 60 |
| REFERÊNCIAS .....                                                                                                                                                                                                | 60 |
| <b>6. CAPÍTULO 2: DESENVOLVIMENTO DE MARCADORES MICROSATÉLITES PARA <i>Oloolygon centralis</i> (Pombal e Bastos, 1996) (ANURA: HYLIDAE)</b> .....                                                                | 65 |
| RESUMO .....                                                                                                                                                                                                     | 65 |
| INTRODUÇÃO .....                                                                                                                                                                                                 | 66 |
| MATERIAL E MÉTODOS .....                                                                                                                                                                                         | 67 |
| RESULTADOS E DISCUSSÃO .....                                                                                                                                                                                     | 71 |
| CONCLUSÃO .....                                                                                                                                                                                                  | 81 |
| REFERÊNCIAS .....                                                                                                                                                                                                | 81 |
| <b>7. CONSIDERAÇÕES FINAIS</b> .....                                                                                                                                                                             | 86 |
| <b>8. MATERIAL SUPLEMENTAR</b> .....                                                                                                                                                                             | 87 |
| <b>9. ANEXOS</b> .....                                                                                                                                                                                           | 93 |

## LISTA DE TABELAS

### CAPÍTULO 1

|                                                                                                                                                              |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 1. Parâmetros da montagem parcial do genoma de <i>Ololygon centralis</i> . ....                                                                       | 49 |
| Tabela 2. Anotação de elementos transponíveis (TEs) identificados nas sequências de <i>Ololygon centralis</i> . ....                                         | 53 |
| Tabela 3. Presença e ausência de elementos transponíveis no genoma de <i>Ololygon centralis</i> , <i>Xenopus tropicalis</i> e <i>Nanorana parkeri</i> . .... | 54 |
| Tabela 4. Frequência e tamanho (pb) dos fragmentos de TEs encontrados nas sequências de <i>Ololygon centralis</i> . ....                                     | 55 |
| Tabela 5. Grupos de elementos repetitivos anotados nas sequências de <i>O. centralis</i> usando RepeatMasker. ....                                           | 55 |
| Tabela 6. Tipos de motivos de repetição e frequência identificadas nas sequências de <i>Ololygon centralis</i> . ....                                        | 57 |
| Tabela 7. Anotação de genes codificadores de RNA transportador identificados nas sequências de <i>Ololygon centralis</i> . ....                              | 58 |

### CAPÍTULO 2

|                                                                                                                                                                                                                                   |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 1. Número de indivíduos amostrados da espécie <i>Ololygon centralis</i> . ....                                                                                                                                             | 67 |
| Tabela 2. Etapas realizadas no QDD para identificação das regiões microssatélites e desenho de <i>primers</i> . ....                                                                                                              | 68 |
| Tabela 3. Características de pares de <i>primers</i> selecionados para porções flanqueadores de regiões microssatélites. Os motivos de repetição e o tamanho da região prevista para amplificação também estão explicitadas. .... | 72 |
| Tabela 4. Características dos 18 locos padronizados em gel de poliacrilamida (6%). ....                                                                                                                                           | 75 |
| Tabela 5. Multiplex de genotipagem do conjunto de <i>primers</i> para <i>Ololygon centralis</i> . ....                                                                                                                            | 76 |
| Tabela 6. Caracterização dos cinco locos microssatélites para 30 indivíduos de <i>Ololygon centralis</i> . ....                                                                                                                   | 79 |
| Tabela 7. Variabilidade genética em três populações de <i>O. centralis</i> usando 5 locos microssatélites. ....                                                                                                                   | 80 |

### MATERIAL SUPLEMENTAR

|                                                                                                                                                     |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela S1. Motivos de repetição e frequência de microssatélites anotadas nas sequências genômicas de <i>Ololygon centralis</i> . ....               | 87 |
| Tabela S2. Características dos indivíduos de <i>Ololygon centralis</i> coletados nos municípios de Silvânia, Pires do Rio e Caldas Novas – GO. .... | 91 |

## LISTA DE FIGURAS

### REVISÃO BIBLIOGRÁFICA

|                                                                                                                                                                                                                                                                                                                   |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1. Fluxo de trabalho do sequenciamento da plataforma Illumina (NGS). Fonte: Mardis, 2013.....                                                                                                                                                                                                              | 20 |
| Figura 2. <i>Slippage</i> durante a replicação de regiões microssatélite. Quando um loop (alça) é formado na fita recém-sintetizada há um aumento na extensão da repetição (a), enquanto que se o loop ocorrer na fita molde há uma diminuição na extensão repetitiva (b). Fonte: Adaptado de Ellegren, 2004..... | 24 |
| Figura 3. Espécime fêmea de <i>Ololygon centralis</i> do município de Novo Gama - GO. Foto: Carlos Eduardo D. Cintra.....                                                                                                                                                                                         | 29 |
| Figura 4. Pontos de ocorrência de <i>Ololygon centralis</i> . 1, Brasília; 2, Silvânia; 3, Orizona; 4, Campo Alegre de Goiás; 5, Ipameri. Fonte: Adaptado de Moura et al., 2010.....                                                                                                                              | 29 |

### CAPÍTULO 1

|                                                                                                                                                                                                                                            |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1. Qualidade das bases da biblioteca de <i>Ololygon centralis</i> antes (a) e depois (b) do controle de qualidade. O eixo y corresponde ao valor da qualidade de phred e o eixo x o tamanho das <i>reads</i> em pares de bases..... | 50 |
| Figura 2. Distribuição das bases dos <i>reads</i> de <i>Ololygon centralis</i> antes (a) e depois (b) do controle de qualidade. O eixo y corresponde ao conteúdo da base e o eixo x o tamanho das <i>reads</i> em pares de bases.....      | 51 |
| Figura 3. Motivos de repetição e número de microssatélites no genoma de <i>O. centralis</i> .....                                                                                                                                          | 56 |
| Figura 4. Top-Hits do Blast de sequências codificadoras de mRNA.....                                                                                                                                                                       | 59 |
| Figura 5. Número de sequências anotadas pela classificação ontologia de genes.....                                                                                                                                                         | 60 |

### CAPÍTULO 2

|                                                                                                                                                    |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1. Padrão de amplificação para 18 locos microssatélites de <i>O. centralis</i> em gel de poliacrilamida (6%).....                           | 74 |
| Figura 2. Locos microssatélites que apresentaram problemas de amplificação quando analisados em eletroforese capilar, A OCE2, B OCE6, C OCE10..... | 76 |
| Figura 3. Padrão de amplificação monomórfico dos locos OCE 24 (A) e OCE 25 (B). O eixo x corresponde ao tamanho do alelo.....                      | 77 |

Figura 4. Frequência alélica obtida para cada marcador microssatélites no conjunto de 30 indivíduos. Nos gráficos, o eixo x representa o tamanho dos alelos e o eixo y, a frequência alélica.....78

## 1. INTRODUÇÃO GERAL

Os anfíbios são o grupo de vertebrados mais ameaçados, cerca de 41% das espécies estão nessa situação ou em perigo de extinção (Hoffmann et al., 2010; Katzenberger et al., 2012). O Brasil é o país com maior diversidade de anuros do mundo e eles estão distribuídos em todos os biomas (Vogt e Bernhard, 2003; Conte et al., 2010; Loebmann e Haddad, 2010; Alho, 2011; Lipinski e Santos, 2014; Oliveira et al., 2015).

No cerrado, bioma que é considerado um dos “hot spots” de biodiversidade do mundo (Ratter et al., 1997; Myers et al., 2000) 51% das espécies de anuros conhecidas são consideradas endêmicas (Valdujo et al., 2012). Fatores relacionados à perda de habitat, aumento de temperatura, umidade e doenças contribuem diretamente para o declínio populacional das espécies desse grupo (Crawford et al., 2010; Katzenberger et al., 2012).

A espécie *Oloolygon centralis* é endêmica do Cerrado, apresenta uma distribuição restrita à Bacia do Rio Paranaíba (Pombal e Bastos, 1996; Brandão e Araújo, 2001; Moura et al., 2010). Apesar de ser uma espécie classificada como não ameaçada pela IUCN (*International Union for Conservation of Nature*) (IUCN, 2017) conhecer sua diversidade genética pode contribuir para análises mais aplicadas, como em estudos de estrutura genética populacional (Telles et al., 2007; Conte et al., 2011).

Os marcadores moleculares são ferramentas que auxiliam nos estudos genéticos, e os microsatélites por serem potencialmente multialélicos tem sido aplicados para conhecer a diversidade genética, possibilitando a orientação de medidas de conservação não só da espécie, mas também do ambiente em que está inserida (Conte et al., 2009; Luquet et al., 2011; Dufresnes e Jeffery, 2011; Nali et al., 2014; Faucher et al., 2016).

A tecnologia de Sequenciamento de Nova Geração (NGS) tem proporcionado o acesso ao genoma de espécies não modelo que juntamente com softwares específicos como o QDD (Meglécz et al., 2014), permitem conhecer e caracterizar as regiões microsatélites. Informações acerca do genoma de *Oloolygon centralis* ainda não foram constatadas, assim o desenvolvimento de ferramentas é fundamental para conhecer a diversidade da espécie (Yuan et al., 2015; Faggioni et al., 2014).

Esse trabalho teve como objetivo geral montar, anotar o genoma parcial de *Oloolygon centralis* e desenvolver marcadores microsatélites para a espécie a fim de subsidiar futuros estudos de genética das populações que contribuirão para a definição de estratégias de conservação de *Oloolygon centralis*.

## 2. OBJETIVOS

### 2.1 OBJETIVO GERAL

- Realizar uma estratégia de sequenciamento de segunda geração com baixa cobertura a fim de montar, anotar e realizar predições de genes no genoma de *Ololygon centralis*, além de desenvolver marcadores moleculares para regiões microssatélites.

### 2.2 OBJETIVOS ESPECÍFICOS

- Produzir bibliotecas genômicas de *Ololygon centralis* e sequenciar com baixa cobertura em plataforma de sequenciamento de segunda geração;
- Montar e anotar as sequências do genoma de *Ololygon centralis*;
- Realizar predição de genes encontrados no genoma de *Ololygon centralis*;
- Identificar, caracterizar e quantificar as regiões microssatélites do genoma de *Ololygon centralis*;
- Desenhar pares de *primers* que flanqueiam regiões microssatélites;
- Padronizar as condições de termociclagem dos *primers* para cada loco microssatélite;
- Caracterizar a variabilidade genética dos locos desenvolvidos para *Ololygon centralis*.

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1 Composição do genoma eucarioto

O tamanho e complexidade que o material genético dos eucariotos apresenta pode ser explicado pela maior parte do DNA nuclear das espécies ser constituída de elementos repetitivos e não pela quantidade de sequências de genes. Ou seja, os elementos repetitivos compõem um grande volume de cópias de fragmentos de DNA homólogos espalhado por todo o genoma (SanMiguel et al., 1996; Pearce et al., 1996; Mehrotra e Goyak, 2014; Biscotti et al., 2015).

As sequências repetitivas podem ser identificadas dispersas no genoma ou dispostas em tandem, que se referem a repetição de uma sequência de DNA contígua na mesma orientação. Dentre os elementos classificados como dispersos estão os elementos transponíveis (TEs), pseudogenes e duplicações segmentares. Os elementos transponíveis têm como principal característica a capacidade de mover-se pelo genoma apresentando grande diversidade de estrutura e tamanho (Charlesworth et al., 1994; Feschotte et al., 2002; Warren et al., 2015).

Por ação dos diferentes mecanismos de transposição, os TEs são divididos em duas classes principais: a classe dos retroelementos que envolve o RNA como molécula intermediária e, por meio do mecanismo de transcriptase reversa, copia e cola a sequência; enquanto que a outra classe DNA transposon é independente de intermediário e o deslocamento é efetuado por cortar e colar a sequência no genoma (Kapitonov e Jurka, 2001; Warren et al., 2015).

Os pseudogenes são regiões do genoma que tem similaridade com genes codificadores de proteínas, mas não possuem potencial de codificação devido a mutações sofridas na sequência que ocasionaram perda de função (Vanin 1985; Mighell et al., 2000; Alexander et al., 2010; Pei et al., 2012). O termo “pseudogene” foi sugerido primeiramente nos estudos do genoma de *Xenopus laevis* (Jacq et al., 1977) e passou a ser detectado em diversos outros organismos (Torrents et al., 2003; Zhang et al., 2010).

Podem ser classificados em: pseudogenes processados, os quais são cópias do gene funcional inseridas no genoma a partir do processo de transcriptase reversa; pseudogenes unitários que surgem a partir de um gene anteriormente funcional; ou pseudogenes duplicados que são oriundos da duplicação de um gene funcional, mas a cópia tornou-se inviável (Weiner et al., 1986; Mighell et al., 2000; Palazzo e Gregory, 2014).

Por último em relação aos elementos dispersos, as duplicações segmentares são sequências que se encontram em duas ou mais regiões genômicas que compartilham 90% de identidade entre elas e seu comprimento pode variar de mil a milhões de nucleotídeos (Alexander et al., 2010).

As sequências em tandem são predominantemente encontradas em regiões intergênicas. Os motivos de repetição com mais de 100 nucleotídeos são denominados de DNA satélite (satDNA), motivos de 7 a 100 nucleotídeos correspondem aos minissatélites e o grupo de microsatélites apresentam motivos de 1 a 6 nucleotídeos (Foulongne-Oriol et al., 2013).

O DNA satélite é estruturado por uma matriz longa de unidades repetitivas (monômeros), é o constituinte principal da heterocromatina constitutiva e são preferencialmente posicionados nas regiões centroméricas, pericentroméricas e subteloméricas (Yunis e Yasmineh, 1971; Schueler et al., 2001; Miga, 2015). Exercem importante função de manutenção da heterocromatina do telômeros e centrômeros, cada espécie apresenta um perfil de DNAsat independente em consequência da alta taxa mutacional dessas sequências (Biscotti et al., 2015).

Os minissatélites são repetições em tandem de números variados (VNTRs), diferindo dos microsatélites pela quantidade de motivos repetidos. Além disso, os minissatélites são caracterizados por um conjunto de sequências heterogêneas que podem chegar a ter 15 kilobases (kb) em tamanho (Jeffreys et al., 1988; Wright, 1994; Abdurakhmono, 2016). São altamente polimórficos e por isso são usados como marcadores moleculares no desenvolvimento de DNA *fingerprint*, diferenciando indivíduos com aplicações na genética forense e em teste de paternidade (Jeffreys et al., 1988; Bentzen e Wright, 1993).

O motivo de repetição dos microsatélites, consiste em arranjos homogêneos em que os motivos podem ser formados de mononucleotídeo (ex: A), dinucleotídeo (ex: AT), trinucleotídeo (ex: ATG), tetranucleotídeo (ex: ATGC), pentanucleotídeo (ex: ATGCT) e hexanucleotídeo (ex: ATGCTA) (Kalia et al., 2011). As sequências de repetição são mais curtas, podendo ter até 1 Kb, ou seja, de 1000 pares de bases (pb) (Tautz, 1989; Abdurakhmonov, 2016).

A partir dos tipos de motivos de repetição, os microsatélites podem ser classificados em: i) perfeito, contendo um único motivo de repetição (ex: CGCGCGCGCGCG); ii) imperfeito, apresentando um par de base que não pertence ao motivo no meio da repetição (ex: CGCGCGTCGCGCG); iii) interrompido, sequência de motivos interrompida por um

pequeno trecho de DNA diferente do motivo de repetição (ex: CGCGCGTGGACGCGCG) e iv) composto, formado por repetições de motivos diferentes (ex: CGCGCGCGATATATATAT) (Oliveira et al., 2006; Mason, 2015).

No geral, os microssatélites são encontrados abundantemente em regiões não codantes, porém também ocorrem nas regiões intergênicas e codificantes do genoma nuclear e organelar. A distribuição dos motivos de repetição em cada região do DNA tende a ser espécie específico (Varshney et al., 2005). Motivos como o de trinucleotídeos e hexanucleotídeos prevalecem nas regiões codificantes, essa permanência pode ser explicada pela força de seleção que atua para manter a janela de leitura da transcrição inalterada (Lawson e Zhang, 2006; Choudhary e Trivedi, 2010).

A origem dos microssatélites está fundamentada em duas hipóteses propostas por Buschiazzi e Gemmell (2006). A primeira propõe o surgimento a partir de um proto-microssatélite que corresponde a uma sequência simples de 3 a 4 unidades repetidas, que não estão em tandem, originados por substituições de bases e em função das mutações surge espontaneamente um microssatélite. O outro modelo considera que os microssatélites são acrescentados no genoma a partir de sítios predispostos nas diferentes classes de elementos transponíveis. A partir da gênese é proposto que os SSR passam por um processo de evolução, completando um ciclo que consiste em expansão, degeneração, contração, morte e um possível renascimento, já que a morte é representada por sequências simples que podem gerar novamente os microssatélites (Buschiazzi e Gemmell, 2006).

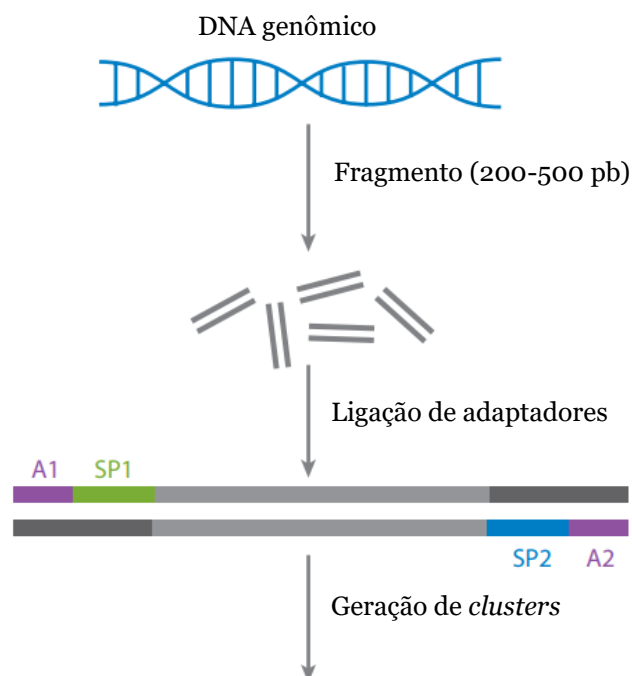
A grande porção do genoma composta por elementos repetitivos já foi considerada sem funcionalidade (Orgel e Crick, 1980) denominadas de DNA “lixo”, termo formalizado por Susumu Ohno em 1972. No entanto, essa percepção já não é mais a mesma e sabe-se que ao DNA repetitivo é atribuído diversas funcionalidades como organização e manutenção do tamanho genômico (Mehrotra e Goyal, 2014; Chalopin et al., 2015), além de serem regiões empregadas para o desenvolvimento de marcadores moleculares (Schlotterer, 2004; Foulongne-Oriol et al., 2013).

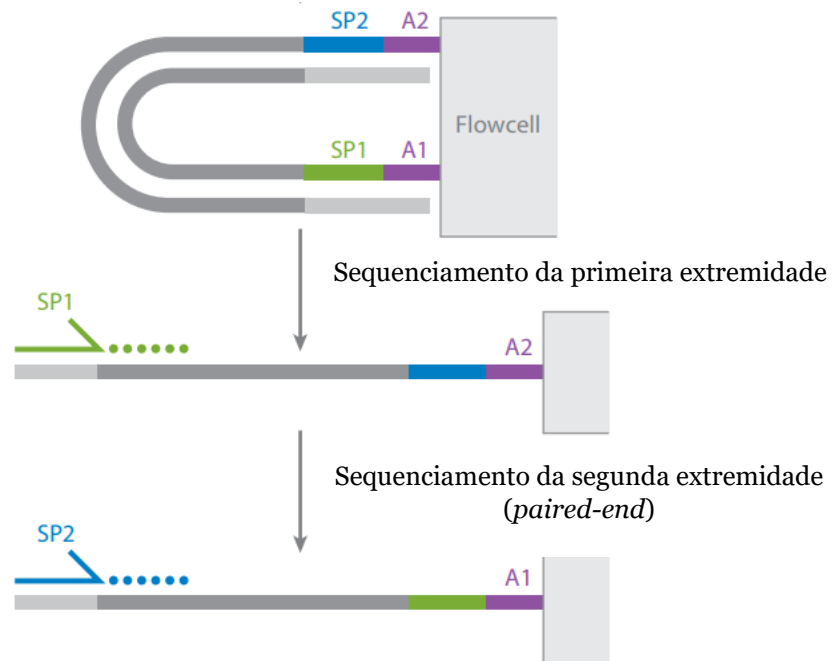
### 3.2 Sequenciamento, montagem e anotação de genomas

Descrever as sequências de nucleotídeos e conhecer a estrutura dos genomas é um campo que vem se tornando cada vez mais acessível em consequência da disponibilidade de ferramentas, o desenvolvimento da técnica de sequenciamento de nova geração (NGS)

permitiu o acesso a genomas de diversos organismos os quais não se tinha conhecimento nenhum de maneira rápida e barata quando comparada a outras tecnologias (Qiao et al., 2013; Sun et al., 2015). A plataforma Illumina é um sequenciador que tem como princípio amplificação por PCR em fase sólida no qual fitas recém-sintetizadas são compostas por nucleotídeos marcados com diferentes fluoróforos que emitem sinais de fluorescência formando imagens as quais são decodificadas nas sequências de nucleotídeos (Metzker, 2010).

O fluxo de trabalho (Figura 1) é iniciado com a preparação de bibliotecas genômicas, em que o material genético a ser sequenciado é fragmentado aleatoriamente em pequenas partes e em ambas extremidades desses fragmentos são ligados adaptadores que serão os *primers* usados para posterior amplificação. Os fragmentos são fixados em uma superfície de vidro, a *flow cell* através da complementariedade de bases dos adaptadores com os oligonucleotídeos presentes nessa superfície. Na *flow cell* acontece a clusterização dos fragmentos, ou seja, um acúmulo de cópias de uma mesma sequência, que emitirão um sinal maximizado a ser detectado com precisão pelo dispositivo de leitura cada base que foi incorporada, os nucleotídeos são adicionados em cada *cluster* em diferentes ciclos (Mardis, 2008; Metzker, 2010; Mardis, 2013).





**Figura 1.** Fluxo de trabalho do sequenciamento da plataforma Illumina (NGS). Fonte: Mardis, 2013.

A partir dos resultados do sequenciamento Illumina que são milhões ou até bilhões de *reads* que podem variar de 25 a 600 bases, é feita a avaliação quanto à qualidade de atribuição de bases nas sequências que é denominado pelo valor de Phred, quanto maior o *score* de Phred maior é a certeza de que as bases que compõem as *reads* foram atribuídas corretamente. FastQC é um exemplo de software que avalia a qualidade das bases e Trimomatic realiza a exclusão das bases sem qualidade (Andrews, 2010; Bolger et al., 2014). A montagem ou *assembly de novo*, que não dispõe de um genoma de referência para orientar as sequências a serem montadas é efetuada com algoritmos baseados em grafos de Bruijn por exemplo, em que as *reads* são decompostas em subsequências de tamanhos K (*K-mer*) e depois é formado um grafo de sobreposição entre o sufixo e prefixo dos *k-mers* (Schatz et al., 2010). Vários montadores são baseados nesse algoritmo como Dipspades (Safonova et al., 2015), SOAPdenovo (Luo et al., 2012), Abyss (Simpson et al., 2009) e Velvet (Zerbino e Birney, 2008) e cada software combina recursos e algoritmos diferentes para operar a montagem.

A metodologia de montagem de genomas consiste em unir *reads* para formar *contigs* que por sua vez se unem para formar os *scaffolds* que correspondem a trechos de sequências mais longas a fim de retratar a estrutura do genoma o mais adequado possível. Depois do alinhamento das *reads* em uma única sequência o próximo passo é anotar e caracterizar a fim de evidenciar os elementos que compõem o genoma da espécie estudada, as sequências

de nucleotídeos representam o ponto inicial para compreensão funcional do organismo e a anotação estimula a interpretação de dados (Stein, 2001; Ekblom e Wolf, 2014).

A anotação pode ser dividida em três níveis, sendo a nível de nucleotídeo que identifica RNAs e elementos repetitivos, nível de proteína que define e atribui função às proteínas encontradas e por último a nível de processo que relaciona o genoma com os processos biológicos (Stein, 2001). A precisão da anotação está diretamente relacionada com a predição dos mecanismos que envolvem o funcionamento do organismo (Zielezinski et al., 2012; Lohse et al., 2014).

Anotação de sequências não codificantes como elementos transponíveis e microssatélites podem ser obtidos através de softwares que utilizam modelos estatísticos ou similaridade de sequências de repetições conhecidas, como o RepeatMasker (Tarailo-Graovac e Chen, 2009). Os dados gerados proporcionam o conhecimento acerca de como o genoma é organizado estruturalmente e são informações que podem ser empregadas no desenvolvimento de marcadores moleculares microssatélites por exemplo (Zalapa et al., 2012).

A predição de genes busca por sequências codantes, identifica códons de início de parada, comprimentos de íntrons e exons visando determinar a estrutura dos genes (Yandell e Ence, 2012), o software Augustus é um preditor de genes amplamente utilizada e a busca pelas regiões gênicas é feita por meio de modelos matemáticos, designado de *ab initio* (Stank e Morgenstern, 2005). O último estágio de anotação consiste em determinar a funcionalidade dos genes baseado na ontologia de genes (*Gene Ontology-GO*) que é realizada com base na similaridade de sequências entre várias espécies no qual os genes são divididos em três classes: componentes celulares, que determina a localização do produto do gene na célula; função molecular, indica a atividade ou função do produto gênico e processos biológicos que designa os processos mais gerais nos quais os produtos gênicos estão envolvidos (Amar et al., 2014).

### 3.3 Marcadores Moleculares

Marcadores moleculares são ferramentas capazes de identificar variações na sequência de DNA entre indivíduos, ou seja, polimorfismos de DNA (Ferreira e Grattapaglia, 1998). Essas variações são utilizadas para estudos de diversidade genética individual e populacional, reconstrução filogenética, mapeamento gênico, teste de

paternidade, aplicações forenses e conservacionistas (Schlotterer, 2004; Telles et al., 2007; Fernandes e Caparroz, 2013; Dufresnes et al., 2014; Decanine, 2016).

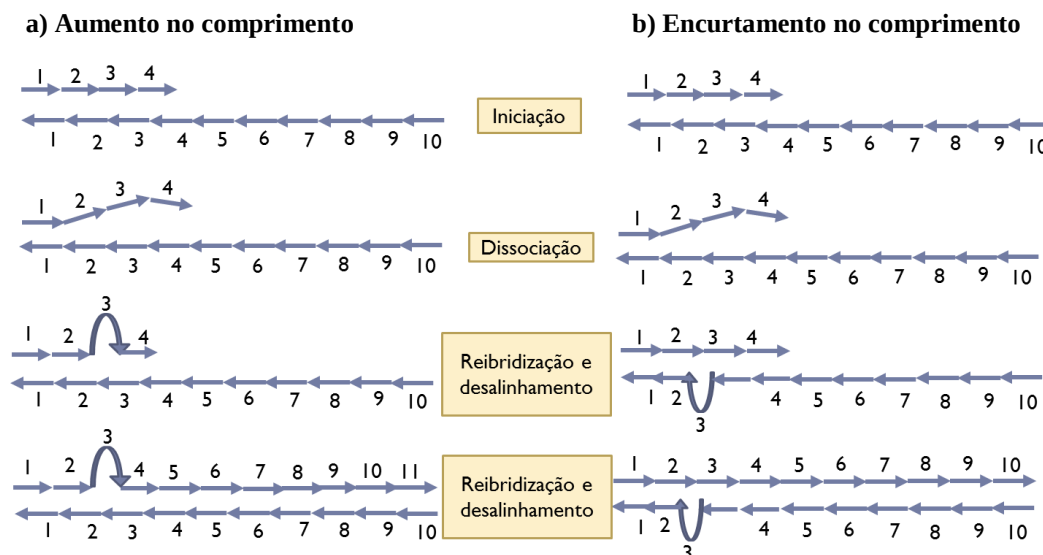
Os marcadores moleculares possuem vantagens em relação a outros marcadores que não se baseiam em DNA porque apresentam: herdabilidade, são transmitidos para a geração seguinte de acordo com os padrões de herança mendeliana; disponibilidade, estão presente em todos os tecidos da amostra; e estabilidade no genoma, uma vez que não são influenciados pelo ambiente, como os marcadores fenotípicos (Agarwal et al., 2008; Duran et al., 2009).

As isoenzimas foram os primeiros marcadores desenvolvidos, baseados nas variantes de proteínas e separadas por tamanho através da técnica de gel de eletroforese, permitindo assim uma caracterização genotípica baseado na sequência de aminoácidos (Ferreira e Grattapaglia, 1998; Faleiro, 2007). Em seguida surgiram os marcadores baseados em DNA, que foram categorizados de acordo com a técnica de obtenção: i) Hibridização, como RFLP (*Restriction Fragment Length Polymorphism*) e Minissatélites; ii) Amplificação por PCR, como RAPD (*Random Amplified Polymorphic DNA*), AFLP (*Amplified Fragment Length Polymorphism*), ISSR (*Inter Simple Sequence Repeats*) e SSR (*Simple Sequence Repeat* - microssatélites); iii) Polimorfismo de sequência de DNA, os SNPs (*Single Nucleotide Polymorphism*) que revela polimorfismo de um único nucleotídeo. (Schlotterer, 2004; Agarwal et al., 2008).

O uso das regiões de microssatélites como marcadores moleculares (SSR) tem sido utilizados em diversas aplicações genéticas em consequência da abundância desse elemento repetitivo no genoma, sua herança co-dominante, natureza multialélica e elevada taxa de mutação que corresponde a cerca de  $10^3$  a  $10^6$  mutações por geração em eucariotos, revelando certa instabilidade nas sequências (Schlotterer, 2004; Sharma et al., 2007; Gemayel et al., 2010).

As variações entre um indivíduo e outro são ocasionadas por erros de recombinação ou *crossing-over* desigual. Esses mecanismos acarretam a perda ou ganho de repetições devido ao mal pareamento dos cromossomos homólogos, e principalmente pelo “deslizamento” da DNA polimerase III (*Strand-slippage*) nas regiões repetitivas durante a replicação. A enzima DNA polimerase III se desassocia temporariamente da fita molde com a recém-sintetizada e ao reassociar ocorre um desalinhamento entre as fitas, que pode resultar na contração ou expansão em uma das cadeias (Figura 2) e consequentemente o

surgimento de novos alelos no loci microsatélite (Levinson e Gutman, 1987; Schlötterer e Tautz 1992; Ellegren, 2004; Bhargava e Fuentes, 2010).



**Figura 2.** *Slippage* durante a replicação de regiões microsatélite. Quando um loop (alça) é formado na fita recém-sintetizada há um aumento na extensão da repetição (a), enquanto que se o loop ocorrer na fita molde há uma diminuição na extensão repetitiva (b). Fonte: Adaptado de Ellegren, 2004.

O sistema de reparo (MMR - *Mismatch repair*) consegue identificar e corrigir a maioria dos erros gerados, desse modo a imutabilidade do microsatélite é dependente do equilíbrio entre o deslizamento, o sistema de reparo e o tamanho da sequência repetitiva (Ellegren, 2000; Kalia et al., 2011). Em um determinado loco microsatélite de região não codante pode existir numerosos alelos com variação neutra e alta taxa mutacional, características que certificam o seu alto poder informativo que somadas ao padrão de herança co-dominante e reprodutibilidade faz dos microsatélites regiões com potencial para o desenvolvimento de marcadores moleculares (Duran et al., 2009; Mason, 2015; Vieira et al., 2016).

### 3.4 Desenvolvimento de marcadores microsatélites

Várias metodologias foram propostas para o isolamento de regiões microsatélites a fim de aumentar a eficiência e especificidade, uma delas é através da construção de uma biblioteca genômica. No método de hibridização seletiva os fragmentos são ligados a adaptadores com sequências conhecidas, desnaturados e hibridizados com as sondas de repetições simples. Em seguida as ligações não específicas são removidas por lavagem e a

sonda hibridizada é restaurado por PCR e clonados para se obter os SSR (Rassmann et al., 1991).

Na técnica de extensão de *primers*, a amostra é submetida a enzimas de restrição que fragmentam o DNA em tamanhos diferentes produzindo a biblioteca primária. Os fragmentos selecionados por tamanho são inseridos no vetor de um fago para a obtenção de DNA circulares de filamentos únicos que irão servir de moldes para a extensão de uma determinada repetição, gerando produtos de fita dupla apenas dos vetores com a repetição alvo (Ostrander et al., 1992; Paetkau et al., 1995). Essas técnicas tornaram-se limitadas pelos processos laboriosos e pelo fato de que apenas uma parte do genoma é clonado. Além disso, durante a construção da biblioteca perdem-se diferentes motivos tornando as abordagens, para espécies com baixa frequência de regiões microssatélites, ineficientes (Zane et al., 2002).

Com o aperfeiçoamento dos métodos de sequenciamento e o surgimento das técnicas de Sequenciamento de Nova Geração (NGS) foi possível obter grande volume de dados, informações sobre milhões ou bilhões de pares de bases, abrangendo a detecção das regiões microssatélites de maneira rápida e eficaz (Mardis, 2008; Delmas et al., 2011). Há várias plataformas NGS disponíveis e com características distintas (Ekblom e Galindo, 2011) que viabilizam estudos com qualquer organismo. A plataforma Illumina tem sido bastante utilizada no desenvolvimento de marcadores microssatélites (Davey et al., 2011; Castoe et al., 2012; Silva et al., 2013).

A plataforma Illumina envolve o preparo de uma biblioteca genômica em que o DNA é submetido a uma tagmentação aleatória, resultando em pequenos fragmentos. Nas extremidades dos fragmentos são adicionados adaptadores que são complementares aos oligonucleotídeos da matriz *flow cell*. Nessa matriz, os fragmentos são imobilizados para a etapa de amplificação por PCR, produzindo várias cópias dos diferentes fragmentos (*cluster*). A amplificação é necessária para que a intensidade do sinal dos nucleotídeos fluorescentes seja maximizado e detectável pelo sistema de imagem do sequenciamento (Metzker, 2010).

Diferente da construção de bibliotecas ou técnicas que exijam muito trabalho prático exaustivo em laboratório na elucidação de sequências, outra maneira de isolar regiões microssatélites é através da mineração *in silico*. Com a disponibilidade de genomas inteiros ou sequências depositadas em banco de dados (como no GenBank) é possível o desenvolvimento de marcadores de baixo custo e com mesmo potencial quando comparada

às outras técnicas, já que as sequências estão prontas e existem ferramentas disponíveis para identificação de marcadores (Sharma et al., 2007).

O genoma ou sequências que estão depositadas em banco de dados são submetidos a ferramentas computacionais específicas, como o software QDD, que tem como principais funções: identificar as sequências repetitivas; analisar a similaridade das sequências, comparando todas contra todas, para manter somente sequências únicas (*singletons*). Além disso, desenham os *primers* e verificam se as regiões microssatélites selecionadas tem associação com elementos transponíveis (Meglécz et al., 2014).

Os *primers* desenhados a partir da identificação das regiões microssatélites correspondem a segmentos complementares às regiões flangeadoras das sequências repetitivas e são específicos para cada loco. É esperado que os *primers* desenhados amplifiquem locos polimórficos, essa característica depende não só da qualidade das sequências de DNA, mas da seleção dos locos das sequências disponíveis (Zalapa et al., 2012).

As regiões flangeadoras adjacentes às sequências microssatélites tem apresentado taxa de mutação menor que a taxa de mutação dos motivos repetitivos pois tem viabilizado a amplificação cruzada ou transferibilidade dos marcadores SSR entre espécies próximas, do mesmo gênero ou de grupos que foram desfeitos recentemente (Varshney et al., 2005; Barbará et al., 2007). É uma alternativa acessível a espécies que não tem sequências disponíveis para o desenvolvimento de *primers* dispor de *primers* já desenhados

### 3.5 Diversidade e conservação de anfíbios

A ordem Anura compreende 6777 espécies descritas no mundo, sendo o maior clado da classe Amphibia (Frost, 2017). São animais ectodérmicos que estão estritamente relacionados com ambientes úmidos. Através da pele permeável que eles possuem é realizada a absorção de água e trocas gasosas, essa característica os tornam vulneráveis a diversos fatores ambientais, como o aquecimento global (Blaustein et al., 1994; Wells, 2007; Vitt e Caldwell, 2013).

Mudanças climáticas relacionadas principalmente à temperatura e umidade são crucias para a manutenção de vida desse grupo, podendo prejudicar aspectos fisiológicos e comportamentais que interferem diretamente na reprodução (Rome et al., 1992; Katzenberger et al., 2012). Dependendo da situação climática, esta pode contribuir para a

proliferação de patógenos o que tem causado grandes declínios populacionais em anuros (Pounds et al., 2006; Verdade et al., 2012)

São inúmeros os impactos causados aos anfíbios tornando-os alvos inevitáveis para a extinção. Cerca de 41% do total das espécies de anfíbios estão ameaçadas ou em perigo de extinção, valor que pode estar subestimado, uma vez que há espécies que não possuem informações suficientes para classificá-las (Hoffmann et al., 2010; Katzenberger et al., 2012). Na fauna brasileira foram descritas 1.080 espécies de anfíbios, destas 1.039 são anuros (Segalla et al., 2016), o Brasil é o país com a maior diversidade de espécies de anuros do mundo. Os anuros estão distribuídos nos Biomas: Pampa (Lipinski e Santos, 2014), Amazônia (Vogt e Bernhard, 2003) Caatinga (Loebmann e Haddad, 2010), Mata Atlântica (Conte et al., 2010), Pantanal (Alho, 2011) e Cerrado (Araújo e Colli, 2000; Oliveira et al., 2015).

O Bioma Cerrado está concentrado principalmente na região central do Brasil e é considerado o segundo maior da América do Sul (MMA, 2017). Diversos estudos indicam o Cerrado como um dos “*hot spots*” de diversidade do mundo, conferindo-lhe um papel fundamental na conservação da biodiversidade do planeta (Ratter et al., 1997; Myers et al., 2000; Diniz-Filho et al., 2006). O Bioma apresenta uma diversificação de ambientes que estão categorizadas em fitofisionomias (Ribeiro e Walter, 2008) nas quais abrigam uma riqueza biológica nos diferentes ecossistemas (Klink e Machado, 2005).

Das espécies de anuros conhecidas no cerrado, 51% são consideradas endêmicas (Valdujo et al., 2012) fator que demonstra o quão é necessário conhecer e estudar a biodiversidade de anuros a favor da conservação já que grande parte da área do Cerrado apresenta a conformação de uma paisagem totalmente fragmentada (WWF, 2012). A desconexão de habitats (*habitat split*) causada por ações antropogênicas pode resultar em declínio populacional, afetando a diversidade em diferentes estágios do ciclo de vida dos anuros (Becker et al., 2007).

O emprego de marcadores moleculares microsatélites em anfíbios tem a finalidade de acessar e explorar a diversidade genética das espécies. Trabalhos como o de Nali et al (2014) que desenvolveram 22 marcadores microsatélites para o gênero *Bokermannohyla* e transferiram para outras espécies com o objetivo de dispor de ferramentas para estudos de genética de paisagem e sistema de acasalamento nas diferentes regiões nas quais as espécies estão inseridas. Wang et al (2013) desenvolveram 113 SSRs para *Nanorana parkeri*, para serem aplicados em pesquisas de genética de população e conservação.

O desenvolvimento e otimização de *primers* auxiliam na compreensão da estrutura genética e viabilidade das populações, permitindo avaliar melhores ações de conservação. A utilização de ferramentas que permitem conhecer a diversidade genética das espécies, como os microssatélites, contribuem para a promoção de medidas de conservação e manutenção da biodiversidade (Conte et al., 2009; Luquet et al., 2011; Dufresnes e Jeffery, 2011; Faucher et al., 2016).

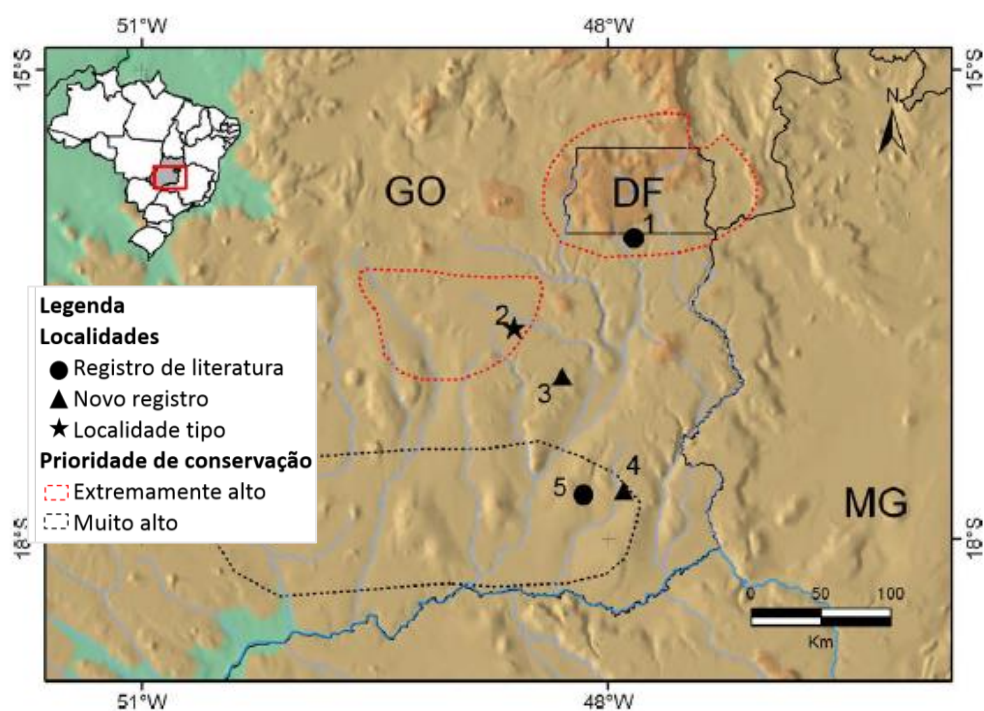
### 3.6 A espécie *Ololygon centralis*

Por meio de estudos filogenéticos, as espécies do gênero *Scinax* foram alocadas em dois grandes clados: *Scinax ruber* e *Scinax catarinae* (Faivovich, 2002; Faivovich et al., 2005). Duellman et al. (2016), ao revisar a família Hylidae baseado em análises filogenéticas, propuseram a formação da subfamília Scinaxinae com quatro gêneros, *Julianus*, *Scinax*, *Sphaenorhynchus* e *Ololygon*. O gênero *Ololygon* passou a corresponder a todas as espécies que pertenciam ao antigo clado *Scinax catarinae* (Duellman et al., 2016).

O gênero *Ololygon* é formado por 45 espécies segundo Frost (2017), a espécie *Ololygon centralis* (Figura 3) habita matas de galeria, é endêmica do Cerrado e foi descrita por Pombal e Bastos em 1996, para a localidade da Floresta Nacional (FLONA) na cidade de Silvânia-GO (localidade-tipo). A presença dessa espécie já foi relatada também em outras localidades como nos municípios de Ipameri, Campo Alegre de Goiás e Orizona, no estado de Goiás e em Brasília, no Distrito Federal (Figura 4). Nos municípios citados as espécies foram encontradas ao redor de cursos d'água e associadas à Bacia do Rio Paranaíba (Pombal e Bastos, 1996; Brandão e Araújo, 2001; Moura et al., 2010).



**Figura 3.** Espécime fêmea de *Ololygon centralis* do município de Novo Gama - GO. Foto: Carlos Eduardo D. Cintra



**Figura 4.** Pontos de ocorrência de *Ololygon centralis*. 1, Brasília; 2, Silvânia; 3, Orizona; 4, Campo Alegre de Goiás; 5, Ipameri. Fonte: Adaptado de Moura et al., 2010.

De acordo com as observações realizadas por Alcantara et al. (2007), o canto dos machos pode ser ouvido de novembro a agosto. A espécie tem reprodução prolongada e durante a época de acasalamento constatou que a quantidade de machos é maior que a de fêmeas resultando em intensa competição. As fêmeas podem depositar uma média de 64 ovos em vegetações marginais a cursos d'água. A precipitação e temperatura mínima do ar

são fatores que estão relacionados com a atividade reprodutiva dessa espécie (Alcantara et al., 2007).

De acordo com a Lista Vermelha da IUCN, *O. centralis* é classificada como espécie pouco preocupante pela ocorrência em áreas protegidas, no entanto atividades como agricultura, pecuária e incêndio são as maiores ameaças à espécie (Bastos et al., 2004). É sabido também que fatores como radiação (UV-B), doenças, introdução de espécies e mudanças climáticas em geral acarretam no declínio de populações dos anfíbios (Beebee e Griffiths, 2005).

Conhecer a diversidade genética de anfíbios é um procedimento primário para a conservação (Faggioni et al., 2014). E o desenvolvimento de marcadores microssatélites torna-se uma ferramenta necessária para gerar informações genéticas que serão úteis em conjunto com outras técnicas para gerenciar a manutenção da espécie (Yuan et al., 2015).

A utilização de marcadores moleculares em espécies como *Ololygon centralis* pode apresentar dados proveitosos para análises genéticas aplicadas como estudos de estrutura genética populacional, processos evolutivos e identificação de espécies crípticas (Igawa et al., 2011; Komaki et al., 2014; Oyamaguchi et al., 2015).

#### 4. REFERÊNCIAS

- ABDURAKHMONOV, I. Y. Introduction to Microsatellites: Basics, Trends and Highlights. In: **Microsatellite Markers**. InTech, 2016.
- AGARWAL, M.; SHRIVASTAVA, N.; PADH, H. Advances in molecular marker techniques and their applications in plant sciences. **Plant cell reports**, v. 27, n. 4, p. 617-631, 2008.
- ALCANTARA, M. B.; LIMA, L. P.; BASTOS, R. P. Breeding activity of *Scinax centralis* (Anura, Hylidae) in central Brazil. **Iheringia. Série Zoologia**, v. 97, n. 4, p. 406-410, 2007.
- ALEXANDER, R. P. et al. Annotating non-coding regions of the genome. **Nature reviews. Genetics**, v. 11, n. 8, p. 559, 2010.
- ALHO, C. J. R. Biodiversity of the Pantanal: its magnitude, human occupation, environmental threats and challenges for conservation. **Brazilian Journal of Biology**, v. 71, n. 1, p. 229-232, 2011.
- AMAR, D. et al. Evaluation and integration of functional annotation pipelines for newly sequenced organisms: the potato genome as a test case. **BMC plant biology**, v. 14, n. 1, p. 329, 2014.
- ANDREWS, S. et al. **FastQC**: a quality control tool for high throughput sequence data. 2010.
- ARAÚJO, A. F. B.; COLLI, G. R. Biodiversidade do Cerrado: Herpetofauna. In Funatura, CI Brasil, Biodiversitas, and UnB (Eds.). Workshop **Ações prioritárias para conservação da biodiversidade do Cerrado e Pantanal**. p. 1–25. Funatura, Anais, Brasília. 2000.
- BARBARA, T. et al. Cross-species transfer of nuclear microsatellite markers: potential and limitations. **Molecular Ecology**, v. 16, n. 18, p. 3759-3767, 2007.
- BASTOS, R.; COLLI, G.; SILVANO, D. *Scinax centralis*. The IUCN Red List of Threatened Species. 2004. Disponível em: <<http://www.iucnredlist.org/details/55945/0>>. Acesso em: 10 jan 2017.

BECKER, C. G. et al. Habitat split and the global decline of amphibians. **Science**, v. 318, n. 5857, p. 1775-1777, 2007.

BEEBEE, T. J. C.; GRIFFITHS, R. A. The amphibian decline crisis: a watershed for conservation biology? **Biological Conservation**, v. 125, n. 3, p. 271-285, 2005.

BENTZEN, P.; WRIGHT, J. M. Nucleotide sequence and evolutionary conservation of a minisatellite variable number tandem repeat cloned from Atlantic salmon, *Salmo salar*. **Genome**, v. 36, n. 2, p. 271-277, 1993.

BHARGAVA, A.; FUENTES, F. F. Mutational dynamics of microsatellites. **Molecular biotechnology**, v. 44, n. 3, p. 250-266, 2010.

BISCOTTI, M. A.; OLMO, E.; HESLOP-HARRISON, J. S. P. Repetitive DNA in eukaryotic genomes. **Chromosome research**, v. 23, n. 3, p. 415-420, 2015.

BLAUSTEIN, A. R.; WAKE, D. B.; SOUSA, W. P. Amphibian declines: judging stability, persistence, and susceptibility of populations to local and global extinctions. **Conservation biology**, v. 8, n. 1, p. 60-71, 1994.

BOLGER, A. M.; LOHSE, M.; USADEL, B. **Trimmomatic**: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, 2014.

BRANDÃO, R. A.; ARAÚJO, A. F. B. de. **A herpetofauna associada às matas de galeria no Distrito Federal. Cerrado: Caracterização e Recuperação de Matas de Galeria**. Planaltina: Embrapa Cerrados, p. 559-604, 2001.

BUSCHIAZZO, E.; GEMMELL, N. J. The rise, fall and renaissance of microsatellites in eukaryotic genomes. **Bioessays**, v. 28, n. 10, p. 1040-1050, 2006.

CASTOE, T. A. et al. Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. **PloS one**, v. 7, n. 2, p. e30953, 2012.

CHALOPIN, D. et al. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. **Genome biology and evolution**, v. 7, n. 2, p. 567-580, 2015.

CHARLESWORTH, B.; SNIEGOWSKI, P.; STEPHAN, W. The evolutionary dynamics of repetitive DNA in eukaryotes. **Nature**, v. 371, n. 6494, p. 215-220, 1994.

CHOUDHARY, O. P.; TRIVEDI, S. Microsatellite or simple sequence repeat (SSR) instability depends on repeat characteristics during replication and repair. **Journal Cell and Molecular Biology**, v. 8, n. 2, p. 21-34, 2010.

CONTE, M. et al. Isolation and characterization of polymorphic microsatellites for the natural populations of barker frog *Physalaemus cuvieri*. **Conservation genetics**, v. 10, n. 6, p. 1849, 2009.

CONTE, C. E. et al. Novos registros na distribuição geográfica de anuros na Floresta com Araucária e considerações sobre suas vocalizações. **Biota Neotropica**, v. 10, n. 2, 2010.

CONTE, M. et al. Study of closely related species within the *Physalaemus cuvieri* group (Anura): contribution of microsatellite markers. **Genetics and molecular research: GMR**, 2011.

CRAWFORD, A. J.; LIPS, K. R.; BERMINGHAM, E. Epidemic disease decimates amphibian abundance, species diversity, and evolutionary history in the highlands of central Panama. **Proceedings of the National Academy of Sciences**, v. 107, n. 31, p. 13777-13782, 2010.

DAVEY, J. W. et al. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. **Nature reviews. Genetics**, v. 12, n. 7, p. 499, 2011.

DECANINE, D. O papel de marcadores moleculares na genética forense. **Revista Brasileira de Criminalística**, v. 5, n. 2, p. 18-27, 2016.

DELMAS, C. E. L. et al. Isolation and characterization of microsatellite loci in *Rhododendron ferrugineum* (Ericaceae) using pyrosequencing technology. **American journal of botany**, v. 98, n. 5, p. e120-e122, 2011.

DINIZ-FILHO, J. A. F. et al. Anuran species richness, complementarity and conservation conflicts in Brazilian Cerrado. **Acta oecologica**, v. 29, n. 1, p. 9-15, 2006.

- DUELLMAN, W. E.; MARION, A. B.; HEDGES, S. B. Phylogenetics, classification, and biogeography of the treefrogs (Amphibia: Anura: Arboranae). **Zootaxa**, v. 4104, n. 1, p. 1-109, 2016.
- DUFRESNE, F.; JEFFERY, N. A guided tour of large genome size in animals: what we know and where we are heading. **Chromosome Research**, v. 19, n. 7, p. 925-938, 2011.
- DUFRESNES, C. et al. Stronger transferability but lower variability in transcriptomic-than in anonymous microsatellites: evidence from Hylid frogs. **Molecular ecology resources**, v. 14, n. 4, p. 716-725, 2014.
- DURAN, C. et al. Molecular genetic markers: discovery, applications, data storage and visualization. **Current Bioinformatics**, v. 4, n. 1, p. 16-27, 2009
- EKBLOM, R.; GALINDO, J. Applications of next generation sequencing in molecular ecology of non-model organisms. **Heredity**, v. 107, n. 1, p. 1, 2011.
- EKBLOM, R.; WOLF, J. B.W. A field guide to whole-genome sequencing, assembly and annotation. **Evolutionary applications**, v. 7, n. 9, p. 1026-1042, 2014.
- ELLEGREN, H. Microsatellite mutations in the germline: implications for evolutionary inference. **Trends in genetics**, v. 16, n. 12, p. 551-558, 2000.
- ELLEGREN, H. Microsatellites: simple sequences with complex evolution. **Nature reviews. Genetics**, v. 5, n. 6, p. 435, 2004.
- FAGGIONI, G. P. et al. Isolation and characterization of microsatellites markers for two South American frogs (*Leptodactylus bufonius* and *L. chaquensis*) using next generation sequencing. **Amphibia-Reptilia**, v. 35, n. 4, p. 405-412, 2014.
- FAIVOVICH, J. A cladistic analysis of *Scinax* (Anura: Hylidae). **Cladistics**, v. 18, n. 4, p. 367-393, 2002.
- FAIVOVICH, J. et al. Systematic review of the frog family Hylidae, with special reference to Hylinae: phylogenetic analysis and taxonomic revision. **Bulletin of the American Museum of natural History**, p. 1-240, 2005.

FALEIRO, F. G. **Marcadores genético-moleculares aplicados a programas de conservação e uso de recursos genéticos**. Embrapa Cerrados, 2007.

FAUCHER, L.; GODÉ, C.; ARNAUD, J. F. Development of nuclear microsatellite loci and mitochondrial single nucleotide polymorphisms for the natterjack toad, *Bufo (Epidalea) calamita* (Bufonidae), using next generation sequencing and Competitive Allele Specific PCR (KASPar). **Journal of Heredity**, v. 107, n. 7, p. 660-665, 2016.

FERNANDES, G. A.; CAPARROZ, R. DNA sequence analysis to guide the release of blue-and-yellow macaws (*Ara ararauna*, Psittaciformes, Aves) from the illegal trade back into the wild. **Molecular biology reports**, v. 40, n. 3, p. 2757-2762, 2013.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introducción al uso de marcadores moleculares en el análisis genético**. Embrapa, 1998.

FESCHOTTE, C.; JIANG, N.; WESSLER, S. R. Plant transposable elements: where genetics meets genomics. **Nature reviews. Genetics**, v. 3, n. 5, p. 329, 2002.

FOULONGNE-ORIOU, M. et al. Genome-wide survey of repetitive DNA elements in the button mushroom *Agaricus bisporus*. **Fungal genetics and biology**, v. 55, p. 6-21, 2013.

FROST, D. R. **Amphibian Species of the World: an Online Reference**. Version 6.0 (Date of access). Electronic Database accessible at <http://research.amnh.org/herpetology/amphibia/index.html>. American Museum of Natural History, New York, USA, 2017.

GEMAYEL, R. et al. Variable tandem repeats accelerate evolution of coding and regulatory sequences. **Annual review of genetics**, v. 44, p. 445-477, 2010.

HOFFMANN, M. et al. The impact of conservation on the status of the world's vertebrates. **Science**, v. 330, n. 6010, p. 1503-1509, 2010.

IGAWA, T. et al. Isolation and characterization of twelve microsatellite loci of endangered Ishikawa's frog (*Odorrana ishikawae*). **Conservation Genetics Resources**, v. 3, n. 3, p. 421-424, 2011.

JACQ, C.; MILLER, J. R.; BROWNLEE, G. G. A pseudogene structure in 5S DNA of *Xenopus laevis*. **Cell**, v. 12, n. 1, p. 109-120, 1977.

JEFFREYS, A. J. et al. Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. **Nucleic Acids Research**, v. 16, n. 23, p. 10953-10971, 1988.

KALIA, R. K. et al. Microsatellite markers: an overview of the recent progress in plants. **Euphytica**, v. 177, n. 3, p. 309-334, 2011.

KAPITONOV, V. V.; JURKA, J. Rolling-circle transposons in eukaryotes. **Proceedings of the National Academy of Sciences**, v. 98, n. 15, p. 8714-8719, 2001.

KATZENBERGER, M. et al. Tolerância e sensibilidade térmica em anfíbios. **Revista da Biologia**, 2012.

KLINK, C. A.; MACHADO, R. B. Conservation of the Brazilian cerrado. **Conservation biology**, v. 19, n. 3, p. 707-713, 2005

KOMAKI, S. et al. Development and characterization of 14 microsatellite markers for *Buergeria japonica* (Amphibia, Anura, Rhacophoridae). **Genes & genetic systems**, v. 89, n. 1, p. 35-39, 2014.

LAWSON, M. J.; ZHANG, L. Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. **Genome biology**, v. 7, n. 2, p. R14, 2006.

LEVINSON, G.; GUTMAN, G. A. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. **Nucleic Acids Research**, v. 15, n. 13, p. 5323-5338, 1987.

LIPINSKI, V. M.; SANTOS, T. G. Estrutura e organização espacial de duas comunidades de anuros do bioma Pampa. **Iheringia, Série Zoologia, Porto Alegre**, v. 104, n. 4, p. 462-469, 2014.

LOEBMANN, D.; HADDAD, C. F. B. Amphibians and reptiles from a highly diverse area of the Caatinga domain: composition and conservation implications. **Biota Neotropica**, v. 10, n. 3, p. 227-256, 2010.

LOHSE, M. et al. Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. **Plant, cell & environment**, v. 37, n. 5, p. 1250-1258, 2014.

LUO, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. **Gigascience**, v. 1, n. 1, p. 18, 2012.

LUQUET, E. et al. Heterozygosity–fitness correlations among wild populations of European tree frog (*Hyla arborea*) detect fixation load. **Molecular Ecology**, v. 20, n. 9, p. 1877-1887, 2011.

MARDIS, E. R. Next-generation DNA sequencing methods. **Annual Review Genomics Human Genetics**, v. 9, p. 387-402, 2008.

MARDIS, E. R. Next-generation sequencing platforms. **Annual review of analytical chemistry**, v. 6, p. 287-303, 2013.

MASON, A. S. SSR genotyping. **Plant Genotyping: Methods and Protocols**, p. 77-89, 2015.

MEGLÉCZ, E. et al. QDD version 3.1: a user-friendly computer program for microsatellite selection and primer design revisited: experimental validation of variables determining genotyping success rate. **Molecular Ecology Resources**, v. 14, n. 6, p. 1302-1313, 2014.

MEHROTRA, S.; GOYAL, V. Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. **Genomics, proteomics & bioinformatics**, v. 12, n. 4, p. 164-171, 2014.

METZKER, M. L. Sequencing technologies--the next generation. **Nature reviews. Genetics**, v. 11, n. 1, p. 31, 2010.

MIGA, K. H. Completing the human genome: the progress and challenge of satellite DNA assembly. **Chromosome research**, v. 23, n. 3, p. 421-426, 2015.

MIGHELL, A. J. et al. Vertebrate pseudogenes. **FEBS letters**, v. 468, n. 2-3, p. 109-114, 2000.

MMA (Ministério do Meio Ambiente). **O bioma Cerrado**. Disponível em: <<http://www.mma.gov.br/biomas/cerrado>>. Acesso em: 17 set. 2017.

MOURA, M. R. et al. Amphibia, Anura, Hylidae, *Scinax centralis* Pombal and Bastos, 1996: distribution extension, geographic distribution map. **Check List**, v. 6, n. 1, p. 173-175, 2010.

MYERS, N. et al. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, n. 6772, p. 853, 2000.

NALI, R. C.; ZAMUDIO, K. R.; PRADO, C. P. A. Microsatellite markers for *Bokermannohyla* species (Anura, hylidae) from the brazilian cerrado and atlantic forest domains. **Amphibia-Reptilia**, v. 35, n. 3, p. 355-360, 2014.

OHNO, S. So much" junk" DNA in our genome. In: **Brookhaven symposia in biology**. p. 366-370, 1972

OLIVEIRA, E. J. et al. Origin, evolution and genome distribution of microsatellites. **Genetics and Molecular Biology**, v. 29, n. 2, p. 294-307, 2006.

OLIVEIRA, S. L. et al. Anfíbios anuros nos covais do município de Jataí, estado de Goiás. 2015.

ORGEL, L. E.; CRICK, F. H. C. Selfish DNA: the ultimate parasite. **Nature**, v. 284, n. 5757, p. 604-607, 1980.

OSTRANDER, E. A. et al. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. **Proceedings of the National Academy of Sciences**, v. 89, n. 8, p. 3419-3423, 1992.

OYAMAGUCHI, H. M.; OKUBO, R. P.; POLLINGER, J. P. Characterization of new polymorphic microsatellite loci for the lesser tree frog (*Dendropsophus minutus*). **Amphibia-Reptilia**, v. 36, n. 1, p. 83-86, 2015.

PAETKAU, D. et al. Microsatellite analysis of population structure in Canadian polar bears. **Molecular ecology**, v. 4, n. 3, p. 347-354, 1995.

PALAZZO, A. F.; GREGORY, T. R. The case for junk DNA. **PLoS genetics**, v. 10, n. 5, p. e1004351, 2014.

PEARCE, S. R. et al. TheTy1-copia group retrotransposons inVicia species: copy number, sequence heterogeneity and chromosomal localisation. **Molecular and General Genetics MGG**, v. 250, n. 3, p. 305-315, 1996.

PEI, B. et al. The GENCODE pseudogene resource. **Genome biology**, v. 13, n. 9, p. R51, 2012.

POMBAL JR, J. P.; BASTOS, R. P. Nova espécie de *Scinax* Wagler, 1830 do Brasil Central (Amphibia, Anura, Hylidae). 1996.

POUNDS, J. A. et al. Widespread amphibian extinctions from epidemic disease driven by global warming. **Nature**, v. 439, n. 7073, p. 161, 2006.

QIAO, L. et al. Transcriptome profile of the green odorous frog (*Odorrana margaretae*). **PloS one**, v. 8, n. 9, p. e75211, 2013.

RASSMANN, K.; SCHLÖTTERER, C.; TAUTZ, D. Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. **Electrophoresis**, v. 12, n. 2-3, p. 113-118, 1991.

RATTER, J. A.; RIBEIRO, J. F.; BRIDGEWATER, S. The Brazilian cerrado vegetation and threats to its biodiversity. **Annals of botany**, v. 80, n. 3, p. 223-230, 1997.

RIBEIRO, J. F.; WALTER, B. M. T. **As principais fitofisionomias do bioma cerrado** in: Sano, SM; Almeida, SP; Ribeiro, JF Cerrado: Ecologia e flora. Embrapa Informação Tecnológica, Brasília, DF, v. 1, p. 406, 2008.

ROME, L. C.; STEVENS, E. Don; JOHN-ALDER, H. B. The influence of temperature and thermal acclimation on physiological function. **Environmental physiology of the amphibians**, p. 205-205, 1992.

SAFONOVA, Y.; BANKEVICH, A.; PEVZNER, P. A. dipSPAdes: assembler for highly polymorphic diploid genomes. **Journal of Computational Biology**, v. 22, n. 6, p. 528-545, 2015.

SANMIGUEL, P. et al. Nested retrotransposons in the intergenic regions of the maize genome. **Science**, v. 274, n. 5288, p. 765-768, 1996.

SCHATZ, M. C.; DELCHER, A. L.; SALZBERG, S. L. Assembly of large genomes using second-generation sequencing. *Genome research*, v. 20, n. 9, p. 1165-1173, 2010.

SCHLÖTTERER, C.; TAUTZ, D. Slippage synthesis of simple sequence DNA. **Nucleic acids research**, v. 20, n. 2, p. 211-215, 1992.

SCHLÖTTERER, C. The evolution of molecular markers—just a matter of fashion? **Nature reviews genetics**, v. 5, n. 1, p. 63, 2004.

SCHUELER, M. G. et al. Genomic and genetic definition of a functional human centromere. **Science**, v. 294, n. 5540, p. 109-115, 2001.

SEGALLA, M. V. et al. Brazilian Amphibians: List of Species. *Herpetologia Brasileira*. v. 5, n. 2, p. 34–46, 2016.

SHARMA, P. C.; GROVER, A.; KAHL, G. Mining microsatellites in eukaryotic genomes. **Trends in biotechnology**, v. 25, n. 11, p. 490-498, 2007.

SILVA, P. I. T. et al. Development and validation of microsatellite markers for *Brachiaria ruziziensis* obtained by partial genome assembly of Illumina single-end reads. **Bmc Genomics**, v. 14, n. 1, p. 17, 2013.

SIMPSON, J. T. et al. ABySS: A parallel assembler for short read sequence data. **Genome research**, v. 19, p. 1117-1123, 2009.

STANKE, M.; MORGENSTERN, B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. **Nucleic Acids Research**, n. 33, 2005.

STEIN, L. Genome annotation: from sequence to biology. **Nature reviews genetics**, v. 2, n. 7, p. 493, 2001.

SUN, Yan-Bo et al. Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. **Proceedings of the National Academy of Sciences**, v. 112, n. 11, p. E1257-E1262, 2015.

TARAILO-GRAOVAC, M.; CHEN, N. Using RepeatMasker to identify repetitive elements in genomic sequences. **Current protocols in bioinformatics**, p. 4.10. 1-4.10. 14, 2009.

TAUTZ, D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. **Nucleic acids research**, v. 17, n. 16, p. 6463-6471, 1989.

TELLES, M. P. C. et al. Landscape genetics of *Physalaemus cuvieri* in Brazilian Cerrado: correspondence between population structure and patterns of human occupation and habitat loss. **Biological Conservation**, v. 139, n. 1-2, p. 37-46, 2007.

TORRENTS, D. et al. A genome-wide survey of human pseudogenes. **Genome research**, v. 13, n. 12, p. 2559-2567, 2003.

VALDUJO, P. H. et al. Anuran species composition and distribution patterns in Brazilian Cerrado, a Neotropical hotspot. **South American Journal of Herpetology**, v. 7, n. 2, p. 63-78, 2012.

VANIN, E. F. Processed pseudogenes: characteristics and evolution. **Annual review of genetics**, v. 19, n. 1, p. 253-272, 1985.

VARSHNEY, R. K.; GRANER, A.; SORRELLS, M. E. Genic microsatellite markers in plants: features and applications. **TRENDS in Biotechnology**, v. 23, n. 1, p. 48-55, 2005.

VERDADE, V. K. et al. A leap further: the Brazilian amphibian conservation action plan. **Alytes**, v. 29, 2012.

VIEIRA, M. L. C. et al. Microsatellite markers: what they mean and why they are so useful. **Genetics and molecular biology**, v. 39, n. 3, p. 312-328, 2016.

VITT, L. J.; CALDWELL, J. P. **Herpetology: an introductory biology of amphibians and reptiles**. Academic Press, 2013.

VOGT, R. C.; BERNHARD, R. Biodiversidade e biogeografia de répteis e anfíbios da Amazônia. **Caderno de Ciência, Instituto Amazônia**, v. 1, p. 1-40, 2003.

WANG, C. et al. Isolation and characterization of 113 polymorphic microsatellite loci for the Tibetan frog (*Nanorana parkeri*) using next generation sequencing. **Conservation Genetics Resources**, v. 5, n. 4, p. 915-924, 2013.

WARREN, I. A. et al. Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. **Chromosome research**, v. 23, n. 3, p. 505-531, 2015.

WEINER, A. M.; DEININGER, P. L.; EFSTRATIADIS, A. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. **Annual review of biochemistry**, v. 55, n. 1, p. 631-661, 1986.

WELLS, K. D. **The Ecology and Behavior of Amphibians**. Chicago and London: The University of Chicago Press, 1148 p., 2007.

WRIGHT, J. M. Mutation at VNTRs: Are minisatellites the evolutionary progeny of microsatellites? **Genome**, v. 37, n. 2, p. 345-347, 1994.

WWF (World Wide Fund for Nature). **Cerrado berço das águas**. 2012. Disponível em: <[http://d3nehc6yl9qzo4.cloudfront.net/downloads/wwf\\_factsheet\\_cerrado\\_pt\\_web.pdf](http://d3nehc6yl9qzo4.cloudfront.net/downloads/wwf_factsheet_cerrado_pt_web.pdf)>.

Acesso em: 17 set. 2017.

YANDELL, M.; ENCE, D. A beginner's guide to eukaryotic genome annotation. **Nature Reviews Genetics**, v. 13, n. 5, p. 329, 2012.

YUAN, S. et al. Development of microsatellite markers for the spiny-bellied frog *Quasipaa boulengeri* (Anura: Dicroglossidae) through transcriptome sequencing. **Conservation genetics resources**, v. 7, n. 1, p. 229-231, 2015.

YUNIS, J. J.; YASMINEH, W. G. Heterochromatin, satellite DNA, and cell function. **Science**, v. 174, n. 4015, p. 1200-1209, 1971.

ZALAPA, J. E. et al. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. **American journal of botany**, v. 99, n. 2, p. 193-208, 2012.

ZANE, L.; BARGELLONI, L.; PATARNELLO, T. Strategies for microsatellite isolation: a review. **Molecular ecology**, v. 11, n. 1, p. 1-16, 2002.

ZERBINO, D. R.; BIRNEY, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. **Genome research**, v. 18, n. 5, p. 821-829, 2008.

ZHANG, Z. D. et al. Identification and analysis of unitary pseudogenes: historic and contemporary gene losses in humans and other primates. **Genome biology**, v. 11, n. 3, p. R26, 2010.

ZIELEZINSKI, A. et al. Annotating a non-model plant genome—a study on the narrow-leafed lupin. **BioTechnologia. Journal of Biotechnology Computational Biology and Bionanotechnology**, v. 93, n. 3, 2012.

## 5. CAPÍTULO 1: MONTAGEM E ANOTAÇÃO DE SEQUÊNCIAS DO GENOMA DE *Ololygon centralis* (Pombal e Bastos, 1996) (ANURA: HYLIDAE) OBTIDAS POR SEQUENCIAMENTO DE SEGUNDA GERAÇÃO COM BAIXA COBERTURA

### RESUMO

O tamanho e a complexidade que o material genético dos eucariotos apresenta está relacionado com a abundância dos diferentes tipos de regiões que o constitui. Essas regiões podem estar relacionadas com funções como organização, funcionamento e manutenção do genoma. As regiões repetitivas podem ser encontradas dispersas no genoma, em locais específicos ou distribuídas em tandem. O objetivo deste trabalho foi sequenciar parcialmente o genoma de *Ololygon centralis*, anotar e quantificar os elementos repetitivos (microssatélites e elementos transponíveis), bem como prever genes presentes no genoma desta espécie de anuro endêmica do bioma Cerrado. Para tanto, o DNA genômico de dois indivíduos de *O. centralis* provenientes da Floresta Nacional de Silvânia, Silvânia, GO, foram utilizados para preparação das bibliotecas genômicas, seguindo o protocolo Nextera DNA (Illumina). As bibliotecas foram validadas e quantificadas utilizando Bioanalyzer e qPCR, respectivamente. As sequências foram obtidas pela plataforma Illumina Miseq (Miseq Reagent v3 kit). Para as análises e cortes de qualidade foram utilizados os programas FastQC e Trimmomatic. As sequências de boa qualidade foram utilizadas para montagem de sequências do genoma utilizando dipSPAdes 1.0. Os elementos repetitivos foram detectados pelo software RepeatMasker por meio da comparação com a biblioteca de *Xenopus silurana tropicalis*, para a predição de genes codificadores de tRNA foi utilizado o software tRNAscan-SE, para rRNA o programa Rfam. E a anotação funcional dos genes foi realizada pelo Blast2go. As sequências genômicas montadas resultou em 13.989 *scaffolds*, correspondente a 4.926.069 pb com N50 de 336. Os elementos repetitivos encontrados no genoma de *O. centralis* (1.795 elementos transponíveis, 957 microssatélites, um DNA satélite e 25 snRNAs) totalizaram 531.337 pb. Os motivos de microssatélites mais frequentes foram os tetranucleotídeos e o de menor frequência os dinucleotídeos com 26,23% e 12,02% respectivamente. Foi localizado somente um elemento de DNA satélite com tamanho de 137 pb. Entre os retrotransposons, o elemento LINE (49,83%) foi o mais abundante, seguido do LTR (48,66%) e SINE (1,56%). As superfamílias DIRS e LINE/L1 foram as mais encontradas nos retrotransposons. Foram identificados genes que codificam as subunidades 5S, 18S e 28S de ribossomos além de 18 tipos de tRNAs no genoma. Quanto as sequências de transcritos em 21 foram atribuídas a classificação de ontologia de gene, as quais foram classificadas em processos biológicos, função molecular e componente celular. A montagem parcial e identificação de TEs, microssatélites, predição de genes nas sequências do anuro *O. centralis* é um importante passo para conhecimento genômico dessa espécie e para os anfíbios em geral.

Palavras-chave: Anura, microssatélites, NGS, TEs.

## INTRODUÇÃO

A espécie *Ololygon centralis* é um hilídeo endêmico do Bioma Cerrado, com distribuição restrita a cursos d'água (Pombal e Bastos, 1996; Brandão e Araújo, 2001; Moura et al., 2010). Apesar da espécie não ser considerada ameaçada de extinção pela União Internacional para a Conservação da Natureza (Bastos et al., 2004) o seu habitat é considerado um dos “hot spots” de biodiversidade do mundo (Myers et al., 2000) o qual sofre com a exploração de recursos e grande perda de habitat devido à expansão agrícola (MMA, 2018).

Pouco se conhece sobre a genômica de *O. centralis*. Até hoje nenhuma espécie, mesmo dentro do gênero, teve o genoma sequenciado. A nível de família (Hylidae) são descritos no banco de dado NCBI (*National Center for Biotechnology Information*) 127 ESTs (*Expressed Sequence Tag*), 200 genes, 75.174 sequências de DNA/RNA e 15.834 sequências de proteínas. Para anuros em geral estão depositados no NCBI o *assembly* de duas espécies de *Xenopus* (Session et al., 2016; Buisine et al., 2015), *Nanorana parkeri* (Sun et al., 2015) e *Rana catesbeiana* (Hammond, 2017).

A utilização do sequenciamento de nova geração (NGS), como a plataforma Illumina, possibilitou o acesso a genomas de diversos organismos os quais não se tinha conhecimento nenhum a nível molecular. Essa técnica aliada a ferramentas de bioinformática permitiu a montagem de genomas *de novo* (Qiao et al., 2013; Sun et al., 2015).

Descartando a obrigatoriedade da utilização de um genoma de referência para o alinhamento das sequências, uma montagem *de novo* consiste em construir a sequência de um genoma baseado apenas nos fragmentos de DNA geradas pelo NGS de forma aleatória (Chaisson et al., 2015) através de algoritmos, baseados em grafos de Bruijn por exemplo, em que as *reads* são decompostas em subsequências de tamanhos K (*K-mer*) e depois é formado um grafo de sobreposição entre o sufixo e prefixo dos *k-mers* (Schatz et al., 2010; Compeau et al., 2011). Essa metodologia tem impulsionado o desenvolvimento de projetos de sequenciamento acarretando na geração de grande quantidade de dados (Metzker 2010; Li et al., 2010; Schatz et al., 2010). Através de um genoma sequenciado pode-se alcançar diversos objetivos como descoberta de polimorfismos, mutações, quantificação de expressão gênica e anotação do genoma (Shendure e Ji, 2008).

Depois que os milhares de nucleotídeos gerados pelo sequenciamento são alinhados e montados em uma única sequência, ainda não se obtêm instantaneamente informações biológicas acerca do organismo. As sequências de DNA representam o início para a compreensão do genoma e a anotação dessas sequências é o primeiro passo para evidenciar características e elementos presentes no genoma da espécie estudada (Ekblom e Wolf, 2014; Stein, 2001).

O método de anotação tem a finalidade de identificar e descrever os elementos que estão presentes no genoma. O processo de anotação pode ser dividido em três níveis, sendo a nível de nucleotídeo que identifica RNAs e elementos repetitivos, nível de proteína que define e atribui função às proteínas encontradas e por último a nível de processo que relaciona o genoma com os processos biológicos (Stein, 2001; Yandell e Ence, 2012; Sun et al., 2015).

Com o intuito de conhecer o genoma de anfíbios esse trabalho teve como objetivos montar parcialmente sequências do genoma de *Oligodon centralis* e realizar a identificação, caracterização, quantificação de elementos repetitivos e predição de genes nas sequências genômicas.

## MATERIAL E MÉTODOS

### Extração de DNA

Dois indivíduos de *Oligodon centralis* (ZUFG9289 e ZUFG9291) usados para a extração de DNA foram coletados na Floresta Nacional de Silvânia (FLONA). As coletas e procedimentos foram realizados com autorização SISBIO/ICMBIO 46522-3 e do Comitê de Ética em Uso de Animais da Universidade Federal de Goiás número 109/14 (ANEXOS). Amostras de fígado foram armazenadas em álcool etílico PA, guardadas em freezer -20 graus Celsius na Coleção Zoológica da Universidade Federal de Goiás. Para extração de DNA os tecidos foram submetidos à fragmentação mecânica e lise em tampão de extração (50mM tris-HCl, pH 8,0; 25mM EDTA e 400mM NaCl) adicionado de SDS 20% e Proteinase K (10mg/μL) e incubados a 56°C por aproximadamente 5 horas para o processo de digestão enzimática. Foram adicionadas às amostras NaCl 5M e após incubação e centrifugação adicionou-se NaAc 3M no sobrenadante. A precipitação foi realizada com Etanol absoluto e a lavagem com Etanol 70%. O *pellet* foi ressuspensão em TE (pH 8,0) + RNase (10 mg/μl). A qualidade do DNA foi analisada em gel de agarose 1% através da eletroforese horizontal.

## **Preparo da biblioteca para o sequenciamento Illumina**

A quantificação do DNA foi realizada em equipamento Qubit® fluorometer 2.0 utilizando o kit Qubit dsDNA HS Assay (ThermoFisher). Para preparação da biblioteca, foi utilizado 50 ng de DNA de cada indivíduo seguindo o protocolo do kit Nextera DNA (Illumina). As bibliotecas foram validadas conforme o tamanho dos fragmentos no Agilent Technologies 2100 Bioanalyzer usando o High Sensitivity DNA kit (Agilent) e quantificadas por PCR em tempo real com o kit Kapa Biosystems (KK4824). Os indivíduos foram sequenciados na plataforma MiSeq da Illumina utilizando os reagentes do MiSeq V3 600 ciclos (Illumina).

A qualidade dos *reads* obtidos no sequenciamento foi analisada utilizando o software FastQC (Andrews, 2010). *Reads* que apresentaram baixa qualidade foram removidos pelo software Trimmomatic (Bolger et al., 2014). O valor mínimo de qualidade utilizado foi o de phred 20 para as duas bibliotecas. Os *reads* com qualidade superior originados dos sequenciamentos dos dois indivíduos foram utilizados na montagem das sequências do genoma de *Ololygon centralis* utilizando o software dipSPAdes 1.0 (Safonova et al., 2015).

## **Identificação de elementos repetitivos e TEs**

A identificação das regiões microssatélites, retroelementos, DNA transposons e *rolling-circles* foi feita utilizando o software RepeatMasker identificando-os por similaridade de sequências (Tarailo-Graovac e Chen, 2009), através do confronto das sequências do genoma de *Ololygon centralis* com a biblioteca de repetições do banco de dados do RepeatMasker para a espécie *Xenopus silurana tropicalis*.

## **Anotação de genes codificantes de tRNA e rRNA**

A predição de genes codificadores de RNAs transportadores (tRNA) foi realizada usando o software versão web tRNAscan-SE que utiliza modelos de covariância e possuem alta sensibilidade e especificidade na definição de sequência de tRNA (Lowe e Chan, 2016). As sequências codificadoras de RNA ribossomais (rRNA) foram identificadas utilizando o serviço web Rfam, que dispõe de um banco de dados de famílias de RNAs estruturais e tem como princípio o alinhamento de sequências por similaridade (Gardner et al., 2008).

## **Anotação funcional de genes codificantes de proteínas**

A predição de RNAs mensageiros (mRNAs) foi realizada por meio do BLAST usando sequências de referência de *Xenopus tropicalis* depositado no NCBI. Foram usados os seguintes parâmetros para esta análise de similaridade: 80% de identidade entre as sequências,  $e\text{-value}^{-16}$ , e  $\text{blast hit} = 1$ . As sequências preditas como genes foram isoladas em um arquivo fasta e submetidas a anotação funcional de genes realizada no software Blast2go (Conesa et al., 2005). A distribuição nas três categorias do Blast2GO componentes celulares, função molecular e processos biológicos, foi realizada usando o nível 5.

## **RESULTADOS E DISCUSSÃO**

### **Montagem do genoma de *Ololygon centralis***

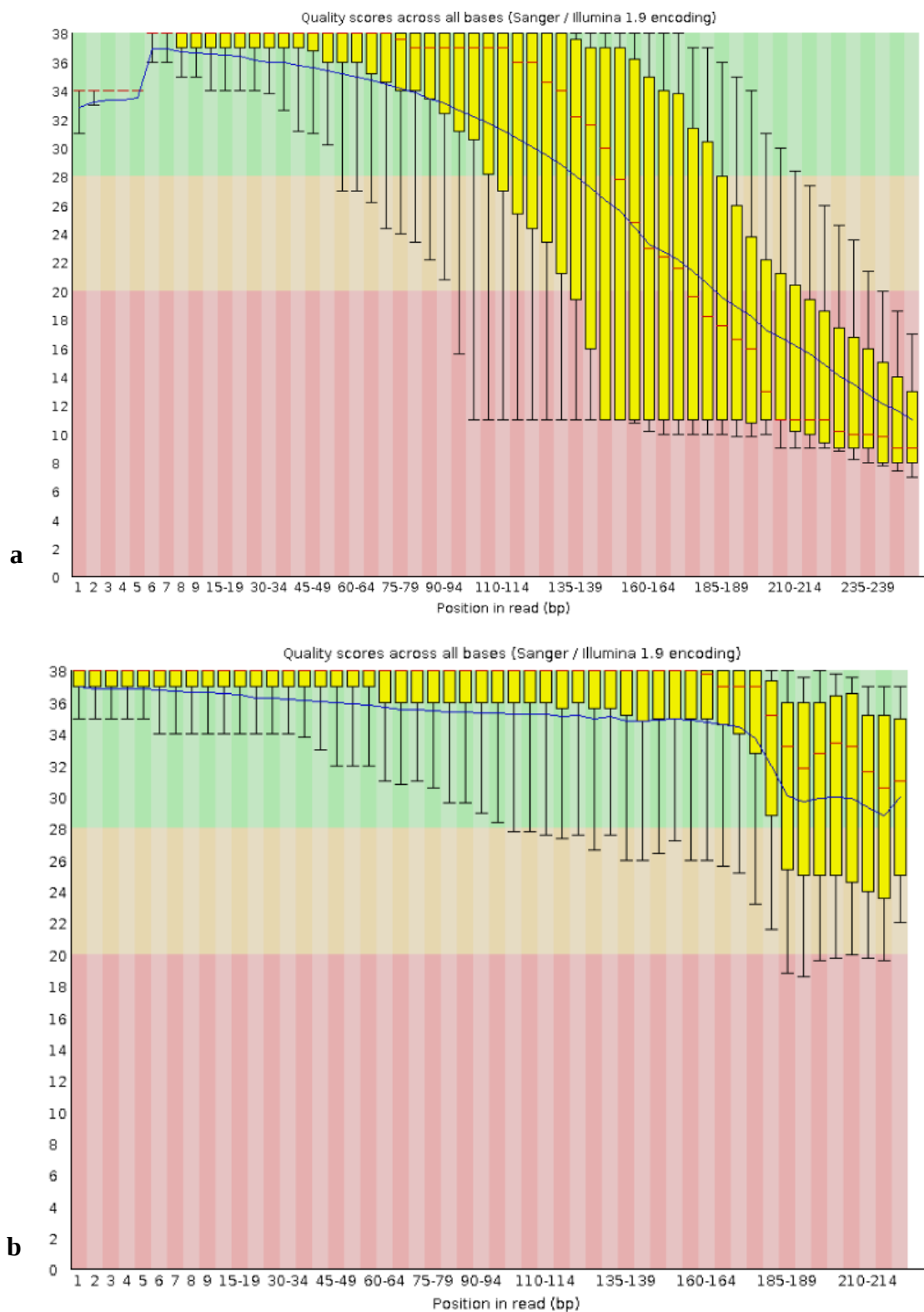
A partir das bibliotecas de *Ololygon centralis*, o sequenciamento resultou em 4.934.289 *reads paired-ends* depois da análise de qualidade das sequências foram mantidos 3.797.560 (76,96%) *reads* de boa qualidade com *phred* acima de 20 (Figura 1), aumentando a frequência dos *reads* curtos e melhorando a distribuição das bases (Figura 2). *Reads* de baixa qualidade comprometem o processo de análise do genoma, pois, a contaminação das sequências diminui a confiabilidade das análises (Del Fabbro et al., 2013; Ekblom e Wolf, 2014).

Na montagem desse genoma foi obtido um total de 4,9 Mega pares de bases (Mpb) que corresponde aos 4.926.069 pb de *scaffolds* os quais tiveram variação de 78 a 12.097 pb em tamanho. Os resultados do sequenciamento representam uma aplicação positiva quanto a geração de dados levando em consideração a limitação de informações genômicas disponíveis dos anfíbios, agregando assim informações para a compreensão do genoma desse grupo que é bastante complexo (Gregory, 2002).

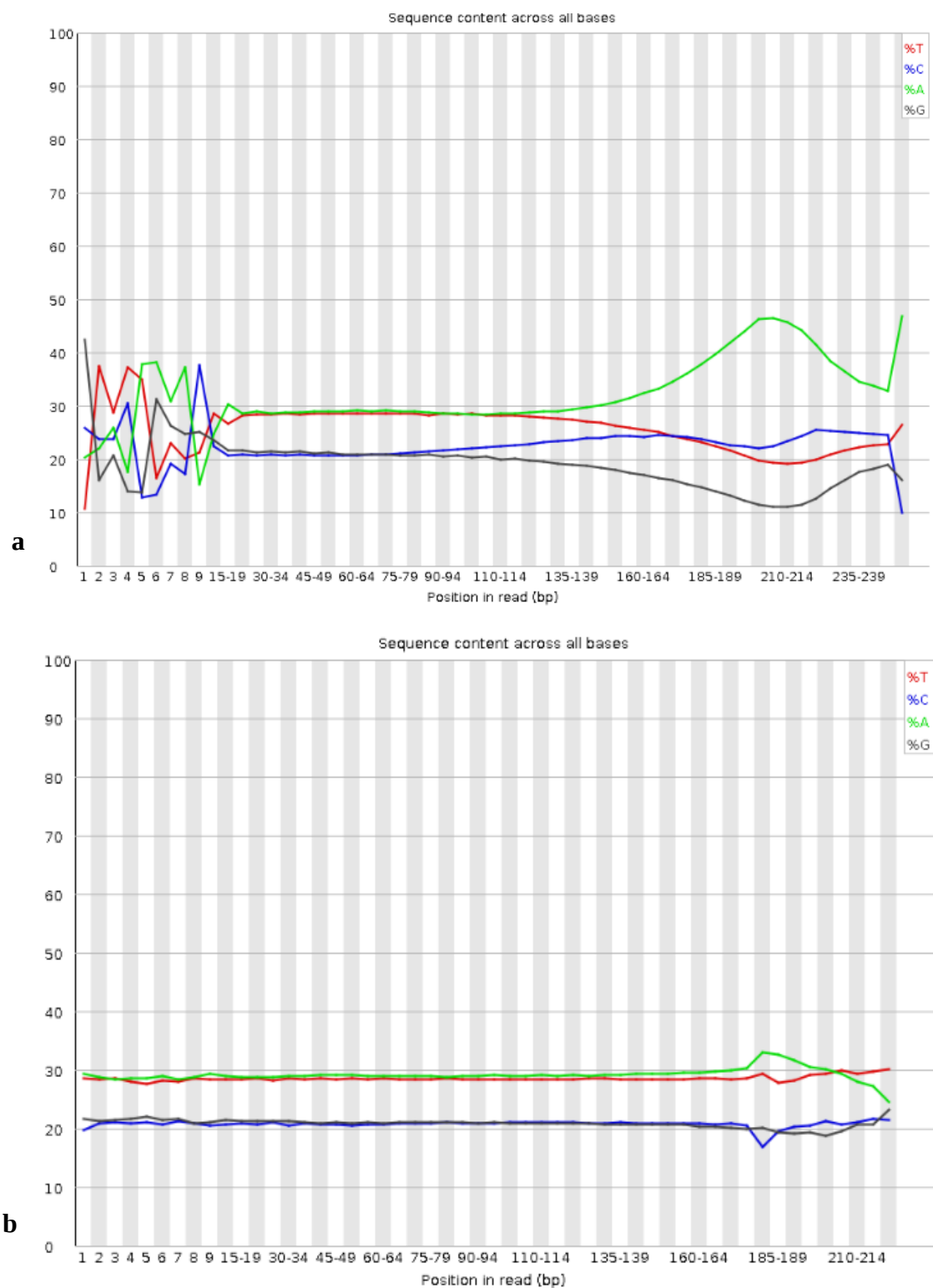
A montagem do genoma não apresentou atribuição de nenhuma base indefinida. Os nucleotídeos A e T representaram 57,62% do total de bases sequenciadas (Tabela 1) e o conteúdo GC do genoma parcial de *O. centralis* sequenciado foi de 42,38%. Essa proporção encontrada no conteúdo de bases se assemelha ao valor caracterizado no genoma do sapo tibetano *Nanorana parkeri*, o qual apresentou 42,5% de conteúdo GC (Sun et al., 2015).

**Tabela 1.** Parâmetros da montagem parcial do genoma de *Ololygon centralis*.

| <b>Parâmetros</b>               | <b>Valores</b> |
|---------------------------------|----------------|
| Número total de scaffolds       | 13.989         |
| Tamanho total dos scaffold (pb) | 4.926.069      |
| Tamanho do maior scaffold (pb)  | 12.097         |
| Tamanho do menor scaffold (pb)  | 78             |
| N50 dos scaffolds               | 336            |
| Scaffolds > 1k nucleotídeos     | 307            |
| Scaffolds > 10k nucleotídeos    | 1              |
| Scaffolds A%                    | 28,91          |
| Scaffolds C%                    | 21,25          |
| Scaffolds G%                    | 21,13          |
| Scaffolds T%                    | 28,71          |
| Scaffolds N%                    | 0,00           |



**Figura 1.** Qualidade das bases da biblioteca de *Ololygon centralis* antes (a) e depois (b) do controle de qualidade. O eixo y corresponde ao valor da qualidade de phred e o eixo x o tamanho das reads em pares de bases.



**Figura 2.** Distribuição das bases dos *reads* de *Ololygon centralis* antes (a) e depois (b) do controle de qualidade. O eixo y corresponde ao conteúdo da base e o eixo x o tamanho das *reads* em pares de bases

### **Anotação de elementos repetitivos no genoma de *O. centralis***

Os elementos transponíveis (TEs) representaram 9,72% (478.424 pb) do total de pares de bases das sequências genômicas de *Ololygon centralis* com 1.795 elementos (Tabela 2). Dentre eles os retroelementos corresponderam a 8,81% (433.382 pb), 0,91% (44.875 pb) DNA transposons e 0,003% (167 pb) não foram identificados. Entre os retrotransposons, a ordem LINE (*Long Interspersed Elements*) (49,83%) foi o mais abundante, seguido do LTR (*Long Terminal Repeat*) (48,66%) e SINE (*Short Interspersed Elements*) (1,56%) sendo as superfamílias LINE/L1 (70,7%) e DIRS (53,41%) as mais encontradas. Uma única ordem foi detectada para DNA transposon, e as famílias mais abundantes foram TcMar-Tc1 (34,98%) e TcMar-Tc2 (17,28%), ambos da superfamília Mariner.

A abundância de retroelementos também foi constatada para o genoma de *Nanorana parkeri*, diferentemente do que foi identificado para a espécie *Xenopus tropicalis* que tem um terço de seu genoma constituído de TEs e apresentou maior abundância de DNA transposon. Apesar da limitação de dados genômicos para essa ordem é interessante mencionar que o genoma de anuros possui uma característica genômica diferente de outros grupos de animais devido à grande quantidade e variedade de elementos transponíveis presentes (Hellsten et al., 2010; Sun et al., 2015). A abundância de TEs podem estar diretamente correlacionados com o aumento do tamanho genômico dada propensão dos TEs de se replicarem (Dufresne e Jeffery, 2011; Canapa et al., 2015; Chalopin et al., 2015).

Como verificado no trabalho de Chalopin et al (2015) há algumas superfamílias que estão presentes em todos os vertebrados e que também foram constatados em nossos resultados como as superfamílias ERVs, Penelope, Line/L1, LINE/CR1 e hAT. Dentre os anuros, DNA transposon Maverick foi identificado somente no genoma de *O. centralis*, já Harbinger foi observado em *O. centralis*, *X. tropicales* e *N. parkeri* (Tabela 3) essas duas superfamílias são ausentes em aves e mamíferos, assim como o LTR copia que foi identificado nos anuros já caracterizados, o que demonstra que há variabilidade quanto a distribuição de TEs em vertebrados no geral.

**Tabela 2.** Anotação de elementos transponíveis (TEs) identificados nas sequências de *Ololygon centralis*.

|                            | Ordem            | Superfamília/<br>Família | Número de<br>elementos | Porcentagem<br>de elementos<br>(%) | Sequências<br>ocupadas<br>(pb) | Percentual<br>de<br>sequências<br>no genoma<br>(%) |
|----------------------------|------------------|--------------------------|------------------------|------------------------------------|--------------------------------|----------------------------------------------------|
| <b>Retroelementos</b>      | LTR<br>elementos | Copia                    | 8                      | 1,1                                | 6.122                          | 0,12                                               |
|                            |                  | Gypsy                    | 229                    | 30,5                               | 83.070                         | 1,69                                               |
|                            |                  | DIRS                     | 400                    | 53,4                               | 109.907                        | 2,23                                               |
|                            |                  | ERV1                     | 25                     | 3,2                                | 7.671                          | 0,15                                               |
|                            |                  | Pao                      | 44                     | 5,8                                | 19.755                         | 0,40                                               |
|                            |                  | Ngaro                    | 45                     | 5,9                                | 10.796                         | 0,22                                               |
|                            |                  | ERVL                     | 1                      | 0,1                                | 53                             | 0,00                                               |
|                            | LINE             | LINE/CR1                 | 38                     | 4,8                                | 15.508                         | 0,32                                               |
|                            |                  | LINE/L1                  | 543                    | 70,7                               | 146.358                        | 2,98                                               |
|                            |                  | LINE/L2                  | 44                     | 5,6                                | 11.514                         | 0,23                                               |
|                            |                  | Penelope                 | 134                    | 17,4                               | 17.798                         | 0,36                                               |
|                            |                  | Rex-Babar                | 11                     | 1,4                                | 2.579                          | 0,05                                               |
|                            | SINE             | tRNA-V-Core-L2           | 25                     | 100                                | 2.451                          | 0,050                                              |
| <b>DNA<br/>transposons</b> | TIR              | PIF-Harbinger            | 9                      | 3,70                               | 2.361                          | 0,05                                               |
|                            |                  | hAT-Ac                   | 7                      | 2,88                               | 1.255                          | 0,03                                               |
|                            |                  | hAT-Charlie              | 26                     | 10,29                              | 2.260                          | 0,05                                               |
|                            |                  | hAT-Tip100               | 9                      | 3,70                               | 458                            | 0,01                                               |
|                            |                  | DNA/Kolobok-T2           | 10                     | 4,12                               | 1.561                          | 0,03                                               |
|                            |                  | Maverick                 | 14                     | 5,76                               | 2.707                          | 0,05                                               |
|                            |                  | PiggyBac                 | 17                     | 7,00                               | 5.007                          | 0,10                                               |
|                            |                  | Mariner-Tc1              | 86                     | 34,98                              | 22.193                         | 0,45                                               |
|                            |                  | Marimer-Tc2              | 43                     | 17,28                              | 4.885                          | 0,10                                               |
|                            |                  | Não classificados        | 25                     | 10,29                              | 2.188                          | 0,04                                               |
|                            |                  |                          |                        |                                    |                                |                                                    |
|                            |                  |                          |                        |                                    |                                |                                                    |
| <b>Não classificados</b>   |                  |                          | 2                      | 100                                | 167                            | 0,003                                              |
| <b>Total</b>               |                  |                          | 1.795                  | 100                                | 478.424                        | 9,72                                               |

**Tabela 3.** Presença e ausência de elementos transponíveis no genoma de *Ololygon centralis*, *Xenopus tropicales* e *Nanorana parkeri*.

|                       | Ordem | Superfamília | <i>Ololygon centralis</i> | <i>*Xenopus tropicales</i> | <i>**Nanorana parkeri</i> |
|-----------------------|-------|--------------|---------------------------|----------------------------|---------------------------|
| <b>Retroelementos</b> | LTR   | Copia        | +                         | +                          | +                         |
|                       |       | Gypsy        | +                         | +                          | +                         |
|                       |       | DIRS         | +                         | +                          | +                         |
|                       |       | ERV1         | +                         | +                          | +                         |
|                       |       | Pao          | +                         | +                          | +                         |
|                       |       | Ngaro        | +                         | +                          | +                         |
|                       |       | BEL          | -                         | +                          | -                         |
|                       | LINE  | LINE/CR1     | +                         | +                          | +                         |
|                       |       | LINE/L1      | +                         | +                          | +                         |
|                       |       | LINE/L2      | +                         | +                          | +                         |
|                       |       | Penelope     | +                         | +                          | +                         |
|                       |       | REX          | +                         | +                          | -                         |
|                       | SINE  |              | +                         | +                          | +                         |
| <b>DNA transposon</b> | TIR   | Harbinger    | +                         | +                          | +                         |
|                       |       | hAT          | +                         | +                          | +                         |
|                       |       | Kolobok      | +                         | +                          | +                         |
|                       |       | Maverick     | +                         | -                          | -                         |
|                       |       | PiggyBac     | +                         | +                          | -                         |
|                       |       | Mariner      | +                         | +                          | +                         |
|                       |       | Merlin       | -                         | +                          | -                         |
|                       |       | Helintrons   | -                         | +                          | -                         |
|                       |       | Polintrons   | -                         | +                          | -                         |

\*Hellsten et al., 2010; \*\*Sun et al., 2015. Presença (+) e ausência (-).

O elemento transponível menos frequente no genoma, com somente 8 cópias, foi o Copia. Este elemento teve média de tamanho de 765,25 pb e tamanho máximo de 3893 pb. No entanto, apesar de menos frequente, foi o elemento com maior tamanho entre todos os encontrados. Já o elemento LINE/L1, apresentou número de cópias elevado (543), com média de tamanho de 269,53 pb e tamanho mínimo de 12 pb (Tabela 4).

Os outros diferentes grupos de elementos repetitivos que foram identificados pelo *RepeatMasker* no genoma de *O. centralis* totalizaram 67.762 pb, compreendendo os microssatélites, elementos de baixa complexidade, pequenos RNAs nucleares e DNA satélite (Tabela 5).

Os microssatélites representaram o maior número de sequências repetitivas encontradas (957 elementos) correspondentes a 51.211 pares de bases, das quais 12,85% (123) foram de mononucleotídeos, 12,02% (115) dinucleotídeos, 17,03% (163) trinucleotídeos, 26,23% (251) tetranucleotídeos 13,58% (130) pentanucleotídeos e 18,29% (175) hexanucleotídeos (Figura 3).

**Tabela 4.** Frequência e tamanho (pb) dos fragmentos de TEs encontrados nas sequências de *Ololygon centralis*.

| Elemento transponível | Tamanho Mínimo (pb) | Tamanho máximo (pb) | Tamanho médio (pb) | Número de elementos |
|-----------------------|---------------------|---------------------|--------------------|---------------------|
| Copia                 | 46                  | 3.893               | 765,25             | 8                   |
| Gypsy                 | 25                  | 3.246               | 362,75             | 229                 |
| DIRS                  | 26                  | 1.654               | 274,77             | 400                 |
| ERV1                  | 80                  | 800                 | 306,84             | 25                  |
| Pao                   | 43                  | 1.166               | 448,98             | 44                  |
| Ngaro                 | 51                  | 502                 | 239,91             | 45                  |
| ERV1                  | 53                  | 53                  | 53                 | 1                   |
| LINE/CR1              | 45                  | 3.835               | 408,11             | 38                  |
| LINE/L1               | 12                  | 3.588               | 269,53             | 543                 |
| LINE/L2               | 31                  | 2.048               | 261,68             | 44                  |
| Penelope              | 29                  | 392                 | 132,82             | 134                 |
| Rex-Babar             | 86                  | 483                 | 234,45             | 11                  |
| tRNA-V-Core-L2        | 38                  | 249                 | 98,04              | 25                  |
| PIF-Harbinger         | 51                  | 708                 | 262,33             | 9                   |
| hAT-Ac                | 17                  | 362                 | 179,28             | 7                   |
| hAT-Charlie           | 33                  | 213                 | 86,92              | 26                  |
| hAT-Tip100            | 44                  | 64                  | 50,88              | 9                   |
| DNA/Kolobok-T2        | 42                  | 771                 | 156,1              | 10                  |
| Maverick              | 88                  | 381                 | 193,36             | 14                  |
| PiggyBac              | 42                  | 774                 | 294,53             | 17                  |
| TcMar-Tc1             | 21                  | 1.651               | 258,06             | 86                  |
| TcMar-Tc2             | 33                  | 604                 | 113,60             | 43                  |

**Tabela 5.** Grupos de elementos repetitivos anotados nas sequências de *O. centralis* usando *RepeatMasker*.

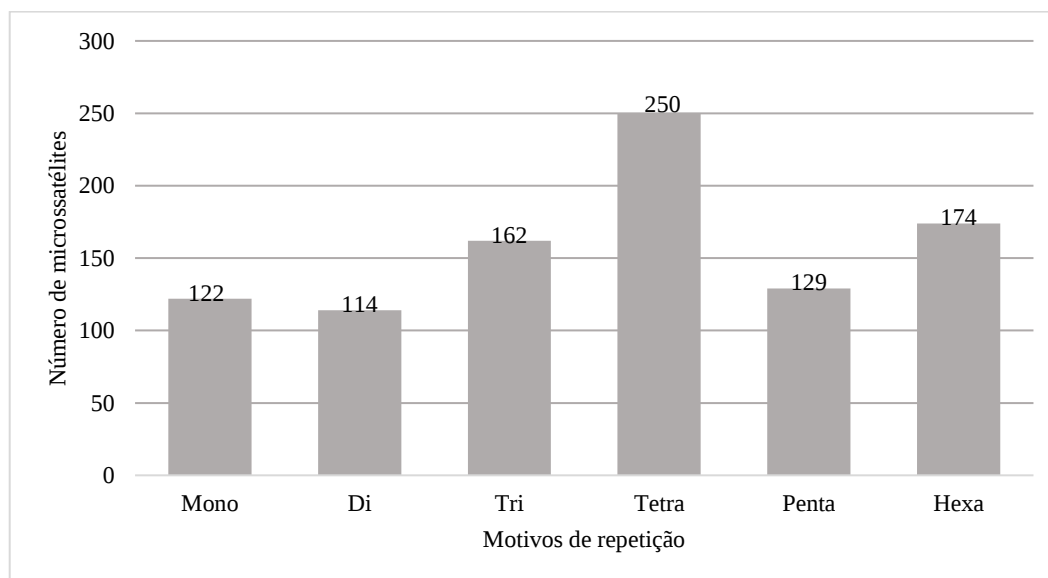
| Tipos de repetição                  | Números de elementos | Sequências ocupadas (pb) | Percentual no genoma (%) |
|-------------------------------------|----------------------|--------------------------|--------------------------|
| Microsatélites                      | 957                  | 51.211                   | 1,040                    |
| Baixa complexidade                  | 254                  | 14.849                   | 0,302                    |
| snRNAs ( <i>Small Nuclear RNA</i> ) | 25                   | 1.565                    | 0,031                    |
| DNA satélite                        | 1                    | 137                      | 0,003                    |
| Total                               | 1.237                | 67.762                   | 1,376                    |

Observando todas as matrizes existentes, os motivos de repetição de regiões microsatélites mais frequentes foram: T (61,79%) e A (36,59%) para mononucleotídeos, AT (54,78%) e AC (16,52%) para dinucleotídeos, AAT (24,54%) e ATT (14,72%) para trinucleotídeos, AGAT (21,51%) e ATTT (15,14%) para tetranucleotídeos, ATTTT (14,62%) e CTTTT (6,15%) para pentanucleotídeos e ATTTTT (6,86%) e ACACAT (4%) para hexanucleotídeos (Tabela 6).

Em vertebrados o dinucleotídeo AC é o motivo de repetição mais encontrado (Tóth et al., 2000), diferentemente do que foi observado em *O. centralis*. O conteúdo rico em AT desse anuro, é uma característica bem descrita para a maioria dos genomas de plantas, e que foi encontrado intimamente relacionados com elementos transponíveis no genoma de arroz (Temnykh et al., 2001; Ellegren, 2004). Todavia, a fração do genoma de *O. centralis* é pequena e pode apresentar-se com proporções diferentes quando uma maior quantidade do genoma for sequenciada.

Os tetranucleotídeos foram os mais frequentes no genoma, representando 31,40% de pb do total dos microssatélites. Esse motivo é encontrado predominantemente em regiões não codificantes assim como a maioria dos motivos, exceto tri- e hexa- que são identificados preferencialmente em regiões codantes (Morgante et al., 2002; Vieira et al., 2016).

Embora exista irregularidade em relação à quantidade das repetições e os motivos de microssatélites identificados no genoma é possível perceber que motivos mais longos como penta e hexa tem a tendência de ter mais diversidade de combinações e menor frequência, ou seja, menos repetições em tandem (Tabela 6 e Material suplementar, Tabela S1) (Zalapa et al., 2012; Li et al., 2017).



**Figura 3.** Motivos de repetição e número de microssatélites no genoma de *O. centralis*.

**Tabela 6.** Tipos de motivos de repetição e frequência identificadas nas sequências de *Ololygon centralis*.

| Matriz           | Motivo | Frequência absoluta | Frequência relativa (%) |
|------------------|--------|---------------------|-------------------------|
| Mononucleotídeo  | T      | 76                  | 61,79                   |
|                  | A      | 45                  | 36,59                   |
|                  | C      | 2                   | 1,63                    |
| Dinucleotídeo    | AT     | 63                  | 54,78                   |
|                  | AC     | 19                  | 16,52                   |
|                  | GT     | 10                  | 8,70                    |
|                  | Outros | 23                  | 20,01                   |
| Trinucleotídeo   | AAT    | 40                  | 24,54                   |
|                  | ATT    | 24                  | 14,72                   |
|                  | CTT    | 20                  | 12,27                   |
|                  | ATG    | 13                  | 7,98                    |
|                  | CCT    | 13                  | 7,98                    |
|                  | ATC    | 10                  | 6,13                    |
|                  | Outros | 43                  | 26,39                   |
| Tetranucleotídeo | AGAT   | 54                  | 21,51                   |
|                  | ATTT   | 38                  | 15,14                   |
|                  | ATCT   | 37                  | 14,74                   |
|                  | AAAT   | 19                  | 7,57                    |
|                  | GTTT   | 11                  | 4,38                    |
|                  | Outros | 92                  | 36,7                    |
| Pentanucleotídeo | ATTTT  | 19                  | 14,62                   |
|                  | CTTTT  | 8                   | 6,15                    |
|                  | AATAT  | 7                   | 5,38                    |
|                  | Outros | 96                  | 73,92                   |
| Hexanucleotídeo  | ATTTTT | 12                  | 6,86                    |
|                  | ACACAT | 7                   | 4,00                    |
|                  | ATATTT | 6                   | 3,43                    |
|                  | Outros | 150                 | 85,55                   |

### Anotação de genes de tRNA e rRNA

Foram preditas 48 sequências de genes que codificam RNAs transportadores distribuídos em 32 contigs, ocupando 3.553 pb das sequências genômicas de *O. centralis*. As sequências dos genes para tRNA tiveram um tamanho médio de 74,02 pb e dentre os 18 tipos de aminoácidos transportados, o mais frequente foi Serina (Ser) representando 12,5% (Tabela 7). No genoma mitocondrial do Hylidae *Hyla tsinlingensis* também foram anotados 18 tRNAs idênticos aos que que identificamos para *O. centralis* (Huang et al., 2016).

Na busca por RNA ribossomal foram identificados genes que codificam os RNAs que formam as subunidades 5S (Bits score: 92.3) com sequências de 114 pb, 18S (Bits score: 1698.1) com 1.867 pb e 28S (Bits score: 1254.5) com 1.765 pb, ocupando 3.746 pb do genoma. Não foi possível identificar DNA ribossomais que codificam as subunidades 12S e 16S que foram caracterizadas no genoma mitocondrial para oito espécies da ordem Anura (Ye et al., 2014; Huang et al., 2016; Machado et al., 2016; Nam et al., 2017).

**Tabela 7.** Anotação de genes codificadores de RNA transportador identificados nas sequências de *Ololygon centralis*.

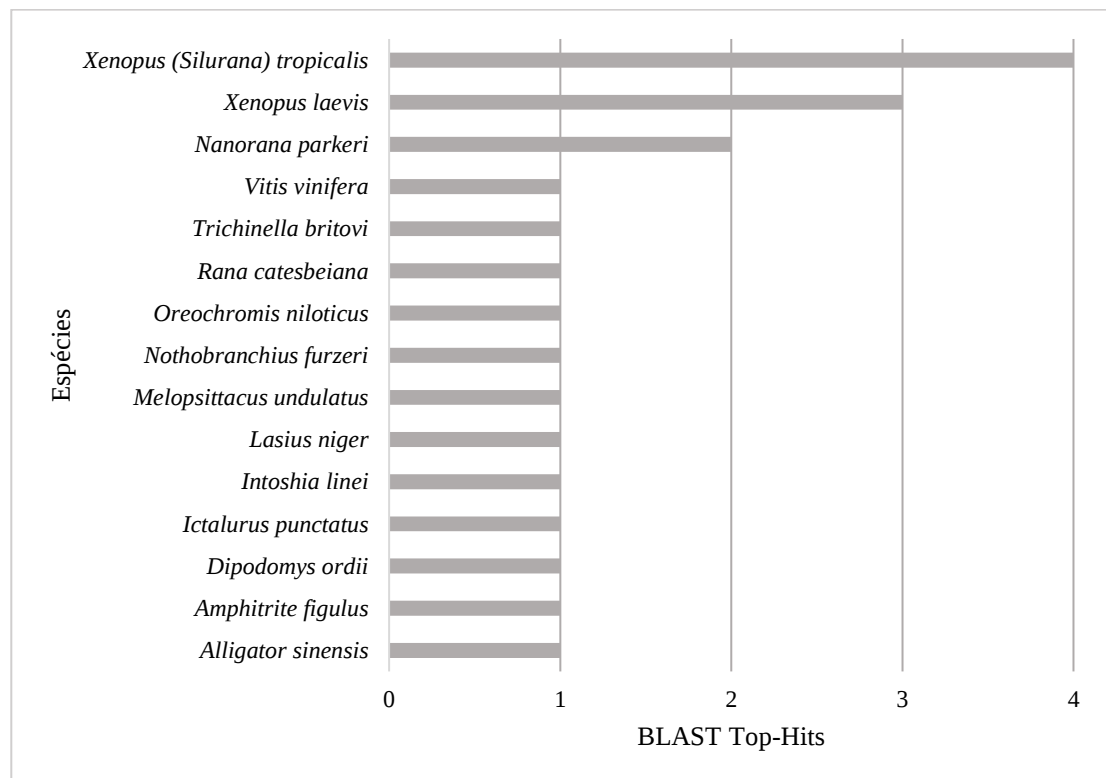
|                          | Frequência absoluta | Tamanho total (pb) | Tipo de aminoácido | Códon           |
|--------------------------|---------------------|--------------------|--------------------|-----------------|
| <b>RNA transportador</b> | 6                   | 457                | Serina             | TGA/GCT/AGA/CGA |
|                          | 4                   | 287                | Alanina            | TGC/CGC/AGC     |
|                          | 4                   | 286                | Prolina            | AGG/CGG         |
|                          | 4                   | 290                | Treonina           | AGT/CGT/TGT     |
|                          | 4                   | 302                | Arginina           | CCT/TCG/CCG     |
|                          | 3                   | 207                | Triptofano         | CCA             |
|                          | 3                   | 239                | Tirosina           | GTA             |
|                          | 3                   | 213                | Glutamato          | TTC/CTC         |
|                          | 3                   | 238                | Isoleucina         | TAT/AAT         |
|                          | 3                   | 244                | Leucina            | AAG/CAG/TAG     |
|                          | 2                   | 142                | Glutamina          | CTG/TTG         |
|                          | 2                   | 145                | Asparagina         | GTT             |
|                          | 2                   | 143                | Metionina          | CAT             |
|                          | 1                   | 71                 | Aspartato          | GTC             |
|                          | 1                   | 72                 | Lisina             | TTT             |
|                          | 1                   | 74                 | Fenilalanina       | GAA             |
|                          | 1                   | 71                 | Glicina            | TCC             |
|                          | 1                   | 72                 | Valina             | TAC             |
| Total                    | 48                  | 3.553              | -                  | -               |

### Anotação de regiões codificantes de proteínas

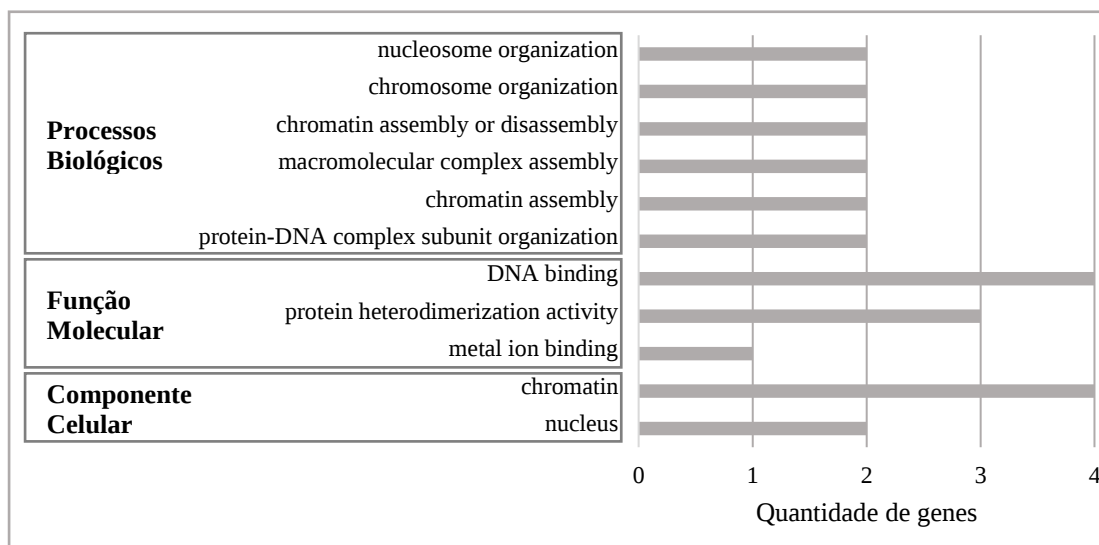
Dos 13.989 *scaffolds* montados para *O. centralis* foram anotados 27 sequências codificadoras de mRNA presentes em 31 *scaffolds*, apresentando um total de 6.851 pb, a menor sequência com 57 pb e o maior 1.034 pb. Nas 27 sequências de regiões codificantes foram encontradas 21 com similaridade em sequências conhecidas pelo Blast2go, a espécie *Xenopus (silurana) tropicalis* foi a mais frequente entre as espécies listadas no BLAST Top-Hits, seguida de *Xenopus laevis* e *Nanorana parkeri* (Rinidae) (Figura 4). Esse resultado era esperado visto que a espécie *Xenopus (silurana) tropicalis* foi usada como referência para a predição das sequências dos transcritos.

Empregando a ferramenta de classificação de ontologia de genes (GO) foi possível atribuir funcionalmente 26 sequências gênicas: 12 genes foram classificados para processos biológicos (PB), 8 genes de função molecular (FM) e 6 na categoria de componentes celulares (CC) (Figura 5). Com base no transcriptoma de três espécies de anuros a maioria dos genes anotados para PB estavam envolvidos em processos metabólicos e celulares, na categoria FM os genes foram classificados com atividade de ligação e para CC estavam incluídos na parte celular (Yang et al., 2012; Qiao et al., 2013).

A espécie *Xenopus tropicalis* tem cerca de 20.000 genes codificadores de proteínas e possui um genoma de 1.7Gpb (Hellsten et al., 2010), já uma outra espécie de anuro *Odorrana margaretae* tem um total de 54,3Mpb de transcritos (Qiao et al., 2013). Ao comparar a quantidade de genes e sequências montadas para *O. centralis* com outros anuros podemos perceber que a nossa montagem foi parcial e que a maior parte das nossas sequências correspondem a regiões não codificantes, já que foi encontrado uma baixa densidade de sequências codificadoras.



**Figura 4.** Top-Hits do Blast: espécies com similaridade de sequências codificadoras de mRNA.



**Figura 5.** Número de sequências codificadoras de genes anotadas segundo a classificação ontologia de genes.

## CONCLUSÃO

A partir do uso do sequenciamento de nova geração com baixa cobertura foi possível proporcionar a produção de dados genômicas inéditos para a espécie *Ololygon centralis*, possibilitando conhecer uma parte do seu conteúdo genômico, contribuindo para acumulação de informações e compreensão acerca da estrutura genômica dos anfíbios, um grupo que ainda é pouco representado no âmbito do conhecimento genômico. Os resultados encontrados sustentam a informação de que os anfíbios possuem uma grande diversidade de elementos transponíveis, a qual podem estar relacionadas com o seu amplo tamanho genômico. A identificação de muitas regiões microssatélites no genoma de *O. centralis* contribuirá para estudos futuros acerca da diversidade genética da espécie pois, são primordiais para o desenvolvimento de marcadores moleculares.

## REFERÊNCIAS

ANDREWS, S. et al. **FastQC**: a quality control tool for high throughput sequence data. 2010. Disponível em: < <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>>. Acesso em: 19 julho 2016.

BASTOS, R.; COLLI, G.; SILVANO, D. *Scinax centralis*. The IUCN Red List of Threatened Species. 2004. Disponível em: <<http://www.iucnredlist.org/details/55945/0>>. Acesso em: 10 jan 2017.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, 2014.

BRANDÃO, R. A.; ARAÚJO, A. F. B. A herpetofauna associada às matas de galeria no Distrito Federal. Cerrado: Caracterização e Recuperação de Matas de Galeria. J. F. Ribeiro, C. E. L. Fonseca, e J. C. Sousa-Silva (eds.). **Empresa Brasileira de Pesquisa Agropecuária-Embrapa cerrados & Ministério do Meio Ambiente**, Planaltina, DF, p. 559-604, 2001.

BUISINE, N. et al. *Xenopus tropicalis* genome re-scaffolding and re-annotation reach the resolution required for in vivo ChIA-PET analysis. **PloS one**, v. 10, n. 9, p. e0137526, 2015.

CANAPA, A. et al. Transposons, genome size, and evolutionary insights in animals. **Cytogenetic and genome research**, v. 147, n. 4, p. 217-239, 2015.

CHAISSON, M. J. P.; WILSON, R. K.; EICHLER, E. E. Genetic variation and the de novo assembly of human genomes. **Nature Reviews Genetics**, v. 16, n. 11, p. 627, 2015.

CHALOPIN, D. et al. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. **Genome biology and evolution**, v. 7, n. 2, p. 567-580, 2015.

COMPEAU, P. E. C.; PEVZNER, P. A.; TESLER, G. How to apply de Bruijn graphs to genome assembly. **Nature biotechnology**, v. 29, n. 11, p. 987, 2011.

CONESA, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, v. 21, n. 18, p. 3674-3676, 2005.

DEL FABBRO, C. et al. An extensive evaluation of read trimming effects on Illumina NGS data analysis. **PloS one**, v. 8, n. 12, p. e5024, 2013.

DUFRESNE, F.; JEFFERY, N. A guided tour of large genome size in animals: what we know and where we are heading. **Chromosome research**, v. 19, n. 7, p. 925-938, 2011.

EKBLOM, R.; WOLF, J. B. W. A field guide to whole-genome sequencing, assembly and annotation, **Evolutionary Application**, v. 7, n. 9, p 1026-1042, 2014.

ELLEGREN, H. Microsatellites: simple sequences with complex evolution. **Nature reviews genetics**, v. 5, n. 6, p. 435, 2004.

GARDNER, P. P. et al. Rfam: updates to the RNA families database. **Nucleic acids research**, v. 37, n. suppl\_1, p. D136-D140, 2008.

GREGORY, T. R. Genome size and developmental complexity. **Genetica**, v. 115, n. 1, p. 131-146, 2002.

HAMMOND, S. A. et al. The North American bullfrog draft genome provides insight into hormonal regulation of long noncoding RNA. **Nature Communications**, v. 8, n. 1, p. 1433, 2017.

HELLSTEN, U. et al. The genome of the Western clawed frog *Xenopus tropicalis*. **Science**, v. 328, n. 5978, p. 633-636, 2010.

HUANG, M. et al. The complete mitochondrial genome of *Hyla tsinlingensis* (Anura: Hylidae). **Mitochondrial DNA Part A**, v. 27, n. 6, p. 4130-4131, 2016.

LI, R. et al. The sequence and de novo assembly of the giant panda genome. **Nature**, v. 463, n. 7279, p. 311, 2010.

LI, Z. et al. Genome-wide mapping and characterization of microsatellites in the swamp eel genome. **Scientific Reports**, v. 7, n. 1, p. 3157, 2017.

LOWE, T. M.; CHAN, P. P. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. **Nucleic acids research**, v. 44, n. W1, p. W54-W57, 2016.

MACHADO, D. J.; LYRA, M. L.; GRANT, T. Mitogenome assembly from genomic multiplex libraries: comparison of strategies and novel mitogenomes for five species of frogs. **Molecular ecology resources**, v. 16, n. 3, p. 686-693, 2016.

METZKER, M. L. Sequencing technologies--the next generation. **Nature reviews. Genetics**, v. 11, n. 1, p. 31, 2010.

MMA (Ministério do Meio Ambiente). **O bioma Cerrado**. Disponível em: <<http://www.mma.gov.br/biomas/cerrado>>. Acesso em: 5 jan 2018.

MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. **Nature genetics**, v. 30, n. 2, p. 194, 2002.

MOURA, M. R. et al. Amphibia, Anura, Hylidae, *Scinax centralis* Pombal and Bastos, 1996: distribution extension, geographic distribution map. **Check List**, v. 6, n. 1, p. 173-175, 2010.

MYERS, N. et al. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, n. 6772, p. 853, 2000.

NAM, D. H. et al. Complete mitochondrial genome of a treefrog, *Hyla* sp. (Anura: Hylidae). **Mitochondrial DNA Part B**, v. 2, n. 1, p. 221-222, 2017.

POMBAL JÚNIOR, J. P.; BASTOS, R. P. Nova espécie de *Scinax* Wagler, 1830 do Brasil Central (Amphibia, Anura, Hylidae). 1996.

QIAO, L. et al. Transcriptome profile of the green odorous frog (*Odorrana margaretae*). **PloS one**, v. 8, n. 9, p. e75211, 2013.

SAFONOVA, Y.; BANKEVICH, A.; PEVZNER, P. A. dipSPAdes: assembler for highly polymorphic diploid genomes. **Journal of Computational Biology**, v. 22, n. 6, p. 528-545, 2015.

SCHATZ, M. C.; DELCHER, A. L.; SALZBERG, S. L. Assembly of large genomes using second-generation sequencing. **Genome research**, v. 20, n. 9, p. 1165-1173, 2010.

SESSION, A. M. et al. Genome evolution in the allotetraploid frog *Xenopus laevis*. **Nature**, v. 538, n. 7625, p. 336, 2016.

SHENDURE, J.; JI, H. Next-generation DNA sequencing. **Nature biotechnology**, v. 26, n. 10, p. 1135, 2008.

STEIN, L. Genome annotation: from sequence to biology. **Nature reviews genetics**, v. 2, n. 7, p. 493, 2001.

SUN, Y. B. et al. Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. **Proceedings of the National Academy of Sciences**, v. 112, n. 11, p. E1257-E1262, 2015.

TARAILO-GRAOVAC, M.; CHEN, N. Using RepeatMasker to identify repetitive elements in genomic sequences. **Current protocols in bioinformatics**, p. 4.10. 1-4.10. 14, 2009.

TEMNYKH, S. et al. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. **Genome research**, v. 11, n. 8, p. 1441-1452, 2001.

TÓTH, G.; GÁSPÁRI, Z.; JURKA, J. Microsatellites in different eukaryotic genomes: survey and analysis. **Genome research**, v. 10, n. 7, p. 967-981, 2000.

VIEIRA, M. L. C. et al. Microsatellite markers: what they mean and why they are so useful. **Genetics and molecular biology**, v. 39, n. 3, p. 312-328, 2016.

YANDELL, M.; ENCE, D. A beginner's guide to eukaryotic genome annotation. **Nature Reviews Genetics**, v. 13, n. 5, p. 329, 2012.

YANG, W. et al. Toward understanding the genetic basis of adaptation to high-elevation life in poikilothermic species: a comparative transcriptomic analysis of two ranid frogs, *Rana chensinensis* and *R. kukunoris*. **BMC genomics**, v. 13, n. 1, p. 588, 2012.

YE, L. et al. The complete mitochondrial genome of *Hyla annectans* (Anura: Hylidae). **Mitochondrial DNA Part A**, v. 27, n. 3, p. 1593-1594, 2014.

ZALAPA, J. E. et al. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. **American journal of botany**, v. 99, n. 2, p. 193-208, 2012.

## 6. CAPÍTULO 2: DESENVOLVIMENTO DE MARCADORES MICROSSATÉLITES PARA *Ololygon centralis* (Pombal e Bastos, 1996) (ANURA: HYLIDAE)

### RESUMO

A espécie de anuro *Ololygon centralis* (Pombal e Bastos 1996) pertence à família Hylidae e é considerada endêmica do Cerrado. Sua distribuição está associada à Bacia do Rio Paranaíba em matas de galeria, habitando regiões de extrema antropização. Estudo a respeito da diversidade genética dessa espécie pode auxiliar no desenvolvimento de medidas de proteção para esse anuro. Nesse contexto, os marcadores microssatélites são utilizados para aplicações em estudos genéticos populacionais, e o sequenciamento de nova geração (NGS) tem possibilitado o acesso a novos marcadores com maior eficiência. Assim, o presente estudo teve como objetivo o desenvolvimento de marcadores microssatélites para *Ololygon centralis*. As sequências de boa qualidade obtidas por meio de NGS plataforma Miseq Illumina, foram utilizadas no software QDD para a identificação de regiões microssatélites e desenho de *primers*. Os *primers* foram submetidos a diferentes condições de termociclagem para amplificação das regiões microssatélites por PCR em 4 indivíduos de *O. centralis*. Dezoito *primers* foram padronizados em gel de poliacrilamida 6%. As temperaturas de anelamento variaram de 50° a 60° C. Sete *primers* apresentaram polimorfismo, sendo que cinco (OCE05, OCE11, OCE20, OCE21, OCE26) foram selecionados para caracterização em 30 indivíduos de *O. centralis* em analisador automático ABI3500. O número de alelos variou de 5 a 17, com uma média de 10 alelos por loco, a amplitude alélica dos locos variou de 148 pb (OCE 20) a 318 pb (OCE 21). A heterozigosidade média esperada entre os cinco locos foi de 0,7 variando de 0,41 a 0,98 e a observada 0,5 variando de 0 a 1. O baixo valor da probabilidade combinada de identidade ( $I=1,13 \times 10^{-6}$ ) e o alto valor de exclusão de paternidade ( $Q = 0,993$ ) demonstraram o bom potencial para utilização desses marcadores microssatélites para discriminação de indivíduos em estudos genético-populacionais e de parentesco, o coeficiente de endogamia foi alto e significativo para os locos OCE 05 e OCE 21. O sequenciamento de baixa cobertura foi eficiente identificação de regiões microssatélites e para o desenho de *primers* específicos para a espécie *O. centralis*. O desenvolvimento desses marcadores contribuirá para estudos de diversidade genética desse anuro e provavelmente contribuirá para estudos com espécies correlacionadas através da transferibilidade de *primers*.

Palavras-chave: Hylidae, Illumina, marcador molecular, microssatélite.

## INTRODUÇÃO

Sequências repetitivas compõem grande parte do genoma eucarioto e dentre elas estão os microssatélites (Biscotti et al., 2015). Essas regiões correspondem a sequências repetidas em tandem e têm sido utilizadas como marcadores moleculares em diversos campos de estudos genéticos por serem multialélicos, neutros, apresentarem co-dominância, e alta taxa mutacional. Essas características os tornam marcadores de grande poder discriminativo e informativo (Ellegren, 2004; Oliveira et al., 2006; Vieira et al., 2016).

Contudo a aplicação dessa ferramenta exige conhecimento prévio das sequências que flanqueiam as regiões microssatélites para que seja viável o desenvolvimento dos iniciadores (*primers*) para amplificação. A técnica de Sequenciamento de Nova Geração (NGS) tem impedido essa limitação pois é um método que tem capacidade de gerar milhares e até milhões de sequências de DNA de maneira eficaz e barata quando comparada a outras técnicas. A detecção de numerosas regiões microssatélites em plataformas de baixa/média cobertura é também interessante pois permite se obter resultados para desenho de *primers* flanqueadores das regiões repetitivas (Mardis, 2008; Zalapa et al., 2012; Zanella et al., 2017).

O desenho de *primers* é uma etapa crítica do processo de desenvolvimento pois refletirá na qualidade das análises dos dados posteriormente, logo ferramentas de bioinformática que auxiliem nesse processo são indispensáveis. O software QDD por exemplo, pesquisa por sequências de microssatélites e desenha *primers* levando em consideração os fatores que poderão atrapalhar na amplificação (Megléczy et al., 2009), com a finalidade de gerar *primers* que amplifiquem apenas um único loco (Zanella et al., 2017).

Com isso, o uso de marcadores moleculares como o microssatélite (SSR) permite acessar e conhecer a variabilidade genética das espécies, visando compreender a estrutura e viabilidade de populações de diversos organismos (Jehle e Arntzen, 2002; Gardner et al., 2011). Em anfíbios, por exemplo, Luquet et al. (2011) avaliou a variabilidade genética de 15 locos microssatélite e correlacionou com o *fitness* das populações de *Hyla arborea*. Também em anuros, Arruda et al. (2010) avaliou a diversidade genética de *Rhinella schneideri* com base em 9 locos microssatélites.

Esses estudos demonstram a aplicabilidade dos SSRs com enfoques ecológicos e conservacionistas, principalmente quando se referem aos anfíbios, vertebrados que possuem cerca de 41% das espécies em ameaça de extinção (Hoffmann et al., 2010). Esses animais apresentam baixa dispersão e são altamente vulneráveis a mudanças climáticas e fragmentação de habitat, fatores que podem levar a extinção diversas espécies se ações de

proteção não forem praticadas (Haddad e Prado, 2005; Katzenberger et al., 2012; Vancine, 2015).

Nesse contexto o objetivo desse trabalho foi desenvolver um conjunto de marcadores microssatélites polimórficos para a espécie *Ololygon centralis*, um anuro endêmico do Bioma Cerrado que tem como habitat matas de galeria em regiões ambientalmente fragmentadas (Pombal e Bastos, 1996). Essa espécie é classificada pela IUCN como espécie pouco preocupante em relação a conservação, apesar de que existe uma lacuna de conhecimento a respeito da estrutura genética dessa espécie. Logo, o desenvolvimento de marcadores microssatélites espécie-específico contribuirá para a exploração da diversidade genética de populações naturais de *Ololygon centralis*.

## MATERIAL E MÉTODOS

### Coleta e extração de DNA

O DNA foi extraído a partir de indivíduos de *Ololygon centralis* coletados nas cidades de Silvânia (Floresta Nacional), Pires do Rio e Caldas Novas localizadas no estado de Goiás (Tabela 1 e Material Suplementar, Tabela S2). As coletas e procedimentos foram realizados com autorização SISBIO/ICMBIO 46522-3 e do Comitê de Ética em Uso de Animais da Universidade Federal de Goiás número 109/14 (Anexos).

**Tabela 1.** Número de indivíduos amostrados da espécie *Ololygon centralis*.

| População                            | Código  | Número | Localidade      | Georeferência                  |
|--------------------------------------|---------|--------|-----------------|--------------------------------|
| <b>Floresta Nacional de Silvânia</b> | OCE-FLO | 16     | Silvânia-GO     | 16°38'04.4"S,<br>48°39'31.2"W  |
| <b>Caldas Novas</b>                  | OCE-CNZ | 7      | Caldas Novas-GO | 17°49'41.74"S,<br>48°39'3.60"W |
| <b>Pires do Rio</b>                  | OCE-PIR | 7      | Pires do Rio-GO | 17°25'18.76"S,<br>48°24'6.91"W |

Todos os procedimentos foram realizados no Laboratório de Genética & Biodiversidade (LGBio) no Instituto de Ciências Biológicas I da Universidade Federal de Goiás. Alíquotas de tecidos como fígado, coração e músculo anteriormente armazenados foram utilizados para a extração do DNA genômico, conforme descrito no capítulo 1.

Para avaliar a qualidade e quantificar o DNA extraído, as amostras foram aplicadas em gel de Agarose 1% corado com Brometo de Etídio (10mg/ml) e submetido a eletroforese horizontal. A quantificação do DNA foi realizada visualmente por meio da comparação das amostras com os marcadores de massa molecular conhecidos de DNA  $\lambda$  10, 20, 50, 100 e 200 ng/ $\mu$ L. A partir dos dados de quantificação, os DNAs foram diluídos para a concentração de 2,5 ng/ $\mu$ L em água Milli-Q e armazenados em freezer -20°C.

### Desenho de *primers*

A partir da montagem do genoma parcial (Capítulo 1), foi realizada a identificação das regiões microssatélites no genoma de *Ololygon centralis*. Essa busca foi realizada no software QDD versão 3.1.2 instalado na máquina virtual VirtualBox 5.0.14 para uso em Windows. Microssatélites perfeitos com matriz de repetição de no mínimo dez para dinucleotídeos, quatro repetições para tetra- e pentanucleotídeos foram identificados no software seguindo o passo 1 (PIPE 1, Tabela 2). Em seguida, foi realizado uma análise de similaridade das regiões identificadas para obter uma única amostra dessa região (PIPE2, Tabela 2). O desenho dos *primers* foi executado pelo mesmo software seguindo o passo 3 (PIPE3, Tabela 2). O software QDD também identificou no genoma as sequências de microssatélites que estavam no contexto de elementos transponíveis utilizando a ferramenta RepeatMasker que rastreou os TEs e pequenos RNAs por meio de similaridade com o banco de dados, para esse estudo utilizamos as sequências do genoma de *Xenopus silurana tropicalis* já disponíveis para essa busca (Megléczy et al., 2009, 2014).

**Tabela 2.** Etapas realizadas no QDD para identificação das regiões microssatélites e desenho de *primers*.

|                        |                                                                                                                     |
|------------------------|---------------------------------------------------------------------------------------------------------------------|
| <b>Passo 1 (PIPE1)</b> | Input: Sequências montadas ou não (arquivo fasta/fastq)<br>Preparação das sequências e detecção dos microssatélites |
| Continua...            |                                                                                                                     |

**Tabela 2.** Continuação

|                        |                                                                                                                                                                                                                                         |
|------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Passo 2 (PIPE2)</b> | Input: _pipe1_for_pipe2.fas<br>Detecção da similaridade entre as sequências microssatélites através da comparação de todos contra todos (BLAST), conservando apenas as sequências únicas (SINGLETONS) para desenho dos <i>primers</i> . |
| <b>Passo 3 (PIPE3)</b> | Input: _pipe2_for_pipe3.fas<br>Desenho dos <i>primers</i> forward e reverse para cada sequência através do software Primer3.                                                                                                            |
| <b>Passo 4 (PIPE4)</b> | Input: _pipe3_primers.tabular<br>Sequências microssatélites são avaliadas pelo software RepeatMasker verificando contaminações e confrontadas com o banco de dados do NCBI para identificação de elementos transponíveis.               |

Para o desenho dos *primers* foram utilizados os seguintes parâmetros: amplificação de fragmento com 150 a 400 pb, variação do tamanho dos *primers* de 20 a 25 pb e tamanho ideal de 22 pb, Temperatura de Melting (T<sub>m</sub>) de 52 a 62 graus Celsius e 30% a 60% de conteúdo GC. Para cada região contendo microssatélite foi selecionado um par de *primers*. Dentre os resultados, foram selecionados 31 pares de *primers* para otimização.

### **Padronização de amplificação**

As reações em Cadeia da Polimerase (PCR) dos pares de *primers* espécie-específicos foram realizadas, primeiramente, com amostras de DNA de quatro indivíduos. A preparação da reação de PCR foi realizada utilizando o seguinte protocolo: 6 µL de DNA (2,5 ng/µL), 1,0 µL de cada *primer forward* e *reverse* (0,9 µM), 1,5 µL de tampão IO 10X (500 mM KCl; 100 mM Tris-HCl pH 8,4; 1% Triton X-100), 1,2 µL de MgCl<sub>2</sub> (50 mM), 1,5 µL de dNTP (2,5 µM), 0,2 µL Taq DNA polymerase Phusion (5 U) e água ultrapura para completar o volume final de 15 µL por reação.

A amplificação foi realizada no termociclador da Applied Biosystems®, e os ciclos ocorreram nas seguintes etapas: (1) Desnaturação inicial a 96°C/5 min.; (2) 35 ciclos que incluiu 96°C/1 min., temperatura de anelamento específica para cada *primer*/1 min., 72°C/1 min.; (3) extensão 72°C/45 min. Os produtos de PCR foram submetidos a eletroforese horizontal em gel de agarose 3% corado com brometo de etídio, onde foi possível a

visualização das bandas amplificadas utilizando transluminador de luz UV. Os tamanhos das bandas foram comparados com marcador molecular de 1Kb plus DNA ladder (Invitrogen).

As amostras amplificadas foram conduzidas à eletroforese vertical com gel desnaturante de poliacrilamida 6%. A coloração do gel de poliacrilamida foi feita com nitrato de prata de acordo com o protocolo proposto por Creste et al. (2001). Depois da secagem, o gel foi colocado sob luz branca para a identificação da qualidade de amplificação e inferir o tamanho dos alelos a partir da comparação com marcadores de peso molecular conhecidos de 10 pb e 50 pb. Os pares de *primers* que não amplificaram ou que estavam com amplificação inespecífica foram submetidos a testes de temperaturas de anelamento utilizando o mesmo protocolo de PCR.

### **Desenvolvimento de multiplex para eletroforese capilar**

*Primers forward*, referentes a dez locos microssatélites, foram selecionados para a marcação com fluoróforos na extremidade 5' (Life Technologies®). Foram utilizadas as seguintes marcas: 6'FAM (azul), NED (amarelo) e VIC (verde) que correspondem a matriz DS33 (*Dye Set*) tendo como *size standard* LIZ (Life Technologies®). Para detecção populacional dos locos, foram realizadas reações de PCR para cada um dos *primers*, com todos os indivíduos das três populações. As reações foram preparadas utilizando o mesmo protocolo da padronização dos *primers*.

Para a montagem da multiplex levou-se em consideração a amplitude alélica dos *primers* assim como a cor da fluorescência para não acarretar a sobreposição dos locos, assim foram distribuídos: *primers* de tamanhos diferentes e de mesma fluorescência ou diferentes fluorescências para mesma amplitude alélica.

Os produtos de PCR foram combinados segundo o sistema de multiplex em uma placa-mãe. Para cada produto de PCR ajustou-se o volume conforme a intensidade da banda visualizada no gel de Agarose. Foi retirado da placa-mãe 1 µL da mistura dos produtos de PCR e dispensado em outra placa, denominada de placa-filha a qual foi adicionado 9 µL de formamida *Hi-Di* + *size standard* na proporção de 8,75 µL: 0,25 µL respectivamente, resultando em 10 µL de reação total.

As amostras foram desnaturadas à temperatura de 95°C por cinco minutos seguido de um choque térmico por um minuto. Após os procedimentos a placa foi inserida no analisador automático de DNA, modelo ABI 3500 (Applied Biosystems®) para a realização da eletroforese capilar e posterior genotipagem. Através da fluorescência emitida que é

reconhecida pelo Genescan o tamanho dos alelos foram estimados e os dados importados para o software GeneMapper v. 5.0 (Applied Biosystems®).

### **Caracterização da variabilidade genética**

Os locos microssatélites foram caracterizados com o auxílio do software FSTAT 2.9.3.2 (Goudet, 2002) no qual estimou as frequências alélicas e genotípicas, o número médio de alelos por loco ( $A$ ); heterozigosidade esperada ( $H_e$ ), heterozigosidade observada ( $H_o$ ), índice de fixação ( $F_{ST}$ ) e significância do coeficiente de endogamia ( $F_{is}$ ). A partir das análises no GDA (*Genetic Data Analysis*) também foi calculado o  $H_o$  para cada população (Lewis e Zaykin, 2002). A probabilidade de exclusão de paternidade ( $Q$ ) e a probabilidade de identidade genética ( $I$ ) foram calculadas com o IDENTITY 1.0 (Wagner e Sefc, 1999). Para a verificação de “erros de genotipagem” como alelos nulo, *stutter* e *drop-out* utilizamos o programa Micro-checker (Van Oosterhout et al., 2004).

## **RESULTADOS E DISCUSSÃO**

### **Desenho de *primers***

Foram identificadas 256 regiões microssatélites no genoma de *Ololygon centralis*. Oitenta e sete pares de *primers* foram desenhados para as regiões identificadas. Do total dos *primers* desenvolvidos, 47 foram para regiões de tetranucleotídeos, 39 dinucleotídeos, e um trinucleotídeo. Ao excluir os *primers* em contexto de elementos transponíveis, selecionamos 31 pares de *primers* com potencial para marcadores moleculares, cada loco foi nomeado OCE, abreviação derivada de *Ololygon centralis* (Tabela 3).

A busca de regiões microssatélites de motivos di- e tetranucleotídeos para desenho de *primers* são frequentes, como pode ser observado nos estudos de Duryea et al. (2009) que otimizaram 12 locos microssatélite sendo 11 tetra- e um trinucleotídeo. Bessa-Silva et al. (2015) também desenvolveram oito locos de marcadores microssatélites, sendo cinco di- e três tetranucleotídeos. O tipo de motivo tem influência nas taxas de mutação, a frequência de mutação que resultará em polimorfismo está diretamente relacionada com o tamanho e a pureza da repetição (Choudhary e Trivedi, 2010; Vieira et al., 2016).

**Tabela 3.** Características dos pares de *primers* selecionados para porções flanqueadoras de regiões microssatélites. Os motivos de repetição e o tamanho da região prevista para amplificação também estão explicitadas.

| Primer | Motivo de repetição  | Sequência do <i>primer</i> (5'-3')                         | Tamanho do fragmento (pb) |
|--------|----------------------|------------------------------------------------------------|---------------------------|
| OCE 1  | (AC) <sub>10</sub>   | F: TTCGATTATCTCCCTCATAGTG<br>R: TTCCCTCTTCTAATTATCTCCC     | 251                       |
| OCE 2  | (AATC) <sub>7</sub>  | F: CTGATACTGCTGCCACTTACAT<br>R: CATTGCAACTTTGTTTCTTGTT     | 277                       |
| OCE 3  | (ATCC) <sub>6</sub>  | F: GAATCACTGACTATGATGGTGTC<br>R: CTGAGGAGATTGTCTGCTG       | 251                       |
| OCE 4  | (AGAT) <sub>12</sub> | F: ATTATGCCCCGAGTGTATTGTAG<br>R: AACCAGGACAGGTAAGTATAAGG   | 290                       |
| OCE 5  | (AGAT) <sub>9</sub>  | F: TCATAGGTAGGCATATAGATGGA<br>R: ATATGCTGAGAACGCATTGTA     | 276                       |
| OCE 6  | (ACTC) <sub>8</sub>  | F: CACCTACACAAAGGATACATTACA<br>R: CTGAAACAGTGCAGGAGAAG     | 378                       |
| OCE 7  | (AC) <sub>5</sub>    | F: ACTGTGTATCATCACAATGCAG<br>R: TATACTACATGAGGCAGTGTGG     | 354                       |
| OCE 8  | (ATC) <sub>7</sub>   | F: GCTTGATCATGAAACACTAACA<br>R: TCATCAGCATTAGCATCAGATA     | 185                       |
| OCE 9  | (AAAG) <sub>9</sub>  | F: TCTCTCGCTTTGTTTGTTC<br>R: AAAGGAAGAAGGAGATGAAGAA        | 202                       |
| OCE 10 | (AC) <sub>10</sub>   | F: TGTGTGTGAAAGCACCCCTTA<br>R: CAGGTATCTGCTCTGATGGTTA      | 232                       |
| OCE 11 | (AC) <sub>10</sub>   | F: AGGTTCTTAAGTATCAGATCCC<br>R: AGTTCTCCTTTAACCCTTTCAG     | 199                       |
| OCE 12 | (AGAT) <sub>5</sub>  | F: CTCATAGAGAGAAAGGTTGCAC<br>R: CTTTATCAAGCAAGCGTCTATC     | 165                       |
| OCE 13 | (AC) <sub>13</sub>   | F: AGCTAAGCTGTATCTGTCAACAT<br>R: CACTGTGTGTCTATAGCACTGTAA  | 242                       |
| OCE 14 | (AG) <sub>14</sub>   | F: ACACTGAGGTGGTCATGTTACT<br>R: TATAAAGGTTGTAGATGGGTGG     | 245                       |
| OCE 15 | (ACTC) <sub>8</sub>  | F: CTGTGCAATCTGTGAGATCAG<br>R: GAGCGTCTGCTGAATATGTAGT      | 178                       |
| OCE 16 | (AGAT) <sub>5</sub>  | F: GTTGTAGCTGTAATTCCTCCA<br>R: CAATACATTGGTAAGTGTGCTTT     | 237                       |
| OCE 17 | (AG) <sub>5</sub>    | F: CAGGCTCTGCATAAACTAAGAC<br>R: CAGCTCACTTAACTTAGGCAC      | 212                       |
| OCE 18 | (ACTC) <sub>5</sub>  | F: AGTGACATGGGACTTTACCC<br>R: CAGCTCACTTAACTTAGGCAC        | 350                       |
| OCE 19 | (AGAT) <sub>12</sub> | F: CCTTTAGACATTGGCTGTATGT<br>R: GAAACGTTGTATTCACCATGA      | 222                       |
| OCE 20 | (AGAT) <sub>10</sub> | F: TCCATCTATCCATCTATATGTCACT<br>R: AAGATTTATTACCCCAATGTGA  | 167                       |
| OCE 21 | (AGAT) <sub>9</sub>  | F: AATCTGCGAGTTGTATATGAGC<br>R: ACTGCTTTCTTCATCTCTAGCA     | 218                       |
| OCE 22 | (AC) <sub>16</sub>   | F: GTTTATACCACCACATGAGGAC<br>R: AGTCATGATGCTTGCACTTTA      | 244                       |
| OCE 23 | (AC) <sub>16</sub>   | F: AAACCACAGTTATCACTTCAGC<br>R: AAATGGGTTACCTAGTCCTCAT     | 253                       |
| OCE 24 | (ACAT) <sub>6</sub>  | F: CAGATTGATTTTCATTGTCTTGT<br>R: CACACATATACATACAGATACTTGC | 255                       |

Continua...

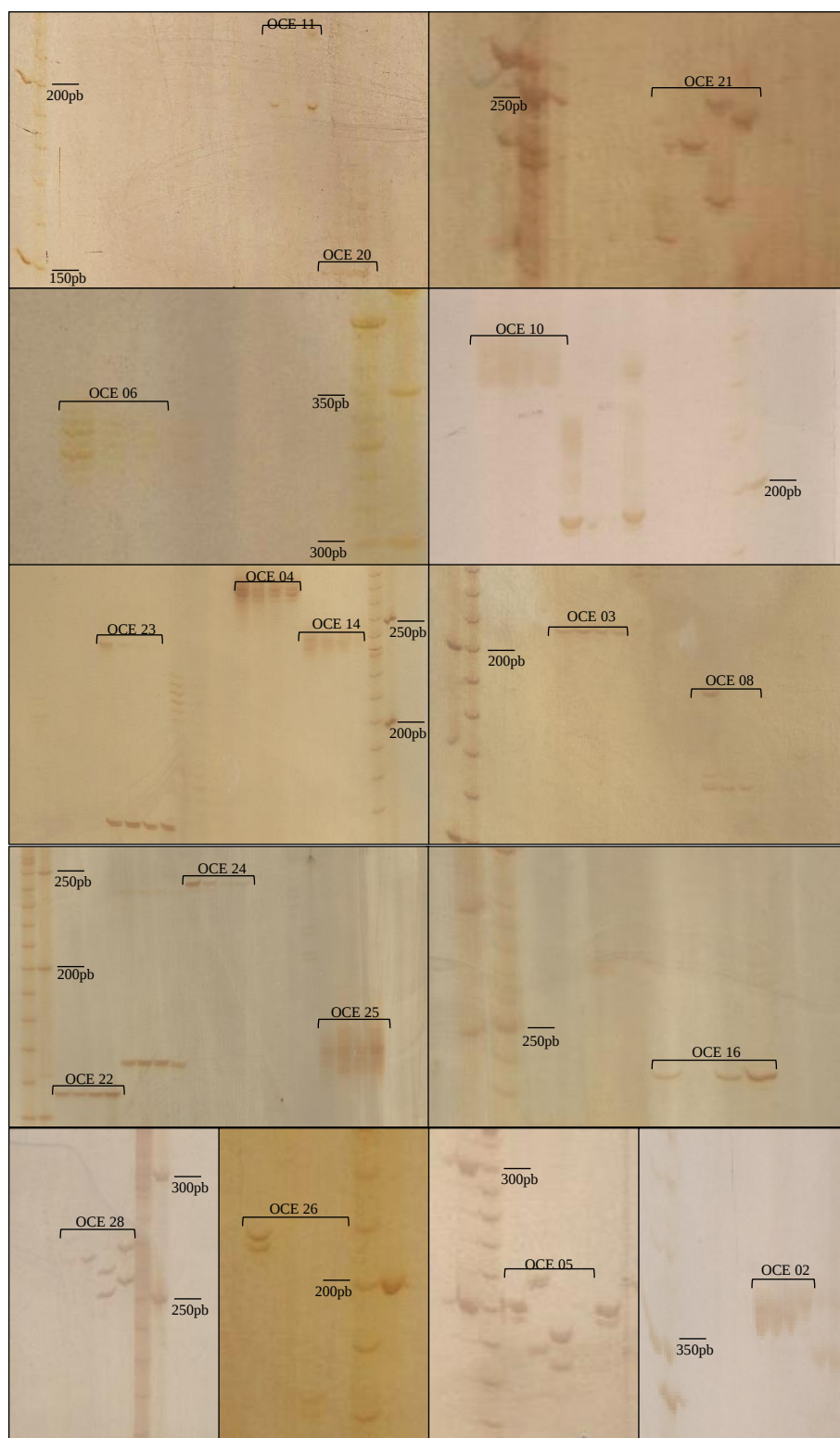
**Tabela 3.** Continuação.

|        |                      |                                                         |     |
|--------|----------------------|---------------------------------------------------------|-----|
| OCE 25 | (AG) <sub>6</sub>    | F: TAATGGTCATATGGATTGCTTT<br>R: CGGTCCCATTTACATACATATT  | 169 |
| OCE 26 | (AAAC) <sub>5</sub>  | F: ATGAATGATTGTTTGACACAGTA<br>R: GCATCAAGAAAGACAACCTAAA | 208 |
| OCE 27 | (AAAG) <sub>5</sub>  | F: AATTTGTTTGTAAATTCTGCTGG<br>R: AGCTGACAAATCAGATCCTTT  | 215 |
| OCE 28 | (AGAT) <sub>7</sub>  | F: GCTCTAGATTGTTGAGCAGAAA<br>R: CGAGTAGAAGGATTCACAGGTA  | 253 |
| OCE 29 | (AGAT) <sub>5</sub>  | F: CTCAGCCTCTTTCCCTATCTAT<br>R: GTTCAGTGACTTCAGACAGAGC  | 195 |
| OCE 30 | (AAAC) <sub>7</sub>  | F: AGAGGTTCTCAGCAACACATT<br>R: CATTGACAGCACCTTAGCTATT   | 207 |
| OCE 31 | (AGAT) <sub>11</sub> | F: GTGCAGCAGGTAAGACATGCTAT<br>R: TTTACCTGGTTGTGCTGCGT   | 179 |

### **Padronização da amplificação dos *primers* de regiões microssatélites**

Dos 31 *primers* selecionados, somente foi possível padronizar a temperatura de anelamento para 18 *primers* (Figura 1). *Primers* que não amplificaram ou não foi observado um padrão de amplificação de PCR foram descartados das análises.

Fatores como concentração de DNA e integridade da amostra podem comprometer a amplificação do produto de PCR em quantidade detectável. Dessa forma, pode ser que a falta de amplificação de determinados *primers* podem estar sujeitos a esses fatores ou simplesmente ao anelamento do *primer*. Diferente dos protocolos usados por Reid et al. (2012), Gutiérrez-Rodríguez e Martínez-Solano (2013), Komaki et al. (2014) em que utilizaram 50 ng, 15 ng e 50 ng de DNA genômico para reações de PCR com volume final de 10 µl, 15 µl e 10 µl respectivamente, o protocolo utilizado nesse trabalho em reações de PCR com 15 µl de volume final usou DNA com concentração de 2,5 ng.



**Figura 1.** Padrão de amplificação para 18 locos microsatélites de *O. centralis* em gel de poliacrilamida (6%).

As temperaturas de anelamento dos *primers* padronizados variaram de 50°C a 60°C (Tabela 4). Dos 18 locos, apenas 7 apresentaram polimorfismo (OCE05, OCE08, OCE11, OCE21, OCE22, OCE26 e OCE28) e 11 foram monomórficos dos quais 10 apresentaram loco homozigoto quando avaliados para quatro indivíduos em eletroforese vertical com gel de poliacrilamida.

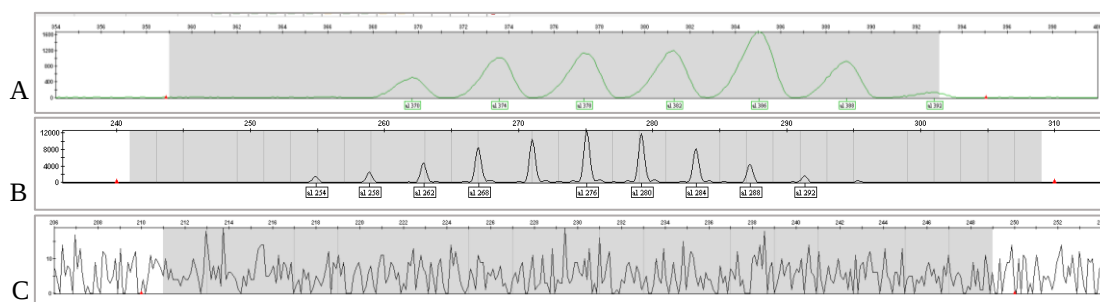
**Tabela 4.** Características dos 18 locos padronizados em gel de poliacrilamida (6%).

| Loco   | Temperatura de anelamento | Amplitude Alélica |
|--------|---------------------------|-------------------|
| OCE02  | 58°C                      | 360               |
| OCE03  | 50°C                      | 205               |
| OCE04  | 60°C                      | 270               |
| OCE05* | 56°C                      | 235-260           |
| OCE06  | 58°C                      | 320               |
| OCE08* | 50°C                      | 160-185           |
| OCE10  | 54°C                      | 220               |
| OCE11* | 54°C                      | 195-223           |
| OCE14  | 60°C                      | 240               |
| OCE16  | 54°C                      | 240               |
| OCE20  | 58°C                      | 150               |
| OCE21* | 54°C                      | 218-260           |
| OCE22* | 54°C                      | 145-155           |
| OCE23  | 56°C                      | 165-240           |
| OCE 24 | 54°C                      | 248               |
| OCE25  | 58°C                      | 170               |
| OCE26* | 56°C                      | 182-218           |
| OCE28* | 54°C                      | 250-275           |

\* Locos polimórficos para os quatro indivíduos analisados.

Entre os *primers* padronizados, foram selecionados 10 (OCE02, OCE05, OCE06, OCE10, OCE11, OCE20, OCE21 OCE24, OCE25 e OCE26) para serem analisados em eletroforese capilar em 30 indivíduos.

Os *primers* OCE02, OCE06 não apresentaram qualidade de amplificação quando analisados pela eletroforese capilar, já o OCE10 marcado não obteve produto de PCR sendo descartado das análises por apresentarem problemas de amplificação em todos indivíduos analisados (Figura 2). Os locos em que não foram observados problema na amplificação foram então distribuídos em duas multiplex (Tabela 5).



**Figura 2.** Locos microsatélites que apresentaram problemas de amplificação quando analisados em eletroforese capilar, A OCE2, B OCE6, C OCE10.

Na multiplex 1, os *primers* OCE11 e OCE21 foram marcados na mesma fluorescência e ao analisar no GeneMapper verificou-se alelos que não estavam na amplitude alélica definida para ambos quando analisado nos testes em gel de poliácridamida 6% só com quatro indivíduos, resultando em sobreposição, assim o *primer* OCE 21 foi analisado em singleplex, restando então na multiplex 1: OCE11, OCE20 e OCE26; multiplex 2: OCE5, OCE24 e OCE25.

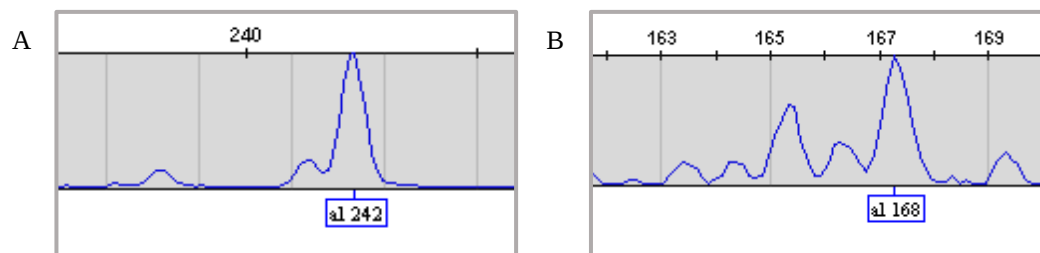
**Tabela 5.** Multiplex de genotipagem do conjunto de *primers* para *Ololygon centralis*

| Primer | Dye   | Amplitude alélica* | Quantidade de reação |
|--------|-------|--------------------|----------------------|
| OCE11  | VIC   | 180-228            | 3 µL                 |
| OCE20  | NED   | 148-164            | 3 µL                 |
| OCE26  | 6'FAM | 182-262            | 2 µL                 |
| OCE05  | VIC   | 218-256            | 4 µL                 |
| OCE24  | 6'FAM | 242                | 2 µL                 |
| OCE25  | 6'FAM | 168                | 2 µL                 |

\* Amplitude alélica dos locos para os trinta indivíduos analisados em eletroforese capilar.

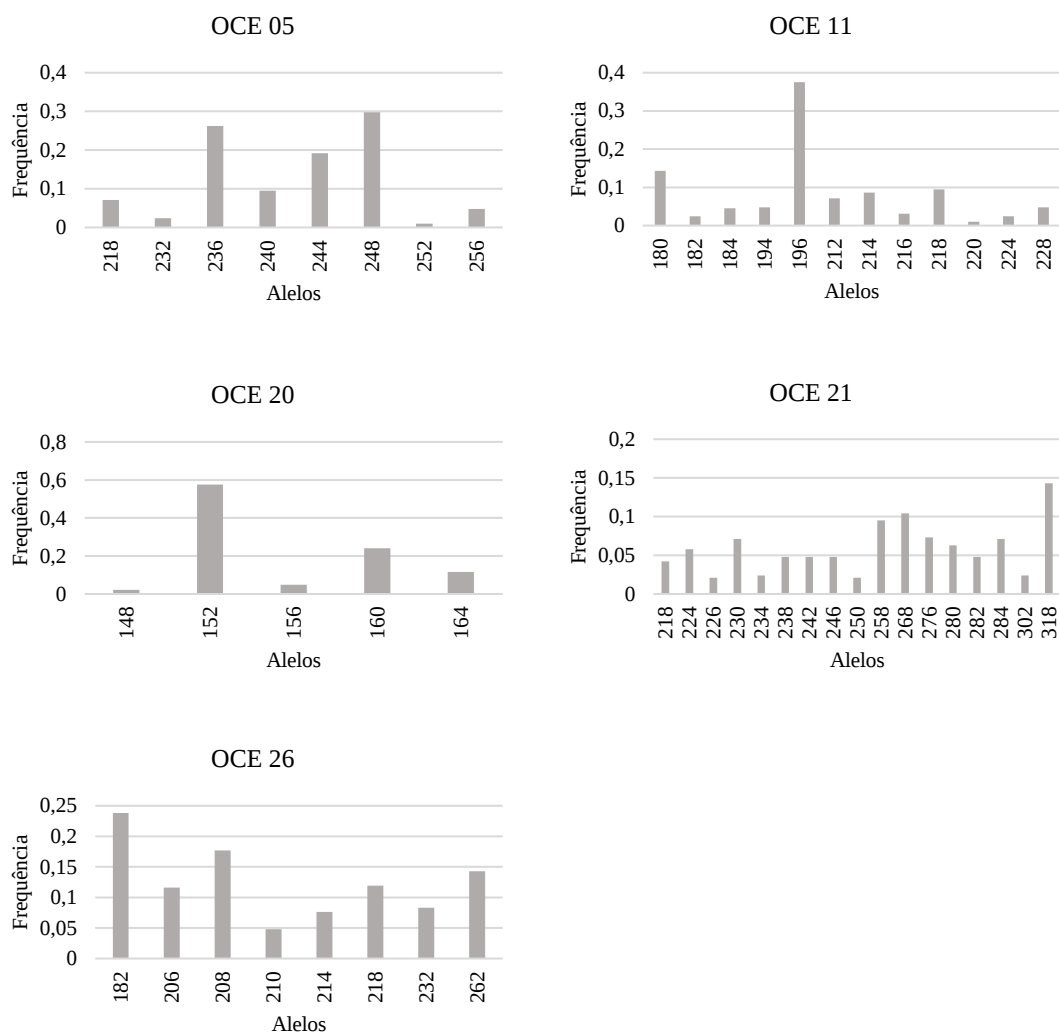
### Caracterização dos marcadores microsatélites

Dos 7 locos avaliados, o OCE 24 e OCE 25 mostraram ser locos monomórficos para todos os indivíduos analisados (Figura 3) o que resultou na exclusão desses locos para análises posteriores e também como para uso como marcadores moleculares.



**Figura 3.** Padrão de amplificação dos locos monomórficos. (A) OCE 24 e (B) OCE 25. O eixo x corresponde ao tamanho do alelo.

A partir das análises dos 5 locos polimórficos foram identificados de cinco a 17 alelos e uma média de 10 alelos por loco indicando um alto polimorfismo. O loco com número máximo de alelos foi o OCE 21 e o com número menor o de OCE 20. A amplitude alélica dos locos variou de 148 pb (OCE 20) a 318 pb (OCE 21) (Figura 4). Para cada loco, foi observado a predominância de pelos menos um alelo. Não houve a presença de alelos raros (frequência menor que 0,005) para nenhum dos locos, a abundância de alelos e frequências para cada loco pode ser visualizada na figura 4.



**Figura 4.** Frequência alélica obtida para cada marcador microsatélites no conjunto de 30 indivíduos. Nos gráficos, o eixo x representa o tamanho dos alelos e o eixo y, a frequência alélica.

Os marcadores apresentaram probabilidade combinada de identidade (I) de  $1,13 \times 10^{-6}$  demonstrando que os locos microsatélites podem ser utilizados para discriminação de indivíduos. A exclusão de paternidade (Q) foi 0,993, valor que expressa uma elevada probabilidade para excluir falsa paternidade (Evet e Weir, 1998) (Tabela 6).

**Tabela 6.** Caracterização dos cinco locos microssatélites para 30 indivíduos de *Ololygon centralis*.

| Loco   | A  | Ho    | He    | I                     | Q     |
|--------|----|-------|-------|-----------------------|-------|
| OCE 05 | 8  | 0,490 | 0,609 | 0,108                 | 0,517 |
| OCE 11 | 12 | 0,448 | 0,780 | 0,075                 | 0,592 |
| OCE 20 | 5  | 0,774 | 0,588 | 0,231                 | 0,336 |
| OCE 21 | 17 | 0,505 | 0,872 | 0,013                 | 0,829 |
| OCE 26 | 8  | 0,394 | 0,642 | 0,046                 | 0,678 |
| Média  | 10 | 0,522 | 0,698 | $1,13 \times 10^{-6}$ | 0,993 |

Legenda: A= número de alelos por loco; Ho = Heterozigosidade observada; He = Heterozigosidade esperada; I = Probabilidade de identidade genética; Q = Probabilidade de exclusão de paternidade.

Considerando as três populações de *O. centralis*, a heterozigosidade média esperada entre os cinco locos foi de 0,7 variando de 0,41 a 0,98 e a observada 0,5 variando de 0 a 1 (Tabela 7). Na população de Caldas Novas não houve nenhum heterozigoto para o loco OCE26 enquanto que na população de Silvânia e Pires do Rio Ho para esse loco foi 0,5 e 0,86, respectivamente. Locos com deficiência de heterozigotos também foram resultados encontrados nos trabalhos de Newman e Squire, 2001 e Garner e Tomio, 2001 demonstrando pouca ou nenhuma variação de alelos entre os indivíduos amostrados em *Rana sylvatica* e *Rana latastei* respectivamente.

A média de heterozigosidade observada e esperada nas populações de Silvânia, Caldas Novas e Pires do Rio foram 0,46, 0,49, 0,68 e 0,60, 0,70, 0,79 respectivamente. No loco OCE20 pode-se verificar um excesso de heterozigotos nas três populações em que o Ho foi maior que o He (Tabela 7). O alto valor encontrado para heterozigosidade observada comparada a esperada ao caracterizar locos microssatélites ocorreram da mesma forma para outras espécies como *Phyllomedusa burmeisteri* (Brunes et al., 2013), *Leptodactylus bufonius* e *L. chaquensis* (Faggione et al., 2014).

A significância do coeficiente de endogamia no loco OCE05 da população de Silvânia e no loco OCE21 na população de Pires do Rio, indica desvio no equilíbrio de Hardy-Weinberg, devido a grande quantidade de homozigotos e consequentemente baixo Ho (Tabela 6). Essa significância pode ser referente a amostragem de indivíduos e presença de alelos nulos que foram certificados para esses locos.

Os alelos nulos correspondem a deficiência de amplificação durante a PCR de um dos alelos de um indivíduo heterozigoto em ocorrência de alterações nas sequências

flanqueadoras, amplificando apenas um dos alelos e formando um “falso” homozigoto (Guichoux et al., 2011). Para os locos que demonstraram alelo nulo é importante que se realize mais testes com populações de maior tamanho amostral.

O índice de fixação global  $F_{ST}$  encontrado foi 0,213 ( $p < 0,05$ ), indicando uma elevada diferenciação genética, uma vez que aproximadamente 20% da variação genética se encontra no componente entre populações. A estrutura genética avaliada par-a-par também foi significativa, sendo que o maior valor encontrado foi entre a população de Silvânia e Pires do Rio ( $F_{ST} = 0,232$ ;  $p < 0,05$ ), seguido de Silvânia e Caldas Novas ( $F_{ST} = 0,227$ ;  $p < 0,05$ ) e por último, a menor diferenciação que foi verificada entre as populações de Caldas Novas e Pires do Rio ( $F_{ST} = 0,144$ ;  $p < 0,05$ ).

O valor de  $F_{ST}$  está diretamente relacionado com fatores ecológicos, microevolutivos e de história de vida da espécie, tais como: fluxo gênico, sistema reprodutivo, capacidade de dispersão, tamanho da população, dentre outros. Espécies que exibem alta estruturação genética entre suas populações devem ser objeto de estudo para definir melhor as estratégias de conservação, uma vez que essas características genéticas interferem na viabilidade e permanência das populações no tempo (Yotoko, 2009; Hartl e Clark, 2010).

**Tabela 7.** Variabilidade genética em três populações de *Ololygon centralis* usando 5 locos microssatélites.

| Locos | Pop. Silvânia<br>(N=16) |                |                |                 | Pop. Caldas Novas<br>(N=7) |                |                |                 | Pop. Pires do Rio<br>(N=7) |                |                |                 |
|-------|-------------------------|----------------|----------------|-----------------|----------------------------|----------------|----------------|-----------------|----------------------------|----------------|----------------|-----------------|
|       | A                       | H <sub>O</sub> | H <sub>E</sub> | F <sub>IS</sub> | A                          | H <sub>O</sub> | H <sub>E</sub> | F <sub>IS</sub> | A                          | H <sub>O</sub> | H <sub>E</sub> | F <sub>IS</sub> |
| OCE05 | 3                       | 0,06           | 0,41           | 0,85*           | 5                          | 0,86           | 0,85           | -0,01           | 4                          | 0,71           | 0,57           | -0,25           |
| OCE11 | 5                       | 0,25           | 0,59           | 0,58            | 5                          | 0,43           | 0,86           | 0,50            | 7                          | 0,71           | 0,89           | 0,20            |
| OCE20 | 4                       | 0,63           | 0,53           | -0,19           | 3                          | 0,71           | 0,52           | -0,36           | 4                          | 1,00           | 0,71           | -0,40           |
| OCE21 | 7                       | 0,88           | 0,82           | -0,07           | 6                          | 0,43           | 0,82           | 0,48            | 7                          | 0,14           | 0,98           | 0,85*           |
| OCE26 | 4                       | 0,50           | 0,64           | 0,23            | 2                          | 0              | 0,48           | 1,00            | 5                          | 0,86           | 0,80           | -0,08           |
| Média | 4,6                     | 0,46           | 0,60           | 0,28            | 4,2                        | 0,49           | 0,70           | 0,32            | 5,4                        | 0,68           | 0,79           | 0,06            |

(N) número de indivíduos, (A) número de alelos, (H<sub>E</sub>) heterozigosidade esperada, (H<sub>O</sub>) Heterozigosidade observada, (F<sub>IS</sub>) coeficiente de endogamia. \*Significativo do Equilíbrio de Hardy-Weinberg para  $P < 0,05$ .

Os cinco locos (Tabela 7) que foram caracterizados para *Ololygon centralis* serão suficientes para realizar estudos genéticos-populacionais da espécie dado que outros trabalhos como de Burnes et al. (2004) e Eterovick et al. (2009) aplicaram análises de diversidade genética, estruturação populacional utilizando quatro e dois locos respectivamente.

## CONCLUSÃO

Apesar da baixa quantidade de sequências microssatélites de boa qualidade para o desenvolvimento de marcadores, obtidas com a construção da biblioteca genômica e sequenciamento na plataforma Illumina (Miseq), foi possível selecionar regiões para desenvolvimento de marcadores para *O. centralis*. Foi possível disponibilizar cinco pares de *primers* padronizados que apresentaram polimorfismos e podem ser utilizados em estudos genético-populacionais dessa espécie de anuro. Esse foi o primeiro trabalho que descreve e padroniza marcadores microssatélites para *Ololygon centralis*. Além do estudo ser importante para avaliar geneticamente as populações da espécie, os marcadores padronizados podem ser úteis no estudo de espécies próximas por meio da transferibilidade de *primers*.

## REFERÊNCIAS

- ARRUDA, M. P. et al. Avaliação da diversidade genética do sapo generalista *Rhinella schneideri* em habitats com perturbações agropecuárias. In: ARRUDA, Maurício Papa de. Avaliação da organização da variabilidade genética em populações de anfíbios de habitats antropizados por meio marcadores microssatélites. 2010. Tese (doutorado) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas, 2010. cap. 1.
- BESSA-SILVA, A. R. et al. Development and characterization of microsatellite loci for *Rhinella marina* (Amphibia, Bufonidae) and their transferability to two closely related species. **Conservation genetics resources**, v. 7, n. 1, p. 247-250, 2015.
- BISCOTTI, M. A.; OLMO, E.; HESLOP-HARRISON, J. S. P. Repetitive DNA in eukaryotic genomes. **Chromosome research**, v. 23, n. 3, p. 415-420, 2015.
- BRUNES, T. O. et al. Characterization of polymorphic microsatellite markers for the Neotropical leaf-frog *Phyllomedusa burmeisteri* and cross-species amplification. **Genetics and Molecular Research**, p. 242-247, 2013.
- BURNS, E. L.; ELDRIDGE, M. D. B.; HOULDEN, B. A. Microsatellite variation and population structure in a declining Australian Hylid *Litoria aurea*. **Molecular Ecology**, v. 13, n. 7, p. 1745-1757, 2004.

- CHOUDHARY, O. P.; TRIVEDI, S. Microsatellite or simple sequence repeat (SSR) instability depends on repeat characteristics during replication and repair. **Journal of Cell and Molecular Biology**, v. 8, n. 2, p. 21-34, 2010.
- CRESTE, S.; NETO, A. T.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reporter**, v. 19, n. 4, p. 299-306, 2001.
- DURYEA, M. C.; BRASILEIRO, C. A.; ZAMUDIO, K. R. Characterization of microsatellite markers for snouted treefrogs in the *Scinax perpusillus* species group (Anura, Hylidae). **Conservation genetics**, v. 10, n. 4, p. 1053-1056, 2009.
- ELLEGREN, H. Microsatellites: simple sequences with complex evolution. **Nature reviews genetics**, v. 5, n. 6, p. 435, 2004.
- ETEROVICK, Paula C. et al. Small Scale Population Structure in the Treefrog *Bokermannohyla saxicola* (Anura, Hylidae). south american **Journal of herpetology**, v. 4, n. 3, p. 235-244, 2009.
- EVETT, I. W.; WEIR, B. S. **Interpreting DNA evidence: statistical genetics for forensic scientists**. Sinauer, 1998.
- FAGGIONI, G. P. et al. Isolation and characterization of microsatellites markers for two South American frogs (*Leptodactylus bufonius* and *L. chaquensis*) using next generation sequencing. **Amphibia-reptilia**, v. 35, n. 4, p. 405-412, 2014.
- GARDNER, M. G. et al. Rise of the machines—recommendations for ecologists when using next generation sequencing for microsatellite development. **Molecular Ecology Resources**, v. 11, n. 6, p. 1093-1101, 2011.
- GARNER, T. W. J.; TOMIO, G. Microsatellites for use in studies of the Italian Agile Frog, *Rana latastei* (Boulenger). **Conservation Genetics**, v. 2, n. 1, p. 77-80, 2001.
- GOUDET, J. **FSTAT**, a program to estimate and test gene diversities and fixation indices, version 2.9. 3. <http://www2.unil.ch/popgen/softwares/fstat.htm>, 2002.

GUICHOUX, E. et al. Current trends in microsatellite genotyping. **Molecular ecology resources**, v. 11, n. 4, p. 591-611, 2011.

GUTIÉRREZ-RODRÍGUEZ, J.; MARTÍNEZ-SOLANO, I. Isolation and characterization of sixteen polymorphic microsatellite loci in the Western Spadefoot, *Pelobates cultripes* (Anura: Pelobatidae) via 454 pyrosequencing. **Conservation Genetics Resources**, v. 5, n. 4, p. 981-984, 2013.

HADDAD, C. F. B.; PRADO, C. P. A. Reproductive modes in frogs and their unexpected diversity in the Atlantic Forest of Brazil. **BioScience**, v. 55, n. 3, p. 207-217, 2005.

HARTL, D.; CLARK, A. **Princípios de genética de populações**. 4 ed. Porto Alegre: Artmed, 660 p. 2010.

HOFFMANN, M. et al. The impact of conservation on the status of the world's vertebrates. **Science**, v. 330, n. 6010, p. 1503-1509, 2010.

JEHLE, R.; ARNTZEN, J. W. Microsatellite markers in amphibian conservation genetics. **Herpetological journal**, v. 12, p. 1-9, 2002.

KATZENBERGER, M. et al. Tolerância e sensibilidade térmica em anfíbios. **Revista da Biologia**, 2012.

KOMAKI, S. et al. Development and characterization of 14 microsatellite markers for *Buergeria japonica* (Amphibia, Anura, Rhacophoridae). **Genes & genetic systems**, v. 89, n. 1, p. 35-39, 2014.

LEWIS, P.; ZAYKIN, D. **Genetic data analysis**: computer programme for the analyses of allelic data - version 1.1, 2002. Disponível em: <<http://http://lewis.eeb.uconn.edu/lewishome/>>. Acesso em: 06 de fev. 2014.

LUQUET, E. et al. Heterozygosity–fitness correlations among wild populations of European tree frog (*Hyla arborea*) detect fixation load. **Molecular ecology**, v. 20, n. 9, p. 1877-1887, 2011.

MARDIS, E. R. Next-generation DNA sequencing methods. **Annual Review of Genomics and Human Genetics**, v. 9, p. 387-402, 2008.

MEGLÉCZ, E. et al. QDD: a user-friendly program to select microsatellite markers and design *primers* from large sequencing projects. **Bioinformatics**, v. 26, n. 3, p. 403-404, 2009.

MEGLÉCZ, E. et al. QDD version 3.1: a user-friendly computer program for microsatellite selection and primer design revisited: experimental validation of variables determining genotyping success rate. **Molecular Ecology Resources**, v. 14, n. 6, p. 1302-1313, 2014.

NEWMAN, R. A.; SQUIRE, T.. Microsatellite variation and fine-scale population structure in the wood frog (*Rana sylvatica*). **Molecular Ecology**, v. 10, n. 5, p. 1087-1100, 2001.

OLIVEIRA, E. J. et al. Origin, evolution and genome distribution of microsatellites. **Genetics and Molecular Biology**, v. 29, n. 2, p. 294-307, 2006.

POMBAL JÚNIOR, J. P.; BASTOS, R. P. Nova espécie de *Scinax* Wagler, 1830 do Brasil Central (Amphibia, Anura, Hylidae). **Boletim do Museu Nacional (N.S.) Zoologia** 371: 1-11, 1996.

REID, K.; HOAREAU, T. B.; BLOOMER, P. High-throughput microsatellite marker development in two sparid species and verification of their transferability in the family Sparidae. **Molecular Ecology Resources**, v. 12, n. 4, p. 740-752, 2012.

VAN OOSTERHOUT, C. et al. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. **Molecular Ecology Notes**, v. 4, n. 3, p. 535-538, 2004.

VANCINE, M. H. **Efeito da fragmentação sobre a persistência de anfíbios anuros (Amphibia: Anura) na Mata Atlântica**. 2015. Trabalho de conclusão de curso (Ecologia) - Universidade Estadual Paulista, Instituto de Biociências de Rio Claro, 2015. Disponível em: <<http://hdl.handle.net/11449/138991>>. Acesso em: 26 fev 2018.

VIEIRA, M. L. C. et al. Microsatellite markers: what they mean and why they are so useful. **Genetics and molecular biology**, v. 39, n. 3, p. 312-328, 2016.

WAGNER, H. W.; SEFC, K. M. **Identity 1.0**. Centre for Applied Genetics - University of Agricultural Sciences Vienna, p. 1-5, 1999.

YOTOKO, K. **Genética de populações**. Universidade Federal de Viçosa. Disponível em: <[ftp://ftp.ufv.br/dbg/Filogenia\\_molecular/usuarios/karla/Apostila/Genetica%20de%20populacoes.pdf](ftp://ftp.ufv.br/dbg/Filogenia_molecular/usuarios/karla/Apostila/Genetica%20de%20populacoes.pdf)>. 78 p. 2009.

ZALAPA, J. E. et al. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. **American journal of botany**, v. 99, n. 2, p. 193-208, 2012.

ZANELLA, C. M. et al. Microsatélites: Metodologias de identificação e análise. In: TURCHETTO-ZOLET, A. C. et al. **Marcadores Moleculares na Era genômica: Metodologias e Aplicações**. Ribeirão Preto: Sociedade Brasileira de Genética, 2017. cap 6.

## 7. CONSIDERAÇÕES FINAIS

A partir das sequências geradas com a técnica de sequenciamento de nova geração de baixa cobertura foi possível montar parcialmente o genoma de *Ololygon centralis* bem como anotar suas características genômicas como elementos transponíveis, microssatélites, genes codificadores de proteínas, de tRNA e de rRNA. Com base nas sequências de microssatélites identificadas foram desenvolvidos pares de *primers* para amplificação dessas regiões. Dentre os primers, cinco foram padronizados e utilizados para genotipagem de indivíduos de três populações de *O. centralis*. A utilização dessas regiões como marcadores genéticos irá auxiliar em estudos genéticos-populacionais da espécie. A utilização dos marcadores em espécies próximas por meio da transferibilidade poderá ser testada. Nossos resultados demonstraram que, pelo sequenciamento de baixa cobertura, é possível caracterizar regiões genômicas e desenvolver marcadores moleculares para análises genômicas e populacionais em anuros.

## 8. MATERIAL SUPLEMENTAR

**Tabela S1.** Motivos de repetição e frequência de microssatélites anotadas nas sequências genômicas de *Oololygon centralis*.

| Matriz           | Motivo | Frequência absoluta | Frequência relativa (%) |
|------------------|--------|---------------------|-------------------------|
| Mononucleotídeo  | T      | 76                  | 61,79                   |
|                  | A      | 45                  | 36,59                   |
|                  | C      | 2                   | 1,63                    |
|                  | AT     | 63                  | 54,78                   |
| Dinucleotídeo    | AC     | 19                  | 16,52                   |
|                  | GT     | 10                  | 8,70                    |
|                  | CT     | 9                   | 7,83                    |
|                  | AG     | 8                   | 6,96                    |
|                  | TT     | 6                   | 5,22                    |
| Trinucleotídeo   | AAT    | 40                  | 24,54                   |
|                  | ATT    | 24                  | 14,72                   |
|                  | CTT    | 20                  | 12,27                   |
|                  | ATG    | 13                  | 7,98                    |
|                  | CCT    | 13                  | 7,98                    |
|                  | ATC    | 10                  | 6,13                    |
|                  | AGC    | 9                   | 5,52                    |
|                  | AGG    | 8                   | 4,91                    |
|                  | AAC    | 6                   | 3,68                    |
|                  | GTT    | 5                   | 3,07                    |
|                  | ACT    | 3                   | 1,84                    |
|                  | GTG    | 3                   | 1,84                    |
|                  | AAG    | 2                   | 1,23                    |
|                  | AGT    | 2                   | 1,23                    |
|                  | CCA    | 2                   | 1,23                    |
|                  | GCT    | 2                   | 1,23                    |
|                  | TTT    | 1                   | 0,61                    |
| Tetranucleotídeo | AGAT   | 54                  | 21,51                   |
|                  | ATTT   | 38                  | 15,14                   |
|                  | ATCT   | 37                  | 14,74                   |
|                  | AAAT   | 19                  | 7,57                    |
|                  | GTTT   | 11                  | 4,38                    |
|                  | CCTT   | 8                   | 3,19                    |
|                  | ACAT   | 7                   | 2,79                    |
|                  | CTTT   | 7                   | 2,79                    |
|                  | AATG   | 6                   | 2,39                    |
|                  | ACAG   | 6                   | 2,39                    |
|                  | GATG   | 6                   | 2,39                    |
|                  | CCCT   | 5                   | 1,99                    |
|                  | AAAC   | 4                   | 1,59                    |
|                  | TCTG   | 4                   | 1,59                    |
|                  | TGTA   | 4                   | 1,59                    |
|                  | AACT   | 3                   | 1,20                    |
|                  | CTCT   | 3                   | 1,20                    |
|                  | TTTT   | 3                   | 1,20                    |
|                  | AGAA   | 2                   | 0,80                    |
|                  | AGGC   | 2                   | 0,80                    |
|                  | CTAC   | 2                   | 0,80                    |
|                  | GTAG   | 2                   | 0,80                    |
|                  | TATA   | 2                   | 0,80                    |
|                  | TTGA   | 2                   | 0,80                    |
|                  | AAGG   | 1                   | 0,40                    |

|                  |       |    |       |
|------------------|-------|----|-------|
|                  | AATT  | 1  | 0,40  |
|                  | ACCA  | 1  | 0,40  |
|                  | ACCC  | 1  | 0,40  |
|                  | ATGC  | 1  | 0,40  |
|                  | CACA  | 1  | 0,40  |
|                  | CATT  | 1  | 0,40  |
|                  | CGAG  | 1  | 0,40  |
|                  | CGGA  | 1  | 0,40  |
|                  | GGA   | 1  | 0,40  |
|                  | GTAC  | 1  | 0,40  |
|                  | GTGA  | 1  | 0,40  |
|                  | TCAC  | 1  | 0,40  |
|                  | TCCA  | 1  | 0,40  |
| Pentanucleotídeo | ATTTT | 19 | 14,62 |
|                  | CTTTT | 8  | 6,15  |
|                  | AATAT | 7  | 5,38  |
|                  | AAAAT | 6  | 4,62  |
|                  | AATGT | 4  | 3,08  |
|                  | ATCTT | 4  | 3,08  |
|                  | ATTAT | 4  | 3,08  |
|                  | CCCTC | 4  | 3,08  |
|                  | AACAT | 3  | 2,31  |
|                  | ATTGT | 3  | 2,31  |
|                  | CACAG | 3  | 2,31  |
|                  | CAGTA | 3  | 2,31  |
|                  | CTCTC | 3  | 2,31  |
|                  | GTTTT | 3  | 2,31  |
|                  | TTCCT | 3  | 2,31  |
|                  | AATAC | 2  | 1,54  |
|                  | ACCGC | 2  | 1,54  |
|                  | CCTGT | 2  | 1,54  |
|                  | CTATA | 2  | 1,54  |
|                  | CTTCT | 2  | 1,54  |
|                  | GTGTT | 2  | 1,54  |
|                  | GTTAT | 2  | 1,54  |
|                  | TCATA | 2  | 1,54  |
|                  | TCTGC | 2  | 1,54  |
|                  | TGTCT | 2  | 1,54  |
|                  | TTGAT | 2  | 1,54  |
|                  | AAATC | 1  | 0,77  |
|                  | AACAA | 1  | 0,77  |
|                  | AATAG | 1  | 0,77  |
|                  | ACACT | 1  | 0,77  |
|                  | ACCTC | 1  | 0,77  |
|                  | AGAAC | 1  | 0,77  |
|                  | AGCGA | 1  | 0,77  |
|                  | AGGTC | 1  | 0,77  |
|                  | ATGGA | 1  | 0,77  |
|                  | ATGTC | 1  | 0,77  |
|                  | ATTAA | 1  | 0,77  |
|                  | ATTGC | 1  | 0,77  |
|                  | CAATG | 1  | 0,77  |
|                  | CACAA | 1  | 0,77  |
|                  | CACAG | 1  | 0,77  |
|                  | CACAT | 1  | 0,77  |
|                  | CATCT | 1  | 0,77  |
|                  | CCACA | 1  | 0,77  |
|                  | CGCTA | 1  | 0,77  |
|                  | CTAAA | 1  | 0,77  |

|                 |        |    |      |
|-----------------|--------|----|------|
|                 | GAAAA  | 1  | 0,77 |
|                 | GACTG  | 1  | 0,77 |
|                 | GCAAC  | 1  | 0,77 |
|                 | GCGAC  | 1  | 0,77 |
|                 | GCGCC  | 1  | 0,77 |
|                 | GTGTA  | 1  | 0,77 |
|                 | TAATT  | 1  | 0,77 |
|                 | TACTT  | 1  | 0,77 |
|                 | TCCCT  | 1  | 0,77 |
|                 | TCCTT  | 1  | 0,77 |
|                 | TGTGG  | 1  | 0,77 |
| Hexanucleotídeo | ATTTTT | 12 | 6,86 |
|                 | ACACAT | 7  | 4,00 |
|                 | ATATTT | 6  | 3,43 |
|                 | CTTTTT | 6  | 3,43 |
|                 | ATGTTT | 5  | 2,86 |
|                 | CCTTTT | 5  | 2,86 |
|                 | TTTCTT | 4  | 2,29 |
|                 | TTTTTA | 4  | 2,29 |
|                 | AAAAGA | 3  | 1,71 |
|                 | ATGAAG | 3  | 1,71 |
|                 | ATTATA | 3  | 1,71 |
|                 | GTATAT | 3  | 1,71 |
|                 | TCTCTT | 3  | 1,71 |
|                 | TTTGTT | 3  | 1,71 |
|                 | AACATA | 2  | 1,14 |
|                 | AATACA | 2  | 1,14 |
|                 | AGATAG | 2  | 1,14 |
|                 | AGTCGG | 2  | 1,14 |
|                 | ATAACA | 2  | 1,14 |
|                 | CACCAT | 2  | 1,14 |
|                 | CTTATT | 2  | 1,14 |
|                 | CTTTAA | 2  | 1,14 |
|                 | CTTTAT | 2  | 1,14 |
|                 | GATACT | 2  | 1,14 |
|                 | GGAGGT | 2  | 1,14 |
|                 | GTTTTA | 2  | 1,14 |
|                 | TATTGT | 2  | 1,14 |
|                 | TCCGAC | 2  | 1,14 |
|                 | TTAATA | 2  | 1,14 |
|                 | TTTATA | 2  | 1,14 |
|                 | AAACAA | 1  | 0,57 |
|                 | AAATAA | 1  | 0,57 |
|                 | AACACA | 1  | 0,57 |
|                 | AACTAC | 1  | 0,57 |
|                 | AAGACA | 1  | 0,57 |
|                 | AAGCAG | 1  | 0,57 |
|                 | AATAGA | 1  | 0,57 |
|                 | AATATA | 1  | 0,57 |
|                 | AATTAA | 1  | 0,57 |
|                 | ACCAAT | 1  | 0,57 |
|                 | ACTATG | 1  | 0,57 |
|                 | ACTATT | 1  | 0,57 |
|                 | ACTTGG | 1  | 0,57 |
|                 | AGCGAG | 1  | 0,57 |
|                 | AGCTAT | 1  | 0,57 |
|                 | AGCTGG | 1  | 0,57 |
|                 | AGTCAA | 1  | 0,57 |
|                 | ATAAGT | 1  | 0,57 |

|        |   |      |
|--------|---|------|
| ATACAT | 1 | 0,57 |
| ATACTA | 1 | 0,57 |
| ATAGAC | 1 | 0,57 |
| ATATAG | 1 | 0,57 |
| ATATGG | 1 | 0,57 |
| ATGAAT | 1 | 0,57 |
| ATGATA | 1 | 0,57 |
| ATGATT | 1 | 0,57 |
| ATGGTG | 1 | 0,57 |
| ATTAGG | 1 | 0,57 |
| ATTATT | 1 | 0,57 |
| CACAGA | 1 | 0,57 |
| CACCAC | 1 | 0,57 |
| CAGATG | 1 | 0,57 |
| CCACTG | 1 | 0,57 |
| CCCGCC | 1 | 0,57 |
| CCGATT | 1 | 0,57 |
| CCTAAC | 1 | 0,57 |
| CCTCTT | 1 | 0,57 |
| CCTTCT | 1 | 0,57 |
| CGATAG | 1 | 0,57 |
| CTACAT | 1 | 0,57 |
| CTCATC | 1 | 0,57 |
| CTTATC | 1 | 0,57 |
| CTTTTA | 1 | 0,57 |
| GAAGAA | 1 | 0,57 |
| GACTTT | 1 | 0,57 |
| GAGTGG | 1 | 0,57 |
| GATATT | 1 | 0,57 |
| GCACCA | 1 | 0,57 |
| GCCGCT | 1 | 0,57 |
| GCCTCA | 1 | 0,57 |
| GCGGGT | 1 | 0,57 |
| GGATCC | 1 | 0,57 |
| GGCAGA | 1 | 0,57 |
| GGTCGT | 1 | 0,57 |
| GTCCCG | 1 | 0,57 |
| GTGATT | 1 | 0,57 |
| GTGTAT | 1 | 0,57 |
| GTGTTT | 1 | 0,57 |
| GTTCCA | 1 | 0,57 |
| GTTGCA | 1 | 0,57 |
| TACAGA | 1 | 0,57 |
| TATACA | 1 | 0,57 |
| TCCACC | 1 | 0,57 |
| TCCGCT | 1 | 0,57 |
| TCTGAA | 1 | 0,57 |
| TCTTCT | 1 | 0,57 |
| TCTTTG | 1 | 0,57 |
| TGAACC | 1 | 0,57 |
| TGAGGA | 1 | 0,57 |
| TGCTCC | 1 | 0,57 |
| TGCTTC | 1 | 0,57 |
| TGTCCT | 1 | 0,57 |
| TTCATT | 1 | 0,57 |

---

**Tabela S2.** Características dos indivíduos de *Ololygon centralis* coletados nos municípios de Silvânia, Pires do Rio e Caldas Novas – GO.

| Espécie             | No. Indivíduo | No. museu | No. tecido | Estado | Município    | Localidade                | Data coleta | Coletores                          | GPS                         |
|---------------------|---------------|-----------|------------|--------|--------------|---------------------------|-------------|------------------------------------|-----------------------------|
| <i>O. centralis</i> | OCE-FLO1      | ZUFG9287  | TA-9-A8    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO3     | ZUFG9289  | TA-9-A10   | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO4     | ZUFG9290  | TA-9-B1    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO5     | ZUFG9291  | TA-9-B2    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO6     | ZUFG9292  | TA-9-B3    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO7     | ZUFG9293  | TA-9-B4    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO9     | ZUFG9295  | TA-9-B6    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO10    | ZUFG9296  | TA-9-B7    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO11    | ZUFG9297  | TA-9-B8    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO12    | ZUFG9298  | TA-9-B9    | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO13    | ZUFG9299  | TA-9-B10   | GO     | Silvânia     | Ponto G, FLONA            | 02/02/2015  | Cintia Targueta e Vinícius Guerra  | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO14    | ZUFG9300  | TA-9-C1    | GO     | Silvânia     | Ponto G, FLONA            | 11/05/2015  | Priscilla Gambale e Rogério Bastos | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO15    | ZUFG9301  | TA-9-C2    | GO     | Silvânia     | Ponto G, FLONA            | 11/05/2015  | Priscilla Gambale e Rogério Bastos | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO16    | ZUFG9302  | TA-9-C3    | GO     | Silvânia     | Ponto G, FLONA            | 11/05/2015  | Priscilla Gambale e Rogério Bastos | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO17    | ZUFG9303  | TA-9-C4    | GO     | Silvânia     | Ponto G, FLONA            | 11/05/2015  | Priscilla Gambale e Rogério Bastos | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -FLO18    | ZUFG9304  | TA-9-C5    | GO     | Silvânia     | Ponto G, FLONA            | 11/05/2015  | Priscilla Gambale e Rogério Bastos | 16°38'04.4"S, 48°39'31.2"W  |
| <i>O. centralis</i> | OCE -PIR1     |           | TA-9-C6    | GO     | Pires do Rio | Fazenda do Plínio         | 13/02/2015  | Vinícius Guerra                    | 17°25'18.76"S, 48°24'6.91"W |
| <i>O. centralis</i> | OCE -PIR2     |           | TA-9-C7    | GO     | Pires do Rio | Fazenda do Plínio         | 13/02/2015  | Vinícius Guerra                    | 17°25'18.76"S, 48°24'6.91"W |
| <i>O. centralis</i> | OCE -PIR3     |           | TA-9-C8    | GO     | Pires do Rio | Fazenda do Plínio         | 13/02/2015  | Vinícius Guerra                    | 17°25'18.76"S, 48°24'6.91"W |
| <i>O. centralis</i> | OCE -PIR4     |           | TA-9-C9    | GO     | Pires do Rio | Fazenda do Plínio         | 13/02/2015  | Vinícius Guerra                    | 17°25'18.76"S, 48°24'6.91"W |
| <i>O. centralis</i> | OCE -PIR5     |           | TA-9-C10   | GO     | Pires do Rio | Fazenda do Plínio         | 13/02/2015  | Vinícius Guerra                    | 17°25'18.76"S, 48°24'6.91"W |
| <i>O. centralis</i> | OCE -PIR6     |           | TA-9-D1    | GO     | Pires do Rio | Fazenda do Plínio         | 13/02/2015  | Vinícius Guerra                    | 17°25'18.76"S, 48°24'6.91"W |
| <i>O. centralis</i> | OCE -PIR7     |           | TA-9-D2    | GO     | Pires do Rio | Fazenda do Plínio         | 13/02/2015  | Vinícius Guerra                    | 17°25'18.76"S, 48°24'6.91"W |
| <i>O. centralis</i> | OCE -CNZ3     | ZUFG10371 | TA-9-D6    | GO     | Caldas Novas | Fazenda do Sr. Zé Machado | 13/03/2015  | Vinícius Guerra                    | 17°49'41.74"S, 48°39'3.60"W |

|                     |            |           |          |    |              |                              |            |                 |                              |
|---------------------|------------|-----------|----------|----|--------------|------------------------------|------------|-----------------|------------------------------|
| <i>O. centralis</i> | OCE -CNZ8  |           | TA-9-E1  | GO | Caldas Novas | Fazenda do Sr. Zé Machado    | 15/05/2015 | Vinícius Guerra | 17°49'41.74"S, 48°39'3.60"W  |
| <i>O. centralis</i> | OCE -CNE3  | ZUFG10372 | TA-9-E6  | GO | Caldas Novas | Fazenda do Sr. Erli (Jacaré) | 18/04/2015 | Vinícius Guerra | 17°49'12.62"S, 48°38'50.75"W |
| <i>O. centralis</i> | OCE -CNE5  | ZUFG10375 | TA-9-E8  | GO | Caldas Novas | Fazenda do Sr. Erli (Jacaré) | 18/04/2015 | Vinícius Guerra | 17°49'12.62"S, 48°38'50.75"W |
| <i>O. centralis</i> | OCE -CNE7  | ZUFG10379 | TA-9-E10 | GO | Caldas Novas | Fazenda do Sr. Erli (Jacaré) | 18/04/2015 | Vinícius Guerra | 17°49'12.62"S, 48°38'50.75"W |
| <i>O. centralis</i> | OCE -CNE8  | ZUFG10380 | TA-9-F1  | GO | Caldas Novas | Fazenda do Sr. Erli (Jacaré) | 18/04/2015 | Vinícius Guerra | 17°49'12.62"S, 48°38'50.75"W |
| <i>O. centralis</i> | OCE -CNE10 | ZUFG10382 | TA-9-F3  | GO | Caldas Novas | Fazenda do Sr. Erli (Jacaré) | 18/04/2015 | Vinícius Guerra | 17°49'12.62"S, 48°38'50.75"W |

## 9. ANEXOS



Ministério do Meio Ambiente - MMA  
Instituto Chico Mendes de Conservação da Biodiversidade - ICMBio  
Sistema de Autorização e Informação em Biodiversidade - SISBIO

### Autorização para atividades com finalidade científica

|                                                                                                                                                                                                                                                                                                                                                  |                                   |                                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|------------------------------------|
| Número: 46522-5                                                                                                                                                                                                                                                                                                                                  | Data da Emissão: 12/02/2017 14:35 | Data para Revalidação*: 14/03/2018 |
| * De acordo com o art. 28 da IN 03/2014, esta autorização tem prazo de validade equivalente ao previsto no cronograma de atividades do projeto, mas deverá ser revalidada anualmente mediante a apresentação do relatório de atividades a ser enviado por meio do Sisbio no prazo de até 30 dias a contar da data do aniversário de sua emissão. |                                   |                                    |

#### Dados do titular

|                                                                                                                                              |                          |
|----------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| Nome: Cintia Polegrini Targueta de Azevedo Brito                                                                                             | CPF: 099.574.837-05      |
| Título do Projeto: Variação genética molecular e cromossômica em duas espécies de <i>Sclerophaga</i> (Amphibia, Anura) do Cerrado Brasileiro |                          |
| Nome da Instituição: UNIVERSIDADE FEDERAL DE GOIÁS                                                                                           | CNPJ: 01.567.601/0001-43 |

#### Cronograma de atividades

| # | Descrição da atividade             | Início (mês/ano) | Fim (mês/ano) |
|---|------------------------------------|------------------|---------------|
| 1 | Coleta e transporte dos indivíduos | 11/2014          | 10/2017       |

#### Observações e ressalvas

|   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | As atividades de campo exercidas por pessoa natural ou jurídica estrangeira, em todo o território nacional, que impliquem o deslocamento de recursos humanos e materiais, tendo por objeto coletar dados, materiais, espécimes biológicos e minerais, peças integrantes da cultura nativa e cultura popular, presente e passado, obtidos por meio de recursos e técnicas que se destinem ao estudo, à difusão ou à pesquisa, estão sujeitas à autorização do Ministério de Ciência e Tecnologia.                                                                        |
| 2 | Esta autorização NÃO exclui o pesquisador titular e os membros de sua equipe da necessidade de obter as anuências previstas em outros instrumentos legais, bem como do consentimento do responsável pela área, pública ou privada, onde será realizada a atividade, inclusive do órgão gestor de terra indígena (FUNAI), da unidade de conservação estadual, distrital ou municipal, ou do proprietário, arrendatário, possessor ou morador de área dentro dos limites de unidade de conservação federal cujo processo de regularização fundiária encontra-se em curso. |
| 3 | Este documento somente poderá ser utilizado para as fins previstas na Instrução Normativa ICMBio nº 03/2014 ou na Instrução Normativa ICMBio nº 10/2010, no que especifica esta Autorização, não podendo ser utilizado para fins comerciais, industriais ou esportivos. O material biológico coletado deverá ser utilizado para atividades científicas ou didáticas no âmbito do ensino superior.                                                                                                                                                                       |
| 4 | A autorização para envio ao exterior de material biológico não consignado deverá ser requerida por meio do endereço eletrônico <a href="http://www.ibama.gov.br">www.ibama.gov.br</a> (Serviço on-line - Licença para importação ou exportação de flora e fauna - CITES e não CITES).                                                                                                                                                                                                                                                                                   |
| 5 | O titular da licença ou autorização e os membros da sua equipe deverão optar por métodos de coleta e instrumentos de captura direcionados, sempre que possível, ao grupo taxonômico de interesse, evitando a morte ou dano significativo a outros grupos, e empregar esforço de coleta ou captura que não comprometa a viabilidade de populações do grupo taxonômico de interesse em condições in situ.                                                                                                                                                                 |
| 6 | O titular da autorização ou de licença permanente, assim como os membros de sua equipe, quando da violação da legislação vigente, ou quando da inadequação, omissão ou falsa descrição de informações relevantes que subsidiaram a expedição do ato, poderá, mediante decisão motivada, ter a autorização ou licença suspensa ou revogada pelo ICMBio, nos termos da legislação brasileira em vigor.                                                                                                                                                                    |
| 7 | Este documento não dispensa o cumprimento da legislação que dispõe sobre acesso a componente do patrimônio genético existente no território nacional, na plataforma continental e na zona econômica exclusiva, ou ao conhecimento tradicional associado ao patrimônio genético, para fins de pesquisa científica, bioprospecção e desenvolvimento tecnológico. Veja maiores informações em <a href="http://www.mma.gov.br/legis">www.mma.gov.br/legis</a> .                                                                                                             |
| 8 | Em caso de pesquisa em UNIDADE DE CONSERVAÇÃO, o pesquisador titular desta autorização deverá contactar a administração da unidade a fim de CONFIRMAR AS DATAS das expedições, as condições para realização das coletas e de uso da infra-estrutura da unidade.                                                                                                                                                                                                                                                                                                         |

#### Outras ressalvas

|   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | 1) Nesta nova Renovação foi novamente alterada a composição da equipe de participantes (exclusão de Murilo Ferreira de Melo e inclusão de Maurivan Vaz Ribeiro e Andreza Arantes Castro);<br>2) Experimento submetido ao CEUA da UFG, conforme consta no Protocolo CEUA/UFGo nº 109/2014;<br>3) Esta Autorização não contempla coleta/sacrifício de anuros em fase larvária (girinos).                                                                                                                                                                |
| 2 | 01 - Evitar que a metodologia de coleta interfira em outras pesquisas, em andamento. 02 - O pesquisador deverá disponibilizar ao ICMBio os resultados das pesquisas desenvolvidas, relatórios e publicações, promovendo, sempre, uma apresentação aos técnicos da UC e ao Conselho Consultivo. 03 - O pesquisador, ao utilizar a Flona, deverá respeitar as normas gerais da Unidade, das Zonas e do Regimento Interno. 04 - As publicações dos resultados de quaisquer pesquisas realizadas na UC deverão fazer constar o nome do ICMBio e da Flona. |

#### Equipe

| # | Nome                         | Função             | CPF            | Doc. Identidade | Nacionalidade |
|---|------------------------------|--------------------|----------------|-----------------|---------------|
| 1 | Alexandro Ribeiro de Moraes  | Coletor            | 014.459.411-04 | 3884442 DOPC-GO | Brasileira    |
| 2 | VINICIUS GUERRA BATISTA      | Coletor            | 019.840.241-40 | 4577001 DOPC-GO | Brasileira    |
| 3 | Priscilla Guedes Gamboa      | Coletor            | 024.762.581-06 | 4738329 SSP-GO  | Brasileira    |
| 4 | Rogério Pereira Bastos       | Coletor/taxonômico | 953.726.797-00 | 1726631 SSP-MG  | Brasileira    |
| 5 | Rafaela Roméria Silva Vieira | Coletor            | 026.511.481-14 | 5106847 SPTC-GO | Brasileira    |

Este documento (Autorização para atividades com finalidade científica) foi expedido com base na Instrução Normativa nº 03/2014. Através do código de autenticação abaixo, qualquer cidadão poderá verificar a autenticidade ou regularidade deste documento, por meio da página do Sisbio/ICMBio na Internet ([www.icmbio.gov.br/sisbio](http://www.icmbio.gov.br/sisbio)).

Código de autenticação: 51753486



Página 1/3



Ministério do Meio Ambiente - MMA  
Instituto Chico Mendes de Conservação da Biodiversidade - ICMBio  
Sistema de Autorização e Informação em Biodiversidade - SISBIO

### Autorização para atividades com finalidade científica

|                 |                                   |                                    |
|-----------------|-----------------------------------|------------------------------------|
| Número: 46522-5 | Data da Emissão: 12/02/2017 14:35 | Data para Revalidação*: 14/03/2018 |
|-----------------|-----------------------------------|------------------------------------|

\* De acordo com o art. 28 da IN 03/2014, esta autorização tem prazo de validade equivalente ao previsto no cronograma de atividades do projeto, mas deverá ser revalidada anualmente mediante a apresentação do relatório de atividades a ser enviado por meio do Sisbio no prazo de até 30 dias a contar da data do aniversário de sua emissão.

#### Dados do titular

|                                                                                                                                         |                          |
|-----------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| Nome: Cintia Pellegrini Targueta de Azevedo Brito                                                                                       | CPF: 099.574.637-05      |
| Título do Projeto: Variação genética molecular e cromossômica em duas espécies de <i>Scinax</i> (Amphibia, Anura) do Cerrado Brasileiro |                          |
| Nome da Instituição: UNIVERSIDADE FEDERAL DE GOIÁS                                                                                      | CNPJ: 01.567.601/0001-43 |

|                            |         |                |                 |            |
|----------------------------|---------|----------------|-----------------|------------|
| 6 Nathane de Queiroz Costa | Coletor | 036.296.211-45 | 5489357 SPTC-GO | Brasileira |
| 7 Maurivan Vaz Ribeiro     | Coletor | 005.407.551-33 | 44/9823 DGPC-GO | Brasileira |
| 8 Andreza Arantes Castro   | Coletor | 045.384.361-19 | 5651668 SSP-GO  | Brasileira |

#### Locais onde as atividades de campo serão executadas

| #  | Município            | UF | Descrição do local            | Tipo               |
|----|----------------------|----|-------------------------------|--------------------|
| 1  | GOIÂNIA              | GO | Goiania                       | Fora de UC Federal |
| 2  | APARECIDA DE GOIÂNIA | GO | Aparecida de Goiania          | Fora de UC Federal |
| 3  | SENADOR CANEDO       | GO | Senador Canedo                | Fora de UC Federal |
| 4  | CALDASINHA           | GO | Caldasinha                    | Fora de UC Federal |
| 5  | ANÁPOLIS             | GO | Anápolis                      | Fora de UC Federal |
| 6  | TEREZÓPOLIS DE GOIÁS | GO | Terezópolis                   | Fora de UC Federal |
| 7  | NERÓPOLIS            | GO | Nerópolis                     | Fora de UC Federal |
| 8  | RIO VERDE            | GO | Rio Verde                     | Fora de UC Federal |
| 9  | JATAI                | GO | Jatã                          | Fora de UC Federal |
| 10 | PIRENÓPOLIS          | GO | Pirenópolis                   | Fora de UC Federal |
| 11 | CORUMBÁ DE GOIÁS     | GO | Corumbá                       | Fora de UC Federal |
| 12 | MINÉRIOS             | GO | Minérios                      | Fora de UC Federal |
| 13 | SERRANÓPOLIS         | GO | Serranópolis                  | Fora de UC Federal |
| 14 | GOIANAPÓLIS          | GO | Goianópolis                   | Fora de UC Federal |
| 15 | NIQUELANDIA          | GO | Niquelândia                   | Fora de UC Federal |
| 16 | BARRO ALTO           | GO | Barro Alto                    | Fora de UC Federal |
| 17 | CRISTIANÓPOLIS       | GO | Cristianópolis                | Fora de UC Federal |
| 18 |                      | GO | FLORESTA NACIONAL DE SILVANIA | UC Federal         |
| 19 | GUARDA-MOR           | MG | Guarda-Mor                    | Fora de UC Federal |
| 20 | ÁGUA FRIA DE GOIÁS   | GO | Água Fria de Goiás            | Fora de UC Federal |

#### Atividades X Taxons

| # | Atividade                                                 | Taxons                                                                     |
|---|-----------------------------------------------------------|----------------------------------------------------------------------------|
| 1 | Coleta/transporte de amostras biológicas in situ          | <i>Scinax centralis</i> , <i>Scinax constrictus</i>                        |
| 2 | Coleta/transporte de espécimes de fauna silvestre in situ | <i>Scinax constrictus</i> ("Oide 30"), <i>Scinax centralis</i> ("Oide 30") |

\* Quantidade de indivíduos por espécie, por localidade ou unidade de conservação, a serem coletados durante um ano.

#### Material e métodos

|   |                                     |                                       |
|---|-------------------------------------|---------------------------------------|
| 1 | Amostras biológicas (Anfibios)      | Outras amostras biológicas (anfibios) |
| 2 | Método de captura/coleta (Anfibios) | Captura manual                        |

#### Destino do material biológico coletado

| # | Nome local destino            | Tipo Destino |
|---|-------------------------------|--------------|
| 1 | UNIVERSIDADE FEDERAL DE GOIÁS | coletor      |

Este documento (Autorização para atividades com finalidade científica) foi expedido com base na Instrução Normativa nº 03/2014. Através do código de autenticação abaixo, qualquer cidadão poderá verificar a autenticidade ou regularidade deste documento, por meio da página do Sisbio/ICMBio na Internet ([www.icmbio.gov.br/sisbio](http://www.icmbio.gov.br/sisbio)).

Código de autenticação: 51753486



Página 2/3



Ministério do Meio Ambiente - MMA  
Instituto Chico Mendes de Conservação da Biodiversidade - ICMBio  
Sistema de Autorização e Informação em Biodiversidade - SISBIO

### Autorização para atividades com finalidade científica

|                                                                                                                                                                                                                                                                                                                                                  |                                   |                                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|------------------------------------|
| Número: 46522-5                                                                                                                                                                                                                                                                                                                                  | Data da Emissão: 12/02/2017 14:35 | Data para Revalidação*: 14/03/2018 |
| * De acordo com o art. 28 da IN 03/2014, esta autorização tem prazo de validade equivalente ao previsto no cronograma de atividades do projeto, mas deverá ser revalidada anualmente mediante a apresentação do relatório de atividades a ser enviado por meio do Sisbio no prazo de até 30 dias a contar da data do aniversário de sua emissão. |                                   |                                    |

#### Dados do titular

|                                                                                                                                        |                          |
|----------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| Nome: Cintia Pelegrineti Targueta de Azevedo Brito                                                                                     | CPF: 099.574.637-05      |
| Título do Projeto: Variação genética molecular e cromossômica em duas espécies de <i>Solax</i> (Amphibia, Anura) do Cerrado Brasileiro |                          |
| Nome da Instituição: UNIVERSIDADE FEDERAL DE GOIAS                                                                                     | CNPJ: 01.567.601/0001-43 |

### Registro de coleta imprevista de material biológico

De acordo com a Instrução Normativa nº 03/2014, a coleta imprevista de material biológico ou de substrato não contemplado na autorização ou na licença permanente deverá ser anotada na mesma, em campo específico, por ocasião da coleta, devendo esta coleta imprevista ser comunicada por meio do relatório de atividades. O transporte do material biológico ou do substrato deverá ser acompanhado da autorização ou da licença permanente com a devida anotação. O material biológico coletado de forma imprevista, deverá ser destinado à instituição científica e, depositado, preferencialmente, em coleção biológica científica registrada no Cadastro Nacional de Coleções Biológicas (CCBIO).

| Taxon* | Qtde. | Tipo de amostra | Qtde. | Data |
|--------|-------|-----------------|-------|------|
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |
|        |       |                 |       |      |

\* Identificar o espécime no nível taxonômico possível.

Este documento (Autorização para atividades com finalidade científica) foi expedido com base na Instrução Normativa nº 03/2014. Através do código de autenticação abaixo, qualquer cidadão poderá verificar a autenticidade ou regularidade deste documento, por meio da página do Sisbio/ICMBio na Internet ([www.icmbio.gov.br/sisbio](http://www.icmbio.gov.br/sisbio)).

Código de autenticação: 51753486



Página 3/3



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DE GOIÁS  
PRÓ-REITORIA DE PESQUISA E INOVAÇÃO  
COMISSÃO DE ÉTICA NO USO DE ANIMAIS/CEUA



Goiânia, 06 de fevereiro de 2017.

**PARECER CONSUBSTANCIADO REFERENTE AO PEDIDO DE EMENDA DO  
PROTOCOLO N°. 109/14.**

**I IDENTIFICAÇÃO:**

1. **Título do projeto:** Variação genética molecular e cromossômica em duas espécies de Scinax
2. (Amphibia, Anura) do Cerrado Brasileiro
3. **Pesquisador Responsável:** Cintia Pelegrineti Targueta de Azevedo Brito
4. **Unidade/Órgão do pesquisador:** ICB/UFG
5. **Pesquisadores Participantes:** Cintia Pelegrineti Targueta de Azevedo Brito; Mariana Pires de Campos Telles; Daniela de Melo e Silva; Rogério Pereira Bastos
6. **Unidade onde será realizado:** ICB/UFG
7. **Data de apresentação do protocolo a CEUA:** 17/11/14
8. **Data da apresentação do Pedido de Emenda:** 16/12/16

**II - Parecer da CEUA:**

Informamos que a *Comissão de Ética no Uso de Animais/CEUA* da Universidade Federal de Goiás, após análise das adequações solicitadas, **APROVOU**, o projeto acima referido, e o mesmo foi considerado em acordo com os princípios éticos vigentes.

Reiteramos a importância deste Parecer Substantiado, e lembramos que o(a) pesquisador(a) responsável deverá encaminhar à CEUA-PRPI-UFG o Relatório Final baseado na conclusão do estudo e na incidência de publicações decorrentes deste, de acordo com o disposto na Lei nº. 11.794 de 08/10/2008, e Resolução Normativa nº. 01, de 09/07/2010 do Conselho Nacional de Controle de Experimentação Animal-CONCEA. O prazo para entrega do Relatório é de até 30 dias após o encerramento da pesquisa, prevista para conclusão em 31/12/2018.

**III - Data da reunião: 06/02/2017.**

**Dra. Marina Pacheco Miguel**  
Coordenadora da CEUA/PRPI/UFG

*Comissão de Ética no Uso de Animais/CEUA*  
Pró-Reitoria de Pesquisa e Inovação/PRPI-UFG, Caixa Postal: 131, Prédio da Reitoria, Piso 1, Campus Samambaia (Campus II) -  
CEP: 74001-970, Goiânia - Goiás, Fone: (55-62) 3521-1876.  
Email: ceua.ufg@gmail.com



## TERMO DE RESPONSABILIDADE

Eu certifico que:

- a) Li os princípios éticos para realização de pesquisas envolvendo animais elaborado pelo Conselho Nacional de Controle de Experimentação Animal (CONCEA) e concordo plenamente com suas exigências durante a duração deste experimento.
- b) Este estudo não é desnecessariamente duplicativo, tem mérito científico e a equipe que participa deste projeto foi treinada e é competente para executar os procedimentos descritos nesse protocolo.
- c) Declaro que os animais serão avaliados e tratados pelo médico veterinário com experiência apropriada e responsável pela pesquisa, quando necessário;|
- d) Declaro ainda que após pesquisas em literatura atualizada, não encontrei métodos alternativos para condução dos experimentos ou aulas no referido projeto.

Desse modo, assino e dou fé.

Goiânia, 15 de dezembro de 2016.

**NOME**

Andreza Arantes de Castro

**ASSINATURA**

*Andreza Arantes Castro*