

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA GERAL
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA



TESE DE DOUTORADO

Isolamento e caracterização de marcadores moleculares de
Hoplias intermedius (Günther, 1864): **genética e forense**

ALUNA: Daniela Núñez Rodriguez

ORIENTADOR: Dr. Evanguedes Kalapothakis

CO-ORIENTADORA: Dra. Susanne Facchin

Belo Horizonte, Fevereiro 2019

Daniela Lidia Núñez Rodriguez

ISOLAMENTO E CARACTERIZAÇÃO DE MARCADORES MOLECULARES DE
Hoplias intermedius (GÜNTHER, 1864): **GENÉTICA E FORENSE**

Tese apresentada ao curso de pós-graduação em Genética do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial para à obtenção do título de Doutor em Genética.

Orientador: Dr. Evanguedes Kalapothakis

Co-orientadora: Dra. Susanne Facchin

Belo Horizonte, Fevereiro 2019

043 Rodriguez, Daniela Lidia Núñez.

 Isolamento e caracterização de marcadores moleculares de *Hoplias intermedius* (GÜNTHER, 1864): genética e forense [manuscrito] / Daniela Lidia Núñez Rodriguez. - 2019.

 103 f. : il. ; 29,5 cm.

 Orientador: Dr. Evanguedes Kalapothakis. Co-orientadora: Dra. Susanne Facchin.

 Tese (doutorado) - Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Departamento de Biologia Geral.

 CDU: 575

Ficha catalográfica elaborada pela Biblioteca do Instituto de Ciências Biológicas da UFMG



Pós-Graduação em Genética
Departamento de Biologia Geral, ICB
Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6627 - C.P. 486 - Pampulha - 31270-901 - Belo Horizonte - MG
e-mail: pg-gen@icb.ufmg.br FAX: (+31) - 3409-2570




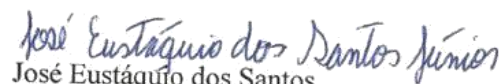
**"Isolamento e caracterização de marcadores moleculares de *Hoplias
intermedius* (Günther, 1864): genética e forense"**

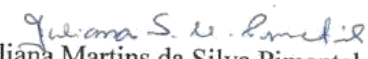
Daniela Lidia Nunez Rodriguez

Tese aprovada pela banca examinadora constituída pelos Professores:



Evanguêdes Kálapothakis, - Orientador
UFMG

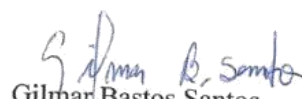

Priscilla Caroline Silva
UFV


José Eustáquio dos Santos
UFMG


Juliana Martins da Silva Pimentel
Pitágoras


Rafael Félix de Magalhães
UFMG


Susanne Facchin
UFMG


Gilmar Bastos Santos
Puc Minas

Belo Horizonte, 15 de fevereiro de 2019.

DEDICATORIA:

À minha avó, Elena de la Torre Bueno, que aprendeu a ler e escrever sozinha, se tornou amante da leitura e educou a dez filhos. Minha vó ensinou a minha mãe a ter a força que uma mulher precisa para alcançar seus sonhos, e ser a mãe amorosa, talentosa e forte que ela é.

Ao meu avô, Tomás Núñez Bazalar, filho de agricultor e primeiro doutor em ciências da cidade de Supe. Meu carinhoso avô ensinou ao meu pai com o seu exemplo a ser o melhor professor, cidadão e pai que eu poderia ter.

AGRADECIMENTOS:

Ao meu orientador, Dr. Evanguedes Kalapothakis, pela oportunidade de pertencer ao seu laboratório e por todos os ensinamentos.

À minha co-orientadora, Dra. Susanne Facchin, por me ensinar as melhores estratégias de desenho experimental que fizeram a diferença no trabalho.

Aos membros da banca, pelo tempo e paciência em avaliar o trabalho.

Aos membros da banca de qualificação, Dra. Bárbara Mendes, Dr. Daniel Carvalho e Dr. Jorge Dergam, por suas contribuições.

Ao INPA, Alessandro Bifi, professor Efrem Ferreira, Camila Ribas, PUCRS, professor Carlos Lucena, PUCMG, professor Daniel Carvalho, Tiago Pessali, Leandra Formentão, Carlos Mascarenhas Alves e Paulo Formagio, por me fornecerem as amostras e me ajudarem na identificação delas.

Aos meus pais, Edith e Tomás por serem meus guias, me ensinar que o esforço e a dedicação por tudo o que fazemos, sempre vale a pena. E porque mesmo a quilômetros de distância, sempre estivemos juntos.

Ao amor da minha vida, Emerson, meu melhor amigo e meu grande amor, porque somos amor, amizade e luta. *“Juntos, codo a codo, somos mucho más que dos”*.

À minhas irmãs, Susana e Irene, porque somos três estrelas que sempre se lembram umas as outras do próprio brilho.

A todos os colegas, amigos e amigas do Laboratório de Biotecnologia e Marcadores Moleculares, por seus ensinamentos e amizade.

Ao Leo, porque ainda sem me conhecer me levou até o rio Cipó, marcando o início do doutorado e de uma grande amizade, Emerson e eu nos sentimos felizes de conhecer você e sua família.

À Hortênsia, pela amizade tão especial e por ter a paciência de ler todas as esquinas do texto, corrigindo e me ensinando a melhorar a escrita em português.

À Renata, Danilo, Renato, Fernanda, Bárbara e Douglas, pelos momentos de paz e felicidade na água, os que me permitiram continuar com alegria, imaginação e força em tudo o que faço.

Ao Instituto de Ciências Biológicas e Universidade Federal de Minas Gerais, por me permitir cursar meus estudos em suas salas e laboratórios.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pela bolsa e financiamento da minha pesquisa.

Minha gratidão a todos.

SUMÁRIO

LISTA DE FIGURAS.....	viii
LISTA DE TABELAS.....	x
LISTA DE ABREVIACÕES	xi
ABSTRACT	1
RESUMO	2
1. INTRODUÇÃO.....	3
1.1. A DIVERSIDADE DA ICTIOFAUNA BRASILEIRA, MANEJO E MEDIDAS DE CONTROLE DA PESCA	3
1.2. O TRÁFICO ILEGAL DE ESPÉCIES E O USO DE NOMES COMERCIAIS PARA PRODUTOS PESQUEIROS	4
1.2.1. Sequenciamento Sanger e Sequenciamento de Segunda Geração.....	5
1.2.2. DNA barcoding e marcadores espécie-específicos	6
1.2.2.1. Citocromo b (Cyt b)	8
1.2.2.2. Citocromo C Oxidase 1 (COI)	8
1.2.2.3. Genes mitocondriais ribossomais: 12SrRNA e 16SrRNA.....	8
1.2.2.4. D-loop.....	9
1.3. FERRAMENTAS BIOINFORMÁTICAS PARA DETECÇÃO DE MARCADORES MOLECULARES	9
1.4. O GÊNERO <i>Hoplias</i> Gill, 1903: EXEMPLO DA COMPLEXIDADE TAXONÔMICA EM RECURSOS	
PESQUEIROS.....	11
1.4.1. Histórico taxonômico do gênero <i>Hoplias</i>	11
1.4.2. Importância comercial e ecológica de <i>Hoplias intermedius</i> e <i>H. malabaricus</i>	15
1.4.3. Diversidade molecular e citogenética das espécies de <i>Hoplias</i>	16
2. OBJETIVOS	18
2.1. Objetivo geral:	18
2.2. Objetivos específicos:.....	18
3. MATERIAIS E MÉTODOS.....	18
3.1. Obtenção de amostras	18
3.2. Isolamento do DNA genômico:.....	20
3.2.1. Método de extração 1: Fenol e clorofórmio	20
3.2.2. Método de extração 2: Kit de Extração de DNA GTS (Pht).....	21
3.3. Quantificação de DNA genômico e análise de integridade	22
3.4. Construção da biblioteca genômica	23
3.5. Sequenciamento de Nova Geração de <i>H. intermedius</i>	25
3.6. Desenho de <i>primers</i> mitocondriais	28

3.7. Padronização de Reações em Cadeia da Polimerase	29
3.8. Sequenciamento de Sanger	29
3.9. Identificação de amostras	31
3.10. Testes de validação dos <i>primers</i>	31
3.10.1. Teste de Sensibilidade	31
3.10.2. Teste de Especificidade	31
3.10.3. Teste de Eficiência	32
4. RESULTADOS	33
4.1. Genoma mitocondrial completo de <i>H. Intermedius</i>	33
4.3. Análise comparativa e agrupamento filogenético de mitogenomas	39
4.4. Localização de marcadores mitocondriais espécie-específicos	42
4.5. Desenho de <i>primers</i> mitocondriais	45
4.6. Padronização de PCR	50
4.7. Testes de Validação	54
4.7.1. Teste de sensibilidade	54
4.7.2. Teste de especificidade	56
4.7.3. Teste de eficiência:	59
A. Teste de eficiência em amostras coletadas em campo:	59
B. Teste de eficiência em amostras obtidas em coleções científicas	61
C. Testes de eficiência dos marcadores específicos para <i>H. intermedius</i> e <i>H. malabaricus</i> usando amostras obtidas em peixarias	62
5. DISCUSSÃO DE RESULTADOS	66
6. CONCLUSÕES	70
7. REFERÊNCIAS BIBLIOGRÁFICAS	71
8. ANEXOS	80
1.1. Sequencias dos amplicons obtidos a partir das amostras do rio doce usando o marcador específico para <i>H. intermedius</i> Hin5F-Hin3R	92
1.2. Sequencias dos amplicons obtidos a partir das amostras do rio doce usando o marcador específico para <i>H. intermedius</i> Hin5F-Hin4R	95

LISTA DE FIGURAS

Figura 1. Esquema adaptado de Mattox et al. (2006) e Oyakawa e Mattox (2009) da região ventral de duas espécies de <i>Hoplias</i> ..	13
Figura 2. Espécies do grupo <i>Hoplias lacerdae</i> , <i>H. aimara</i> e <i>H. microlepis</i> representando o grupo <i>H. malabaricus</i> ..	14
Figura 3. Mapa de distribuição geográfica das amostras analisadas.....	19
Figura 4. Eletroforese de DNA usado na montagem da biblioteca genômica de <i>H. intermedius</i> ..	33
Figura 5. Distribuição do Q score do Sequenciamento de Nova Geração de <i>H. intermedius</i>	34
Figura 6. Gráfico de distribuição da qualidade das <i>reads</i> obtidas de <i>H. intermedius</i>	35
Figura 7. Mapa de anotação do mitogenoma de <i>H. intermedius</i>	36
Figura 8. Árvore de consenso Neighbour-joining usando 38 mitogenomas completos de Characiformes e Siluriformes..	40
Figura 9. Comparação de mapas de anotação dos mitogenomas de 24 espécies da ordem Characiformes.....	42
Figura 10. Representação gráfica dos nucleotídeos diagnósticos localizados pelo SPIDER usando o mitogenoma de <i>H. intermedius</i> e seis amplitudes de janela diferentes.....	43
Figura 11. Representação gráfica dos nucleotídeos diagnósticos localizados pelo SPIDER usando o mitogenoma de <i>H. malabaricus</i> e seis amplitudes de janela diferentes.....	44
Figura 12. Mapa de anelamento dos <i>primers</i> espécie-específicos desenhados para <i>H. intermedius</i>	46
Figura 13. Mapa de anelamento dos <i>primers</i> espécie-específicos desenhados para <i>H. malabaricus</i>	47
Figura 14. Padronização da PCR usando o marcador Hin5F-Hin3R específico para <i>H. intermedius</i> ..	53
Figura 15. Padronização da PCR usando o marcador Hin5F-Hin4R específico para <i>H. intermedius</i> ..	53
Figura 16. Padronização da PCR usando o marcador Hin5F-Hin5R específico para <i>H. malabaricus</i> ..	54
Figura 17. Teste de sensibilidade dos <i>primers</i> específicos para <i>H. intermedius</i> Hin5F-Hin3R.....	55
Figura 18. Teste de sensibilidade dos <i>primers</i> específicos para <i>H. intermedius</i> Hin5F-Hin4R.....	55
Figura 19. Teste de Especificidade dos <i>primers</i> específicos para <i>H. intermedius</i>	57
Figura 20. Teste de Especificidade dos <i>primers</i> Homa5F-Homa5R específicos para <i>H. malabaricus</i> ..	58

Figura 21. Teste de controle de qualidade do DNA molde usado nos testes de especificidade.....	58
Figura 22. Identificação molecular de amostras coletadas no rio Doce usando o marcador específico para <i>H. intermedius</i> Hin5F-3R.....	59
Figura 23. Identificação molecular de amostras coletadas no rio Doce usando o marcador específico para <i>H. intermedius</i> Hin5F-Hin4R.....	59
Figura 24. Controle de qualidade do DNA usando o marcador universal para peixes Fish16S e as amostras coletadas no rio Doce.	60
Figura 25. Teste de eficiência do marcador específico para <i>H. intermedius</i> (Hin5F-Hin4R) usando amostras do rio São Francisco.....	61
Figura 26. Teste de eficiência do marcador específico para <i>H. malabaricus</i> (Homa5F-Homa5R) usando amostras do rio São Francisco.....	61
Figura 27. Controle de qualidade da integridade do DNA das amostras coletadas no rio São Francisco usando o marcador universal para peixes Fish16S..	61
Figura 28. Identificação molecular de amostras coletadas em peixarias do rio Grande usando o marcador específico para <i>H. intermedius</i> Hin5F-Hin3R..	62
Figura 29. Identificação molecular de amostras coletadas em peixarias do rio Grande usando o marcador específico para <i>H. intermedius</i> Hin5F-Hin4R..	63
Figura 30. Controle de qualidade da integridade do DNA das amostras coletadas no rio Grande usando o marcador universal para peixes Fish16S..	63
Figura 31. Identificação molecular usando amostras de filés congelados de “trairão” e o marcador específico para <i>H. intermedius</i>	64
Figura 32. Identificação molecular de <i>H. malabaricus</i> usando o marcador específico Homa5F-Homa5R e amostras de filés congelados de “trairão”	65
Figura 33. Controle de qualidade da integridade do DNA das amostras obtidas a partir de filés congelados de “trairão” usando o marcador universal para peixes Fish16S.	65

LISTA DE TABELAS

Tabela 1. Volumes dos reagentes utilizados na qPCR de controle de qualidade da biblioteca genômica.	25
Tabela 2. Sequências mitogenômicas de peixes recuperadas do NCBI.....	27
Tabela 3. Quantificação do DNA para sequenciamento de Sanger.....	30
Tabela 4. Dados quantitativos e qualitativos da anotação do mitogenoma de <i>H. intermedius</i>	37
Tabela 5. Características das sequências de <i>primers</i> espécie específicos para <i>H. intermedius</i> e <i>H. malabaricus</i>	48
Tabela 6. <i>Primers</i> espécie-específicos para <i>H. intermedius</i> e <i>H. malabaricus</i>	50
Tabela 7. Volumes utilizados nos testes de padronização das reações de PCR usando DMSO como agente adstringente.....	51
Tabela 8. Volumes utilizados nos testes de padronização das reações de PCR usando KCl como agente adstringente.....	51
Tabela 9. Volumes utilizados nos testes de padronização das reações de PCR sem usar agente adstringente.....	52
Tabela 10. Condições ótimas de amplificação de <i>primers</i> espécie-específicos de <i>H. intermedius</i> ..	52
Tabela 11. Sensibilidade dos <i>primers</i> específicos para <i>H. intermedius</i> e <i>H. malabaricus</i>	54
Tabela 12. Especificidade dos <i>primers</i> específicos para <i>H. intermedius</i> e <i>H. malabaricus</i>	56
Tabela 13. Lista de tecidos analisados a partir de coleções científicas.	81
Tabela 14. Localização de possíveis marcadores específicos para <i>H. intermedius</i>	87
Tabela 15. Localização de possíveis marcadores específicos para <i>H. malabaricus</i>	90

LISTA DE ABREVIÇÕES

°C	Graus Celsius
BLAST	<i>Basic Local Alignment Search Tool</i>
BOLD	Barcode of Life Datasystems
CE	Eletoforese capilar <i>Convention on International Trade in Endangered Species of Wild Fauna and Flora</i>
CITES	<i>and Flora</i>
CTAB	Brometo de Cetil Trimetilamonio
dATP	Desoxiadenosina trifosfato
dCTP	Desoxicitosina trifosfato
dDNTP	Dideoxinucleotídeos
dGTP	Desoxiguanina trifosfato
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
dNTP	Desoxi (nucleotídeo) 5' -trifosfato
dTTP	Desoxitimina trifosfato
EDTA	Ácido etilenodiamino tetra-acético
F	<i>Primer</i> tipo forward
GC	Guanina-Timina
HCl	Ácido clorídrico
IBAMA	Instituto Brasileiro do Meio Ambiente e Recursos Renováveis
KCl	Cloreto de potássio
kPa	quilopascal
ml	mililitro
mM	milimolar
mtDNA	DNA mitocondrial
NaCl	Cloreto de sódio
NCBI	<i>National Center for Biotechnology Information</i>
Ng	nanograma
NGS	<i>Next Generation Sequencing</i>
NPL	<i>Nextera Library Plate</i>
NPM	<i>PCR Master Mix</i>
OL	Origem de replicação da cadeia leve
PB	Pares de base
PB	<i>Binding Buffer</i>

PCGs	Regioes codificadoras de proteina
PCR	Reação em Cadeia da Polimerase
PCR-CE	PCR-eletroforese capilar
PPC	<i>Primer coquetel</i>
R	<i>Primer tipo reverso</i>
RBS	<i>Resuspension Buffer</i>
RNA	Ácido ribonucleico
rRNAs	RNA ribossomais
SNP	<i>Single nucleotide polymorphism</i>
TDE1	<i>Tagment DNA Enzyme</i>
TE	Tampão Tris e EDTA
uL	microlitro

1 **ABSTRACT**

2

3 The conservation of Brazilian ichthyofauna is facing a growing threat, which highlights
4 the importance of new tools for the enrichment of conservation measures. Thus, Forensic
5 Genetics and the design of molecular markers, emerge to improve illegal fishing control
6 activities and species detection on poorly preserved samples. The choice of the target species
7 of this study, *Hoplias intermedius* (Günther, 1864), responds to the environmental impact of
8 the basins where species is distributed, its increasing importance as a fishing resource and
9 the great taxonomic complexity of the genus. The objectives of our study include the
10 sequencing and assembly of the whole mitochondrial genome, design of mitochondrial
11 species-specific primers for *H. intermedius* and its validation as a mitochondrial molecular
12 marker for its applications in forensic aims. DNA was extracted and a genomic library was
13 constructed by using new generation sequencing. Fragments were obtained with Nextera DNA
14 Sample Preparation kit and MiSeq Reagent kit V3-600 paired-end. The mtDNA assembly was
15 made with CLC Workbench software v8.5.1 obtaining coverage of 304.45% and its annotation
16 was performed on the MITOfish platform, showing a typical mitogenome arrangement of
17 vertebrates (13 protein coding genes, two ribosomal RNA genes, 22 RNA transporter genes
18 and one control region). MEGA6 software was used for the alignment of mitochondrial
19 sequences of *H. intermedius* and other mtDNA with 99% of identity. The detection of
20 informative sequence regions for the target species was validated by SPIDER R-package.
21 Sensitivity, specificity and efficiency tests were performed by using samples collected in the
22 field, from scientific collections and from fishmongers. PCR reactions were standardized for
23 each primer combination using 5U/ul Taq Polymerase, 2mM dNTPs, and specific combinations
24 of buffer solutions IB, IC, IIC and 10X IVB, 2M KCl (Phoneutria) and DMSO. The annealing
25 temperature range varied among 59°, 62° and 64° C and the number of cycles between 30
26 and 35. Results showed that the mtDNA of *H. intermedius* is a circular DNA molecule of 16629
27 bp, with 43.97% GC content. The ATG start codon was found in 13 non-coding protein genes
28 encoded in the (+) strand, and as expected the ND6 gene encoded in the (-) strand. *Primer*
29 *pairs* Hin5F-Hin3R and Hin5F-Hin4R were highly specific for *H. intermedius*, both design for
30 the 16S rRNA region. Amplicon sequences were obtained by using bidirectional sequencing,
31 and subsequently constructed an NJ tree for each marker. Our results demonstrate that
32 molecular markers for *H. intermedius* pursue a huge potential to identify the target species,
33 and its application includes integrative taxonomy for *Hoplias* species and wildlife forensic
34 identification of poor DNA quality samples.

35 Key words: Characiform, forensic DNA, mitogenome, neotropical fish, NGS, "trairão".

1 RESUMO

2

3 A conservação da ictiofauna brasileira está enfrentando uma ameaça crescente, que
4 destaca a importância de novas ferramentas para o enriquecimento das medidas de
5 conservação de peixes. Assim, a Genética Forense e o desenho de marcadores moleculares
6 surgem para melhorar as atividades de controle da pesca ilegal e a detecção de espécies em
7 amostras mal preservadas. A escolha da espécie alvo deste estudo, *Hoplias intermedius*, está
8 relacionada ao impacto ambiental das bacias onde a espécie se encontra distribuída, sua
9 crescente importância como recurso pesqueiro e a grande complexidade taxonômica do
10 gênero. Os objetivos do nosso estudo incluem sequenciamento e montagem do genoma
11 mitocondrial completo, desenho de *primers* específicos mitocondriais para *H. intermedius* e
12 sua validação como marcador molecular mitocondrial visando sua aplicação com fins
13 forenses. O DNA foi isolado e uma biblioteca genômica foi construída com o uso de
14 sequenciamento de nova geração. Os fragmentos foram obtidos com o kit Nextera DNA
15 Sample Preparation e com o kit de reagentes MiSeq V3-600. A montagem do mtDNA foi feita
16 com o software CLC Workbench v8.5.1, obtendo uma cobertura de 304,45%. A anotação do
17 genoma mitocondrial foi realizada na plataforma MITOfish, mostrando um arranjo
18 mitogenômico típico de vertebrados (13 genes codificadores de proteínas, dois genes de RNA
19 ribossômicos, 22 genes transportadores de RNA e uma região de controle). O software
20 MEGA6 foi utilizado para o alinhamento de sequências mitocondriais de *H. intermedius*. A
21 detecção de regiões informativas para a espécie-alvo foi validada pelo pacote do R, SPIDER.
22 Foram realizados também testes de sensibilidade, especificidade e eficiência utilizando
23 amostras coletadas em campo, amostras de coleções científicas e amostras de peixarias. As
24 reações de PCR foram padronizadas para cada combinação de *primers* utilizando Taq
25 polimerase 5U/ul, 2 mM dNTPs, 2M KCl e combinações específicas de soluções tampão IB,
26 IC, IIC e 10X IVB, (Phoneutria) e DMSO. A faixa de temperatura de anelamento variou entre
27 59°, 62° e 64°C e o número de ciclos de extensão entre 30 e 35. Os resultados mostraram
28 que o mtDNA de *H. intermedius* é uma molécula de DNA circular de 16629 pb, contendo GC
29 de 43,97%. O códon de início ATG foi encontrado em 13 genes não codificantes de proteínas,
30 codificados na cadeia (+) e, como esperado, o gene *ND6* codificado na cadeia (-). Os pares
31 de *primers* Hin5F-Hin3R e Hin5F-Hin4R foram altamente específicos para *H. intermedius*,
32 ambos projetados na região 16S rRNA. As sequências de amplificação foram obtidas para
33 ambos os marcadores usando sequenciamento bidirecional e posteriormente foi construída
34 uma árvore NJ para cada marcador. Nossos resultados demonstram que os marcadores
35 moleculares para *H. intermedius* têm um enorme potencial para identificar a espécie-alvo, e
36 sua aplicação inclui taxonomia integrativa para espécies de *Hoplias* e identificação forense
37 usando amostras com DNA de baixa qualidade.

38

39 Palavras chave: Characiformes, DNA forense, Mitogenoma, ictiofauna Neotropical NGS,
40 "Trairão".

1. INTRODUÇÃO

1.1. A DIVERSIDADE DA ICTIOFAUNA BRASILEIRA, MANEJO E MEDIDAS DE CONTROLE DA PESCA

Estima-se que a ictiofauna continental e costeira da América do Sul atualmente está compreendida em aproximadamente 9100 espécies, representando o 27% da diversidade de peixes do mundo (Reis *et al.*, 2016). De acordo com o grau de endemismo e diversidade, as bacias da América do Sul foram divididas e classificadas em 13 complexos. Entre os complexos de bacias do Amazonas, Orinoco e Paraná-Paraguay, estão registradas 3599 espécies de peixes (Reis *et al.*, 2016). Em quarto lugar, a bacia do rio São Francisco representa a quarta maior e mais diversa bacia do continente, com uma extensão de 593 mil km² (Paredes *et al.*, 1983) e sendo registradas 211 espécies de peixe, sendo 64% registros endêmicos para a região (Reis *et al.*, 2016).

De acordo com a classificação em regiões biogeográficas proposta por *Abell et al.*, (2008), as bacias do complexo leste da costa Atlântica incluem os rios Vaza-Barris, Itapicuru, Paraguaçu, Contas, Pardo, Jequitinhonha, Mucuri, São Mateus, Doce, Paraíba do Sul, Macaé e Macacu, as quais se encontram completamente contidas em território brasileiro. Sua extensão, a diversidade da ictiofauna e o alto grau de endemismo das bacias que a compõem, a posiciona como uma importante unidade biogeográfica (Abell *et al.*, 2008). Estas regiões de grande diversidade enfrentam fortes impactos que minoram a sobrevivência de espécies, entre essas fontes de impacto estão a construção de barragens, o lançamento de esgotos urbanos e industriais sem tratamento, os projetos de irrigação, as atividades agropecuárias sem manejo adequado do solo, a eliminação da vegetação ripária, as atividades extrativistas, os projetos de navegação, a pesca ilegal e a introdução de espécies não nativas (Agostinho *et al.*, 2005).

Atualmente a ictiofauna se mantém como um recurso de grande importância, abastecendo as necessidades alimentares mundialmente. Um grande número de espécies de peixes é comercializado para o consumo e produção de alimentos, representando 17% do total de proteína animal consumida a nível mundial. Além disso, segundo a FAO (2016) foi registrado para esse ano o maior incremento de consumo per-capita em países como Brasil, Chile, China, México e Peru (FAO, 2016). Na Amazônia, a pesca representa a principal fonte econômica e alimentar estando entre as principais ameaças à fauna da região, pois apesar do comércio de peixes estar restrito a mercados locais, a captura tem registrado 260 mil toneladas por ano (Ardura *et al.*, 2010).

1 A ictiofauna possui um papel fundamental na dinâmica da cadeia trófica de
2 ecossistemas aquáticos naturais, e participa ativamente no entorno socioeconômico e na
3 pesca recreativa. Porém, a manutenção destas atividades econômicas muitas vezes facilita a
4 introdução de espécies não nativas, gerando impactos negativos na diversidade genética de
5 populações nativas e tornando o ecossistema vulnerável (Fragoso-Moura *et al.*, 2016).

6 A fim de apoiar legalmente a preservação da diversidade biológica surgiram
7 legislações para a proteção de espécies, representada internacionalmente pela Convenção
8 sobre Comércio Internacional de Espécies Ameaçadas de Extinção (CITES) e localmente pelo
9 Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA). Desta
10 forma, são formuladas as leis ambientais que regulam a pesca, incluindo medidas de
11 conservação como a interdição temporária da pesca e a inspeção e multas por danos
12 ambientais (Vieira *et al.*, 2009).

13 **1.2. O TRÁFICO ILEGAL DE ESPÉCIES E O USO DE NOMES COMERCIAIS PARA** 14 **PRODUTOS PESQUEIROS**

15 Publicações científicas recentes apontam que algumas espécies dos gêneros *Brycon*
16 Müller & Troschel, 1844, *Leporinus* Agassiz, 1829, *Pimelodus* Lacepède, 1803 e *Prochilodus*
17 Agassiz, 1829, foram categorizados pelo Ministério do Meio Ambiente no ano de 2014
18 (portaria nº 444 e 445) como vulneráveis ou ameaçados de extinção (Chagas *et al.*, 2015).
19 Porém, a última estatística pesqueira foi publicada em 2011 pelo extinto Ministério da Pesca
20 e Aquicultura, sendo ainda escassos os estudos sobre a regulação da pesca que informem
21 dados oficiais a respeito da biomassa e das características geográficas e biológicas da pesca
22 ilegal de água doce no Brasil.

23
24 O uso incorreto das chaves de identificação e caracteres diagnósticos de espécies,
25 muitas vezes gera o agrupamento errôneo de espécies diferentes dentro de um mesmo nome
26 comum, representando uma ameaça para a gestão da pesca. Como agravante a esta
27 problemática, frequentemente, as evidências das apreensões ilegais são amostras de
28 natureza diversa. Os produtos apreendidos variam desde filés congelados, espécimes
29 decompostos ou inclusive amostras já processadas pela indústria de alimentos (alimentos em
30 conserva, medicamentos, etc.), tornando sua identificação um desafio.

31 As medidas de controle da pesca são dificultadas devido a esta identificação
32 imprecisa, onde as espécies são usualmente comercializadas usando nomes incorretos. Não
33 é incomum o agrupamento de gêneros e inclusive famílias diferentes dentro de um mesmo
34 táxon específico (Ardura *et al.*, 2010; de Azevedo Chagas *et al.*, 2015). Os erros de

1 identificação e a rotulagem incorreta de produtos derivados da pesca ocorrem regularmente
2 com a finalidade de mascarar uma espécie de valor módico por outra de maior qualidade e
3 maior valor no mercado (Logan *et al.*, 2008; Carvalho *et al.*, 2011). Essas substituições ilegais
4 são favorecidas pela dificuldade ou impossibilidade de reconhecimento das espécies usando
5 caracteres morfológicos em produtos de músculo congelado ou processado pela indústria
6 alimentar. No Brasil o Programa de Proteção e Defesa do Consumidor (PROCON) tem
7 penalizado economicamente, utilizando técnicas moleculares como *DNA barcoding* para
8 realizar a identificação das produtos apreendidos.(Carvalho *et al.*, 2015).

9
10 Assim, é fundamental reconhecer quais espécies ocorrem em uma determinada área
11 geográfica e quais delas são alvo de comércio, ressaltando a urgência em desenhar, validar
12 e aplicar ferramentas que permitam identificar os recursos pesqueiros, principalmente em
13 áreas geográficas que apresentam um rápido crescimento econômico, como é o caso das
14 grandes bacias hidrográficas brasileiras (Jenkins, 2003). Neste contexto, o controle e
15 detecção de espécies exóticas, o monitoramento de indivíduos reproduzidos em cativeiro pela
16 aquicultura, a identificação imprecisa de espécies nativas, a fraude alimentar e a dificuldade
17 de identificação de amostras processadas ou degradadas, representam um grande desafio
18 para o controle da pesca (Karlsson *et al.*, 2014).

19 De acordo com a revisão publicada pela FAO, "*Fish Identification Tools for Fishery*
20 *Biodiversity and Fisheries Assessments*", as ferramentas utilizadas para identificação de
21 peixes variam desde genética molecular, tecnologias de imagens, hidro acústica e
22 morfometria, sendo a identificação taxonômica o eixo fundamental de todas elas. Porém,
23 devido a um escasso financiamento em pesquisa básica, cada vez com menor frequência são
24 oferecidos treinamentos em identificação taxonômica, provocando identificações pouco
25 exatas. A identificação correta é relevante não unicamente para fins científicos, mas também
26 econômicos, sendo fundamental inclusive para o sucesso da reprodução de um recurso
27 determinado para fins de aquicultura (Fischer, 2013).

28 29 **1.2.1. Sequenciamento Sanger e Sequenciamento de Segunda Geração**

30
31 O Sequenciamento de Sanger em eletroforese capilar (Sanger *et al.*, 1977) é um
32 método de sequenciamento baseado na terminação de cadeia. A reação de sequenciamento
33 é composta de um único *primer* (*forward* ou *reverso*), DNA molde, DNA polimerase, dNTPs
34 (dATP, dTTP, dGTP, and dCTP) e dideoxynucleotídeos (ddNTPs) marcados com um corante
35 diferencial para cada um. Logo os produtos da reação de sequenciamento são analisados
36 usando eletroforese capilar.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

O Sequenciamento de Nova Geração, também chamado de Sequenciamento Massivo em Paralelo ou Sequenciamento de Segunda Geração (Raza and Ahmad, 2016), têm proporcionado tecnologias, que em comparação ao método de Sanger, não requerem por exemplo uma metodologia baseada na clonagem bacteriana para fragmentos de DNA. Os métodos de *Next Generation Sequencing* (NGS) se baseiam na preparação de bibliotecas, as quais são moléculas de DNA ou RNA desnaturadas que se ligam a uma superfície sólida que contém os adaptadores. Logo é realizada a síntese da fita complementar a cada fragmento ligado ao adaptador. São feitos de milhares a milhões ciclos desta amplificação, obtendo diretamente o produto do sequenciamento (Yang *et al.*, 2014).

Uma desvantagem da tecnologia de NGS é o tamanho curto dos fragmentos gerados, o que resulta em dificuldades de montagem, anotação e outras análises bioinformáticas (Jaszczyszyn *et al.*, 2014), embora essa condição não represente uma grande desvantagem para estudos onde se analisam marcadores mitocondriais. Um dos motivos disto é que inclusive uma pequena amostra de tecido representa um grande número de cópias de DNA mitocondrial por cada célula contida na amostra.

1.2.2. DNA barcoding e marcadores espécie-específicos

A identificação de espécies usando marcadores genéticos busca isolar regiões no genoma que sejam diferenciais e conservadas dentro do grupo taxonômico alvo. Em geral, a escolha entre as várias possibilidades de marcadores a serem usados depende de alguns fatores entre os quais se destacam o custo e a eficácia para identificar o indivíduo, espécie ou população para o qual foi desenhado. A abordagem para a identificação genética de espécies geralmente inclui a amplificação de um fragmento mitocondrial, podendo ser feita com *primers* universais ou com *primers* espécie-específicos.

O genoma mitocondrial é uma molécula circular, que nos vertebrados possui um tamanho variável entre 16 e 17 kb. Sua estrutura é altamente conservada e inclui 13 genes codificadores de proteína (PCGs), dois genes ribossomais (rRNAs), 22 genes de transferência e duas regiões não codificadoras que regulam a replicação e a transição do mtDNA: a região controle (D-loop) e a origem de replicação da cadeia de leve (OL) (Bedore *et al.* 2017; Brandao-Dias *et al.* 2014; Carmo *et al.* 2014; Chagas *et al.* 2015; Núñez-Rodriguez *et al.* 2016; Pimentel *et al.* 2014; Resende *et al.* 2016; Satoh *et al.* 2016; Siqueira *et al.* 2014). A molécula de DNA mitocondrial não possui histonas o que faz com que a reparação da molécula seja

1 limitada e a taxa de mutação 10 vezes maior que no DNA nuclear (Brown *et al.*, 1979). Devido
2 a que cada gene possui uma taxa de evolução particular, o genoma mitocondrial é
3 frequentemente usado para testar relações filogenéticas. Por outro lado, comparado com o
4 genoma nuclear, o tamanho de sua molécula é menor e possui um maior número de cópias
5 por célula, o que permite o uso de pequenas quantidades de tecido e obter uma alta qualidade
6 nas análises (Arif *et al.*, 2011).

7
8 Em muitas espécies existem mecanismos que eliminam o mtDNA do esperma,
9 garantizando que o DNA mitocondrial seja herdado somente da mãe (Chan e Schon, 2012).
10 Em muitos vertebrados, incluindo os peixes, esses mecanismos incluem a digestão
11 enzimática do mtDNA do esperma (Nishimura *et al.*, 2006), reduzindo a diversidade
12 intraespecífica (Dawnay *et al.*, 2007; Rojas *et al.*, 2011; Yang *et al.*, 2014).

13
14 Os marcadores universais desenhados no mitogenoma têm sido amplamente usados
15 para identificar a biodiversidade e testar suas relações filogenéticas a partir da técnica do DNA
16 barcoding. Estes *primers* são desenhados em regiões conservadas entre um grupo alvo que
17 pode incluir espécies, gêneros, famílias ou até ordens diferentes, sendo a identificação do
18 produto da amplificação analisado posteriormente por sequenciamento, análise filogenética e
19 teste de similaridade da sequência obtida em bancos de dados públicos como NCBI e BOLD
20 (Farrell *et al.*, 2000).

21
22 Por sua vez, os *primers* espécie-específicos irão anelar em regiões exclusivas para
23 uma única espécie, sendo o desenho desse tipo de *primer* feito em regiões de alta
24 concentração de *mismatches* entre a sequência alvo e as sequências filogeneticamente mais
25 próximas a ela. A vantagem dos *primers* específicos para uma espécie única consiste na
26 identificação na etapa de amplificação, sem a necessidade da reação de sequenciamento.
27 Isto só é feito após os *primers* passar por um processo de validação, diminuindo os custos
28 das análises de identificação.

29
30 Frequentemente, os *primers* desenhados no genoma mitocondrial estão baseados na
31 sequência completa ou parcial do citocromo oxidase c subunidade 1 (*Cox 1* ou *COI*), porém
32 sua escolha não necessariamente representa o marcador mais apropriado e informativo para
33 o grupo taxonômico alvo (Hebert *et al.*, 2003; Hebert *et al.*, 2003; Morgan *et al.*, 2005; Moritz
34 & Cicero, 2004).

1 Por esse motivo, listamos brevemente algumas das principais características
2 moleculares dos genes mitocondriais que tem demonstrado uma grande eficácia em estudos
3 que envolvem identificação molecular de espécies:

4 5 **1.2.2.1. Citocromo b (Cyt b)**

6 É codificado na fita pesada, e compreende uma região de tamanho aproximado de
7 1140 pares de bases, sendo sua posição na molécula dependente da espécie e relativa à
8 origem de replicação. É a única proteína sintetizada pelo genoma mitocondrial que pertence
9 ao conjunto das 10 proteínas do Complexo III do Sistema mitocondrial de fosforilação
10 oxidativa, conferindo um alto grau de identidade entre diferentes organismos. Esta
11 característica permite inferir relações filogenéticas, apresentando ao mesmo tempo uma
12 grande variação interespecífica e menor variação intraespecífica (Parson *et al.*, 2000). Tobe
13 *et al.* (2009) realizaram uma análise comparativa utilizando as sequências de 236 mamíferos,
14 onde se mostrou que os genes *COI* e cytochrome b (*Cyt b*) são igualmente informativos para
15 o grupo, porém, na região do *Cyt b* é possível utilizar fragmentos menores para identificar
16 amostras a nível específico (Palomares *et al.*, 2002; Ciavaglia *et al.*, 2015).

17 18 **1.2.2.2. Citocromo C Oxidase 1 (COI)**

19 Pertence ao Complexo IV do sistema mitocondrial de fosforilação oxidativa. É uma das
20 regiões mitocondriais mais conservadas e possui um tamanho de aproximadamente 1550 pb.
21 Sua taxa de evolução reduzida se deve ao fato deste gene codificar proteínas necessárias
22 para a produção de energia celular, fazendo com que esta região seja altamente conservada
23 (Hebert *et al.*, 2003).

24 A técnica do DNA barcoding foi inicialmente proposta por Hebert *et al.* (2003), e usa
25 frequentemente um fragmento de 650 pb do Citocromo oxidase c subunidade 1. Esta
26 metodologia tem demonstrado alta eficiência em muitos grupos biológicos, obtendo-se 90%
27 de sucesso na identificação da ictiofauna marinha e continental (Carvalho *et al.*, 2011; Dawnay
28 *et al.*, 2007; Pereira *et al.*, 2013). Esta técnica tem sido aprimorada para diversas espécies,
29 sendo desenhados *primers* para fragmentos curtos do gene *COI* (entre 200 pb e 300 pb),
30 chamados mini-barcodes (Hajibabaei *et al.*, 2006; Meusnier *et al.*, 2008), porém sua
31 performance é melhor em produtos processados pela indústria alimentar (Hellberg *et al.*,
32 2017).

33 **1.2.2.3. Genes mitocondriais ribosomais: 12SrRNA e 16SrRNA**

34 Os genes ribosomais estão codificadas pelo seu tamanho, sendo a subunidade 16S
35 maior que a subunidade 12S, com um tamanho de aproximadamente de 1559 pares de bases

1 (pb) e 959 pb, respectivamente. O papel deles no metabolismo celular é realizar a tradução
2 do RNA mensageiro em proteínas mitocondriais. Embora possuam muitas substituições, a
3 estrutura de ambos os genes compartilha funções e estruturas em diversos organismos (Yang
4 *et al.*, 2014). Foran *et al.*, (2015) testaram diferentes marcadores mitocondriais (12S rRNA,
5 16S rRNA, *Cyt b*, *COI* e região controladora) em diversas espécies de mamíferos, peixes,
6 aves, répteis não aviários, anfíbios, insetos, artrópodes e anelídeos, nos quais os marcadores
7 12S rRNA e 16S rRNA obtiveram maior sucesso de amplificação esses genes não se
8 mostraram eficientes para identificar um grande número de espécies usando marcadores
9 universais.

11 **1.2.2.4. D-loop**

12 No DNA mitocondrial existem duas regiões não codificantes que abrangem menos de
13 7% do total da sequência, das quais, a maior e principal delas é o loop de deslocamento ou
14 D-loop, contendo a característica tripla fita (Larizza *et al.*, 2002). Nesta região encontram-se
15 os promotores de transcrição e de o origem de replicação, servindo assim como região de
16 controle da expressão do mtDNA (Branch *et al.*, 2011). A região não codificante evolui de duas
17 a cinco vezes mais rápido que o citocromo b (*Cyt b*) e o citocromo oxidase I (*COI*) (Jamandre
18 *et al.*, 2014).

19 Devido a sua grande variabilidade, a região terminal da extremidade 5' do D-loop é
20 amplamente utilizada para resolver questões evolutivas e populacionais nos níveis inter e
21 intraspecífico em espécies da ictiofauna, embora essa alta variabilidade não ocorra em todos
22 os táxons (Jamandre *et al.*, 2014). Devido a sua alta taxa de variabilidade, quando a região d-
23 Loop é usada como método de identificação, é esperada uma eficiência limitada para
24 identificar espécies filogeneticamente mais próximas (Yang e Speller, 2006). Porém, regiões
25 altamente variáveis possuem também potencial informativo, sendo que a região controladora
26 D-loop tem sido usada para amplificar e identificar com sucesso seis espécies de peixe a partir
27 de DNA antigo extraído de ossos (Flamingh *et al.*, 2018).

28 **1.3. FERRAMENTAS BIOINFORMÁTICAS PARA DETECÇÃO DE MARCADORES** 29 **MOLECULARES**

30
31 Para facilitar e otimizar a localização de potenciais marcadores moleculares existem
32 ferramentas bioinformáticas como o *Species Identity and Evolution in R (SPIDER)* que localiza
33 potenciais regiões informativas para o grupo alvo, além do PRIMERblast que permite o
34 desenho de *primers*.

1 O SPIDER é um pacote em linguagem R que localiza fragmentos informativos com
2 base na teoria de DNA barcode (Brown *et al.*, 2012). Desta forma, são analisados dados
3 filogenéticos gerados por outros pacotes de análises filogenéticas (*Adgenet*, *MASS*, *Pegas* e
4 *Ape*). Logo, as sequências de DNA são divididas em fragmentos de tamanho predeterminado
5 pelo pesquisador. Estes fragmentos são chamados pelo *SPIDER* de “janelas” (Brown *et al.*,
6 2012). Após o alinhamento do mitogenoma alvo com mitogenomas de outras espécies, o
7 *SPIDER* analisa o número de *mismatches* presentes em cada fragmento. Embora o *SPIDER*
8 não forneça as sequências que amplifiquem estes fragmentos detectados, a o programa
9 permite incrementar a confiabilidade das regiões mais informativas escolhidas como
10 marcador.

11
12 O desenho dos *primers* que irão amplificar o fragmento escolhido pode ser feito no
13 *PRIMERblast*, onde as sequências do *primer* são avaliadas em base a sua capacidade para
14 amplificar o fragmento, o tamanho do fragmento esperado, características físico-químicas das
15 sequências, probabilidade de autocomplementaridade e especificidade *in silico* com as
16 sequências do banco de dados do NCBI. As características físico-químicas das sequências
17 de *primers* podem ser analisadas também em plataformas digitais gratuitas como *Primer3Plus*
18 (Untergasser *et al.*, 2007) e *Oligo Analyzer* (site da *IDT DNA Technologies*), *Sequence*
19 *Manipulation Suite – PCR Primer Stats* (Stothard, 2000), *Multiple Primer Analyzer* (site da
20 *ThermoFischer Scientific*) e *PrimerDigital*.

21
22 Após a obtenção de sequências de *primers* que amplifiquem um fragmento informativo
23 para o grupo alvo, esta metodologia precisa ser validada. O conceito de validação abrange o
24 processo pelo qual se comprova que a metodologia usada é confiável, robusta e reproduzível
25 para alcançar os objetivos pelos quais foi desenhada. Esta validação é feita de modo tal que
26 quando a metodologia validada for utilizada, não sejam necessárias repetições para confirmar
27 os resultados, e ao mesmo tempo estes sejam precisos e exatos (Linacre e Tobe, 2013). Os
28 estudos de validação testam a reprodutibilidade do método, usando condições diversas como
29 DNA de diferentes espécies, distintas concentrações de DNA molde e diferentes parâmetros
30 físico-químicos de amplificação (Dawnay *et al.*, 2007).

31
32 A validação ocorre após a coleta de amostras e é fundamental para uma correta
33 identificação forense, assim como para sua aceitação como evidência no sistema judiciário
34 (Linacre e Tobe, 2013). De acordo com a segunda recomendação proposta por Linacre *et al.*,
35 (2011) é fundamental contar com “*espécimes voucher*” (exemplares designados para
36 documentar a identidade dos organismos coletados que representam a fonte de DNA em
37 estudo) e que os mesmos sejam identificados com precisão taxonômica, que estejam

1 depositados em coleções científicas, acompanhados de dados de coleta (nome da coleção,
2 número de tombo, localidade de coleta especificando as coordenadas geográficas, coletor,
3 número de exemplares, etc.), dados morfométricos e dados merísticos (contagem de
4 escamas, espinhos, raios, dentes, etc.) (Moore e Kornfield, 2011). Embora os mesmos autores
5 acrescentem que, caso não seja possível disponibilizar exemplares *voucher*, a autenticação
6 pode ser feita utilizando sequências de genes mitocondriais validados, como é o caso do *Cyt*
7 *b* ou do *COI*, cujas sequências estejam depositadas em bases de dados como o GenBank.

8 **1.4. O GÊNERO *Hoplias* Gill, 1903: EXEMPLO DA COMPLEXIDADE TAXONÔMICA EM** 9 **RECURSOS PESQUEIROS**

10 **1.4.1. Histórico taxonômico do gênero *Hoplias***

11

12 Taxonomicamente os peixes continentais sul-americanos estão representados
13 principalmente por três grandes grupos: Siluriformes (3936 spp.), Characiformes (2496 spp.)
14 e Gymnotiformes (252 spp.) (Fricke *et al.*, 2019). A ordem Characiformes agrupa 24 famílias
15 (Fricke *et al.*, 2019) entre as quais se encontram as popularmente conhecidas como “traíras”,
16 “trairão”, “jejus” ou “marobás”. Estes peixes são parte da família Erythrinidae e comumente
17 habitam rios, lagos e bosques inundados (Oyakawa e Mattox, 2009).

18

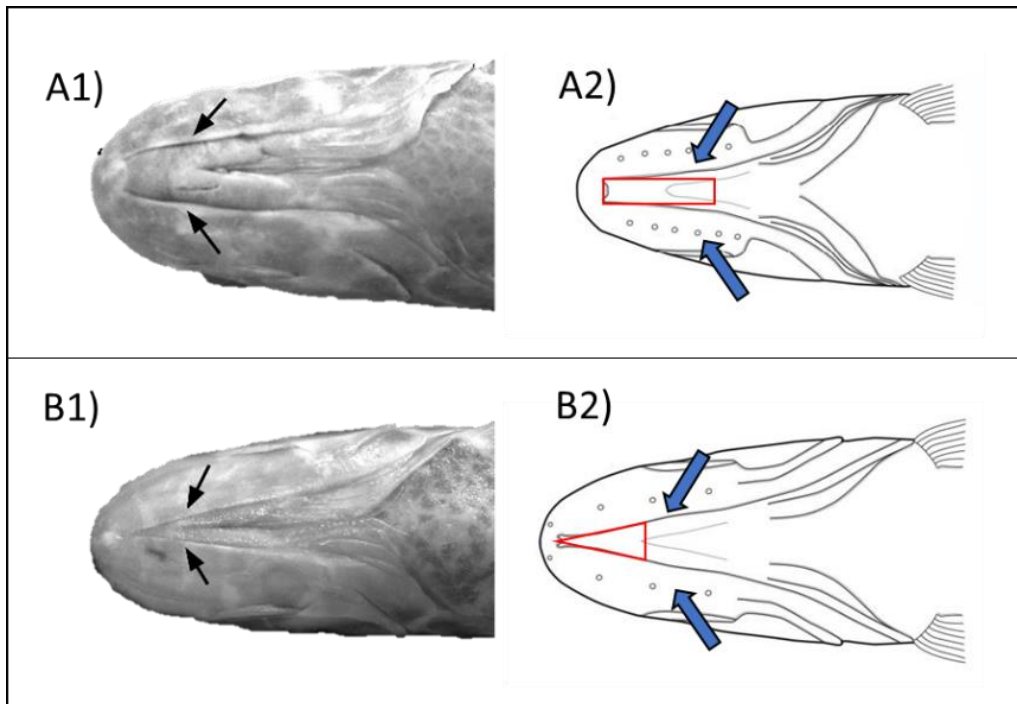
19 Os representantes da família Erythrinidae se encontram distribuídos unicamente na
20 região Neotropical, ocorrendo a maior diversidade na região amazônica (Reis *et al.*, 2003). De
21 acordo com a morfologia externa, os três gêneros que representam a família (*Erythrinus*
22 Scopoldi, 1777; *Hoplerythrinus* Gill, 1985 e *Hoplias* Gill, 1903) se caracterizam por ter o corpo
23 cilíndrico, nadadeira caudal arredondada, ausência de nadadeira adiposa e linha lateral com
24 37 a 47 escamas (Reis *et al.*, 2003). Os gêneros *Erythrinus* e *Hoplerythrinus* são peixes que
25 alcançam aproximadamente 60 cm na vida adulta, no entanto *Hoplias* consegue atingir os 100
26 cm de comprimento padrão (Mattox *et al.*, 2006). Entre os três gêneros que compõem a família
27 Erythrinidae, o gênero *Hoplias* possui a maior diversidade com 14 espécies descritas
28 atualmente (Mattox *et al.*, 2006; Oyakawa e Mattox, 2009; Mattox *et al.*, 2014), sendo algumas
29 delas economicamente importantes em mercados locais (Mattox *et al.*, 2006).

30 Na última revisão do gênero (Oyakawa e Mattox, 2009), as espécies de *Hoplias* foram
31 classificadas em três grandes grupos usando caracteres morfológicos externos: grupo *H.*
32 *lacerdae*, grupo *H. macrophthalmus* e grupo *H. malabaricus*. As espécies que compõem o
33 grupo *H. lacerdae* (*H. australis*, *H. brasiliensis*, *H. curupira*, *H. lacerdae* e *H. intermedius*) foram
34 caracterizadas pela ausência de placas dentárias na região dorsal do basiial e basibranquial
35 e pela convergência quase paralela das bordas do dentário (Oyakawa e Mattox, 2009). A

1 espécie alvo *H. intermedius* se diferencia das espécies que compõem o grupo *H. lacerdae*,
2 com exceção de *H. australis* e *H. lacerdae*, pelo número de escamas na linha lateral entre 42-
3 46 (vs. 38-43 em *H. brasiliensis* e 34-39 em *H. curupira*). Adicionalmente, se diferencia a *H.*
4 *intermedius* pelo número de poros do canal laterosensorial da região ventral do dentário entre
5 4-6 (vs. 5 em *H. australis*, e 6-8 em *H. lacerdae*). Distingue-se também a *H. intermedius* pelo
6 perfil anterior da cabeça em forma angular (vs. perfil arredondado em *H. australis*).

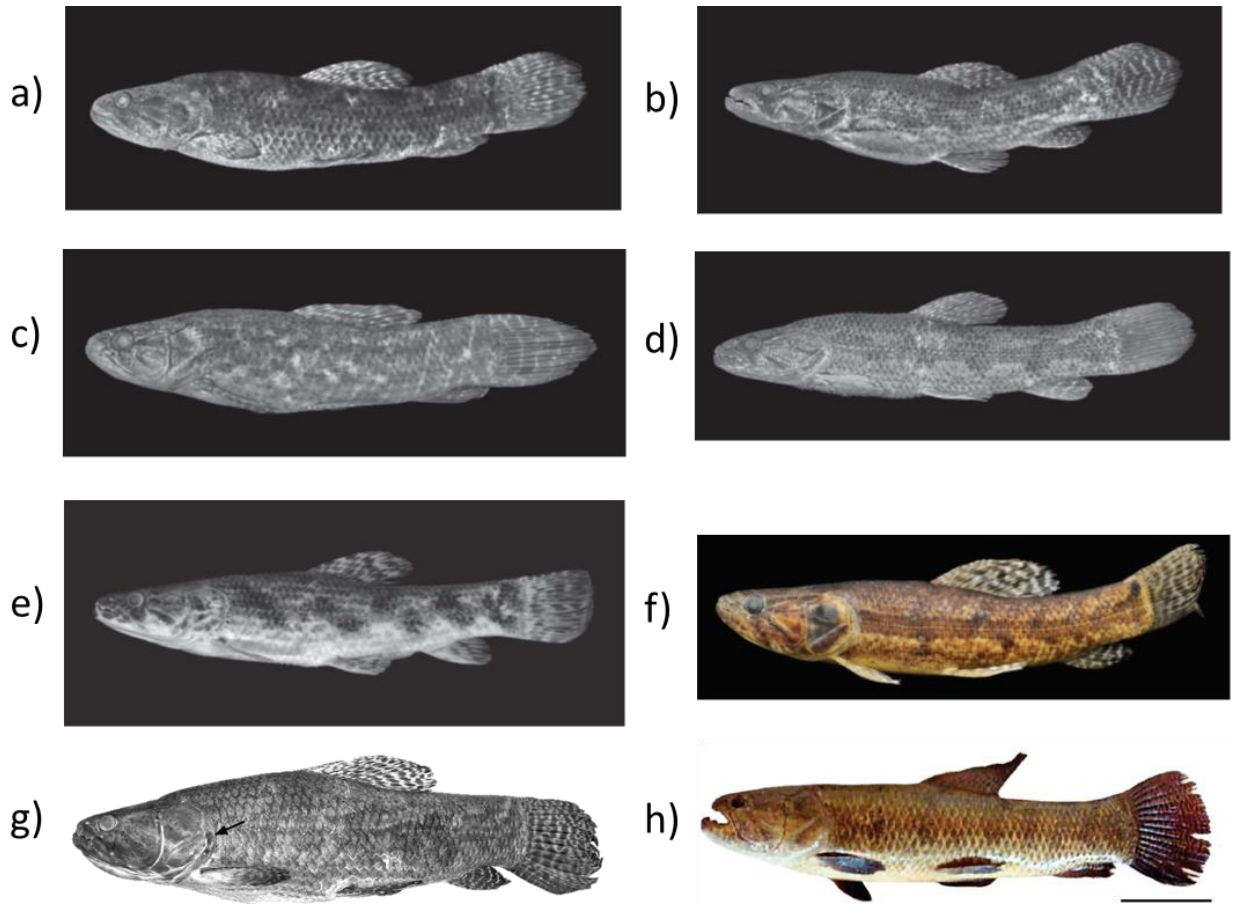
7 Baseadas em informação de descrições taxonômicas originais, as chaves dicotômicas
8 são uma série ordenada de alternativas as quais representam caracteres diagnósticos, onde
9 sua presença ou ausência de tais caracteres permitem tomar uma decisão sobre a
10 identificação do organismo. É fornecida neste trabalho uma versão resumida da chave de
11 identificação disponibilizada originalmente por Oyakawa e Mattox (2009), na qual é possível
12 identificar os caracteres utilizados pelos autores para diferenciar as espécies do grupo *H.*
13 *lacerdae*, *H. aimara* e do grupo *H. malabaricus* (Figura 1 e Figura 2).

14



1

2 **Figura 1. Esquema adaptado de Mattox et al. (2006) e Oyakawa e Mattox (2009) da região ventral**
 3 **de duas espécies de *Hoplias*.** As setas indicam o dentário. **A1)** *Hoplias malabaricus*, confluência do
 4 dentário formando uma região retangular e **A2)** poros na superfície do dentário de *H. malabaricus*. **B1)**
 5 *Hoplias intermedius* com dentário confluindo formando uma região triangular e **B2)** poros na superfície
 6 do dentário de *H. intermedius*.



1

2 **Figura 2. Espécies do grupo *Hoplias lacerdae*, *H. aimara* e *H. microlepis* representando o grupo *H.***
 3 ***malabaricus*. a)** *H. australis*, holótipo, MCP 40175, 220 mm SL; **b)** *H. brasiliensis*, neótipo, MZUSP
 4 45483, 162.9 mm SL; **c)** *H. curupira*, holótipo, MZUSP 45582 157.9 mm SL; **d)** *H. intermedius*, MZUSP
 5 69370, 169.3 mm SL; **e)** *H. lacerdae*, MZUSP 61748, 150.5 mm SL; **f)** *H. microlepis*, USNM 293250,
 6 176.2 mm SL; **g)** *H. aimara* MZUSP 87964, 355.3 mm SL; **h)** *H. malabaricus*. (Grassi et al., 2017; Mattox
 7 et al., 2006; Oyakawa e Mattox, 2009; Mattox e Oyakawa, 2014).

8

1 **1.4.2. Importância comercial e ecológica de *Hoplias intermedius* e *H.*** 2 ***malabaricus***

3
4 Para um manejo adequado dos recursos pesqueiros se requiere também conhecer os
5 aspectos biológicos, geográficos e evolutivos das espécies alvos da pesca. Geograficamente,
6 as espécies que compõem o gênero *Hoplias* em sua maioria ocorrem em áreas mais restritas,
7 como é o caso das espécies do grupo *H. lacerdae*. Assim, *H. brasiliensis* está restrita às bacias
8 costeiras do nordeste do Brasil, desde o rio Paraguaçu até o rio Jequitinhonha; *H. curupira*
9 está restrita aos rios amazônicos, incluindo rio Capim, Tocantins, Xingu, Tapajós e tributários
10 e rios costeiros do Suriname; *H. lacerdae*, distribuído no rio Ribeira de Iguapé, rio Uruguai e
11 rio Negro no Uruguai (Oyakawa e Mattox, 2009). Um caso particular é o representado por *H.*
12 *intermedius*, espécie que possui uma distribuição disjunta nas bacias do rio São Francisco,
13 rio Doce e cabeceiras do rio Paraná, incluindo o rio Grande, rio Paranaíba e rio Piquiri
14 (Oyakawa e Mattox, 2009). Excepcionalmente, *H. malabaricus* encontrasse amplamente
15 distribuído na região neotropical (Oyakawa, 2003). Esta distribuição abrangente tem permitido
16 a crescente importância econômica da indústria pesqueira de “traíras” e “trairões”. Desta
17 forma a pesca de subsistência tem colocado as espécies do gênero *Hoplias* como a sétima
18 espécie mais consumida. A produção de “traíra” em piscicultura tem registrado valores
19 flutuantes sempre acima das 7100 toneladas, apresentando no ano de 2010 uma produção
20 crescente que chegou a 9821 toneladas (Dias-Neto e Dias, 2015).

21 As espécies do gênero *Hoplias* por sua natureza predadora possui uma alimentação
22 baseada em peixes, camarões e insetos aquáticos (Dias-Neto e Dias, 2015). Sua maturação
23 sexual ocorre após um ano, produzindo em torno de 2500 ovos, o que embora seja
24 considerado como uma fecundidade baixa (Dias-Neto e Dias, 2015) é compensado por um
25 período reprodutivo extenso, permitindo vários ciclos reprodutivos ao ano (Chaves *et al.*,
26 2011). No estágio larval a alimentação é feita a base de plâncton, mudando para insetos em
27 uma etapa posterior e finalmente com hábitos piscívoros na fase adulta (Chaves *et al.*, 2011).
28 Assim como no caso de *Pseudoplatystoma corruscans* (Spix & Agassiz, 1829) ou *P.*
29 *reticulatum* Eigenmann & Eigenmann, 1889 (conhecidas popularmente como “surubim” ou
30 “pintado”), o comportamento predador constitui uma desvantagem para sua criação em
31 cativeiro. Nesse sentido, para o cultivo de espécies predadoras como as do grupo *H. lacerdae*
32 e *H. malabaricus*, é recomendada uma dieta de composição variável e a homogeneização do
33 tamanho dos indivíduos de um mesmo ambiente, fatores fundamentais para diminuir o
34 canibalismo dentro das populações (Luz *et al.*, 2000; Luz e Portella, 2002).

1 **1.4.3. Diversidade molecular e citogenética das espécies de *Hoplias***

2
3 Para explicar a distribuição disjunta de *H. intermedius*, Oyakawa e Mattox (2009) citam
4 a teoria proposta por Potter (1997) e Ribeiro (2006), em que a distribuição abrangente de
5 alguns táxons e as relações filogenéticas estreitas entre eles sugerem uma conexão com
6 eventos geológicos da margem continental da costa sudeste da América do Sul. Da mesma
7 forma, para o gênero *Hoplias*, Dergam et al. (2002) com base em uma análise filogeográfica
8 mitocondrial e evidência citogenética, concluíram que as populações costeiras de *H.*
9 *malabaricus* evoluíram em conexão com as populações do rio Doce, atribuindo a baixa
10 divergência entre elas a eventos geológicos do pré Holoceno.

11 Estudos de diversidade populacional em espécies do grupo *H. lacerdae* têm sido feitos
12 usando marcadores citogenéticos, concluindo que as populações de *H. aimara*, *H.*
13 *brasiliensis*, *H. intermedius* e *H. lacerdae* não mostraram variações cromossômicas
14 significantes (Oliveira et al., 2015). Esses resultados concordam com o proposto por Blanco
15 et al., (2010), no qual se encontrou que populações de *H. malabaricus* ao serem populações
16 de menor tamanho e possuir uma menor dispersão, fortalecem a fixação de arranjos
17 cromossômicos. Contrariamente, em espécies de maior dispersão (como Anostomidae,
18 Curimatidae, Prochilodontidae e espécies do grupo *H. lacerdae*) a diversidade cromossômica
19 inter-populacional resulta em mais homogênea (Bertollo et al., 1986).

20 Padrões de distribuição populacional e filogeografia de peixes têm sido elucidados
21 através do DNA mitocondrial, sendo o gene *COI* frequentemente usado para entender a
22 diversidade genética intra-específica (Hajibabaei et al., 2006). Assim, recentemente a
23 diversidade e distribuição geográfica das espécies do gênero *Hoplias* foram analisadas
24 usando a técnica do DNA barcoding a partir do gene *COI* (Cardoso et al., 2018). A
25 amostragem inclui todas as espécies do gênero ao longo da região Neotropical. Como
26 resultado foi obtido que todas as sequências inicialmente identificadas como *H. malabaricus*
27 representaram 15 linhagens mitocondriais com alto valor de *bootstrap*. Do total de linhagens
28 obtidas, apenas quatro correspondiam espécies válidas, permanecendo dentro do complexo
29 *H. malabaricus*. As demais 11 linhagens foram consideradas espécies ainda não descritas.
30 Para o grupo *H. lacerdae*, a identificação taxonômica e molecular mostraram resultados
31 similares, enquanto que *H. curupira* mostrou duas linhagens diferentes para a mesma espécie
32 (Cardoso et al., 2018).

33 A revisão bibliográfica anterior, evidencia a importância do gênero *Hoplias* na pesca,
34 aquicultura, assim como sua complexidade taxonômica, citogenética e evolutiva. Desta forma,
35 o presente estudo busca fornecer ferramentas genéticas que possam auxiliar no manejo de

1 estoques pesqueiros, otimizar a reprodução em cativeiro e aprimorar resultados imprecisos
2 nas pesquisas realizadas no gênero. Embora a espécie alvo do trabalho tenha sido *H.*
3 *intermedius*, foi necessário desenvolver adicionalmente marcadores para *H. malabaricus*,
4 devido à ocorrência simpátrica e sintópica entre as duas espécies.

5

1 2. OBJETIVOS

2 2.1. Objetivo geral:

3 Desenvolvimento de marcadores moleculares para identificação e caracterização de
4 *H. intermedius*.

5 2.2. Objetivos específicos:

- 6 a. Sequenciar parcialmente o genoma de *H. intermedius*.
- 7 b. Montar totalmente o genoma mitocondrial de *H. intermedius*.
- 8 c. Identificar, isolar e caracterizar regiões mitocondriais espécie-específicas de *H.*
9 *intermedius*.
- 10 d. Identificar, isolar e caracterizar regiões mitocondriais espécie-específicas de *H.*
11 *malabaricus*.
- 12 e. Testar a especificidade dos marcadores específicos para *H. intermedius* e *H.*
13 *malabaricus*.
- 14 f. Testar a sensibilidade dos marcadores específicos para *H. intermedius*.
- 15 g. Testar a eficiência dos marcadores específicos para *H. intermedius* em populações
16 naturais.
- 17 h. Validar os marcadores específicos para identificar *H. intermedius*.

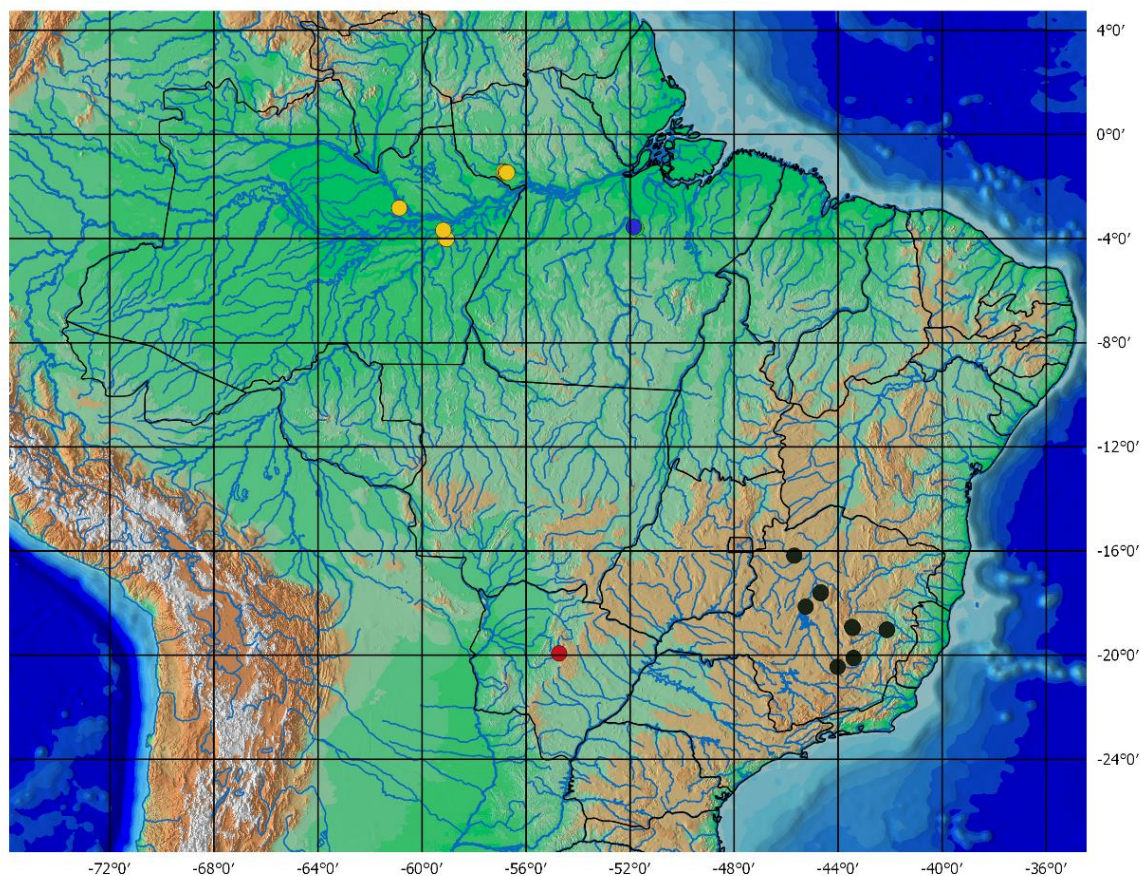
18 3. MATERIAIS E MÉTODOS

19 3.1. Obtenção de amostras

20 As amostras coletadas em campo foram de tecido muscular ou fragmentos de
21 nadadeira. Cada amostra foi rotulada e preservada em etanol 70% para seu transporte e
22 processamento imediato no Laboratório de Biotecnologia e Marcadores Moleculares da
23 Universidade Federal de Minas Gerais.

24 Inicialmente foram obtidos dois exemplares de *H. intermedius* para o sequenciamento
25 do genoma mitocondrial, para o qual foram coletados dois exemplares (uma fêmea e um
26 macho) na localidade tipo da espécie. Os exemplares foram identificados seguindo os
27 caracteres de diagnose propostos por Oyakawa e Mattox (2009). Esta coleta foi realizada em
28 colaboração com a equipe de pesquisa do projeto Manuelzão, coordenada pelo biólogo Carlos
29 Bernardo Mascarenhas Alves nas seguintes coordenadas: 19°20'01.1"S, 43°39'17.0"W.
30 Foram obtidas amostras de 30 indivíduos de *H. intermedius* na bacia do rio Doce, coletados
31 pelo biólogo Tiago Pessali (Museu de Ciências Naturais da Pontifícia Universidade Católica

1 de Minas Gerais). Foram requeridos empréstimos de amostras de coleções de tecidos do
2 grupo de espécies *H. lacerdae* do Instituto Nacional de Pesquisas Amazônicas (INPA),
3 Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) e Pontifícia Universidade
4 Católica de Minas Gerais (PUC-MG). Todas as coordenadas, localidades de coleta, número
5 e instituição de depósito podem ser consultadas na seção de anexos, Tabela 13. Para a
6 análise de amostras comercializadas como “traíra” ou “trairão”, foram adquiridos 11 filés
7 congelados da região metropolitana de Belo Horizonte (comercializadas como “traíra”) e 39
8 amostras frescas obtidas de pescadores da região do Rio Grande (comercializadas como
9 “trairão”). O mapa (Figura 3) foi construído no programa QGis 2.28.27 seguindo o tutorial
10 proposto por Calegari *et al.*, (2016) e representa as localidades de coleta para cada espécie
11 analisada, sendo que mais de um exemplar por espécie foi coletado na mesma localidade.



12

13 **Figura 3. Mapa de distribuição geográfica das amostras analisadas.** Os círculos representam a
14 localidade de coleta das espécies analisadas. **Círculo vermelho:** *H. australis*, círculo preto: *H.*
15 *intermedius*, **círculo amarelo:** *H. malabaricus*, **círculo azul:** *H. aimara* e *H. curupira*.

16

17

1 **3.2. Isolamento do DNA genômico:**

2 **3.2.1. Método de extração 1: Fenol e clorofórmio**

3 Esta metodologia foi aplicada em amostras coletadas em campo e amostras obtidas
4 como filé congelado.

5 ***Fase 1- Isolamento:***

6 Foi pesada 1g de amostra de músculo e macerada em gelo seco, usando um gral e
7 pistilo. Foram adicionados 1,5 mL de solução TEN 9 (50 mM Tris-HCl a pH 9, 200 mM NaCl
8 e 100 mM EDTA) cuja função é neutralizar as cargas negativas da molécula de DNA
9 permitindo a elas se agregarem, atuando como um agente quelante que remove os íons de
10 Fe^{2+} , Ca^{2+} e Mg^{2+} inativando a ação degradadora de DNase (elementos frequentemente
11 presentes em ossos, escamas, cascas, etc) (Dominguez e Ward, 2010).

12 O produto anterior foi transferido para um tubo de 50 ml e imediatamente adicionados
13 150 μ L de solução SDS (dodecil sulfato de sódio) a 20%, que rompe membranas celulares. A
14 solução resultante foi homogeneizada sutilmente por 10 minutos e logo foram adicionados
15 150 μ L de proteinase K (Phoneutria), a qual inativa DNases que poderiam degradar o DNA
16 purificado, e auxilia na liberação do DNA nuclear e mitocondrial, fragmentando proteínas
17 estruturais das membranas celulares, organelas e histonas, liberando assim o DNA na
18 solução (Hilz *et al.*, 1975). A solução final foi deixada em banho-maria a 56°C *overnight*.

19 ***Fase 2- Purificação:***

20 Depois da incubação, foram adicionados 1/10 do volume total da solução (40uL) de
21 RNAse e incubados novamente em banho-maria a 56°C por 30 minutos. Em seguida foram
22 adicionados 440 μ L de fenol e logo a solução foi homogeneizada por 2 minutos observando a
23 formação de duas fases bem definidas. Logo após, foram adicionados 440 μ L de clorofórmio,
24 homogeneizados durante 2 minutos e finalmente centrifugados por 10 minutos a velocidade
25 máxima de 13.200 rpm (centrífuga Eppendorf 5415D, rotor F45-24-11). Sendo a função do
26 fenol e do clorofórmio a de desnaturar os agentes contaminantes deixando-os na fase
27 orgânica (fase inferior), alocar os ácidos nucleicos na fase aquosa (fase superior), e auxiliar
28 na remoção de carboidratos e lipídeos (Butler, 2011).

29 A fase superior (contendo o DNA) foi separada em outro tubo, no qual foram
30 adicionados 500 μ L de clorofórmio para em seguida homogeneizar a solução por 2 minutos. A
31 solução resultante foi centrifugada por 7 minutos a 13.200 rpm e a fase superior foi novamente
32 separada em outro tubo.

1 **Fase 3- Precipitação:**

2 Foram utilizados 1 ml de solução de precipitação de DNA (Phoneutria) e o produto final
3 foi homogeneizado por 5 minutos. Seguido de centrifugação por 10 minutos a velocidade
4 13.200 rpm, foi descartado o sobrenadante e adicionados 1 ml de álcool 70% ao pellet
5 restante. Finalmente foi centrifugada a solução por 5 minutos a velocidade 13.200 rpm,
6 descartado o sobrenadante e preservado até todos os resíduos de álcool evaporarem. A
7 amostra foi finalmente armazenada na geladeira a -20°C.

8 **Fase 4 - Ressuspensão:**

9 Foram utilizados 100 µL de solução TE e a solução foi incubada em banho-maria a
10 50°C durante 15 minutos.

11 **3.2.2. Método de extração 2: Kit de Extração de DNA GTS (Pht)**

12 O segundo protocolo aplicado foi o GTS, baseado na utilização de sílica, com o objetivo
13 de isolar DNA de amostras de coleção de tecido.

14 **Fase 1- Extração:**

15 Foi utilizada 1 g de amostra sólida ou líquida, adicionados 1,2 µl de Proteinase K (10
16 mg/ml) para iniciar a digestão do tecido muscular e 20 µl de tampão Lifton (0,2 M de sacarose,
17 0,05 M de EDTA, 0,1 M de Tris, 0,5% de SDS) que preserva a manutenção do pH em valor
18 neutro e assim beneficia a estabilidade da molécula de DNA. A solução foi incubada a 60°C
19 na estufa *overnight*.

20 **Fase 2- Purificação:**

21 Foi utilizado o kit GTS (Phoneutria) ajustando os volumes indicados pelo fabricante.
22 Foram adicionados 100 µl de solução I, que realiza a segunda fase de rompimento de
23 membranas intracelulares expondo finalmente as moléculas de DNA em solução. A solução
24 resultante foi homogeneizada, incubada em banho-maria a 60-70°C por 10 minutos, e
25 centrifugada a 3000 g por 15 minutos com a finalidade de separar compostos residuais
26 (proteínas, lipídeos e demais restos celulares) na fase inferior, mantendo as moléculas de
27 DNA na fase sobrenadante.

28 O volume de sobrenadante obtido foi transferido para um novo tubo, ao qual se
29 adicionou 5 µl da solução II, contendo moléculas de sílica com uma carga elétrica
30 determinada, o qual permite as moléculas de DNA se unirem com as moléculas de sílica
31 (Teixeira *et al.*, 2012). A solução foi homogeneizada, permitindo que a sílica presente na
32 solução II se una às moléculas de DNA. Após uma nova centrifugação o sobrenadante foi

1 transferido para um novo tubo com solução II, a solução foi homogeneizada e incubada em
2 banho-maria por 20 minutos a 55 °C. Novamente a solução foi centrifugada por 5 minutos a
3 3.000 g com a finalidade de excluir na fase superior possíveis restos celulares. O
4 sobrenadante foi descartado.

5 O precipitado obtido foi lavado com 50 µl de solução III A, que permitire separar
6 impurezas que podem ter permanecido. Em seguida a solução foi homogeneizada
7 manualmente, centrifugada por 5 minutos a 1.500 g e o sobrenadante novamente descartado.
8 Logo após foram adicionados 50 µl de solução IV B, conhecida também como *Solução de*
9 *lavagem*, a qual retira possíveis inibidores de DNA polimerase residuais e outros agentes
10 químicos presentes na solução de lise (solução I) e solução de lavagem A (solução III). A
11 solução obtida foi homogeneizada, centrifugada por 5 minutos a 1.500 g e o sobrenadante foi
12 descartado. A solução foi lavada com 50 µl de acetona P.A que acelera a remoção de álcool.
13 Foi feita uma centrifugação por 5 minutos a 1.500 g e o sobrenadante completamente retirado.
14 O pellet foi secado por 30 minutos a 50°C (ou *overnight* a 37°C) mantendo o tubo aberto.

15 ***Fase 3 - Resuspensão:***

16 Foram adicionados 25 µl de TE pH 8 do kit GTS-pht, que funciona como meio
17 levemente básico que brinda estabilidade às moléculas de DNA. A solução foi incubada em
18 banho-maria por 30 minutos a 50°C, e depois foi centrifugada por 10 minutos a 3.300 g com
19 a finalidade de separar o DNA molde da sílica e o expor na fase sobrenadante.

20 **3.3. Quantificação de DNA genômico e análise de integridade**

21 Antes de iniciar todas as reações de PCR, foi feita uma quantificação da concentração
22 de DNA (ng/µL) utilizado. Esta análise foi realizada no fluorímetro Qubit® 2.0 (Life) usando o
23 Qubit® dsDNA BR (Broad-Range) Assay Kit para intervalo de leitura de [1 ng/µL; 10 ng/µL] e
24 Qubit® dsDNA HS (High Sensitive) Assay Kit para intervalo de leitura de [0,01 ng/µL; 1 ng/µL]
25 seguindo as especificações do fabricante. As leituras foram obtidas em ng/ µL.

26

3.4. Construção da biblioteca genômica

3.4.1. Marcação da PCR-Nextera

Nesta etapa foi usado o DNA molde obtido pelo método de extração Fenol e clorofórmio a partir do indivíduo macho coletado em campo. As moléculas de DNA foram marcadas com os transposons Nextera que fragmentam o DNA genômico e fazem a ligação dos adaptadores no final da sequência de cada fragmento.

Assim, com o DNA previamente quantificado, foram utilizados 50 ng de amostra de DNA em volume de 20 µL, obtendo uma concentração final de 2,5 ng/µL. Em seguida foram adicionados 25 µL de solução tampão TD Buffer e 5 µL TDE1 (enzima contendo os marcadores de transposons). A solução, com um volume final de 50 µL, foi homogeneizada 10 vezes por 1 minuto, e centrifugada a 280g por 1 minuto.

3.4.2. Purificação do DNA marcado

Nesta etapa através de soluções tampão e centrifugações sucessivas, o DNA marcado é purificado e separado dos transposons Nextera. Assim, foram adicionados 250 µL de PB (Binding Buffer) e homogeneizados manualmente. A solução foi transferida para uma coluna e centrifugada por 1 minuto a 13000 g, descartando o volume filtrado. A coluna foi lavada com 750 µL de solução tampão PB Buffer e centrifugada por 1 minuto a 13000g, descartando novamente o volume filtrado, em seguida a solução foi novamente centrifugada para secar o volume restante. Posteriormente a solução foi eluída com 25 µL de RBS (buffer de ressuspensão) e novamente a solução foi centrifugada por 1 minuto a 13000g.

3.4.3. PCR Illumina Genoma Nextera

Nesta etapa o DNA purificado é ligado aos index e adaptadores (P5 e P7), através de uma PCR de cinco ciclos. Esses adaptadores permitem a ligação do DNA à *Flow cell* e através de uma amplificação em ponte formam-se os *clusters* onde ocorre o sequenciamento por síntese.

Foi preparada uma solução contendo 15 µL de NPM (PCR Master Mix, contendo solução tampão, dinucleotídeos e Taq polimerase) e 5 µL de PPC (*primer coquetel*). Foram adicionados 5 µL do INDEX 501, 5 µL do INDEX 701 e 20 µL da solução obtida na etapa anterior. A solução foi homogeneizada manualmente de 10 a 15 vezes.

Posteriormente estes fragmentos foram amplificados utilizando o protocolo indicado pelo fabricante, que consta de um programa de 72°C por 3 minutos, 98° por 30 segundos, 8

1 ciclos de: 98°C por 10 segundos, 63°C por 30 segundos e 72°C por 3 minutos e armazenados
2 a 10°C.

3 **3.4.4. Purificação da PCR**

4 Nesta etapa foram usados AMPure XP beads para purificar a biblioteca de DNA,
5 permitindo remover fragmentos de DNA de tamanho muito curto. O volume final de 50 µL
6 obtido anteriormente contendo os amplicons foi purificado utilizando 25 µL de *beads*
7 magnéticas (AMPure XP beads), as quais se acoplam aos fragmentos de DNA em presença
8 de PEG e solução tampão de NaCl. A solução foi transferida para um eppendorf de 600 µL,
9 incubada por 5 minutos a temperatura ambiente e logo colocada em uma raque magnética
10 por 2 minutos, descartando posteriormente o sobrenadante. Foi feita uma lavagem usando
11 200 µL de etanol 80%, mantido por 30 segundos no tubo e descartando posteriormente o
12 sobrenadante. O processo de lavagem foi repetido mais uma vez e então a solução foi
13 ressuspensa em 32,5 µL de RBS (Resuspension Buffer), homogeneizado manualmente
14 com a pipeta, mantido a temperatura ambiente por 2 minutos. A solução foi novamente
15 colocada em uma raque magnética por 2 minutos, transferindo logo 30 µL do sobrenadante a
16 uma nova placa NPL (Nextera Library Plate) e finalmente adicionados 30 µL de álcool 80%.

17 **3.4.5. Análise quantitativo: PCR em tempo real**

18 Para se obter uma ótima qualidade de dados do sequenciamento na plataforma
19 Illumina é necessário obter uma densidade ótima de cluster em cada espaço da *flow cell*.
20 Desta forma, se faz necessária realizar uma quantificação e análise de qualidade da biblioteca
21 de DNA.

22 Foi utilizado o Library Quantification Kit - Illumina/Universal (Kapa), que amplifica
23 sequências que possuam os adaptadores P5 e P7 e permite detectar uma possível
24 contaminação da biblioteca. A solução foi preparada utilizando Master Mix, água ultrapura e
25 DNA da biblioteca em duas concentrações diferentes (1:1000 e 1:2000) em tubos individuais
26 cada uma. A solução do controle positivo foi preparada utilizando Master Mix, água ultrapura
27 e cinco soluções padrão com concentrações já conhecidas e fornecidas pelo kit em tubos
28 individuais cada uma. A solução de controle negativo foi preparada utilizando Master Mix e
29 água (volumes detalhados na Tabela 1).

30

1 **Tabela 1. Volumes dos reagentes utilizados na qPCR de controle de qualidade da biblioteca**
 2 **genômica.**

	Amostra (µL)	Controle positivo (µL)	Controle negativo (µL)
	Master Mix	12	12
	Água	4	8
Amostra	1:1000	4	---
	1:2000	4	---
Padrão	20 pM	---	---
	2 pM		4
	0,2 pM		4
	0,02 pM		4
	0,002 pM		4

3

4 **3.5. Sequenciamento de Nova Geração de *H. intermedius***

5 O sequenciamento foi realizado no sistema integrado MiSeq para geração automática
 6 de clusters de DNA por amplificação em ponte. A leitura *paired-end* foi de 2x300, o kit de
 7 leitura utilizado foi o Nextera e o kit de corrida o MiSeq Reagent kit V3-600.

8 **3.5.1. Verificação da qualidade das sequências geradas**

9 Foram utilizados os programas PRINSEQ 0.20.4 (Schmieder e Edwards, 2015) e
 10 FastQC 0.11.3. Estes programas podem ser usados para *trimar* e formatar as sequências,
 11 removendo sequências curtas, sequências de baixa qualidade ou sequências que contém Ns
 12 (base nucleotídica indeterminada), de forma tal que sejam detectados artefatos que poderiam
 13 ter sido gerados na preparação da biblioteca ou nos ciclos de sequenciamento. Os resultados
 14 são fornecidos como um resumo estatístico dos dados permitindo visualizar graficamente os
 15 resultados do sequenciamento, como por exemplo, o comprimento de sequências, conteúdo
 16 de GC, complexidade e distribuição do índice de qualidade, número de *reads* duplicadas, etc.

17 As *reads* geradas no sequenciamento foram avaliadas pela ferramenta de controle de
 18 qualidade PRINSEQ. O PRINSEQ é uma ferramenta que fornece um resumo com parâmetros
 19 estatísticos e qualidade dos dados, os quais são usados para filtrar, formatar e trimar
 20 sequências obtidas pela metodologia NGS. Um desses parâmetros é o valor do índice de
 21 qualidade PHRED, que mede a qualidade de cada base nucleotídica. O valor PHRED é
 22 definido como uma propriedade logaritmicamente relacionada com a probabilidade de erro de
 23 sequenciamento P . Desta forma, um valor Q10 significa uma probabilidade de erro de 1/10
 24 na leitura de cada nucleotídeo, Q20 significa uma probabilidade de erro de 1/100, e Q30
 25 significa uma probabilidade de erro de 1/1000 (Ewing and Green, 1998).

26

3.5.2. Montagem *De novo*

Usando os resultados do sequenciamento, foi montado o genoma mitocondrial de *H. intermedius*. A reconstrução da sequência mitocondrial sem uma sequência de referência, ou montagem *De novo* do genoma mitocondrial, foi feita usando o programa CLC Workbench software v8.5.1 (CLC Bio-Qiagen, Aarhus, Denmark). Para reconstruir a sequência do mitogenoma, foram montadas *paired-end reads* não duplicadas utilizando a ferramenta "De Novo Assembly" no CLC Genomics Workbench. Foram mantidos os parâmetros predefinidos, obtendo um total de 1068 contigs, filtrando tamanhos mínimos de *contigs* de 600pb, a fim de diferenciar sequências nucleares de mitocondriais. A escolha de uma montagem *de novo* responde a ausência de uma sequência mitocondrial completa da espécie alvo de espécies do gênero *Hoplias*.

3.5.3. Anotação do Genoma Mitocondrial

Após a obtenção da sequência foi feita a anotação do genoma mitocondrial, a qual consiste na identificação de regiões não codificadoras de proteínas e identificação de genes (Stein, 2001). A anotação foi realizada utilizando a plataforma MitoFish e a ferramenta Mitoannotator (Iwasaki *et al.*, 2013) e logo depositada na base de dados GenBank com número de acesso KU523584. MitoFish é uma base de dados gratuita e standardizada que contem 450.205 sequências (parciais e completas), representando 24.661 espécies. A ferramenta usada foi o Mitoannotator (Iwasaki *et al.*, 2013), sendo verificados manualmente os códons iniciais e de finalização. Posteriormente o genoma mitocondrial anotado foi submetido na base de dados GenBank (número de acesso KU523584).

3.5.4. Alinhamento e comparação de sequências mitocondriais

Logo após a montagem, foram recuperadas 38 espécies que possuíam genoma mitocondrial completo (Tabela 2) e logo as sequências foram comparadas mediante a ferramenta BLASTn, sendo analisadas as porcentagens de identidade e similaridade de sequência e *query cover*.

As sequências foram alinhadas e foi desenhada uma árvore de distância consenso usando o algoritmo Neighbor-Joining no programa MEGA6 (Tamura *et al.*, 2013) para visualizar o agrupamento das sequências. Com a finalidade de obter uma visualização gráfica das moléculas circulares do mtDNA de todas as sequências, se utilizou o programa CGView Comparison Tool (Grant *et al.*, 2012) o qual permite observar a estrutura e disposição de regiões codificadoras e não codificadoras de proteína de cada sequência dispostas em um mesmo mapa circular comparativo.

1 Tabela 2. Sequencias mitogenomicas de peixes recuperadas do NCBI

Ordem	Familia	Espécie	Número de acesso	
Characiformes	Alestidae	<i>Phenacogrammus interruptus</i>	NC_004699.1	
	Anostomidae	<i>Leporinus piavussu</i>	NC_025767.1	
		<i>Leporinus elongatus</i>	NC_0342801.1	
	Bryconidae	<i>Brycon orbignyanus</i>	KM245044	
		<i>Brycon henni</i>	NC_026873.1	
		<i>Salminus brasiliensis</i>	NC_024941.1	
	Chalceidae	<i>Chalceus macrolepidotus</i>	NC_004700.1	
	Characidae	<i>Astyanax mexicanus</i>	AP011982.1	
		<i>Hasemania nana</i>	NC_022724.1	
		<i>Hemigrammus bleheri</i>	NC_027946.1	
		<i>Paracheirodon axelrodi</i>	NC_023270.1	
		<i>Paracheirodon innesi</i>	KT783482.1	
		<i>Crenuchus spirulus</i>	AP011986.1	
		<i>Chilodus punctatus</i>	NC_015801.1	
	Citharinidae	<i>Citharinus congicus</i>	NC_015805.1	
	Cynodontidae	<i>Hydrolycus scomberoides</i>	NC_015813.1	
	Ctenoluciidae	<i>Boulengerella maculata</i>	AB070207.1	
	Curimatidae	<i>Curimata mivartii</i>	NC_025951.1	
		<i>Curimatopsis evelynae</i>	NC_011988.1	
	Erythrinidae	<i>Hoplias intermedius</i>	KU523584	
		<i>Hoplias malabaricus</i>	AP011992.1	
	Distichodontidae	<i>Distichodus sexfasciatus</i>	NC_015836.1	
	Hemiodontidae	<i>Hemiodopsis gracilis</i>	NC_015816.1	
	Hepsetidae	<i>Hepsetus odoe</i>	NC_015819.1	
	Lebiasinidae	<i>Lebiasina astrigata</i>	NC_015750.1	
	Parodontidae	<i>Apareiodon affinis</i>	NC_015834.1	
	Prochilodontidae	<i>Ichthyoelephas longirostris</i>	NC_025950.1	
		<i>Prochilodus argenteus</i>	NC_027689.1	
		<i>Prochilodus costatus</i>	NC_027690.1	
		<i>Prochilodus lineatus</i>	NC_024939.1	
		Serrasalmididae	<i>Colossoma macropomum</i>	KP188830.1
			<i>Piaractus brachypomus</i>	NC_025315.2
	<i>Piaractus mesopotamicus</i>		NC_024940.1	
<i>Pygocentrus nattereri</i>	NC_015840.1			
Siluriformes	Pseudopimelodidae	<i>Lophiosilurus alexandri</i>	NC_026845.1	
	Pimelodidae	<i>Pseudoplatystoma magdaleniatum</i>	NC_026845.1	
		<i>Pseudoplatystoma corruscans</i>	NC_026526.1	
		<i>Pimelodus pictus</i>	NC_015707.1	

2

3

4

3.5.5. Localização de marcadores mitocondriais espécie-específicos

Utilizando as bases de dados públicas como NCBI e BOLD, foi feita uma busca das sequências mitocondriais completas ou parciais de todas as espécies do gênero *Hoplias*, assim como as sequências de outros gêneros com 99% de similaridade com a espécie alvo. Logo após, utilizando o programa MEGA6 e o algoritmo Muscle, foi feito um alinhamento do mitogenoma completo de *H. intermedius* e as sequências anteriormente recuperadas nas bases de dados.

A localização de marcadores foi feita a partir de uma pesquisa manual de regiões altamente variáveis que poderiam ter potencial de marcadores moleculares para *H. intermedius*. A pesquisa localizou inserções, deleções ou alterações em comparação com a sequência das outras espécies alinhadas.

A validação dos genes escolhidos para o isolamento de marcadores moleculares foi realizada utilizando o pacote de R: *Species Identity and Evolution in R* (SPIDER). O programa realiza uma busca por regiões informativas para a espécie alvo, a qual se baseia em análises estatísticas de dados filogenéticos gerados por outros pacotes (Adgenet, MASS, Pegas e Ape). A partir do alinhamento de sequências de DNA similares grupo alvo, é feita uma análise em fragmentos de tamanho determinado (janelas), que permitem localizar a região do genoma que possui maior número de *mismatches* (Brown, 2012). Para marcadores espécie-específicos, é usada a função *slideNucDiag* para análise de *nucleotídeos diagnósticos*, ou *mismatches*. É possível indicar ao programa o tamanho de marcador desejado, chamado no programa de “*window width*” ou “*amplitude de janela*”. A partir de amplitudes de janela de 100, 200, 300, 400, 500 e 600 pb, foram gerados gráficos nos quais se avaliaram as regiões com maior número de *mismatches*, chamados no programa como “*nucleotídeos diagnósticos*”.

Os resultados são apresentados em forma de matriz, mostrando as posições dos nucleotídeos diagnósticos na sequência do táxon alvo. Para cada uma destas posições, o SPIDER fornece o número de *mismatches* (n) contabilizados em uma determinada posição (x), para o tamanho de width (y) escolhido.

3.6. Desenho de primers mitocondriais

Posteriormente, todas as possíveis regiões foram localizadas no genoma mitocondrial anotado e escolhidas as sequências flanqueadoras a estes marcadores. Foi feito um novo alinhamento usando sequências do genoma mitocondrial completo de 38 espécies de Characiformes. Foram selecionadas como primers unicamente fragmentos de sequências que

1 possuem pelo menos alta quantidade de *mismatches* entre a sequência de *H. intermedius* e
2 as outras sequências do alinhamento.

3 Previamente à síntese das sequências de *primers*, um primeiro controle de qualidade
4 dos *primers* foi feito a partir de uma análise *in silico* usando a ferramenta Primer-BLAST do
5 NCBI (Ye *et al.*, 2012), a qual permitiu verificar a especificidade de cada conjunto de *primers*,
6 o número de *mismatches* com a sequência alvo, o tamanho do amplicon, a temperatura de
7 *melting* (T_m), a porcentagem de GC e o grau de auto complementaridade.

8 **3.7. Padronização de Reações em Cadeia da Polimerase**

9 O objetivo de padronizar as reações de PCR consistiu em estabelecer, conservar e
10 garantir um alto nível de qualidade da técnica de identificação usando *primers* mitocondriais.
11 O objetivo é encontrar os parâmetros ideais de amplificação da sequência alvo e constitui um
12 dos passos fundamentais da validação do uso de *primers* mitocondriais usando regiões do
13 genoma que são específicas para a espécie em questão.

14 Entre os parâmetros avaliados, foram testadas variações do número de ciclos (30, 35
15 ciclos), temperatura de anelamento (50°, 53°, 56°, 59°, 62° e 64°), concentrações de
16 dimetilsulfóxido (DMSO) (3%, 5% e 7%), cloreto de potássio (KCl) (30mM, 50mM e 100mM) e
17 diferentes soluções tampão IB, IC, IIC e IVB (Phoneutria). Foram mantidos constantes os
18 volumes e concentrações de dNTP, Taq polimerase, *primers* e DNA molde. Os volumes de
19 cada reagente utilizado estão detalhados no Anexo (Tabela 12, 13 e 14).

20 Usando o termociclador Applied Biosystems (modelo Veriti), foram realizadas reações
21 que consistiram em uma fase inicial a 94°C durante 3 minutos, desnaturação a 92°C durante
22 45 segundos, anelamento durante 45 segundos (variação de temperaturas usadas no Anexo),
23 extensão a 72°C durante 45 segundos, extensão final a 72°C durante 10 minutos, e
24 manutenção a 10°C até o infinito. Os tamanhos dos produtos de PCR foram visualizados em
25 gel de agarose 1,5%, corado com brometo de etídio, utilizando um marcador de peso
26 molecular (Phoneutria) de 700, 500 e 200 pb.

27 **3.8. Sequenciamento de Sanger**

28 Para análise de sequências curtas foi utilizado o Sequenciamento de Sanger em
29 eletroforese capilar como método de verificação da sequência dos produtos de PCR obtidos
30 a partir dos *primers* espécie-específicos desenhados.

1 Para a reação de sequenciamento, os produtos de PCR foram quantificados e diluídos
2 em H₂O obedecendo a Tabela 3 (fornecida pela Applied Biosystems) a qual sugere a
3 concentração de amplicon requerido de acordo ao tamanho do mesmo:

4 **Tabela 3. Quantificação do DNA para sequenciamento de Sanger. Tamanho do produto em pares**
5 **de bases e quantidade teórica de DNA em nanogramas.**

Tamanho do produto de PCR (pb)	Quantidade teórica (ng)
100-200	1-3
200-500	3-10
500-1000	5-20
1000-2000	10-40
>2000	40-100

6

7 A reação de sequenciamento foi preparada em microtubos independentes para cada
8 *primer*. Foram adicionados 0,5 µL de Reaction Mix do kit BigDye Terminator v3.1 Cycle
9 Sequencing, 1,5 µL de Tampão de diluição 5X, 2 µL de *primer* (0,8 pmol/ µL), e os volumes
10 de DNA de interesse e de água ultrapura segundo a Tabela 2. O volume final foi de 10 µL. O
11 programa de amplificação seguiu os seguintes parâmetros: uma desnaturação inicial de 96°C
12 por um minuto, 44 ciclos de: 96°C por 15 segundos, 50°C por 15 segundos, e 60° por 4
13 minutos, e finalmente conservação a 4°C até a purificação.

14 Foi feita uma precipitação e lavagem dos produtos de extensão com o objetivo de
15 retirar os nucleotídeos marcados que não foram incorporados aos produtos de extensão, a
16 fim de minimizar os fragmentos de baixa qualidade de sequenciamento. Para isso, foram
17 adicionados em cada microtubo 1µL de EDTA 125 mM e 26 µL de solução de precipitação,
18 logo homogeneizados e encubados a temperatura ambiente por 15 minutos. Posteriormente
19 cada solução foi centrifugada a 14000 g por 45 minutos, o sobrenadante retirado e foram
20 adicionados 35 µL de etanol 70% em cada tubo. Os tubos foram centrifugados novamente a
21 16500 g por 15 minutos a temperatura de 4° C e retirado o sobrenadante. Cada microtubo foi
22 aberto até evaporar todo o volume de etanol e finalmente foram tampados e armazenados a
23 -20°C.

24 A análise bioinformática dos amplicons obtidos foi realizada com os programas BioEdit
25 (Hall, 1999) e Genious (Kearse *et al.*, 2003). As sequências foram editadas individualmente e
26 logo foi criada uma sequência consenso baseada no *forward* e reverso do mesmo par de
27 *primers*. A comprovação de cada sequência consenso foi realizada através de um
28 alinhamento com o genoma mitocondrial completo de *H. intermedius* usando o programa

1 MEGA6. Foi verificado o tamanho do fragmento, e posteriormente sua especificidade foi
2 analisada com a ferramenta BLAST do NCBI.

3 **3.9. Identificação de amostras**

4 Usando os marcadores mitocondriais desenhados para detectar *H. intermedius*, o
5 marcador para detectar *H. malabaricus* e o marcador universal Fish16S, foram analisadas
6 amostras obtidas em campo, amostras obtidas em coleções de tecido do INPA, PUC-MG e
7 PUCRS e amostras provenientes de mercados locais da região de Belo Horizonte e próximas
8 ao Rio Grande, no Alto Paraná. Todas as amostras foram preservadas em etanol 70%
9 imediatamente após a coleta.

10 A identificação morfológica dos espécimes vouchers foi feita usando os caracteres
11 diagnósticos propostos por Oyakawa e Mattox (2009) como a ausência de placas dentárias
12 na região dorsal do basiial e basibrânquial (vs. presença de placas dentárias em *H.*
13 *malabaricus*), convergência quase paralela das bordas do dentário, número de escamas na
14 linha lateral entre 42 - 46 (vs. 38 - 43 em *H. brasiliensis* e 34 - 39 em *H. curupira*), número de
15 poros do canal laterosensorial da região ventral do dentário entre 4 - 6 (vs. 5 em *H. australis*,
16 e de 6 - 8 em *H. lacerdae*) e o perfil anterior da cabeça em forma angular (vs. perfil
17 arredondado em *H. australis*).

18 **3.10. Testes de validação dos primers**

19 **3.10.1. Teste de Sensibilidade**

20 A sensibilidade de todos os conjuntos de primers foi avaliada separadamente fazendo
21 diluições logarítmicas seriadas de base 10, a partir de uma concentração inicial de 10 ng/uL
22 de DNA até uma concentração final de 10⁻⁵ ng/uL.

23 **3.10.2. Teste de Especificidade**

24 Para excluir a possibilidade de amplificação de falso-positivo, realizamos um teste de
25 especificidade baseado em DNA extraído de amostras obtidas a partir de músculo ou
26 nadadeiras de 11 espécies, incluindo *H. aimara*, *H. curupira*, *H. australis*, *H. intermedius* e *H.*
27 *malabaricus*. As seis espécies restantes foram testadas, pois possuem distribuição
28 atualmente descrita no rio São Francisco e alta similaridade de sequência com a espécie alvo
29 (*Brycon orbignyanus*, *Prochilodus argenteus*, *P. costatus*, *P. lineatus*, *Lophosilurus alexandri*,
30 *Piaractus mesopotamicus*).

1 Foi feita uma PCR utilizando as condições de amplificação padronizadas
2 anteriormente para cada par de *primers* (Tabela 3). O marcador 16S foi utilizado como controle
3 positivo da integridade do DNA utilizado, verificando que as amplificações negativas não
4 correspondam a uma possível degradação ou fragmentação da molécula de DNA, e como
5 controle negativo a ausência de DNA.

6 **3.10.3. Teste de Eficiência**

7 De acordo com a localidade de ocorrência do grupo alvo, é preciso obter exemplares
8 coletados em campo. Os indivíduos devem possuir características morfológicas conservadas,
9 coordenadas geográficas, dados de coleta e número de tomo fornecido por coleções
10 zoológicas credenciadas, de forma tal que seja possível corroborar a identificação taxonômica
11 através de diagnoses atualizadas (Vink *et al.*, 2012).

12 A eficácia dos *primers* espécie-específicos foi avaliada usando o maior número de
13 indivíduos possível da espécie alvo de populações diferentes, com o objetivo de alcançar uma
14 significância estatística (Miot *et al.*, 2011). Para garantir a representatividade geográfica das
15 localidades de ocorrência de *H. intermedius*, os *primers* foram testados em populações
16 diferentes correspondentes ao Rio São Francisco, Rio Doce e afluentes. Todos os espécimes
17 vouchers de *H. intermedius* estão depositados na coleção de ictiologia do Museo de Ciências
18 Naturais da Pontifícia Universidade Católica de Minas Gerais (Tabela 13).

19

1 4. RESULTADOS

2 4.1. Genoma mitocondrial completo de *H. Intermedius*

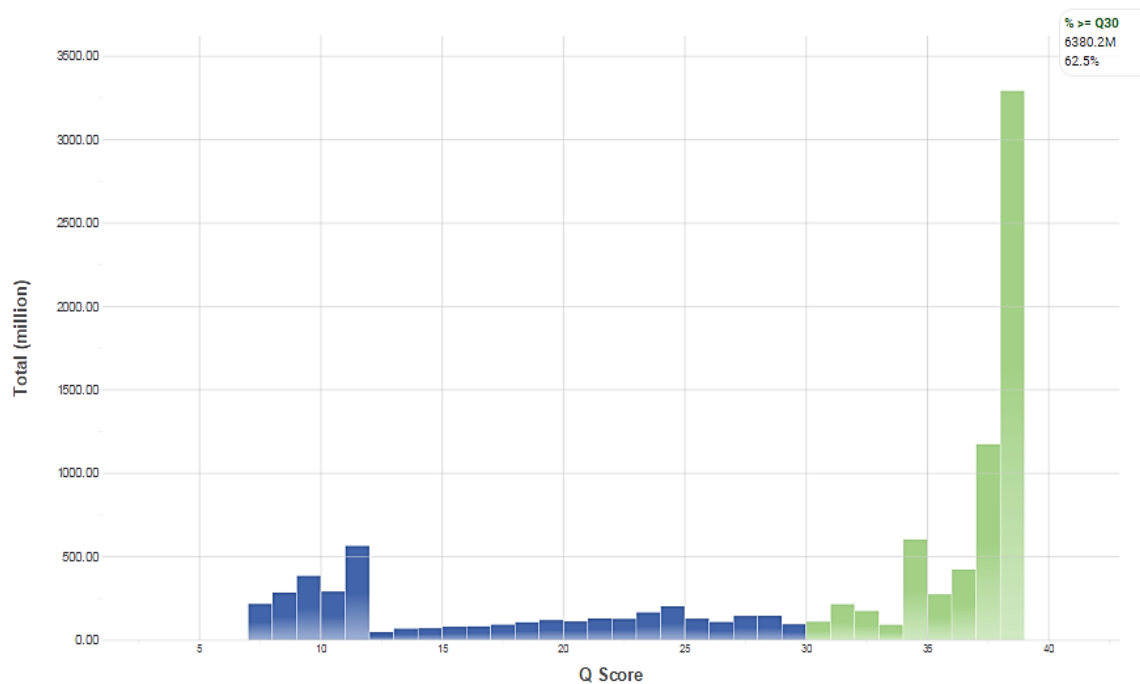
3 A preparação da biblioteca foi feita com uma concentração de DNA molde obtida a
4 partir do exemplar macho coletado em campo. A quantificação obtida no fluorímetro Qubit 2.0
5 (Life) usando uma leitura dsDNA BR foi de 34 ng/μL. Posteriormente foi feita uma análise de
6 integridade em gel de agarose 0.7% (Figura 4).



7

8 **Figura 4. Eletroforese de DNA usado na montagem da biblioteca genômica de *H. intermedius*.** Foi
9 feito um gel de agarose 0,7%, corado com brometo de etídio. PM: Marcador de peso molecular
10 Phoneutria com fragmentos de 700, 500 e 200 pares de base.

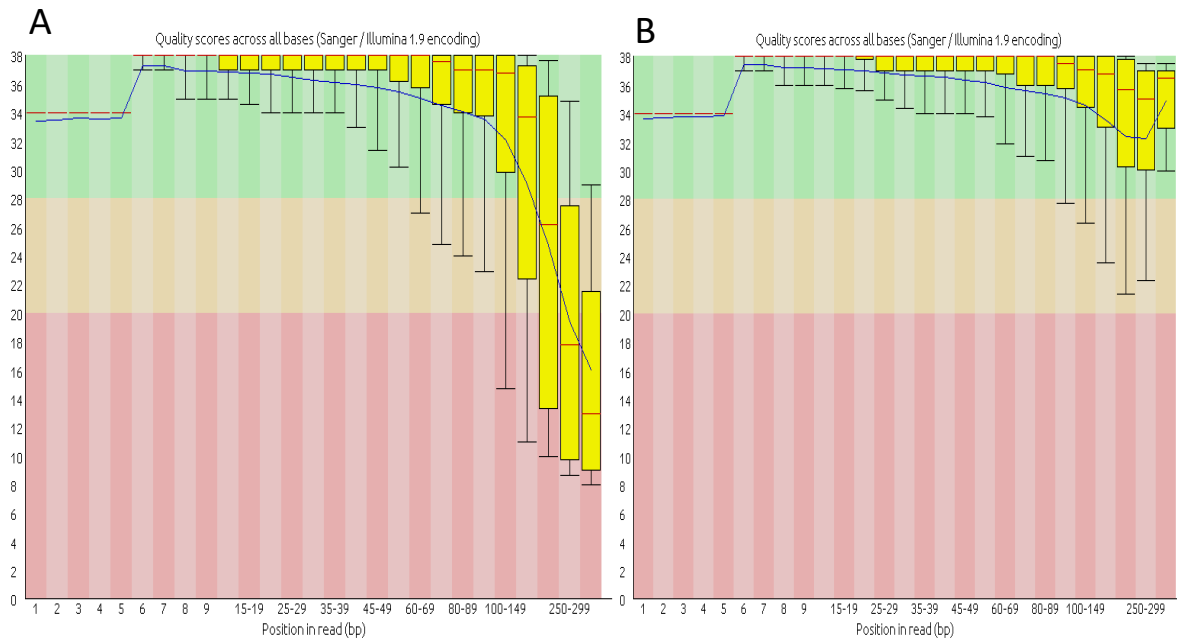
11 O sequenciamento NGS permitiu a obtenção de um total de 33.899.624 *reads*, com
12 um tamanho médio de 168 pb por *read*, e um total de 9.477.718.801 nucleotídeos. No total,
13 foram geradas 6,38 Gb *reads*, das quais 62,5% do total de sequências possuíram qualidade
14 maior ou igual a Q30 (Figura 5).



1

2 **Figura 5. Distribuição do Q score do Sequenciamento de Nova Geração de *H. intermedius*.** Gráfico
 3 cumulativo indicando a porcentagem de sequências e probabilidade de erro em sua leitura. **Eixo X:**
 4 índice de qualidade. **Eixo Y:** total de sequências obtidas. **Cor azul:** reads com qualidade de Q1 a Q29;
 5 **cor verde:** reads com qualidade de Q30 a Q40.

6 O resultado original do sequenciamento resultou em que 69,8% das reads
 7 apresentavam um índice PHRED = Q28 (aproximadamente dois erros de leitura por cada 1000
 8 bases), sendo encontradas 33.899.624 de reads com um tamanho médio de 300 pb por read.
 9 Foi realizada uma trimagem dos dados originais e obtido que o novo total de reads foi de
 10 25.422.830, das quais 99,9% apresentaram uma qualidade PHRED Q28. Obteve-se assim
 11 um total de 4.901.748.093 nucleotídeos, com um tamanho médio de 192,81 pb por read
 12 (Figura 6).



1

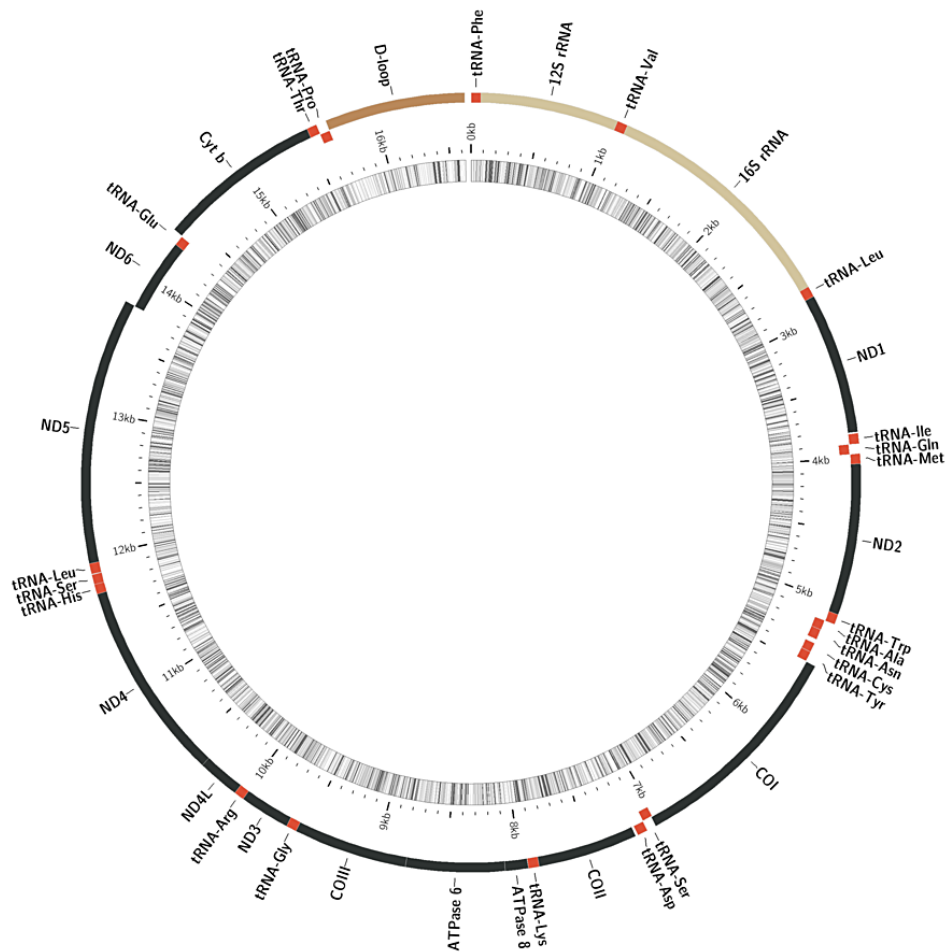
2 **Figura 6. Gráfico de distribuição da qualidade das reads obtidas de *H. intermedius*. A) total de**
 3 **seqüências antes da trimagem; B) total de seqüências após a trimagem. Eixo X:** posição da base na
 4 **read. Eixo Y:** índice de qualidade dividido verticalmente em três regiões, verde (qualidade excelente),
 5 **laranja (qualidade média) e vermelha (qualidade baixa). Para cada posição é feito um diagrama de**
 6 **caixa ilustrando a distribuição de valores de qualidade. No gráfico B (seqüências trimadas), a maioria**
 7 **das seqüências tiveram valores de qualidade de PHRED maiores que 30.**

8

9 **4.2. Montagem e anotação do genoma mitocondrial**

10 Usando os resultados do sequenciamento, foi montado o genoma mitocondrial de *H.*
 11 *intermedius* a partir de uma montagem *De novo*. Foram obtidos um total de 1.068 *contigs*, dos
 12 quais o de maior cobertura (304,45%) e tamanho (16.629 pb) foi identificado como a
 13 seqüência correspondente ao mitogenoma completo de *H. intermedius*.

14 As frequências nucleotídicas para cada base foram de 29,70% de Adenina, 26,33% de
 15 Timina, 15,45% de Guanina e 28,52% de Citosina, sendo o conteúdo GC um total de 43,97%.
 16 As regiões do genoma mitocondrial de *H. intermedius* apresentaram um arranjo típico de
 17 vertebrados: 13 genes codificadores de proteínas, 02 genes de RNA ribossômico, 22 genes
 18 de RNA transportador e 01 região controladora (D-loop) (Figura 7).



1

2 **Figura 7. Mapa de anotação do mitogenoma de *H. intermedius*.** São mostrados dois anéis
 3 concêntricos. **Anel interno:** diferencia regiões com maior conteúdo de GC (linhas escuras) ou menor
 4 conteúdo de GC (linhas claras) por cada 5pb. Os números alocados no anel pontilhado indicam o
 5 tamanho do genoma em pares de base (pb). **Anel externo:** mostra a anotação dos genes: genes
 6 codificadores de proteína (preto), genes de rRNA (marrom claro), genes de tRNA (vermelho) e região
 7 D-loop (marrom escuro). Os genes (tRNA-Gln, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, tRNA-Ser,
 8 *ND6*, tRNA-Glu, tRNA-Pro) alocados na fita leve (L) aparecem deslocados do anel externo, e os genes
 9 da fita pesada (H) permanecem no mesmo anel.

10 O códon de início ATG foi encontrado em 13 genes não codificadores de proteína
 11 codificados na fita H, e em *ND6* codificado na fita L. Foram encontrados quatro genes
 12 codificadores de proteína (PCGs): (*ND1*, *ATPase8*, *NDL4*, *ND5*) contendo o códon de
 13 terminação completo TAA, similar a *Salminus brasiliensis* (Brandao-Dias *et al.*, 2014),
 14 *Prochilodus lineatus* (Carmo *et al.*, 2014), *P. argenteus* e *P. costatus* (Chagas, Carmo, *et al.*,
 15 2015). Foi encontrado um gene (*COII*) contendo o códon de terminação AGG, similar a *Brycon*
 16 *orbignyanus* (Siqueira *et al.*, 2014) e *Piaractus mesopotamicus* (Pimentel *et al.*, 2014). Em
 17 quatro genes (*ND3*, *ND4*, *COII*, *Cytb*) se observou o códon de terminação incompletos T--, e
 18 dois genes (*ATPase6*, *COIII*) mostrando TA- como códon de terminação incompleto. Os
 19 códons de terminação incompletos foram provavelmente completados como TAA por
 20 poliadenilação pós-transcricional (Ojala *et al.*, 1981) (Tabela 4).

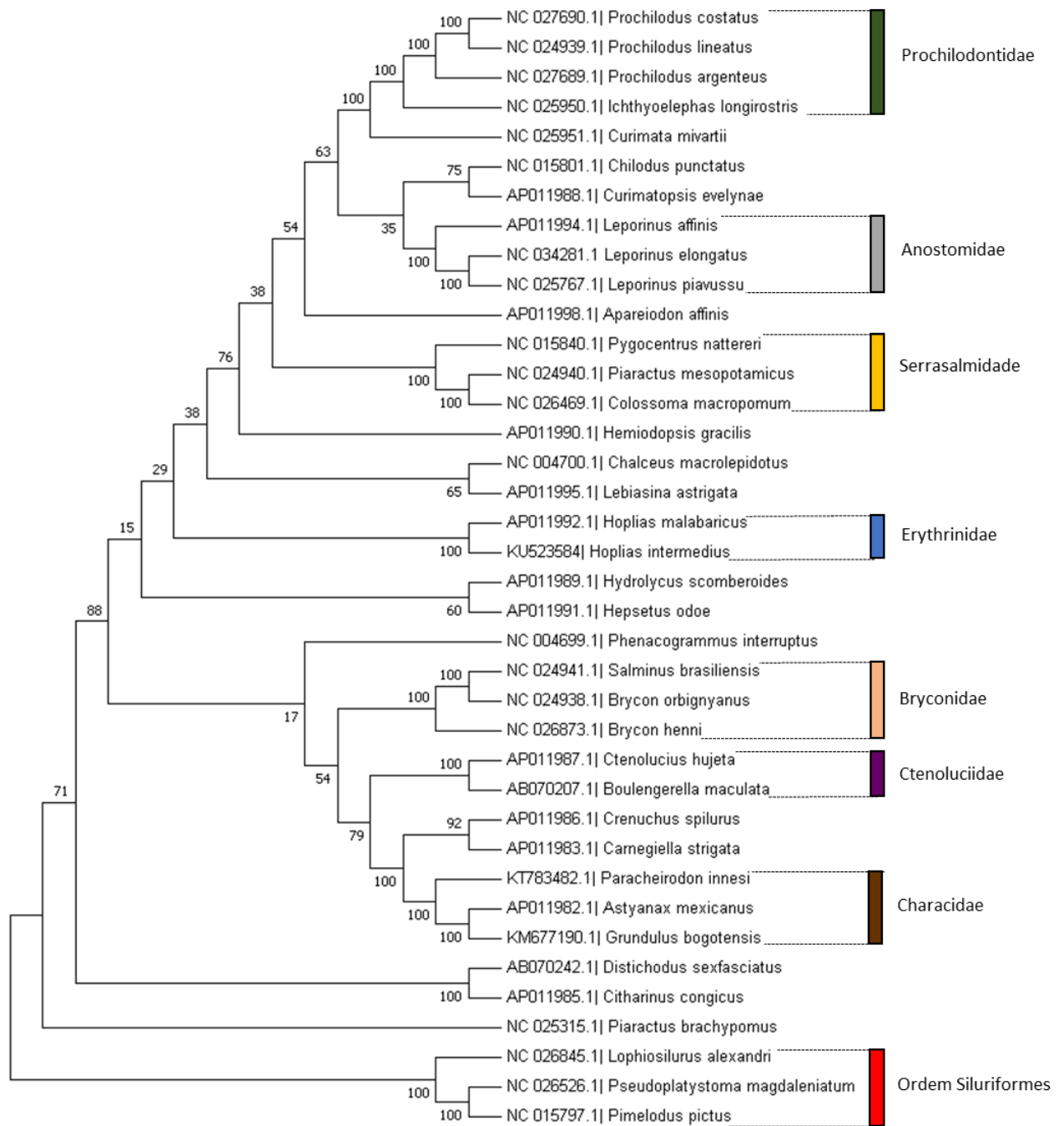
Tabela 4. Dados quantitativos e qualitativos da anotação do mitogenoma de *H. intermedius*.

Nome do Gene	Produto	Tipo de sequência	Número da primeira base	Número da última base	Tamanho da região	Tipo de Fita	Códon de início	Códon de parada
tRNA^{Phe}		tRNA	1	68	67	H		
12S rRNA		rRNA	69	1023	954	H		
tRNA^{Val}		tRNA	1024	1095	71	H		
16S rRNA		rRNA	1096	2771	1675	H		
tRNA^{Leu}		tRNA	2772	2846	74	H		
Nd1	NADH dehydrogenase subunidade 1	CDS	2847	3821	974	H	ATG	TAA
tRNA^{Ile}		tRNA	3831	3902	71	H		
tRNA^{Gln}		tRNA	3901	3971	70	L		
tRNA^{Met}		tRNA	3971	4039	68	H		
Nd2	NADH dehydrogenase subunidade 2	CDS	4040	5084	1044	H	ATG	T--
tRNA^{Trp}		tRNA	5085	5155	70	H		
tRNA^{Ala}		tRNA	5157	5225	68	L		
tRNA^{Asn}		tRNA	5227	5299	72	L		
tRNA^{Cys}		tRNA	5331	5396	65	L		
tRNA^{Tyr}		tRNA	5398	5468	70	L		
COI	cytochrome c oxidase subunidade I	CDS	5470	7026	1556	H	GTG	AGG
tRNA^{Ser}		tRNA	7018	7088	70	L		
tRNA^{Asp}		tRNA	7092	7164	72	H		

COII	cytochrome c oxidase subunit II	CDS	7179	7869	690	H	ATG	T--
tRNA^{Lys}		tRNA	7870	7943	73	H		
ATPase 8	ATPase subunidade 8	CDS	7945	8112	167	H	ATG	TAA
ATPase 6	ATPase subunidade 6	CDS	8103	8785	682	H	ATG	TA-
COIII	cytochrome c oxidase subunidade III	CDS	8786	9570	784	H	ATG	TA-
tRNA^{Gly}		tRNA	9571	9643	72	H		
Nd3	NADH dehydrogenase subunidade 3	CDS	9644	9992	348	H	ATG	T--
tRNA^{Arg}		tRNA	9993	10062	69	H		
Nd4L	NADH dehydrogenase subunidade 4L	CDS	10063	10359	296	H	ATG	TAA
Nd4	NADH dehydrogenase subunidade 4	CDS	10353	11733	1380	H	ATG	T--
tRNA^{His}		tRNA	11734	11802	68	H		
tRNA^{Ser}		tRNA	11803	11870	67	H		
tRNA^{Leu}		tRNA	11875	11947	72	H		
Nd5	NADH dehydrogenase subunidade 5	CDS	11948	13789	1841	H	ATG	TAA
Nd6	NADH dehydrogenase subunidade 6	CDS	13786	14304	518	L	ATG	TGA
tRNA^{Glu}		tRNA	14305	14373	68	L		
Cyt b	cytochrome b	CDS	14379	15519	1140	H	ATG	T--
tRNA^{Thr}		tRNA	15520	15590	70	H		
tRNA^{Pro}		tRNA	15589	15658	69	L		
D-Loop		D-loop	15659	16629	970	H		

4.3. Análise comparativa e agrupamento filogenético de mitogenomas

Com o objetivo de verificar a montagem do mitogenoma de *H. intermedius*, foi construída uma árvore filogenética usando o algoritmo Neighbor Joining (Figura 8). Foram recuperadas a partir da base de dados do NCBI, 38 sequências mitocondriais de espécies de peixes incluindo a Ordem Characiformes e Siluriformes. Devido a utilização de poucas espécies representativas de cada família, os valores de suporte de alguns clados tem um valor baixo. Como resultado mostra-se, como era esperado, o agrupamento em um mesmo clado dos mitogenomas de *H. intermedius* e *H. malabaricus* representando a Família Erythrinidae. Da mesma forma foi validado o clado da Família Prochilodontidae (Chagas et al., 2015). A árvore valida também o proposto por Mirande (2010), onde os gêneros *Brycon* e *Salminus* pertencem à Família Bryconidae. O grupo externo, aqui representado por duas Famílias da Ordem Siluriformes, Pimelodidae e Pseudopimelodidae, foi mantido como um único clado.

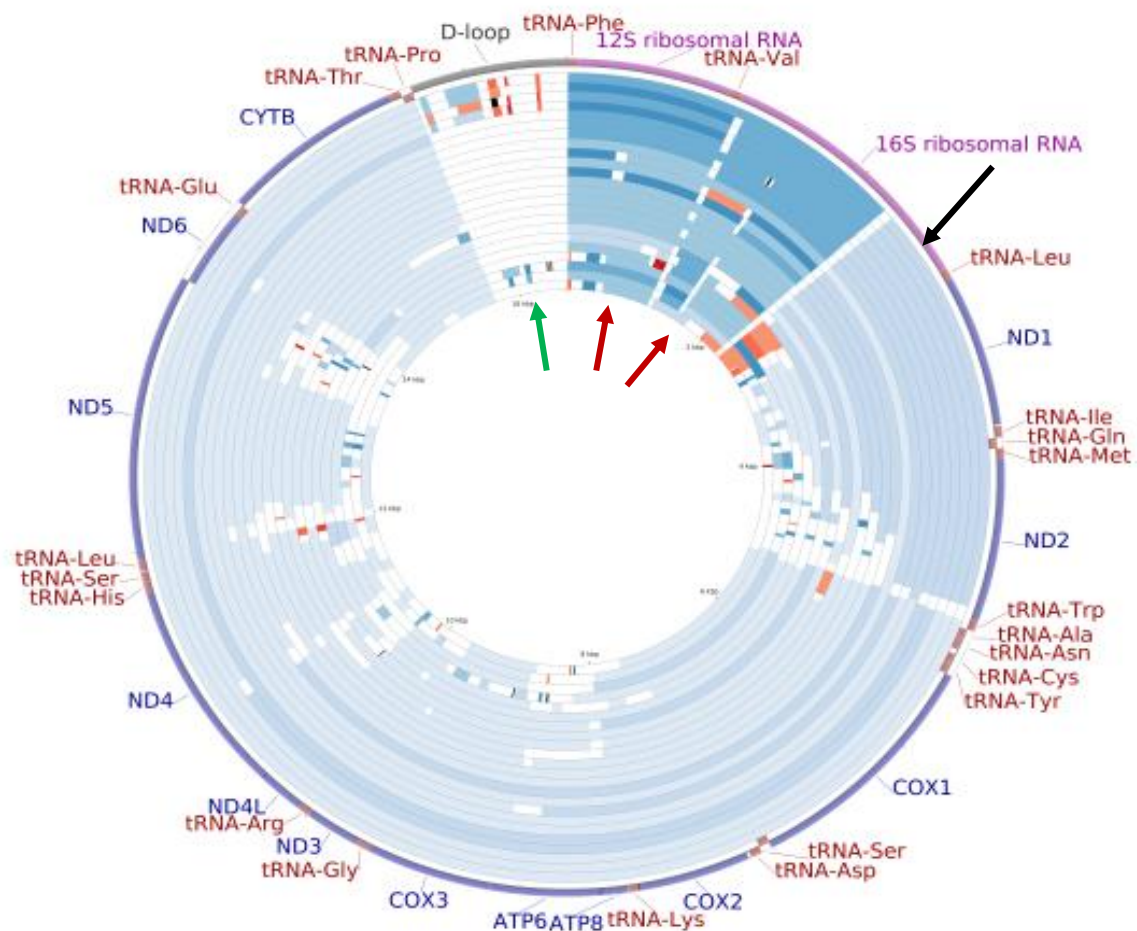


1

2 **Figura 8. Árvore de consenso Neighbour-joining usando 38 mitogenomas completos de**
3 **Characiformes e Siluriformes.** A árvore foi obtida usando bootstrap de 1000 réplicas e distâncias
4 evolutivas calculadas usando o método Kimura 2 parâmetros. Foram incluídos 38 mitogenomas,
5 sendo excluída a região D-loop. O grupo interno é representado por famílias da Ordem
6 Characiformes (Prochilodontidae, Anostomidae, Serrasalminidae, Erythrinidae, Bryconidae,
7 Ctenoluciidae e Characidae) e o grupo externo por duas famílias da Ordem Siluriformes
8 (Pimelodidae e Pseudopimelodidae, barra lateral vermelha). A família Erythrinidae (barra lateral
9 azul), agrupa em um único clado os mitogenomas de duas espécies do gênero *Hoplias*: *H.*
10 *intermedius* e *H. malabaricus*, com mitocondrial completo e quase completo, respectivamente. Os
11 números na base dos ramos, indicam os valores de suporte para cada ramo. Cada espécie está
12 acompanhada do número de acesso GenBank.

1 Usando o programa CGView Comparison Tool (Stothard e Wishart, 2005) foi feita
2 uma comparação da sequência mitocondrial completa de *H. intermedius* (Núñez-Rodríguez
3 *et al.*, 2016) com 23 mitogenomas completos de peixes. O mapa comparativo das
4 moléculas circulares (Figura 9) relaciona as sequências analisadas usando a porcentagem
5 de identidade obtidas pelo algoritmo do BLAST com base na molécula de mtDNA inteira.
6 O anel mais externo representa o mitogenoma de *H. intermedius*, e os anéis restantes
7 estão ordenados por grau de identidade. Quanto maior a identidade com a sequência alvo,
8 mais externa será a posição do anel, e da mesma forma, quanto menor a identidade, mais
9 interna a posição do anel. Para ilustrar a identidade por regiões gênicas, são atribuídas
10 cores diferentes, sendo as regiões representadas por cores claras as de baixo grau de
11 identidade e as regiões escuras as que possuem alto grau de identidade.

12 Assim, usando a ferramenta BLAST do NCBI, obteve-se que 6 espécies (*H.*
13 *malabaricus*, *Colossoma macropomum*, *Prochilodus argenteus*, *P. costatus*, *P. lineatus* e
14 *Pygocentrus nattereri*) possuem 82% de similaridade entre suas sequências. Como
15 esperado, a sequência com maior similaridade com *H. intermedius* foi a de *H. malabaricus*.
16 A similaridade dentro de regiões gênicas mostrou que existe similaridade em fragmentos
17 muito curtos dentro da região D-loop entre algumas sequências, enquanto uma
18 similaridade média é apresentada nas regiões 12S e 16S rRNA entre todas as sequências
19 analisadas. No entanto, os outros genes mitocondriais apresentaram uma similaridade
20 média-baixa também para todas as sequências analisadas.



1

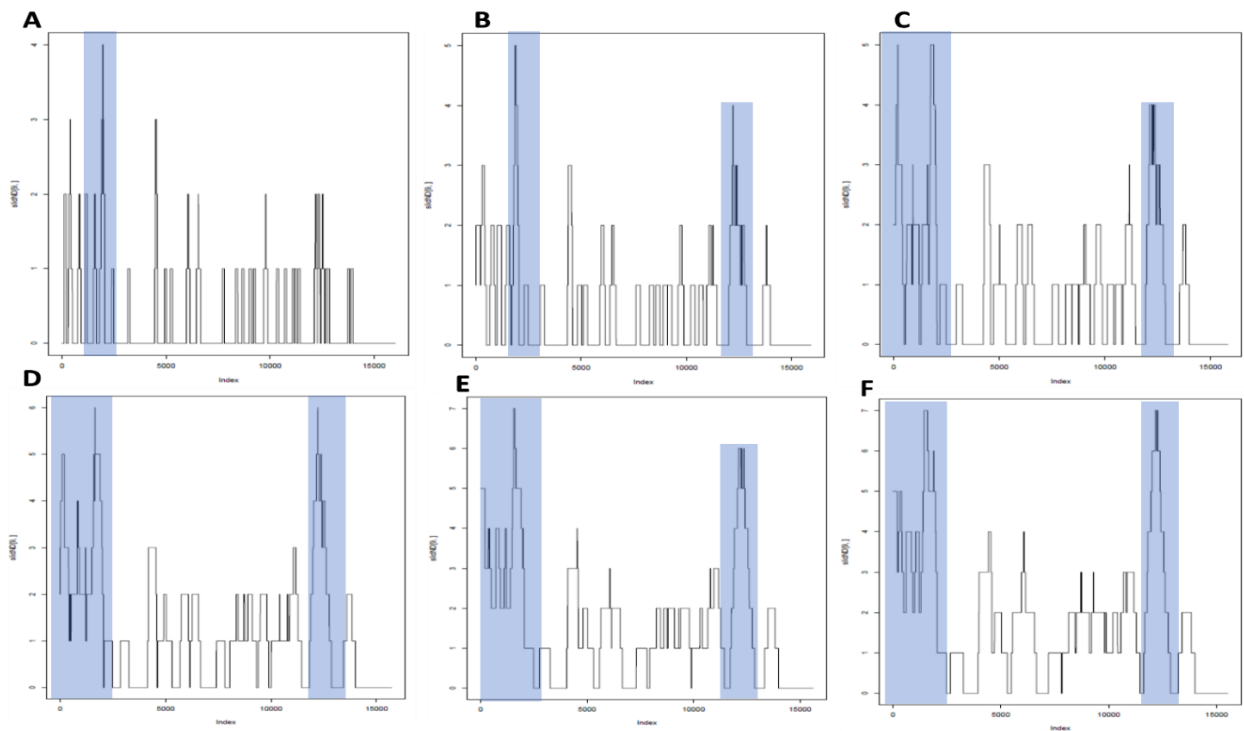
2 **Figura 9. Comparação de mapas de anotação dos mitogenomas de 24 espécies da ordem**
 3 **Characiformes. Seta preta:** anel externo representa a espécie alvo *H. intermedius*. **Seta verde:**
 4 região D-loop apresentando a menor similaridade entre seqüências, com exceção de fragmentos
 5 curtos internos. **Seta vermelha:** regiões 12S rRNA e 16S rRNA apresentam genes com a maior
 6 similaridade entre todas as regiões das seqüências comparadas.

7 **4.4. Localização de marcadores mitocondriais espécie-específicos**

8 O alinhamento prévio dos 38 mitogenomas foi utilizado para construir a árvore de
 9 consenso Neighbour-joining, assim como para a análise no pacote SPIDER, sendo todos
 10 os mitogenomas trimados para um tamanho de 16.120 pb, como requerido pelo programa.
 11 Nos gráficos para *H. intermedius* (Figura 10) e *H. malabaricus* (Figura 11) o eixo Y indica
 12 o número de *mismatches*, enquanto o eixo X nos indica a posição onde estão localizados
 13 os nucleotídeos diagnósticos no mitogenoma de *H. intermedius*. Foram detectadas um total
 14 de 11 regiões informativas (com maior número de *mismatches* ou *nucleotídeos*
 15 *diagnósticos*) distribuídas principalmente nos genes 16S rRNA e ND5 para *H. intermedius*
 16 e 9 regiões informativas para *H. malabaricus* localizadas no gene 16S rRNA. Os maiores
 17 valores de *mismatches* (de 4 a 7 *mismatches*) foram nas janelas de tamanho de 200 pb,

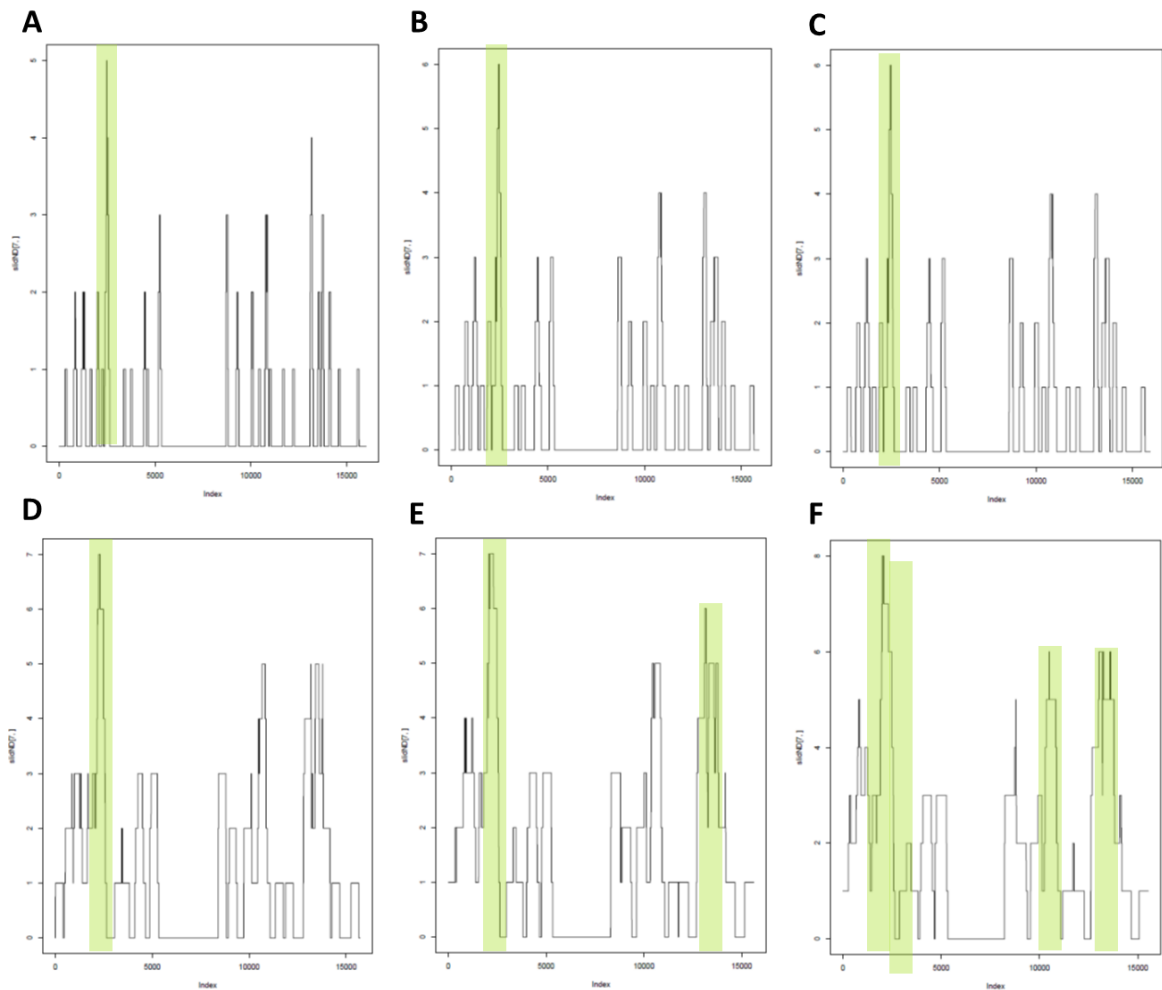
1 500 pb e 600 pb (consultar Anexos para ver o resultado completo na Tabela 13 para *H.*
2 *intermedius* e Tabela 14 para *H. malabaricus*).

3



4

5 **Figura 10. Representação gráfica dos nucleotídeos diagnósticos localizados pelo SPIDER**
6 **usando o mitogenoma de *H. intermedius* e seis amplitudes de janela diferentes.** Para cada
7 amplitude de janela são ressaltados os picos maiores (quadro azul) que correspondem a uma região
8 gênica do mitogenoma com maior quantidade de *mismatches*. **Eixo X:** posição numérica dos
9 nucleotídeos na sequência do mitogenoma completo da espécie alvo. **Eixo Y:** Quantidade de
10 nucleotídeos diagnósticos (*mismatches*). **A:** amplitude de janela de 100 pb, picos com máximo de 4
11 *mismatches*; **B:** amplitude de janela de 200 pb, picos com máximo de 5 e 4 *mismatches*; **C:**
12 amplitude de janela de 300 pb, picos com máximo de 5 e 4 *mismatches*; **D:** amplitude de janela de
13 400 pb, picos com máximo de 6 *mismatches*; **E:** amplitude de janela de 500 pb, picos com máximo
14 de 6 e 7 *mismatches*; **F:** amplitude de janela de 600 pb, picos com máximo de 7 *mismatches*.

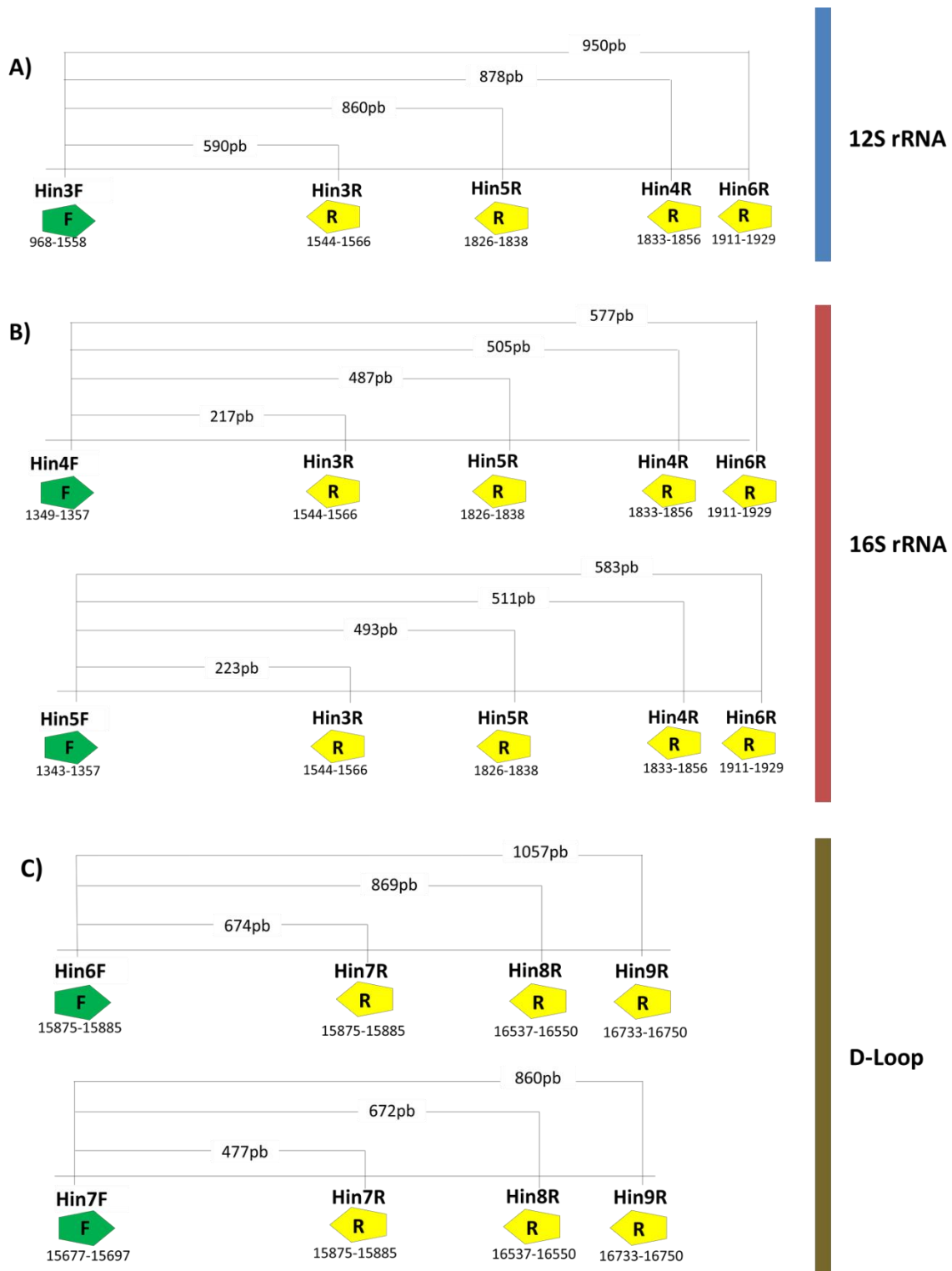


1

2 **Figura 11. Representação gráfica dos nucleotídeos diagnósticos localizados pelo SPIDER**
 3 **usando o mitogenoma de *H. malabaricus* e seis amplitudes de janela diferentes.** Para cada
 4 amplitude de janela são ressaltados os picos maiores (quadro verde) que correspondem a uma
 5 região gênica do mitogenoma com maior quantidade de *mismatches*. **Eixo X:** posição numérica dos
 6 nucleotídeos na sequência do mitogenoma completo da espécie alvo. **Eixo Y:** Quantidade de
 7 nucleotídeos diagnósticos (*mismatches*). **A:** amplitude de janela de 100 pb, picos com máximo de 4
 8 *mismatches*; **B:** amplitude de janela de 200 pb, picos com máximo de 5 e 4 *mismatches*; **C:**
 9 amplitude de janela de 300 pb, picos com máximo de 5 e 4 *mismatches*; **D:** amplitude de janela de
 10 400 pb, picos com máximo de 6 *mismatches*; **E:** amplitude de janela de 500 pb, picos com máximo
 11 de 6 e 7 *mismatches*; **F:** amplitude de janela de 600 pb, picos com máximo de 7 *mismatches*.

1 **4.5. Desenho de *primers* mitocondriais**

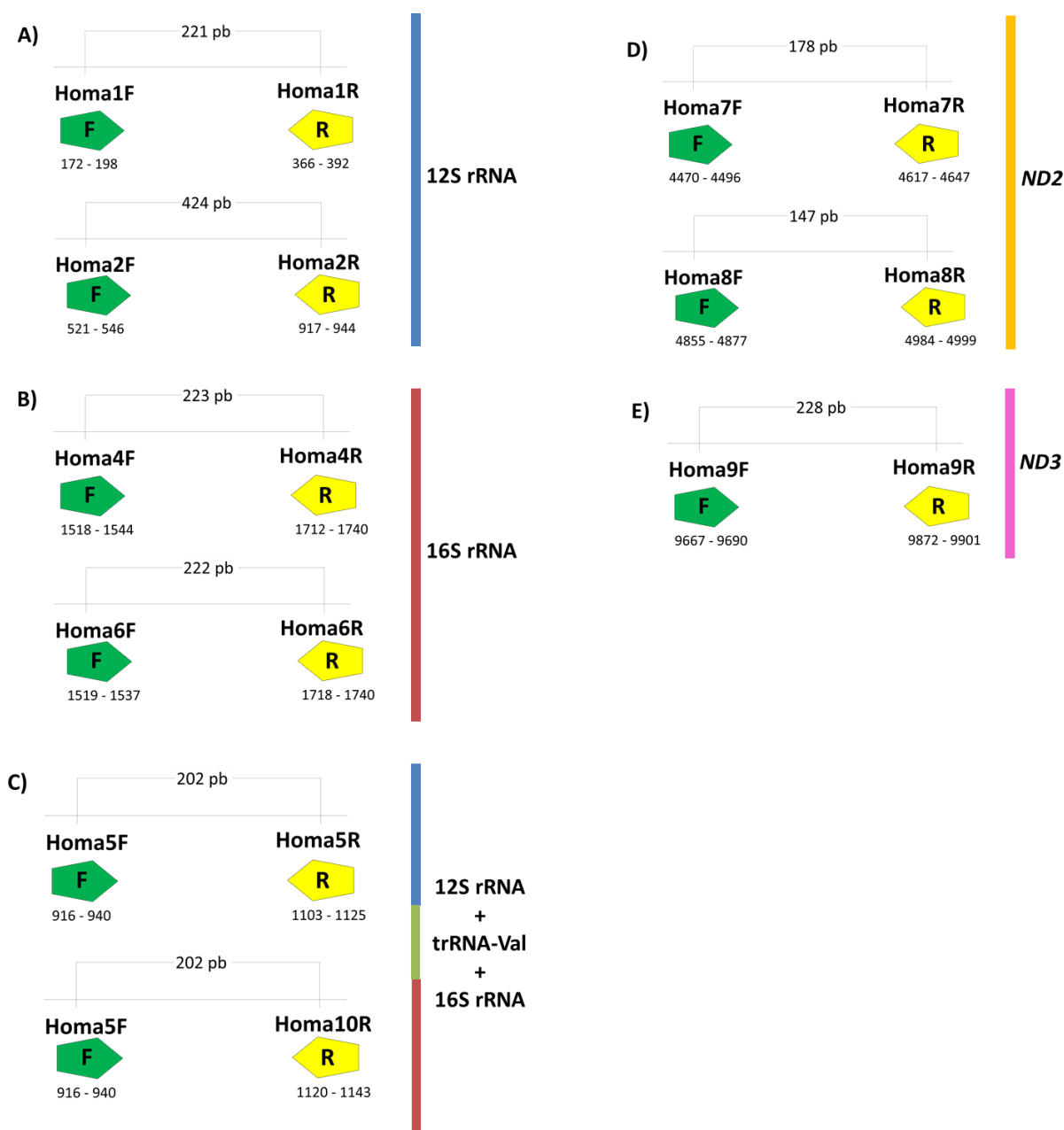
2 Após a localização das regiões de potenciais marcadores, pela busca manual e pelo
3 SPIDER, foi feita a busca pelos primers que amplificariam estas regiões. A partir do
4 alinhamento das 38 sequências de mitogenoma completo de peixes, foi realizada uma
5 busca manual de sequencias de primers, regiões altamente variáveis de tamanho
6 aproximado de 20 pb usando o programa BioEdit. No total foram detectadas 18 regiões
7 para *H. malabaricus* e 24 regiões de alta variabilidade para *H. intermedius*. Logo após,
8 foram escolhidos os *primers forward* e reversos, desenhado o mapa de anelamento e
9 realizada a localização da região gênica do marcador para *H. intermedius* (Figura 13) e *H.*
10 *malabaricus* (Figura 14). Cada par de *primer* foi analisado *in silico* na plataforma
11 PrimerBLAST (Tabela 5).



1

2 **Figura 12. Mapa de anelamento dos primers espécie-específicos desenhados para *H.***
 3 ***intermedius*. Seta verde: primer forward. Seta amarela: primer reverso. Acima de cada seta**
 4 **está descrito o nome do primer e, os números debaixo delas, indicam o intervalo de nucleotídeos do**
 5 **mitogenoma de *H. intermedius* no qual foi desenhado o primer. Os números entre as linhas**
 6 **horizontais indicam o tamanho do amplicon esperado. A) Barra azul: marcadores específicos para**
 7 ***H. intermedius* isolado na região do gene 12S rRNA, B) barra vermelha: marcadores específicos**
 8 **para *H. intermedius* isolado na região do gene 16S rRNA, C) barra marrom: marcadores específicos**
 9 **para *H. intermedius* isolado na região do do D-Loop.**

1



2

3 **Figura 13.** Mapa de anelamento dos *primers* espécie-específicos desenhados para *H.*
 4 *malabaricus*. **Seta verde:** *primer forward*. **Seta amarela:** *primer reverso*. Acima de cada seta está
 5 descrito o nome do *primer* e, os números debaixo delas, indicam o intervalo de nucleotídeos da
 6 sequência do mitogenoma de *H. malabaricus* no qual foi desenhado o *primer*. Os números entre as
 7 linhas horizontais indicam o tamanho do amplicon esperado. **A) Barra azul:** marcador específico de
 8 *H. malabaricus* isolado na região do gene 12S rRNA, **B) barra vermelha:** marcadores específicos para
 9 *H. malabaricus* isolado na região do gene 16S rRNA, **C) barra de 3 cores:** marcadores específicos
 10 para *H. malabaricus* isolado no final da região 16SrRNA até o início da região 12S rRNA, **D) barra**
 11 **amarela:** marcadores específicos para *H. malabaricus* isolado na região do gene *ND2*, **E) barra rosa:**
 12 marcadores específicos para *H. malabaricus* isolado na região do gene *ND3*.

Tabela 5. Características das sequências de *primers* espécie específicos para *H. intermedius* e *H. malabaricus*.

Primer	Espécie-específico	Região	Sequência 5' 3'	T° de anelamento (°C)	Tamanho (pb)	Conteúdo de GC (%)
Hin3F	<i>H. intermedius</i>	12S rRNA	CACATATTCCAGCGGG	51,22	16	56,25
Hin5F	<i>H. intermedius</i>	16S rRNA	GCCAGTCGTA CT TGA	48,76	15	53,33
Hin3R	<i>H. intermedius</i>	16S rRNA	GGTGTGGTGGCCCTGGTGAG	65,19	20	70,00
Hin4F	<i>H. intermedius</i>	16S rRNA	CGTACTTGAGTGAG	42,00	14	50,00
Hin4R	<i>H. intermedius</i>	16S rRNA	ACGGGCTATAGGAGAAGGGTACAC	63,06	24	54,17
Hin5R	<i>H. intermedius</i>	16S rRNA	GTACACCTTCTTA	33,91	13	38,46
Hin6R	<i>H. intermedius</i>	D-loop	GTTTTTGTCATTTTA	35,89	15	20,00
Hin6F	<i>H. intermedius</i>	D-loop	TTCCCATGATT	27,55	11	36,36
Hin7F	<i>H. intermedius</i>	D-loop	CTCACCCTAACTCCCAAAGC	58,57	21	52,38
Hin7R	<i>H. intermedius</i>	D-loop	GCTAGCCCCTTTTAACAAGCAGG	62,40	24	50,00
Hin8R	<i>H. intermedius</i>	D-loop	TTCCTCGGGGAGCG	53,56	14	71,43
Hin9R	<i>H. intermedius</i>	D-loop	GGGTAAAGCTGCTGATGG	55,40	18	55,56
Homa1F	<i>H. malabaricus</i>	16S rRNA	AGCTGGCATCAGGCACCAAATTTAAG	64,70	27	44,44
Homa1R	<i>H. malabaricus</i>	16S rRNA	TTTGGCTTTATTACGGGAGTCCTCTAA	61,94	27	40,74
Homa2F	<i>H. malabaricus</i>	16S rRNA	CACAGATGTTGTCATTACACAAACAT	58,86	26	34,62
Homa2R	<i>H. malabaricus</i>	16S rRNA	AATTACTTGGGTTAAGTTAATGAATGGG	58,54	28	32,14
Homa3F	<i>H. malabaricus</i>	16S rRNA	CCCATTCTTA ACTTAACCCAAGTAATT	58,54	28	32,14
Homa3R	<i>H. malabaricus</i>	16S rRNA	TATTTAAATATGTGGGTGGG	50,04	20	35,00
Homa4F	<i>H. malabaricus</i>	16S rRNA	ACTAGAATAATTTCCACACATGAGAAT	57,09	27	59,63

Homa4R	<i>H. malabaricus</i>	16S rRNA	GATAAAGGGGTGTAGGATAAGATATCAGG	60,39	29	41,38
Homa5F	<i>H. malabaricus</i>	12S rRNA+ tRNA-Val+ 16S rRNA	CCCCATTCATTAACCTAACCCAAGT	59,28	25	40,00
Homa5R	<i>H. malabaricus</i>	12S rRNA+ tRNA-Val+ 16S rRNA	TAAATATGTGGGTGGGCTAGC	57,50	21	47,62
Homa6F	<i>H. malabaricus</i>	16S rRNA	CTAGAATAATTTCCACACA	46,99	19	31,58
Homa6R	<i>H. malabaricus</i>	16S rRNA	GATAAAGGGGTGTAGGATAAGA	54,12	22	40,91
Homa7F	<i>H. malabaricus</i>	16S rRNA	CCGTTAATCCAACAATTCTTATCACTC	58,88	27	37,04
Homa7R	<i>H. malabaricus</i>	16S rRNA	GTGCTTAGGGTTAGTATTGTAAGTTGG	59,93	27	40,74
Homa8F	<i>H. malabaricus</i>	16S rRNA	CTTACACTAGGACTTCTACTATAT	52,12	24	33,33
Homa8R	<i>H. malabaricus</i>	16S rRNA	TCCTTGATCTAGGAAGAC	49,29	18	44,44
Homa9F	<i>H. malabaricus</i>	16S rRNA	TGCCATTACATTCTCCCTATATATC	55,86	25	36,00
Homa9R	<i>H. malabaricus</i>	16S rRNA	TGATATGGGTGGAGTAGCCC	58,27	20	55,00
Homa10R	<i>H. malabaricus</i>	12S rRNA+ tRNA-Val+ 16S rRNA	CCGATGGTTACCTTTTCTG	53,38	19	47,37

1 4.6. Padronização de PCR

2 Foram realizadas as reações de padronização para cada um dos conjuntos de
3 *primers* que se mostraram específicos para à espécie alvo nos testes *in silico* (Tabela 6,
4 Tabela 7, Tabela 8 e Tabela 9).

5
6 **Tabela 6. Primers espécie-específicos para *H. intermedius* e *H. malabaricus***

Conjunto de primers	Região gênica	Possíveis espécies alvo
Hin3F-Hin3R	12S rRNA	<i>Hoplias intermedius</i>
Hin3F-Hin4R	12S rRNA	<i>Hoplias intermedius</i>
Hin5F-Hin3R	16S rRNA	<i>Hoplias intermedius</i> <i>Kryptolebias marmoratus</i> <i>Clupea harengus</i>
Hin5F-Hin4R	16S rRNA	<i>Hoplias intermedius</i>
Hin7F-Hin7R	D-Loop	<i>Hoplias intermedius</i>
Hin7F-Hin7R	D-Loop	<i>Hoplias intermedius</i>
Hin7F-Hin7R	D-Loop	<i>Hoplias intermedius</i>
Homa5F-Homa5R	12S rRNA+ tRNA-Val+ 16S rRNA	<i>Hoplias malabaricus</i>
Homa5F-Homa10R	12S rRNA+ tRNA-Val+ 16S rRNA	<i>Hoplias malabaricus</i>

7

1 Tabela 7. Volumes utilizados nos testes de padronização das reações de PCR usando DMSO
 2 como agente adstringente.

Tipo de Tampão	DMSO 3%		DMSO 5%		DMSO 7%	
	IB, IC, IIC	IVB	IB, IC, IIC	IVB	IB, IC, IIC	IVB
H ₂ O (µL)	12.6	10.6	12.2	10.2	11.8	9.8
DMSO (µL)	0.6	0.6	1.0	1.0	1.4	1.4
Tampão (µL)	2.0	4.0	2.0	4.0	2.0	4.0
dNTPs (µL)	2.0	2.0	2.0	2.0	2.0	2.0
Taq polimerase (µL)	0.2	0.2	0.2	0.2	0.2	0.2
<i>Primer Forward</i> (µL)	0.8	0.8	0.8	0.8	0.8	0.8
<i>Primer Reverso</i> (µL)	0.8	0.8	0.8	0.8	0.8	0.8
DNA (µL)	1.0	1.0	1.0	1.0	1.0	1.0
Vf por Reação (µL)	20.0	20.0	20.0	20.0	20.0	20.0

3

4 Tabela 8. Volumes utilizados nos testes de padronização das reações de PCR usando KCl
 5 como agente adstringente.

Tipo de Tampão	KCl 30 mM		KCl 50 mM		KCl 100 mM	
	IB, IC, IIC	IVB	IB, IC, IIC	IVB	IB, IC, IIC	IVB
H ₂ O (µL)	12.9	10.9	12.7	10.7	12.2	10.2
DMSO (µL)	0.3	0.3	0.5	0.5	1.0	1.0
Tampão (µL)	2.0	4.0	2.0	4.0	2.0	4.0
dNTPs (µL)	2.0	2.0	2.0	2.0	2.0	2.0
Taq polimerase (µL)	0.2	0.2	0.2	0.2	0.2	0.2
<i>Primer Forward</i> (µL)	0.8	0.8	0.8	0.8	0.8	0.8
<i>Primer Reverso</i> (µL)	0.8	0.8	0.8	0.8	0.8	0.8
DNA (µL)	1.0	1.0	1.0	1.0	1.0	1.0
Vf por Reação (µL)	20.0	20.0	20.0	20.0	20.0	20.0

6

1 **Tabela 9. Volumes utilizados nos testes de padronização das reações de PCR sem usar**
 2 **agente adstringente.**

	IB, IC, IIC	IVB
H₂O (µL)	13.2	11.2
Tampão (µL)	2.0	4.0
dNTPs (µL)	2.0	2.0
Taq polimerase (µL)	0.2	0.2
Primer Forward (µL)	0.8	0.8
Primer Reverso (µL)	0.8	0.8
DNA (µL)	1.0	1.0
Vf por Reação (µL)	20.0	20.0

3

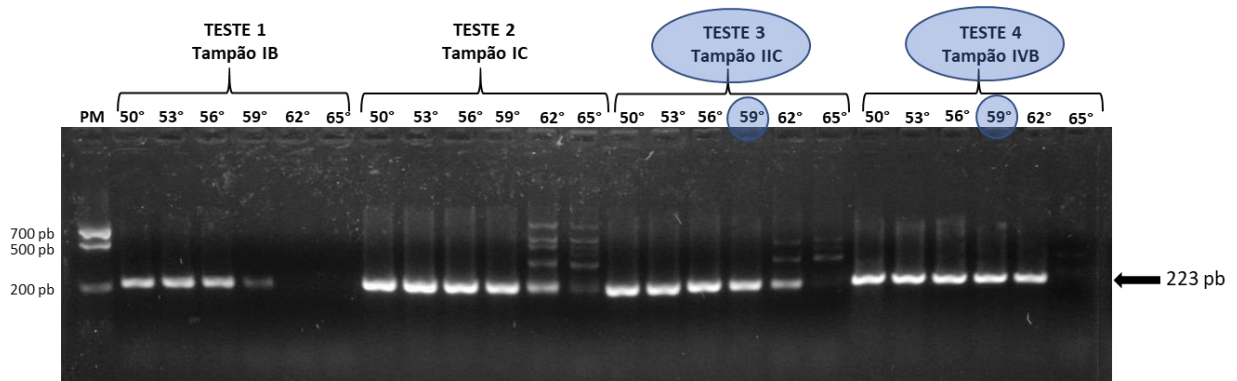
4 Após 1587 reações de PCR, foram obtidas as melhores condições de amplificação
 5 para dois pares de *primers* específicos para *H. intermedium*: Hin5F-Hin3R (Figura 14), e
 6 Hin5F-Hin4R (Figura 15); e um par de *primers* para *H. malabaricus*: Homa5F-Homa5R
 7 (Figura 16). Os volumes das reações padronizadas são detalhados na Tabela 10.

8 **Tabela 10. Condições ótimas de amplificação de primers espécie-específicos de *H.***
 9 ***intermedium*.**

Combinação de Primers	Região	Tamanho do Fragmento (pb)	Condições ótimas de amplificação
Hin5F - Hin3R	16Sr RNA	223	Tampão IVB 59°C 30 ciclos
Hin5F - Hin4R	16Sr RNA	511	Tampão IIC 62°C 30 ciclos
Homa5F-Homa5R	12S rRNA+ tRNA-Val+ 16S rRNA	202	Tampão IC DMSO 3% 56°C 30 ciclos

10

1

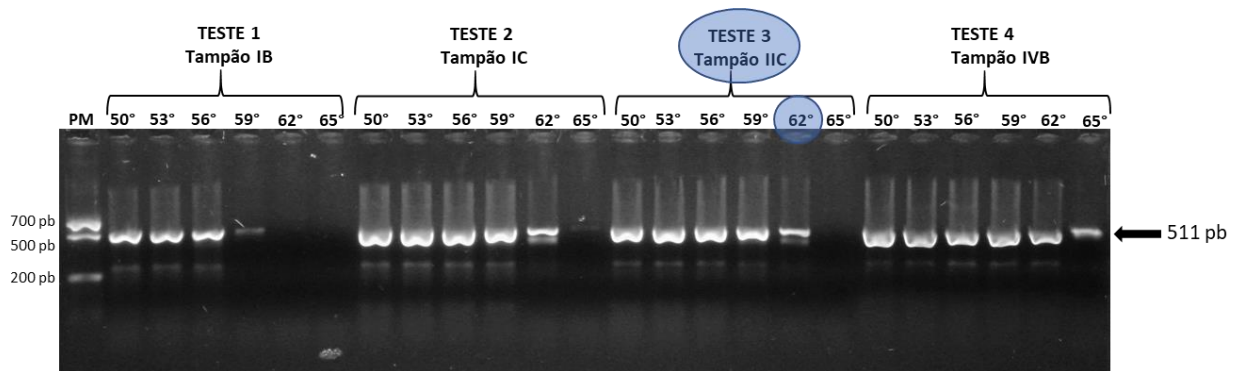


2

3 **Figura 14. Padronização da PCR usando o marcador Hin5F-Hin3R específico para *H.***
4 ***intermedius*.** Os *primers* amplificam uma região parcial do gene 16S rRNA com tamanho esperado
5 de 223pb. Foi usada uma concentração de DNA alvo de 10ng/uL. Foram realizados quatro testes:
6 **Teste 1:** solução tampão IB e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
7 **Teste 2:** solução tampão IC e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
8 **Teste 3:** solução tampão IIC e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
9 **Teste 4:** solução tampão IVB e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
10 **Círculo azul:** melhores parâmetros de amplificação do fragmento esperado no teste 3 (tampão IIC
11 e 59°C de temperatura de anelamento) e teste 4 (tampão IVB com 59°C de temperatura de
12 anelamento). **Seta preta:** Tamanho de amplicon esperado. Os amplicons obtidos foram aplicados
13 em uma electroforese em gel de agarose 1.5% corado com brometo de etídio, usando um
14 marcador de peso molecular (PM) de 700, 500 e 200 pb.

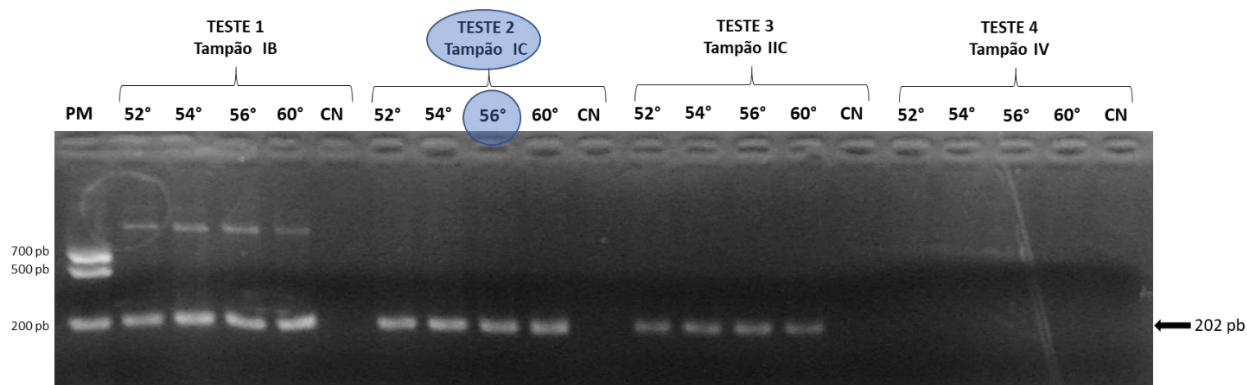
15

16



17

18 **Figura 15. Padronização da PCR usando o marcador Hin5F-Hin4R específico para *H.***
19 ***intermedius*.** Os *primers* amplificam uma região parcial do gene 16S rRNA com tamanho esperado
20 de 511 pb. Foi usada uma concentração de DNA alvo de 10ng/uL. Foram realizados quatro testes:
21 **Teste 1:** solução tampão IB e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
22 **Teste 2:** solução tampão IC e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
23 **Teste 3:** solução tampão IIC e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
24 **Teste 4:** solução tampão IVB e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
25 **Círculo azul:** melhor parâmetro de amplificação do fragmento esperado no teste 3 (tampão IIC e
26 62°C de temperatura de anelamento). **Seta preta:** Tamanho de amplicon esperado. Os amplicons
27 obtidos foram aplicados em uma electroforese em gel de agarose 1.5% corado com brometo de
28 etídio, usando um marcador de peso molecular (PM) de 700, 500 e 200 pb.



1

2 **Figura 16. Padronização da PCR usando o marcador Hin5F-Hin5R específico para *H.***
 3 ***malabaricus*.** Os *primers* amplificam uma região parcial do gene 16S rRNA com tamanho esperado
 4 de 202 pb. Foi usada uma concentração de DNA alvo de 10ng/uL. Foram realizados quatro testes:
 5 **Teste 1:** solução tampão IB e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
 6 **Teste 2:** solução tampão IC e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
 7 **Teste 3:** solução tampão IIC e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
 8 **Teste 4:** solução tampão IVB e seis temperaturas de anelamento (50°C, 53°C, 56°C, 62°C, 65°C).
 9 **Círculo azul:** melhor parâmetro de amplificação do fragmento esperado no Teste 2 (solução
 10 tampão IC e 56°C de temperatura de anelamento). **Seta preta:** Tamanho de amplicon esperado.
 11 Os amplicons obtidos foram aplicados em uma electroforese em gel de agarose 1.5% corado com
 12 brometo de etídio, usando um marcador de peso molecular (PM) de 700, 500 e 200 pb.

13 4.7. Testes de Validação

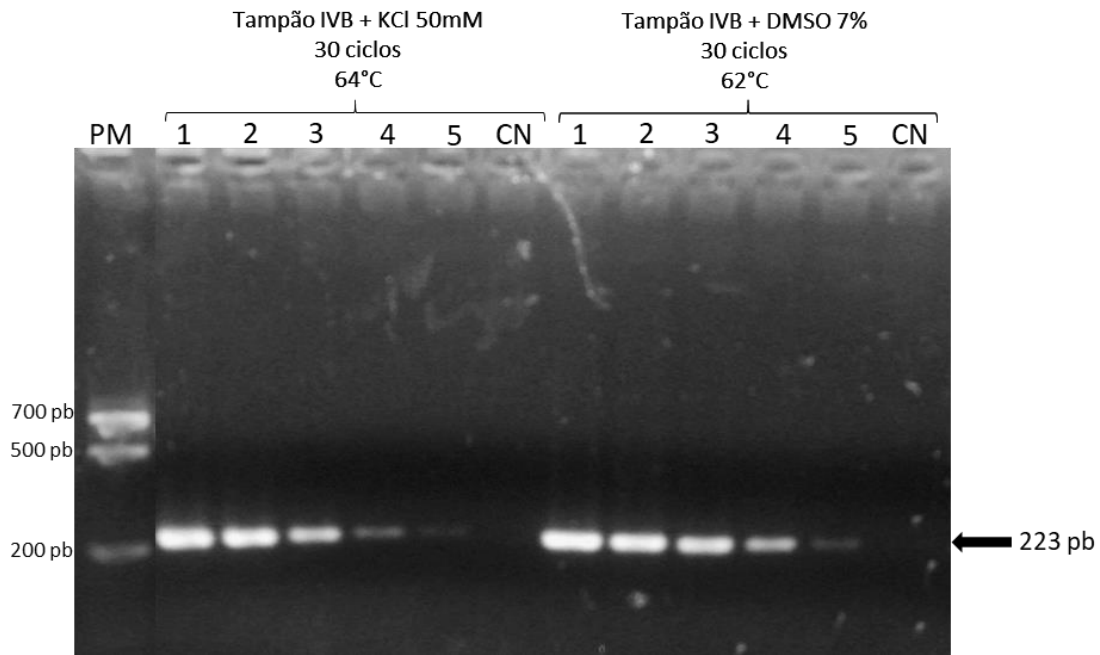
14 4.7.1. Teste de sensibilidade

15 A sensibilidade de todos os conjuntos de *primers* foi avaliada separadamente
 16 fazendo diluições seriadas logarítmicas de base 10, a partir de uma quantidade inicial de
 17 10 ng/μL de DNA até a não visualização da banda de amplificação. Após três repetições
 18 de cada teste, foi obtido que o conjunto de *primers* Hin5F-Hin3R amplificou até a
 19 quantidade de 10⁻⁴ ng/μL de DNA molde (Figura 17 e Tabela 11), enquanto que Hin5F-
 20 Hin4R até 10⁻³ ng/μL (Figura 18 e Tabela 11).

21 **Tabela 11. Sensibilidade dos *primers* específicos para *H. intermedius* e *H. malabaricus*. O**
 22 **número de símbolos (+) indica a intensidade da banda para cada concentração de DNA**
 23 **usada. O símbolo (-) indica ausência de banda.**

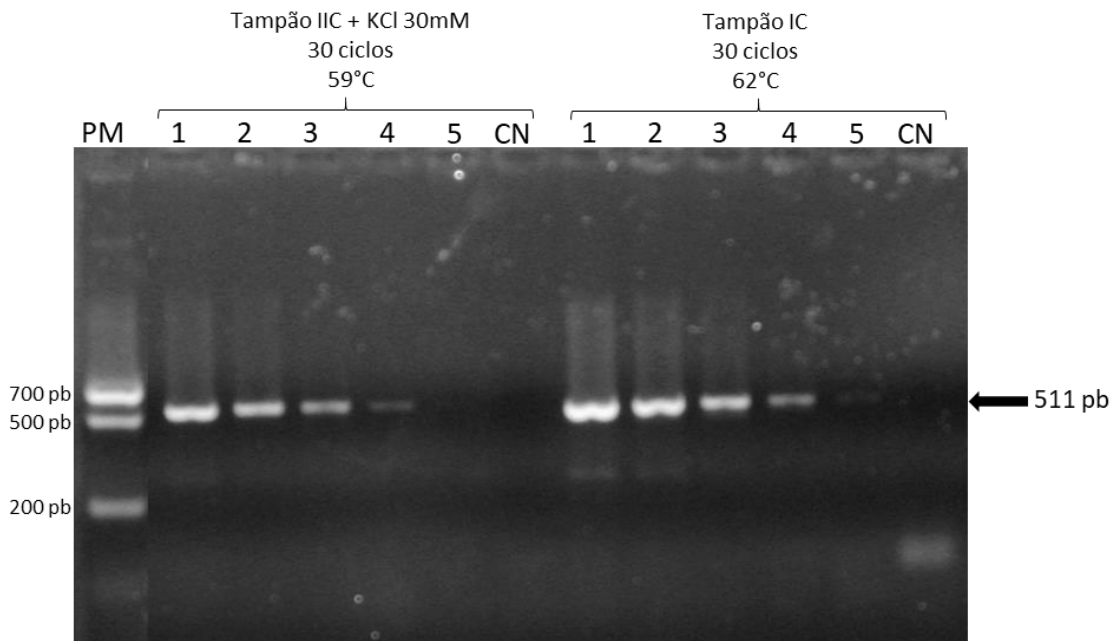
<i>Primers</i>	Hin5F - Hin3R	Hin5F - Hin4R
Localização do amplicon	16S rRNA	16S rRNA
Tamanho do fragmento	223 pb	511 pb
Concentração de DNA		
10 ng	++++	+++
1 ng	++++	++
0.1 ng	+++	+
0.01 ng	++	+
0.001 ng	+	-
controle negativo	-	-

1



2

3 **Figura 17. Teste de sensibilidade dos primers específicos para *H. intermedius* Hin5F-Hin3R.**
4 Tamanho de amplicon esperado de 223 pb. As letras indicam a quantidade de DNA molde, **A:** 10
5 ng/uL, **B:** 1ng/uL, **C:** 0.1 ng/uL, **D:** 0.01 ng/uL, **E:** 0.001 ng/uL, **CN:** Controle negativo (ausência de
6 DNA na reação). **PM:** Padrão de peso molecular (700 pb, 500 pb e 200 pb). **Seta preta:** indica o
7 tamanho do amplicon esperado. Como resultado o *primer* se mostrou sensível até de 0.001 ng/uL
8 de DNA alvo.



9

10 **Figura 18. Teste de sensibilidade dos primers específicos para *H. intermedius* Hin5F-Hin4R**
11 Tamanho de amplicon esperado de 511 pb. Foram utilizadas duas condições de amplificação
12 previamente testadas na etapa de padronização. As letras indicam quantidade de DNA molde, **A:**
13 10 ng/uL, **B:** 1ng/uL, **C:** 0.1 ng/uL, **D:** 0.01 ng/uL, **E:** 0.001 ng/uL, **CN:** Controle negativo (ausência
14 de DNA na reação). **PM:** Padrão de peso molecular (700 pb, 500 pb e 200 pb). **Seta preta:** indica
15 o tamanho do amplicon esperado. Como resultado o *primer* se mostrou sensível até 0.01 ng/uL de
16 DNA alvo.

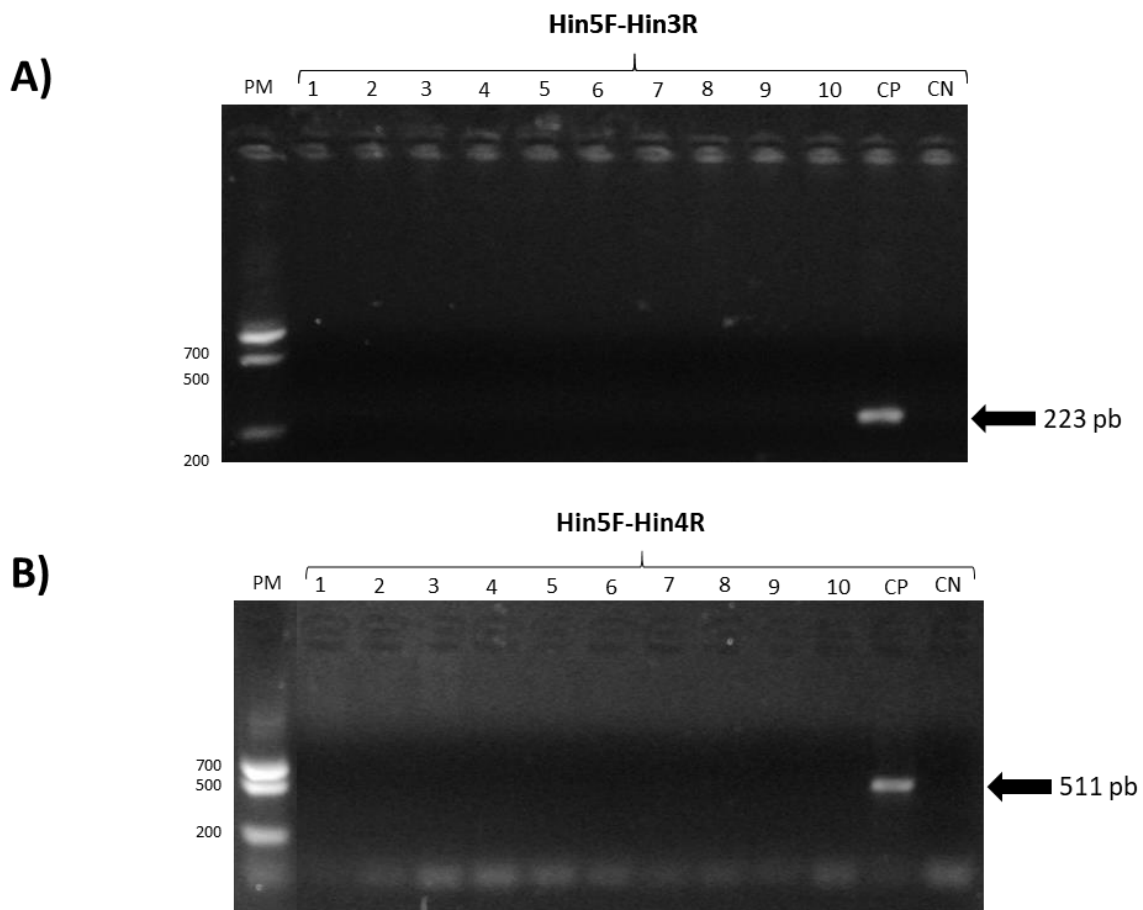
1 **4.7.2. Teste de especificidade**

2 Unicamente os *primers* desenhados na região 16S rRNA se mostraram altamente
 3 específicos para *H. intermedius* (Figura 19) e *H. malabaricus* (Figura 20). O marcador
 4 universal Fish16S com tamanho de 121 pb (Figura 21) foi utilizado como controle positivo
 5 da integridade do DNA utilizado, verificando que as ampliações negativas não
 6 correspondiam a uma possível degradação da molécula de DNA, ou erro de manipulação
 7 (Tabela 12).

8 **Tabela 12. Especificidade dos *primers* específicos para *H. intermedius* e *H. malabaricus*. O**
 9 **símbolo (+) indica amplificação positiva e o símbolo (-) amplificação negativa para cada**
 10 **espécie analisada.**

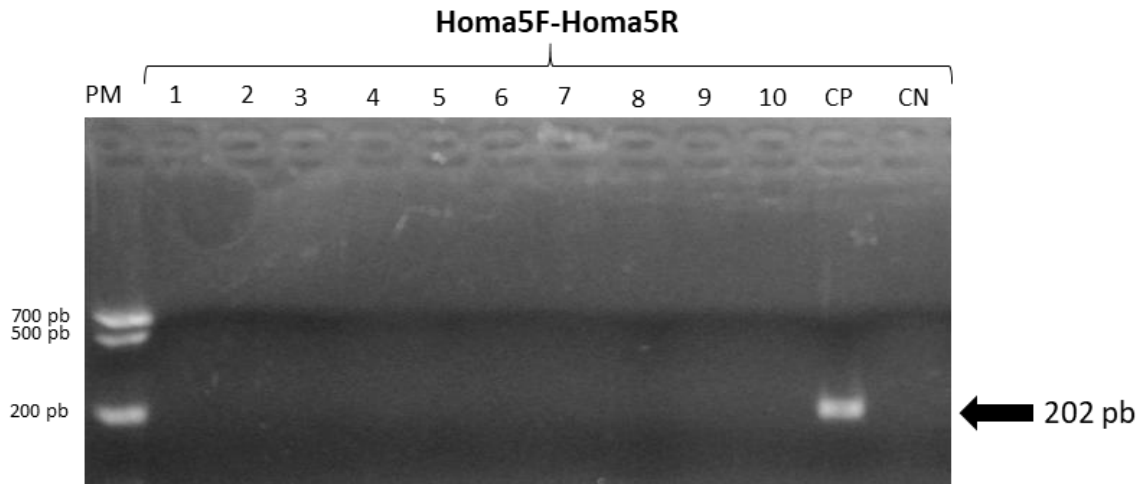
<i>Primers</i>	Hin5F - Hin3R	Hin5F - Hin4R	Homa5F – Homa5R	Universal
Localização do amplicon	16S rRNA	16S rRNA	12S rRNA+ tRNA-Val+ 16S rRNA	16S rRNA
Tamanho do fragmento	223 pb	511 pb	202 pb	121 pb
1 <i>Prochilodus argenteus</i>	-	-	-	+
2 <i>Prochilodus costatus</i>	-	-	-	+
3 <i>Prochilodus lineatus</i>	-	-	-	+
4 <i>Lophosilurus alexandri</i>	-	-	-	+
5 <i>Piaractus mesopotamicus</i>	-	-	-	+
6 <i>Brycon orbygnianus</i>	-	-	-	+
7 <i>Hoplias aimara</i>	-	-	-	+
8 <i>Hoplias curupira</i>	-	-	-	+
9 <i>Hoplias australis</i>	-	-	-	+
10 <i>Hoplias malabaricus</i>	-	-	+	+
11 <i>Hoplias intermedius</i>	+	+	-	+

11



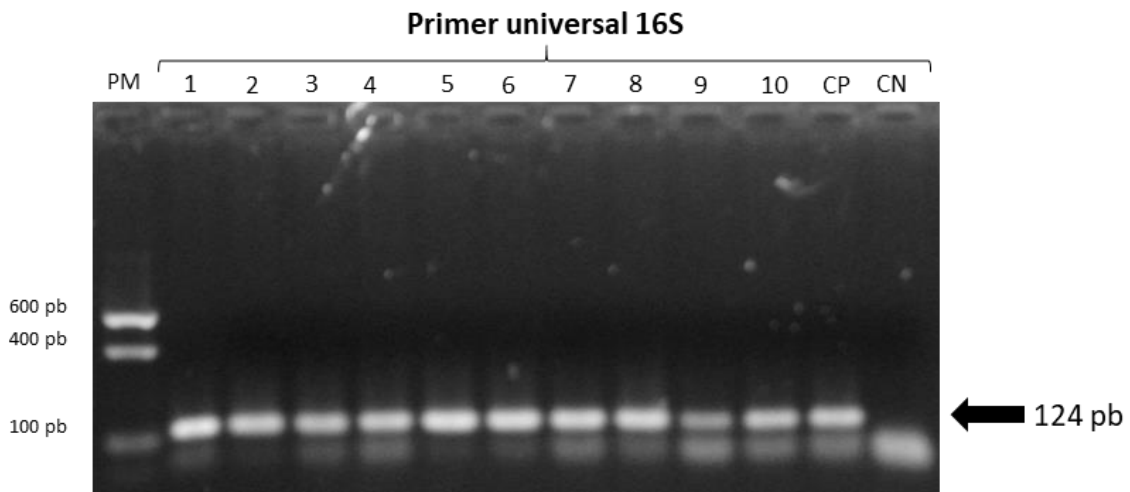
1

2 **Figura 19. Teste de Especificidade dos primers específicos para *H. intermedius*.** **A)** Marcador
3 Hin5F-Hin3R, com tamanho de amplicon de 223 pb e condições de amplificação: 10ng/uL de DNA
4 molde, solução tampão IVB, 7% de DMSO, 59°C e 30 ciclos de amplificação. **B)** Marcador Hin5F-
5 Hin4R com tamanho de amplicon de 511 pb e condições de amplificação: 10ng/uL de DNA molde,
6 solução tampão IC, 62°C e 30 ciclos de amplificação. Os números indicam a espécie que foi usada
7 na reação, **1:** *Prochilodus argenteus*, **2:** *Prochilodus costatus*, **3:** *Prochilodus lineatus*, **4:**
8 *Lophiosilurus alexandri*, **5:** *Piaractus mesopotamicus*, **6:** *Brycon orbygnianus*, **7:** *Hoplias aimara*,
9 **8:** *H. curupira*, **9:** *H. australis*, **10:** *H. malabaricus*, **CP:** controle positivo usando DNA de *H.*
10 *intermedius*, **CN:** controle negativo (ausência de DNA na reação), **PM:** marcador de peso
11 molecular.



2 **Figura 20. Teste de Especificidade dos primers Homa5F-Homa5R específicos para *H.***
 3 ***malabaricus*.** O tamanho de amplicon esperado foi de 202 pb. As condições de amplificação:
 4 10ng/uL de DNA molde, solução tampão IC, 56° de temperatura de anelamento e 30 ciclos de
 5 amplificação. Os números indicam a espécie que foi usada na reação, **1:** *Prochilodus argenteus*,
 6 **2:** *Prochilodus costatus*, **3:** *Prochilodus lineatus*, **4:** *Lophiosilurus alexandri*, **5:** *Piaractus*
 7 ***mesopotamicus*, 6:** *Brycon orbygnianus*, **7:** *Hoplias aimara*, **8:** *H. curupira*, **9:** *H. australis*, **10:** *H.*
 8 ***intermedius*, CP:** controle positivo usando DNA de *H. malabaricus*, **CN:** controle negativo
 9 (ausência de DNA na reação), **PM:** marcador de peso molecular (700 pb, 500 pb, 200 pb).

10

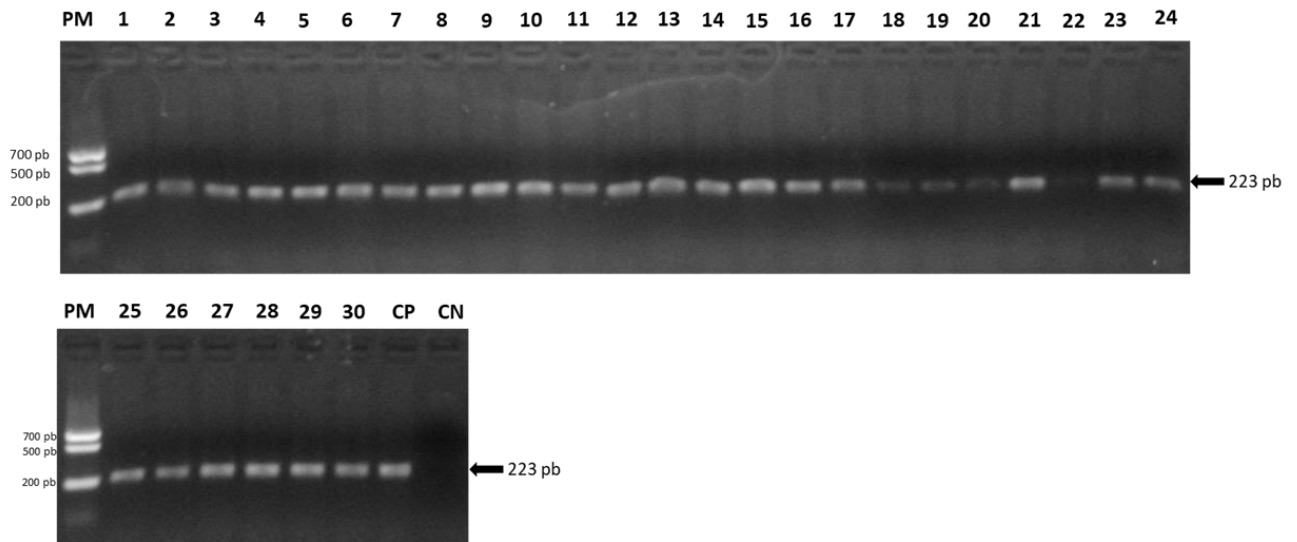


12 **Figura 21. Teste de controle de qualidade do DNA molde usado nos testes de especificidade.**
 13 Foi usado o conjunto de primers Fish16S (tamanho de amplicon de 124 pb) e 10ng/uL de DNA
 14 molde. Os números indicam a espécie que foi usada na reação, **1:** *Prochilodus argenteus*, **2:**
 15 *Prochilodus costatus*, **3:** *Prochilodus lineatus*, **4:** *Lophiosilurus alexandri*, **5:** *Piaractus*
 16 ***mesopotamicus*, 6:** *Brycon orbygnianus*, **7:** *Hoplias aimara*, **8:** *H. curupira*, **9:** *H. australis*, **10:** *H.*
 17 ***intermedius*, 11:** *H. malabaricus*, **CP:** controle positivo (DNA da espécie alvo), **CN:** controle
 18 negativo (ausência de DNA na reação), **PM:** marcador de peso molecular (600 pb, 400 pb, 100
 19 pb).

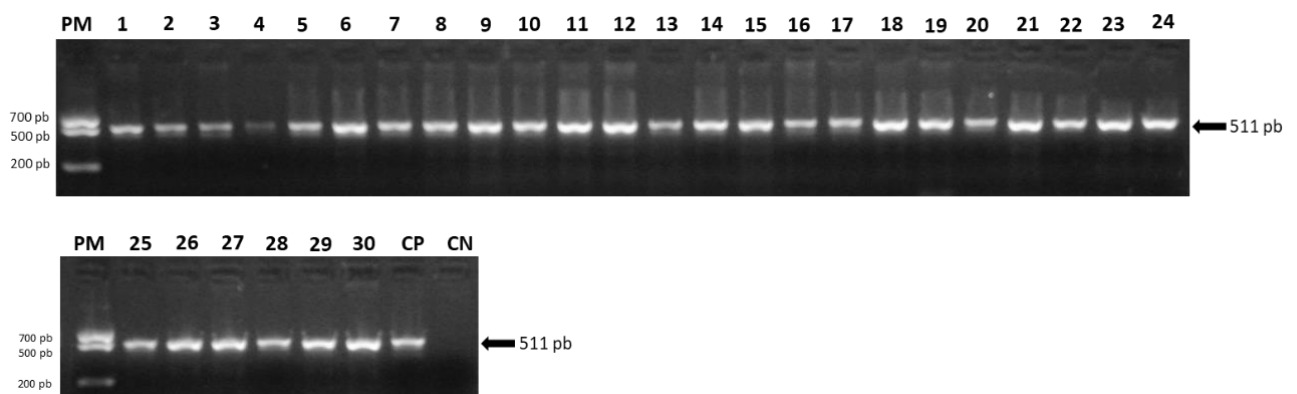
1 **4.7.3. Teste de eficiência:**

2 **A. Teste de eficiência em amostras coletadas em campo:**

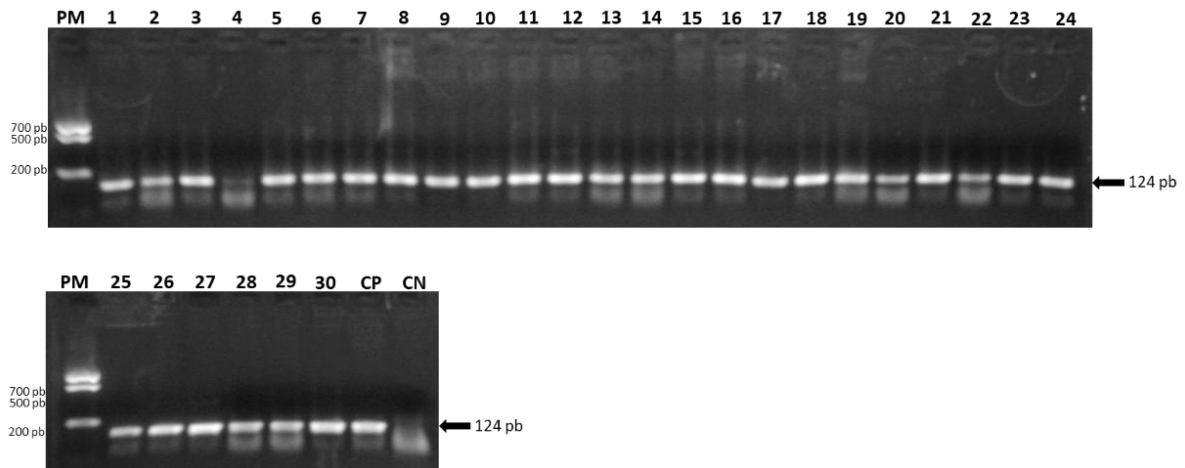
3 Foram analisadas 30 amostras do rio Doce, das quais todas foram positivamente
4 amplificadas pelos marcadores específicos para *H. intermedius* (Figura 22 e Figura 23).
5 A integridade do DNA utilizado foi comprovado com o *primer* universal 16S que amplifica
6 um tamanho de fragmento de 121 pb da região 16S rRNA (Figura 24).



7
8 **Figura 22. Identificação molecular de amostras coletadas no rio Doce usando o marcador**
9 **específico para *H. intermedius* Hin5F-3R.** Tamanho de fragmento esperado de 223 pb.
10 **Canaletas de 1 ao 30:** Amplificação positiva das amostras de *H. intermedius* coletadas no rio
11 Doce. **CP:** Controle positivo. **CN:** Controle negativo (ausência de DNA na reação). **Seta preta:**
12 indica o resultado positivo sendo visualizada uma banda esperada de tamanho 223 pb. **PM:**
13 Padrão de peso molecular (700pb, 500pb e 200pb) que permite identificar o tamanho das bandas
14 amplificadas.



15
16 **Figura 23. Identificação molecular de amostras coletadas no rio Doce usando o marcador**
17 **específico para *H. intermedius* Hin5F-Hin4R.** Tamanho de fragmento esperado de 511 pb.
18 **Canaletas de 1 ao 30:** Amplificação positiva das amostras de *H. intermedius* coletadas no rio
19 Doce. **CP:** Controle positivo. **CN:** Controle negativo (ausência de DNA na reação). **Seta preta:**
20 indica o resultado positivo sendo visualizada uma banda esperada de tamanho 511pb. **PM:** Padrão
21 de peso molecular (700pb, 500pb e 200pb) que permite identificar o tamanho das bandas
22 amplificadas.



1

2 **Figura 24. Controle de qualidade do DNA usando o marcador universal para peixes Fish16S**
 3 **e as amostras coletadas no rio Doce.** Tamanho de fragmento esperado de 124 pb. **Canaletas**
 4 **de 1 ao 30:** Amplificação positiva das amostras de *H. intermedius* coletadas no rio Doce. **CP:**
 5 **Controle positivo.** **CN:** Controle negativo (ausência de DNA na reação). **Seta preta:** indica o
 6 **resultado positivo** sendo visualizada uma banda esperada de tamanho 124 pb. **PM:** Padrão de
 7 **peso molecular** (700pb, 500pb e 200pb) que permite identificar o tamanho das bandas
 8 **amplificadas.**

B. Teste de eficiência em amostras obtidas em coleções científicas

Foram analisadas 22 amostras catalogadas como *H. intermedius* coletadas na bacia do rio São Francisco (número de tombo na Tabela 13 da seção de anexos). Como resultado, 10 amostras foram positivamente amplificadas pelos marcadores específicos para *H. intermedius* (Figura 25 e Figura 26) e 11 amostras foram amplificadas pelo marcador específico para *H. malabaricus* (Figura 26). As duas amostras restantes não amplificaram devido à degradação da molécula de DNA.

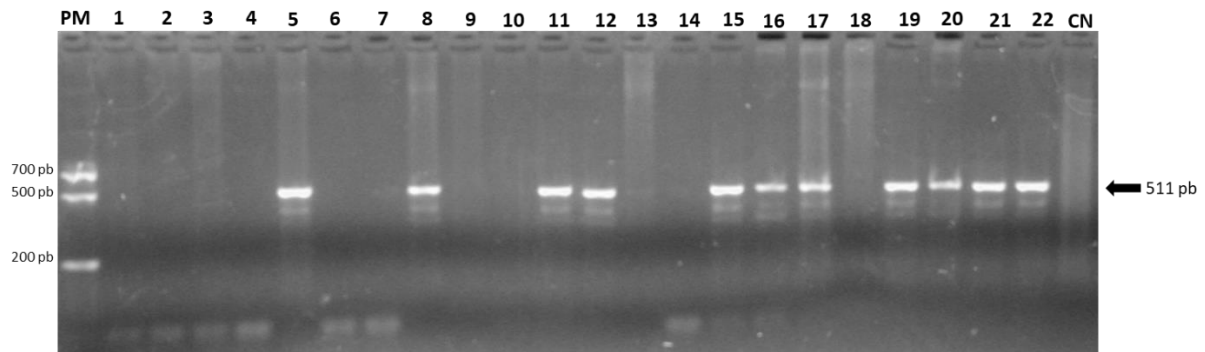


Figura 25. Teste de eficiência do marcador específico para *H. intermedius* (Hin5F-Hin4R) usando amostras do rio São Francisco. A seta indica o resultado positivo sendo visualizada uma banda esperada de tamanho 511 pb. 22: Controle positivo. CN: controle negativo (ausência de DNA na reação). PM: Padrão de peso molecular (700pb, 500pb e 200pb).

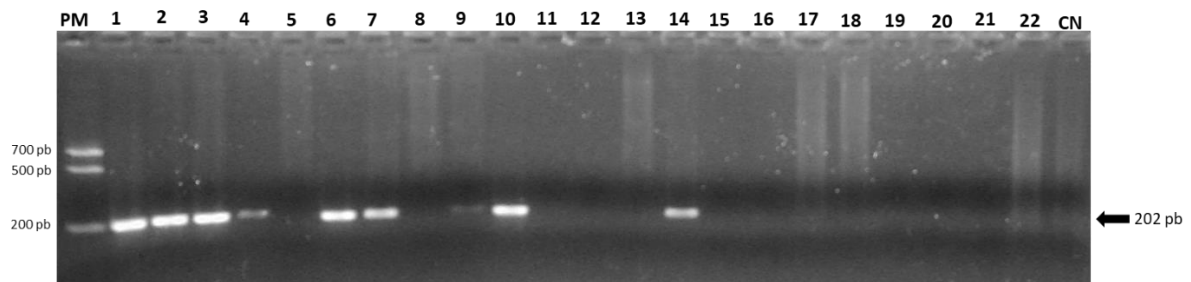


Figura 26. Teste de eficiência do marcador específico para *H. malabaricus* (Homa5F-Homa5R) usando amostras do rio São Francisco. A seta indica o resultado positivo sendo visualizada uma banda esperada de tamanho 202 pb. CP: Controle positivo. CN: controle negativo (ausência de DNA na reação). PM: Padrão de peso molecular (700pb, 500pb e 200pb).

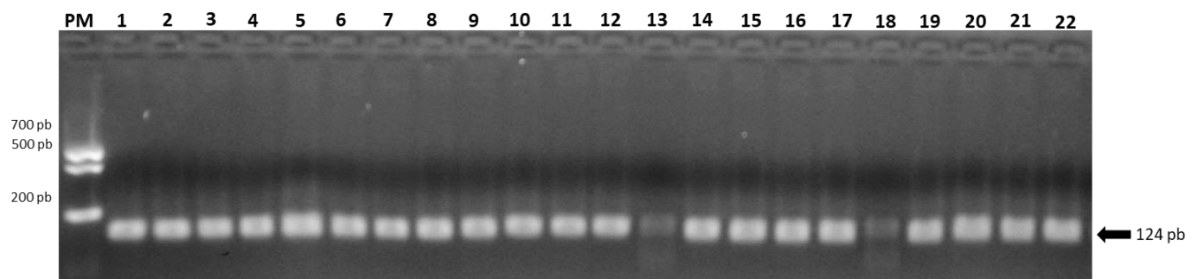


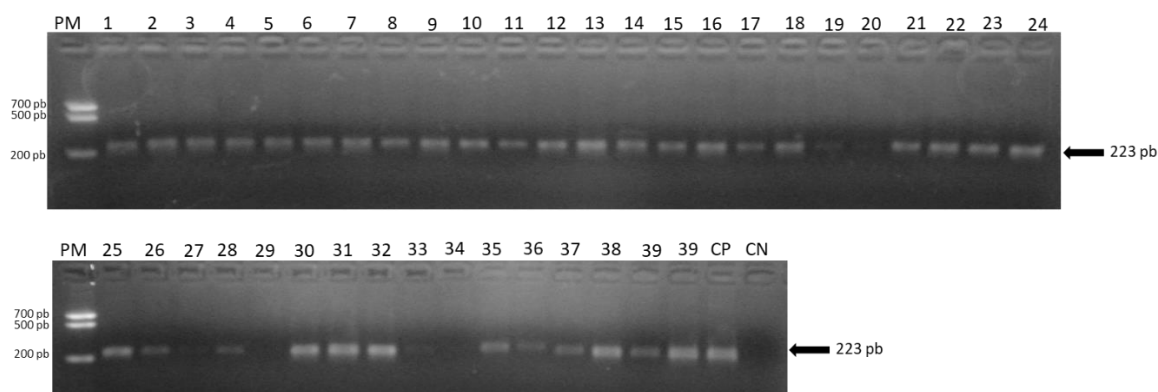
Figura 27. Controle de qualidade da integridade do DNA das amostras coletadas no rio São Francisco usando o marcador universal para peixes Fish16S. A seta indica o resultado positivo sendo visualizada uma banda esperada de tamanho 124 pb. CP: Controle positivo. CN: controle negativo (ausência de DNA na reação). PM: Padrão de peso molecular (700pb, 500pb e 200pb).

1 **C. Testes de eficiência dos marcadores específicos para *H.***
2 ***intermedius* e *H. malabaricus* usando amostras obtidas em peixarias**

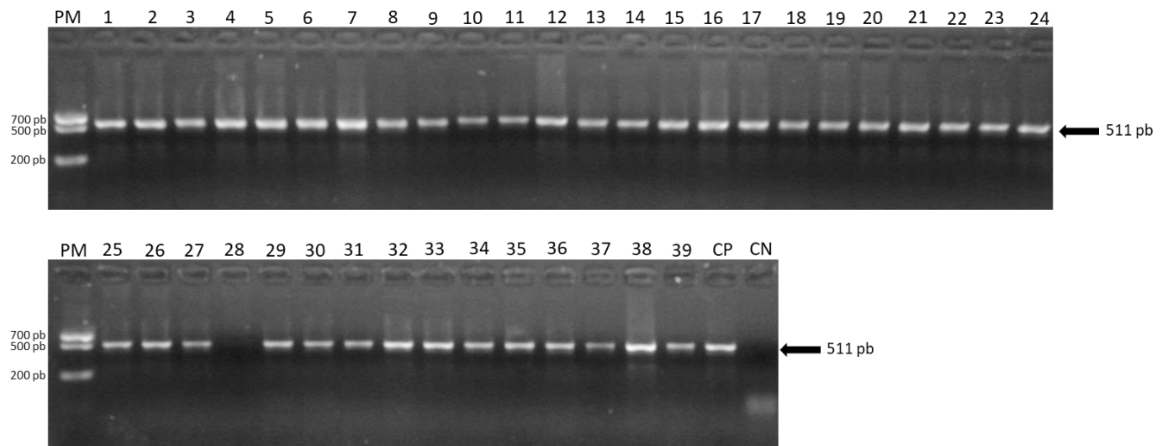
3 Das 39 amostras obtidas a partir de uma peixaria do rio Grande (alto Paraná),
4 todas foram positivamente amplificadas (exceto pela amostra 28) pelos marcadores
5 específicos para *H. intermedius* (Hin5F-3R e Hin5F-Hin4R) (Figura 28 e Figura 29).

6 Foram analisadas amostras provenientes de 11 filés congelados comercializados
7 como “trairão” na região metropolitana de Belo Horizonte. Foram usados marcadores
8 específicos para *H. intermedius* (Hin5F-Hin3R e Hin5F-Hin4R) (Figura 31), e o marcador
9 específico para *H. malabaricus* (Homa5F-Homa5R) (Figura 32). Como resultado, se
10 obteve que os 11 filés congelados foram identificados como *H. malabaricus*.

11 A integridade do DNA utilizado foi comprovado com o *primer* universal 16S que
12 amplifica um tamanho de fragmento de 121 pb da região 16S rRNA (Figura 30 e Figura
13 33).

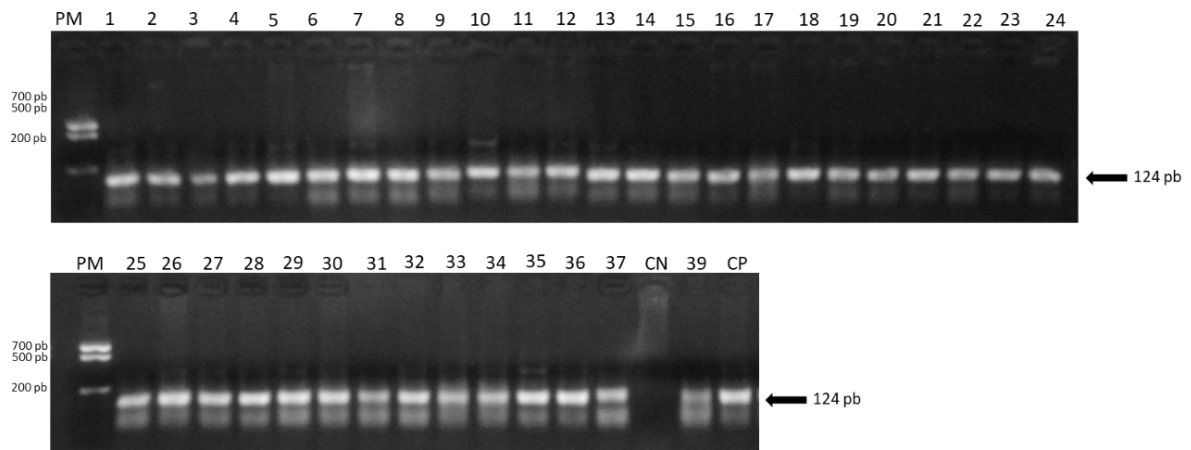


14
15 **Figura 28. Identificação molecular de amostras coletadas em peixarias do rio Grande**
16 **usando o marcador específico para *H. intermedius* Hin5F-Hin3R.** Tamanho de fragmento
17 esperado de 223 pb. **Canaletas de 1 ao 39:** Amplificação positiva das amostras de *H. intermedius*
18 coletadas no rio Grande. **CP:** Controle positivo. **CN:** Controle negativo (ausência de DNA na
19 reação). **Seta preta:** indica o resultado positivo sendo visualizada uma banda esperada de
20 tamanho 223 pb. **PM:** Padrão de peso molecular (700pb, 500pb e 200pb) que permite identificar o
21 tamanho das bandas amplificadas.



1

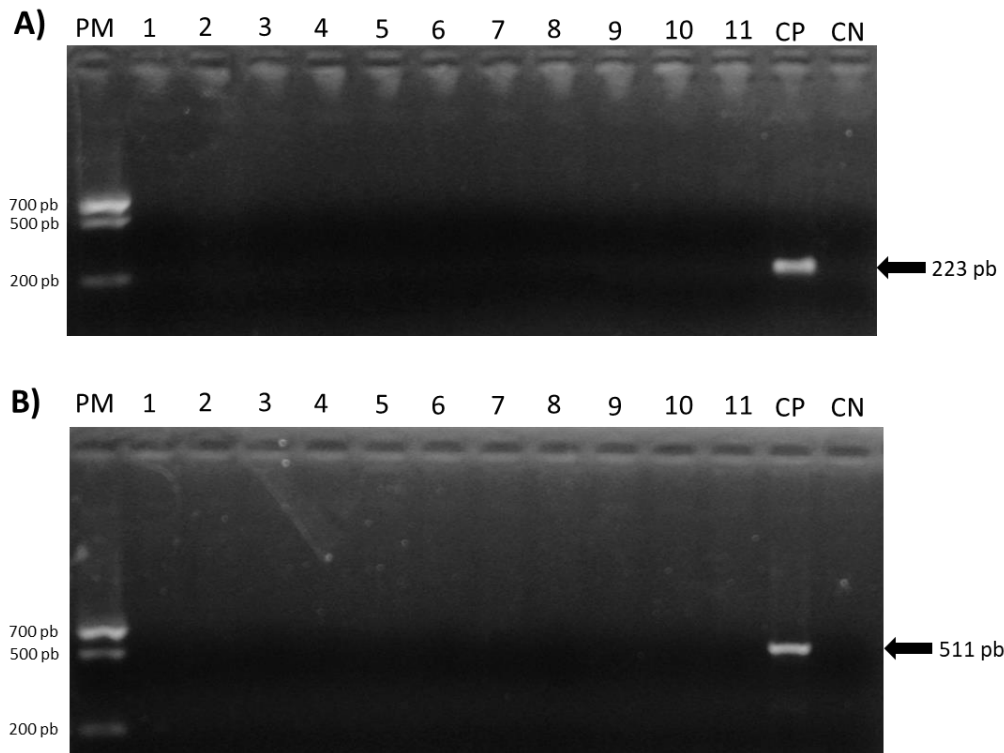
2 **Figura 29. Identificação molecular de amostras coletadas em peixarias do rio Grande**
 3 **usando o marcador específico para *H. intermedius* Hin5F-Hin4R.** Tamanho de fragmento
 4 esperado de 511 pb. **Canaletas de 1 ao 39:** Amplificação positiva das amostras de *H. intermedius*
 5 coletadas no rio Grande. **CP:** Controle positivo. **CN:** Controle negativo (ausência de DNA na
 6 reação). **Seta preta:** indica o resultado positivo sendo visualizada uma banda esperada de
 7 tamanho 511 pb. **PM:** Padrão de peso molecular (700pb, 500pb e 200pb) que permite identificar o
 8 tamanho das bandas amplificadas.



9

10 **Figura 30. Controle de qualidade da integridade do DNA das amostras coletadas no rio**
 11 **Grande usando o marcador universal para peixes Fish16S.** Tamanho de fragmento esperado
 12 de 124 pb. **Canaletas de 1 ao 39:** Amplificação positiva das amostras de *H. intermedius* coletadas
 13 no rio Doce. **CN:** Controle negativo (ausência de DNA na reação). **Seta preta:** indica o resultado
 14 positivo sendo visualizada uma banda esperada de tamanho 124 pb. **PM:** Padrão de peso
 15 molecular (700pb, 500pb e 200pb) que permite identificar o tamanho das bandas amplificadas.

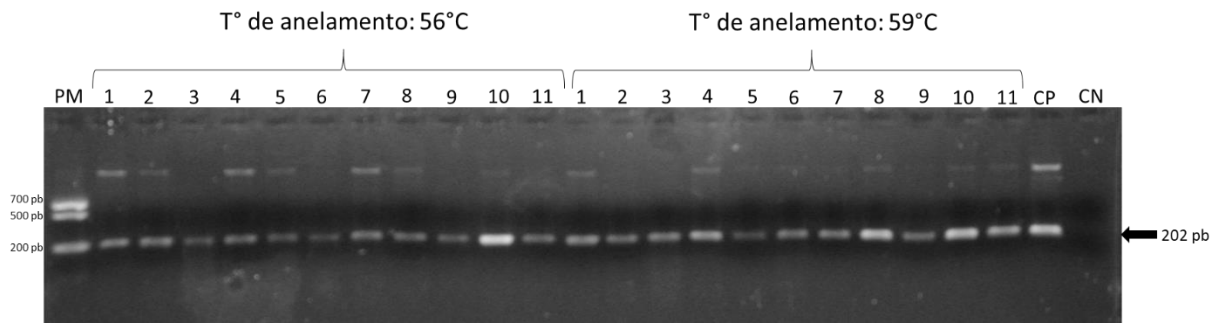
16



1

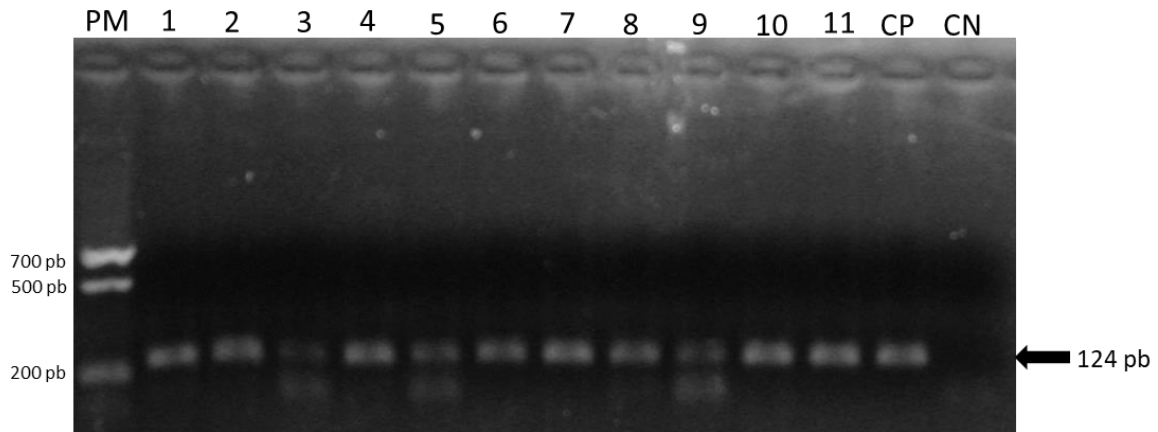
2 **Figura 31. Identificação molecular usando amostras de filés congelados de “trairão” e o**
 3 **marcador específico para *H. intermedius*. A) Marcador Hin5F-Hin3R, tamanho de fragmento**
 4 **esperado de 223 pb. B) Marcador Hin5F-Hin4R, tamanho de fragmento esperado de 511 pb.**
 5 **Canaletas de 1 ao 11: DNA de filés congelados. CP: Controle positivo. Seta preta: indica o**
 6 **resultado positivo sendo visualizada uma banda esperada de tamanho 223pb para Hin5F-Hin3R**
 7 **e 511pb para Hin5F-Hin4R. CN: controle negativo (ausência de DNA na reação). PM: Padrão de**
 8 **peso molecular (700pb, 500pb e 200pb) que permite identificar o tamanho das bandas**
 9 **amplificadas.**

10



1

2 **Figura 32. Identificação molecular de *H. malabaricus* usando o marcador específico Homa5F-**
 3 **Homa5R e amostras de filés congelados de “trairão”.** Tamanho de fragmento esperado de 202
 4 pb. As amostras foram amplificadas usando duas temperaturas de anelamento (56°C e 59°).
 5 **Canaletas de 1 ao 11:** DNA de filés congelados. **CP:** Controle positivo. Seta branca: indica o
 6 resultado positivo sendo visualizada uma banda esperada de tamanho 202 pb. **CN:** controle
 7 negativo (ausência de DNA na reação). **PM:** Padrão de peso molecular (700pb, 500pb e 200pb)
 8 que permite identificar o tamanho das bandas amplificadas.



9

10 **Figura 33. Controle de qualidade da integridade do DNA das amostras obtidas a partir de**
 11 **filés congelados de “trairão” usando o marcador universal para peixes Fish16S.** A seta
 12 indica o resultado positivo sendo visualizada uma banda esperada de tamanho 124 pb. CP:
 13 Controle positivo. CN: controle negativo (ausência de DNA na reação). PM: Padrão de peso
 14 molecular (700pb, 500pb e 200pb).

1 5. DISCUSSÃO DE RESULTADOS

2 Os *primers* mitocondriais representam a melhor alternativa metodológica quando
3 se analisam amostras processadas ou degradadas, nas quais o DNA encontra-se
4 fragmentado. A metodologia utilizada no presente trabalho busca diminuir a margem de
5 erro nas identificações moleculares de *H. intermedius* “trairão”, usando como pontos
6 críticos a detecção de regiões mitocondriais altamente informativas e a posterior
7 validação delas a partir de espécimes com identificação taxonômica corroborada. Cada
8 etapa do desenvolvimento destes marcadores específicos para *H. intermedius* busca
9 aumentar a probabilidade de que o marcador seja capaz de detectar unicamente a
10 espécie alvo.

11 Seguindo as recomendações de McDonald e Sarre (2017) e Collins and
12 Cruickshank, (2013), temos seguido um processo de validação que inclui a análise de
13 distância genética entre a sequência da amostra, a sequência esperada do marcador e
14 os genes 16S rRNA disponíveis das espécies de *Hoplías*. Desde a proposta do DNA
15 barcoding por Herbert et al. (2003), diversos trabalhos têm aprimorado o uso desta técnica
16 em muitos grupos animais e mais recentemente em sua aplicação para estudos de DNA
17 ambiental (Bohmann *et al.*, 2014; Goldberg *et al.*, 2016; Macdonald e Sarre, 2017). Desta
18 forma, a análise molecular utilizando vouchers depositados em coleções zoológicas,
19 possibilita aos marcadores acompanhar as mudanças nas categorias taxonômicas ao
20 longo do tempo. Este processo de validação deve ser feito para todo novo marcador
21 proposto, diminuindo assim identificações incertas (Vink *et al.*, 2012; Collins e
22 Cruickshank, 2013).

23 O mitogenoma completo da espécie alvo *H. intermedius* “trairão” foi o primeiro
24 mitogenoma completo disponibilizado para o gênero *Hoplías*, o qual inclui atualmente 14
25 espécies descritas. A localização de regiões altamente informativas para o grupo alvo se
26 incrementa em número e qualidade devido à obtenção do mitogenoma completo,
27 permitindo delimitar marcadores eficientes para identificação de fauna em qualquer região
28 gênica ou intergênica. Neste sentido, o sequenciamento massivo tem sido amplamente
29 utilizado na genética forense e DNA antigo (Yang *et al.*, 2014), assim como na obtenção
30 de mitogenomas completos de espécies da fauna nativa (Brandao-Dias *et al.* 2014;
31 Carmo *et al.* 2014; Pimentel et al. 2014; Chagas et al. 2015; Núñez-Rodríguez *et al.* 2016;
32 Resende *et al.*, 2016; Bedore *et al.*, 2017; Lemos Queiroz *et al.*, 2018; Pereira *et al.*,
33 2019).

34 A disposição dos genes e dos tRNAs de *H. intermedius* coincide com o encontrado
35 por (Sato *et al.*, 2016), os que compararam 250 mitogenomas de espécies diferentes
36 (entre peixes cartilaginosos e ósseos), sendo a estrutura gênica conservada em 214

1 espécies: 12S, 16S, Nd1, COI, COII, ATPase 8, ATPase 6, COIII, Nd3, Nd4L, Nd4, Nd5,
2 Nd6, Cyt b e D-loop). Da mesma forma, a proporção nucleotídica do mtDNA de *H.*
3 *intermedius* foi igualmente conservada: (% de Citosina \approx % de Adenina) > (% de Timina)
4 > (% de Guanina) (Asakawa *et al.*, 1991). Predominantemente, o códon de início foi ATG
5 e os códons de parada mais comuns foram TAA e o códon incompleto T--. Tais resultados
6 concordam com Satoh *et al.*, (2016), sendo códon de início GTG e o códon de parada
7 AGG foram unicamente encontrados no gene citocromo oxidase subunidade 1 (COI).

8 A disponibilidade do genoma mitocondrial completo de *H. intermedius* permitiu
9 aumentar a busca por marcadores específicos e informativos para a espécie alvo, sendo
10 avaliadas todas as regiões gênicas do mitogenoma. Porém, para *H. malabaricus*, não foi
11 possível avaliar a região do D-loop, pois o mitogenoma *quase completo* (número de
12 acesso GenBank AP011992) da espécie disponível, não inclui esta região. Para as 14
13 espécies descritas de *Hoplias*, estão disponibilizadas unicamente sequências parciais
14 para cinco espécies: *H. aimara*, *H. intermedius*, *H. malabaricus*, *H. microlepis* e *H.*
15 *malabaricus*, as quais correspondem aos genes 12SrRNA, 16SrRNA e COI. As maiores
16 porcentagens de identidade de *H. intermedius* em comparação com espécies congêneres
17 mostraram que *H. intermedius* é 100% idêntico com *H. malabaricus* e *H. microlepis* para
18 o gene 12SrRNA. Da mesma forma *H. intermedius* mostrou de 100% de identidade com
19 *H. aimara* e *H. microlepis* para o gene COI. O gene com menor identidade foi o gene
20 16SrRNA, o qual mostrou 94% de identidade entre *H. intermedius*, *H. aimara* e *H.*
21 *malabaricus*. Como apontado por Deagle *et al.*, (2014), os genes rRNA mitocondriais
22 fornecem uma maior resolução taxonômica e possibilitam o desenho de *primers* mais
23 conservados, assim, em concordância com nossos resultados, para *H. intermedius* e *H.*
24 *malabaricus*, os marcadores mais específicos foram detectados no gene 16S rRNA.

25
26 Foram recuperadas a partir do NCBI as sequências de mitogenoma completo de
27 peixes da ordem Characiformes, das quais aquelas com maior similaridade com as
28 sequências da espécie alvo *H. intermedius* foram de *Colossoma macropomum*, com
29 localidade tipo na Amazônia e rio Orinoco, mas também presente em piscicultura no Rio
30 São Francisco (Barbosa e Soares, 2009), *Prochilodus costatus* originalmente endêmico
31 do rio São Francisco e introduzido no rio Jequitinhonha (Reis *et al.*, 2003), e de
32 *Pygocentrus nattereri* com distribuição no rio Amazonas e introduzido intencionalmente
33 nas lagoas marginais do rio São Francisco e bacia do rio Doce. Embora o SPIDER tenha
34 sugerido regiões nos genes 16S rRNA e ND5 como possíveis marcadores específicos
35 para *H. intermedius*, a busca manual de *primers* detectou fragmentos com variabilidade
36 suficiente apenas para o gene 16S rRNA. Após análises *in silico* e *in vitro*, unicamente
37 dois conjuntos de *primers* desenhados manualmente alocados na região 16S rRNA

1 mostraram-se altamente específicos, sensíveis e eficientes para *H. intermedius*. No
2 entanto, todos os 50 pares de *primers* sugeridos pelo PrimerBLAST (Ye *et al.*, 2012) se
3 mostraram inespecíficos para *H. intermedius* nas análises *in silico*, não sendo
4 sintetizada nenhuma destas sequências para análise *in vitro*. No caso de *H.*
5 *malabaricus*, a região mais informativa para a espécie foi uma região intergênica,
6 incluindo o gene 16S rRNA, tRNA^{Val} e 12S rRNA, a qual não foi detectada pelas
7 ferramentas bioinformáticas utilizadas, como SPIDER ou PrimerBLAST.

8 De acordo com os resultados obtidos pelo *Primer-BLAST*, a melhor especificidade
9 seria obtida pelos conjuntos de *primers* Hin5F-Hin4R (região 16SrRNA), Hin7F-Hin9R e
10 Hin7F-Hin8R (região D-loop), os quais potencialmente amplificariam unicamente a *H.*
11 *intermedius*. Para os conjuntos de *primer* Homa 5R-Homa5F e Homa5F-Homa10R
12 (região 16SrRNA) os testes *in silico* no *Primer-BLAST* indicaram alta especificidade para
13 *H. malabaricus*. Porém, nos testes *in vitro*, apenas os conjuntos de *primers* Hin5F-Hin3R
14 e Hin5F-Hin4R (16SrRNA) amplificaram unicamente *H. intermedius*, no entanto o
15 conjunto Hin7F-Hin9R (região D-loop) revelou uma alta inespecificidade, amplificando
16 *Prochilodus argenteus*, *P. costatus*, *P. lineatus*, *Lophosilurus alexandri*, *Brycon*
17 *orbignianus*, e *Pirinampus pirinampus*. No caso dos *primers* para *H. malabaricus*, os dois
18 pares de *primers* se mostraram específicos, embora o *primer* Homa5F-10R amplificou um
19 fragmento de 500 pb aproximadamente, sendo muito maior ao esperado (202 pb).

20 A análise *in silico* permitiu ainda observar a variação dos parâmetros
21 físicoquímicos entre as sequências de *primers*, mostrando que a eficiência era maior para
22 tamanhos de sequência entre 16 a 24pb e conteúdo de GC entre 50 e 70%, não sendo
23 eficiências as sequências com tamanhos curtos de 14 a 19pb ou tamanhos maiores de
24 25 a 29 pb. As condições de padronização obtidas, demonstram que para as sequências
25 desenhadas, as soluções tampão IC, IIC e IVB (Phoneutria) facilitaram a amplificação
26 devido a presença de KCl (500 mM) e NaCl (400 mM), diminuindo a repulsão das fitas e
27 aumentando a estabilidade. Essas soluções tampão possuem uma concentração de
28 MgCl₂ (2,0 mM), com exceção do tampão IVB que possui uma concentração variável,
29 que embora tenha facilitado a amplificação, aumentou o aparecimento de bandas
30 inespecíficas.

31 A análise usando o algoritmo *blastn* mostrou uma diferença nos resultados de
32 acordo com o tamanho do *word size* (tamanho de fragmento que o algoritmo blast usa
33 para realizar a comparação entre a sequência alvo e as sequências disponíveis no banco
34 de dados do NCBI). Os amplicons obtidos a partir do marcador Hin5F-Hin3R (223 pb)
35 foram identificados como *H. intermedius* unicamente para *word sizes* de 11 ou valores

1 menores, no entanto para o marcador maior (Hin5F-Hin4R, 511pb) a identificação era
2 feita inclusive para *word sizes* de 28.

3 Nossos resultados apresentam dois marcadores específicos para *H. intermedius*
4 com tamanhos de 223 pb (Hin5F-Hin3R) e 511 pb (Hin5F-Hin4R), os que embora não
5 mostraram um suporte elevado entre os ramos que continham as sequências das
6 amostras do rio Doce de *H. intermedius*, ambos marcadores se mostraram altamente
7 eficientes e específicos na identificação da espécie alvo. Como mostrado por Hajibabaei
8 *et al.* (2006b) em primatas e por Xiang e Hickey (2007) em fungos, nos quais foram
9 usados marcadores mini-barcodes de 300 pb, os fragmentos curtos possuem também um
10 alto poder de identificação, no entanto não sejam suficientemente informativos para
11 entender as relações evolutivas entre espécies e construir filogenias moleculares
12 (Hajibabaei *et al.*, 2006b). É importante considerar que quando se analisam amostras
13 produto da evidência de crimes ambientais ou amostras de museus, frequentemente o
14 DNA está altamente degradado ou são disponibilizadas quantidades muito pequenas de
15 tecido que tornam limitado o número de experimentos a serem realizados (Spencer *et*
16 *al.*, 2010). Desta forma, o uso de marcadores desenhados especificamente para uma
17 espécie alvo, assim como de *primers* que amplifiquem regiões curtas (aproximadamente
18 200 pb), permitem otimizar a utilização da amostra (Spencer *et al.*, 2010).

19 A complexa diversidade do gênero *Hoplias* e a distribuição ampla de algumas de
20 suas espécies deve-se à sua grande capacidade de adaptação a ambientes elevados,
21 como são as cabeceiras do Paraná. Assim, em comparação à *H. intermedius*, outras
22 espécies de Characiformes têm apresentado uma distribuição disjunta (Menezes, 1988;
23 Vari & Harold, 2001). A ampla distribuição de *H. malabaricus* foi evidenciada a través de
24 estudos envolvendo cariótipos, os quais sugeriram um ancestral comum e fluxos recentes
25 entre as bacias do alto rio Doce, rio Paraná e rio Paraíba do Sul (Dergam *et al.*, 2002).
26 Uma vez que *H. intermedius* e *H. malabaricus* são espécies simpátricas e sintópicas, é
27 ressaltada a importância dos estudos que compreendem abordagens moleculares,
28 citogenéticas e filogeográficas, que auxiliem no entendimento da origem, da distribuição
29 das espécies de *Hoplias* e do número de espécies que compõem o gênero atualmente.
30 Marcadores que consigam identificar a espécie alvo e diferenciá-las das espécies mais
31 similares a nível molecular é altamente informativo desde o ponto de vista filogenético,
32 ecológica e econômica, já que auxilia na delimitação de espécies e permite o
33 monitoramento da pesca e o controle de fraudes alimentares.

34 Essa metodologia multidisciplinar foi desenvolvida a partir de coleções científicas,
35 genética molecular e ferramentas de bioinformática, as quais atualmente são essenciais
36 para o monitoramento da biodiversidade. Assim, sua aplicação não se limita apenas para

1 identificar tecidos animais altamente degradados por processos industriais, pesca ou
2 caça ilegal, mas abrange todas as amostras que não possuem caracteres diagnósticos
3 suficientes.

4 **6. CONCLUSÕES**

5

6 As três metodologias de extração de DNA mostraram grande eficiência para o
7 isolamento do DNA de *H. intermedius*, permitindo a construção de uma biblioteca
8 genômica de alta qualidade, e por consequência, a obtenção da sequência mitocondrial
9 da espécie alvo usando sequenciamento NGS. O mitogenoma de *H. intermedius* coincide
10 em estrutura e organização, não somente com o genoma mitocondrial de Osteichthyes e
11 Chondrichthyes.

12 Os dados obtidos do genoma mitocondrial completo permitiram abranger todas as
13 regiões mitogenômicas para a busca de marcadores específicos para *H. intermedius*. Em
14 comparação com a região do gene 16S, o gene *COI* não representa uma região altamente
15 informativa para o desenho de marcadores específicos para *H. intermedius* e *H.*
16 *malabaricus*. Desta forma, para o desenho de marcadores espécie-específico é altamente
17 recomendável avaliar todas as regiões gênicas e intergênicas possíveis.

18 Dentro dos parâmetros considerados, os *primers* e marcadores moleculares
19 desenvolvidos foram específicos para *H. intermedius* e *H. malabaricus*. Assim, os
20 marcadores específicos aqui desenvolvidos e validados, são altamente aplicáveis na
21 identificação de amostras de tecidos que não possuem espécimes voucher.

22 Finalmente, os marcadores que foram desenvolvidos neste trabalho, podem ser
23 usados para fins econômicos e pesqueiros, como são a transposição, controle da pesca
24 e dos recursos pesqueiros. Da mesma forma propõem-se os marcadores moleculares
25 espécie-específicos como uma ferramenta complementar à estudos de taxonomia
26 integrativa do gênero *Hoplías*, aprimorando a avaliação da biodiversidade,
27 monitoramento de populações e a qualidade de bancos de tecido e bibliotecas de dados
28 genéticos já disponíveis.

1 7. REFERÊNCIAS BIBLIOGRÁFICAS

- 2 Abe, K. T. et al. (2014) 'Systematic and historical biogeography of the Bryconidae (
- 3 Ostariophysi : Characiformes) suggesting a new rearrangement of its genera and an old
- 4 origin of Mesoamerican ichthyofauna', *BMC Evolutionary Biology*, 14, p. 152.
- 5 Abell, R. et al. (2008) 'Freshwater Ecoregions of the World : A New Map of Biogeo-
- 6 graphic Units for Freshwater Biodiversity Conservation', 58(5).
- 7 Agostinho, A., Thomaz, S. and Gomes, L. (2005) 'Conservação da biodiversidade em
- 8 águas continentais do Brasil', *Algae*, 1(1), pp. 70–78. doi:
- 9 10.1017/CBO9781107415324.004.
- 10 Ardura, A., Pola, I. G., et al. (2010) 'Application of barcoding to Amazonian
- 11 commercial fish labelling', *Food Research International*. Elsevier Ltd, 43(5), pp. 1549–
- 12 1552. doi: 10.1016/j.foodres.2010.03.016.
- 13 Ardura, A., Rosa, A., et al. (2010) 'DNA barcoding for conservation and management
- 14 of Amazonian commercial fish', *Biological Conservation*. Elsevier Ltd, 143(6), pp. 1438–
- 15 1443. doi: 10.1016/j.biocon.2010.03.019.
- 16 Arif, I. A. et al. (2011) 'DNA marker technology for wildlife conservation', *Saudi Journal*
- 17 *of Biological Sciences*, 18(3), pp. 219–225. doi: 10.1016/j.sjbs.2011.03.002.
- 18 Asakawa, S. et al. (1991) 'Strand-specific Nucleotide Composition bias in Echinoderm
- 19 and Vertebrate Mitochondrial Genomes', *Journal of Molecular Evolution*, 32, pp. 511–520.
- 20 de Azevedo Chagas, A. T. et al. (2015) 'Illegal hunting and fishing in Brazil: A study
- 21 based on data provided by environmental military police', *Natureza e Conservacao*.
- 22 *Associação Brasileira de Ciência Ecológica e Conservação*, 13(2), pp. 183–189. doi:
- 23 10.1016/j.ncon.2015.11.002.
- 24 Barbosa, J. M. and Soares, E. C. (2009) 'Perfil da ictiofauna da bacia do Sao
- 25 Francisco: estudo preliminar', *Revista Brasileira de Engenharia de Pesca*, 4(1).
- 26 Bedore, A. G. et al. (2017) 'Leporinus elongatus (Characiformes , Anostomidae):
- 27 complete mtDNA sequence of an economically important fish from the Paraná and La
- 28 Plata river basins', *Mitochondrial DNA Part B: Resources*, 2(1), pp. 261–263. doi:
- 29 10.1080/23802359.2017.1318685.
- 30 Bertollo LAC, Moreira-Filho O, Galetti Jr PM. (1986). Cytogenetics and taxonomy:
- 31 consideration based on chromosome studies of freshwater fish. *J Fish Biol.*;28:153–9.
- 32 Blanco, D. R. et al. (2010) 'Karyotypic diversity between allopatric populations of the
- 33 group *Hoplias malabaricus* (Characiformes: Erythrinidae): evolutionary and
- 34 biogeographic considerations', 8(2), pp. 361–368.
- 35 Bohmann, K. et al. (2014) 'Environmental DNA for wildlife biology and biodiversity
- 36 monitoring', *Trends in Ecology & Evolution*. Elsevier Ltd, 29(6), pp. 358–367. doi:
- 37 10.1016/j.tree.2014.04.003.

- 1 Børsting, C. and Morling, N. (2015) 'Next generation sequencing and its applications
2 in forensic genetics', *Forensic Science International: Genetics*, 18, pp. 78–89. doi:
3 10.1016/j.fsigen.2015.02.002.
- 4 Branch, S., Dehkordi, P. G. and Branch, S. (2011) 'Genetic Polymorphisms of
5 Mitochondrial Genome D-loop Region in Bakhtiarian Population by PCR-RFLP',
6 *International Journal of Biology*, 3(4), pp. 41–46. doi: 10.5539/ijb.v3n4p41.
- 7 Brandao-Dias, P. F. P. et al. (2014) 'Complete mitochondrial genome of *Salminus*
8 *brasiliensis*', 1736, pp. 1–2. doi: 10.3109/19401736.2014.958676.
- 9 Brown WM, George M, Wilson AC (1979) 'Rapid evolution of animal mitochondrial
10 DNA', *Proc Natl Acad Sci USA* 76:1967–1971
- 11 Brown, S. D. J. et al. (2012) 'Spider: An R package for the analysis of species identity
12 and evolution, with particular reference to DNA barcoding', *Molecular Ecology Resources*,
13 12(3), pp. 562–565. doi: 10.1111/j.1755-0998.2011.03108.x.
- 14 Butler, J.M. (2011). 'Advanced topics in forensic dna typing methodology. National
15 Institute of Standards and Technology, Gaithersburg, Maryland, USA. Academic Press,
16 pp. 29–47.
- 17 Calegari, B. B., Delapieve, M. L. S. and Sousa, L. M. (2016) 'Tutorial para preparação
18 de mapas de distribuição geográfica', *Boletim SBI*, (September), pp. 15–30. doi:
19 10.2307/3183698.
- 20 Camelier, P. and Zanata, A. M. (2014) 'Biogeography of freshwater fishes from the
21 Northeastern Mata Atlântica freshwater ecoregion: Distribution, Endemism, and area
22 relationships', *Neotropical Ichthyology*, 12(4), pp. 683–698. doi: 10.1590/1982-0224-
23 20130228.
- 24 Cardoso, Y. P. et al. (2018) 'A continental-wide molecular approach unraveling
25 mtDNA diversity and geographic distribution of the Neotropical genus *Hoplias*', *PLoS*
26 *ONE*, 13(8), pp. 1–25. doi: 10.1371/journal.pone.0202024.
- 27 Carmo, A. O. do et al. (2014) 'Complete mitochondrial genome sequence of
28 *Prochilodus lineatus*', *Mitochondrial DNA*, 1736, pp. 1–2. doi:
29 10.3109/19401736.2014.971300.
- 30 Carvalho, D. C. et al. (2011) 'DNA barcoding unveils a high rate of mislabeling in a
31 commercial freshwater catfish from Brazil', 22(October), pp. 97–105. doi:
32 10.3109/19401736.2011.588219.
- 33 Carvalho, D. C. et al. (2015) 'DNA Barcoding identification of commercialized seafood
34 in South Brazil: A governmental regulatory forensic program', *Food Control*, 50, pp. 784–
35 788. doi: 10.1016/j.foodcont.2014.10.025.
- 36 Carvalho, D. C. De et al. (2011) 'Deep barcode divergence in Brazilian freshwater
37 fishes: the case of the São Francisco River basin', 1736(March 2016). doi:
38 10.3109/19401736.2011.588214.
- 39 Chagas, A. T. de A., Carmo, A. O., et al. (2015) 'Description and comparison of two
40 economically important fish species mitogenomes: *Prochilodus argenteus* and

- 1 Prochilodus costatus', Mitochondrial DNA, 1736, pp. 1–2. doi:
2 10.3109/19401736.2015.1053125.
- 3 Chagas, A. T. de A., Costa, M. A. da, et al. (2015) 'Illegal hunting and fishing in Brazil:
4 a study based on data provided by environmental military police', *Natureza &*
5 *Conservação*, 3, pp. 183–189.
- 6 Chan, D., Schon E. (2012) ' Eliminating Mitochondrial DNA from Sperm',
7 *Developmental Cell*, 22, pp. 469-470.
- 8 Chaves, M. F. et al. (2011) 'Dinâmica reprodutiva e estrutura populacional de *Hoplias*
9 *aff. malabaricus* (Bloch, 1794) (Characiformes, Erythrinidae), em açude da Bacia do Rio
10 Taperoá, Paraíba', *Biotemas*, 22(2), pp. 85–89. doi: 10.5007/2175-7925.2009v22n2p85.
- 11 Ciavaglia, S. A. et al. (2015) 'Molecular identification of python species: Development
12 and validation of a novel assay for forensic investigations', *Forensic Science International:*
13 *Genetics*. Elsevier Ireland Ltd, 16, pp. 64–70. doi: 10.1016/j.fsigen.2014.12.002.
- 14 Civera T. 2003. Species identification and safety of fish products. *Vet Res Commun*
15 27:481–489.
- 16 Collins, R. A. and Cruickshank, R. H. (2013) 'The seven deadly sins of DNA
17 barcoding', *Molecular Ecology Resources*, 13(6), pp. 969–975. doi: 10.1111/1755-
18 0998.12046.
- 19 Dawnay, N. et al. (2007) 'Validation of the barcoding gene COI for use in forensic
20 genetic species identification', 173, pp. 1–6. doi: 10.1016/j.forsciint.2006.09.013.
- 21 Deagle, B. E. et al. (2014) 'DNA metabarcoding and the cytochrome c oxidase subunit
22 I marker: Not a perfect match', *Biology Letters*, 10(9). doi: 10.1098/rsbl.2014.0562.
- 23 Dergam, J. A. et al. (2002) 'Phylogeography and RAPD-PCR variation in *Hoplias*
24 *malabaricus* (Bloch , 1794) (Pisces , Teleostei) in southeastern Brazil', 387, pp. 379–
25 387.
- 26 DeSalle R, Birstein VJ. 1996. PCR identification of black caviar. *Nature* 381:197–198.
- 27 Dias-Neto, J. and Dias, F. (2015) *O uso da Biodiversidade Aquática No Brasil: uma*
28 *avaliação com foco na pesca*. Edited by Ibama. Brasília.
- 29 van Dijk, E. L. et al. (2014) 'Ten years of next-generation sequencing technology',
30 *Trends in Genetics*, 30(9), pp. 418–426. doi: 10.1016/j.tig.2014.07.001.
- 31 Dominguez, K. and Ward, W. S. (2010) 'A novel nuclease activity that is activated by
32 Ca²⁺ chelated to EGTA', *Biology of Reproduction*, 55, pp. 193–199. doi:
33 10.3109/19396360903234052.A.
- 34 Fischer, J. (2013) *Fish identification tools for biodiversity and fisheries assessments.*
35 *Review and guidance for decision-makers.*, FAO fisheries and aquaculture technical
36 paper. doi: 1932.
- 37 FAO (2016) *The State of World Fisheries and Aquaculture 2016*. Contributi. Rome.

1 Farrell, L. E., Roman, J. and Sunquist, M. E. (2000) 'Dietary separation of sympatric
2 carnivores identified by molecular analysis of scats', *Molecular Ecology*, 9(10), pp. 1583–
3 1590. doi: 10.1046/j.1365-294X.2000.01037.x.

4 Flamingh, A. et al. (2018) 'Species identification and mitochondrial genomes of
5 ancient fish bones from the Riverine Kachemak tradition of the Kenai Peninsula, Alaska',
6 *Mitochondrial DNA Part B: Resources*. Informa UK Ltd., 3(1), pp. 409–411. doi:
7 10.1080/23802359.2018.1456371.

8 Foran, D. R., Fischer, A. B. and Stoloff, M. E. (2015) 'A Comparison of Mitochondrial
9 DNA Amplification Strategies for Species Identification', *Journal of Forensic Investigation*,
10 3(2). doi: 10.13188/2330-0396.1000025.

11 Frago-Moura, E. N. . et al. (2016) 'Loss of biodiversity in a conservation unit of the
12 Brazilian Atlantic Forest : the effect of introducing non-native fish species', 76(1), pp. 18–
13 27.

14 Fricke, R., Eschmeyer, W. N. & Van der Laan, R. (eds) 2019. *ESCHMEYER'S*
15 *CATALOG OF FISHES: GENERA, SPECIES,*
16 *REFERENCES.*([http://researcharchive.calacademy.org/research/ichthyology/catalog/fish](http://researcharchive.calacademy.org/research/ichthyology/catalog/fish_catmain.asp)
17 [catmain.asp](http://researcharchive.calacademy.org/research/ichthyology/catalog/fish_catmain.asp)). Electronic version accessed 21 02 2019.

18 Goldberg, C. S. et al. (2016) 'Critical considerations for the application of
19 environmental DNA methods to detect aquatic species', pp. 1299–1307. doi:
20 10.1111/2041-210X.12595.

21 Grant, J. J. R. et al. (2012) 'Comparing thousands of circular genomes using the
22 CGView Comparison Tool.', *BMC genomics*, 13(1), p. 202. doi: 10.1186/1471-2164-13-
23 202.

24 Hajibabaei, M. Singer, G.A., Hickey, D.A. (2006b) 'Benchmarking DNA barcodes: An
25 assessment using available primate sequences', *Genomes*, 49 (7): 851-854. doi:
26 10.1139/g06-025

27 Hebert, P. D. N. et al. (2003) 'Biological identifications through DNA barcodes',
28 *Proceedings of Royal Society of London*, (270), pp. 313–321. doi:
29 10.1098/rspb.2002.2218.

30 Hebert, P. D. N. et al. (2003) 'Biological identifications through DNA barcodes',
31 *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), pp. 313–321. doi:
32 10.1098/rspb.2002.2218.

33 Hebert, P., Ratnasingham, S. and Jeremy, R. deWaard (2003) 'Barcoding animal life :
34 cytochrome c oxidase subunit 1 divergences among closely related species', *Biology*
35 *Letters*, 270, pp. 96–99. doi: 10.1098/rsbl.2003.0025.

36 Hellberg, R. S., Hernandez, B. C. and Hernandez, E. L. (2017) 'Identification of meat
37 and poultry species in food products using DNA barcoding', *Food Control*. Elsevier Ltd,
38 80, pp. 23–28. doi: 10.1016/j.foodcont.2017.04.025.

39 Herrmann B.G., Frischauf A.M. (1987). Isolation of genomic DNA. *Methods in*
40 *Enzymology*. 152:180–183.

- 1 Hilz, H., Wiegers, U. and Adamietz, P. (1975) 'Stimulation of proteinase K action by
2 denaturing agents: application to the isolation of nucleic acids and the degradation of
3 masked proteins', *Biochem*, 108, pp. 103–108.
- 4 Iwasaki, W. et al. (2013) 'MitoFish and MitoAnnotator: A Mitochondrial Genome
5 Database of Fish with an Accurate and Automatic Annotation Pipeline', 30(11), pp. 2531–
6 2540. doi: 10.1093/molbev/mst141.
- 7 Jamandre, B. W., Durand, J. and Tzeng, W. (2014) 'High Sequence Variations in
8 Mitochondrial DNA Control Region among Worldwide Populations of Flathead Mullet *Mugil*
9 *cephalus*', *International Journal for Zoology*, 2014, pp. 1–9.
- 10 Jane E. Huffman and Wallace, J. R. (eds) (2012) *Wildlife Forensics: methods and*
11 *applications*. 1th edn. New Delhi, India: John Wiley & Sons, Ltd.
- 12 Karlsson, S. et al. (2014) 'A standardized method for quantifying unidirectional genetic
13 introgression', pp. 3256–3263. doi: 10.1002/ece3.1169.
- 14 Kearse, M. et al. (2003) 'Geneious Basic: An integrated and extendable desktop
15 software platform for the organization and analysis of sequence data', *Bioinformatics*,
16 18(12), pp. 69–83. doi: 10.1093/bioinformatics/bts199.
- 17 Kimura, M. (1980) 'A simple method for estimating evolutionary rates of base
18 substitutions through comparative studies of nucleotide sequences', *Journal of Molecular*
19 *Evolution* ©, 16(1330).
- 20 Larizza, A. et al. (2002) 'Lineage Specificity of the Evolutionary Dynamics of the
21 mtDNA D-Loop Region in Rodents', *Journal of Molecular Evolution*, pp. 145–155. doi:
22 10.1007/s00239-001-0063-4.
- 23 Lemos Queiroz, A. L. et al. (2018) 'Complete mitochondrial genome sequence of
24 *Myrmecophaga tridactyla* from Brazilian Savanna', *Mitochondrial DNA Part B: Resources*.
25 Informa UK Ltd., 3(2), pp. 681–682. doi: 10.1080/23802359.2018.1481785.
- 26 Linacre, A. et al. (2011) 'ISFG: Recommendations regarding the use of non-human
27 (animal) DNA in forensic genetic investigations', *Forensic Science International: Genetics*,
28 5(5), pp. 501–505. doi: 10.1016/j.fsigen.2010.10.017.
- 29 Linacre, A. and Tobe, S. (2013) *Wildlife DNA Analysis*. 1th edn. Glasgow, UK: Wiley-
30 Blackwell.
- 31 Logan, C. A. et al. (2008) 'An impediment to consumer choice: Overfished species
32 are sold as Pacific red snapper', *Biological Conservation*, 141(6), pp. 1591–1599. doi:
33 10.1016/j.biocon.2008.04.007.
- 34 Luz, R. K. et al. (2000) 'Avaliação de canibalismo e comportamento territorial de
35 alevinos de trairão (*Hoplias lacerdae*)', 22(2), pp. 465–469.
- 36 Luz, R. K. and Portella, M. C. (2002) 'Larvicultura de Trairão (*Hoplias lacerdae*) em
37 Água Doce e Água Salinizada. Trairao (*Hoplias lacerdae*) Larviculture in Slightly Saline
38 Freshwater', 2002, pp. 829–834.

- 1 Macdonald, A. and Sarre, S. (2017) 'A framework for developing and validating taxon-
2 specific primers for specimen identification from environmental DNA', *Molecular Ecology*
3 *Resources*, 17, pp. 708–720. doi: 10.1111/1755-0998.12618.
- 4 Marko PB, Lee SC, Rice AM, Gramling JM, Fitzhenry TM, McAlister JS, Harper GR,
5 Moran AL. 2004. Fisheries: Mislabelling of a depleted reef fish. *Nature* 430:309–310.
- 6 Mattox, G. M. T., Bifi, A. G. and Oyakawa, O. T. (2014) 'Taxonomic study of *Hoplias*
7 *microlepis* (Günther, 1864), a trans-Andean species of trahiras (*Ostariophysi*:
8 *Characiformes*: *Erythrinidae*)', 12(2), pp. 343–352. doi: 10.1590/1982-0224-20130174.
- 9 Mattox, G. M. T., Toledo-Piza, M. and Oyakawa, O. T. (2006) 'Taxonomic Study of
10 *Hoplias* *Aimara* (Valenciennes, 1846) and *Hoplias* *macrophthalmus* (Pellegrin, 1907)
11 (*Ostariophysi*, *Characiformes*, *Erythrinidae*)', *Copeia*, 2006(3), pp. 516–528.
- 12 Menezes, N. A. (1988). Implications of the distribution patterns of the species of
13 *Oligosarcus* (Teleostei, Characidae) from central and southern South America. Pp. 295-
14 304. In: Hyer, W. R. & P. E. Vanzolini (Eds.). *Proceedings of a workshop on Neotropical*
15 *distribution patterns*. Rio de Janeiro, Academia Brasileira de Ciências, 488p.
- 16 Meusnier, I. et al. (2008) 'A universal DNA mini-barcode for biodiversity analysis',
17 *BMC Genomics*, 9, pp. 4–7. doi: 10.1186/1471-2164-9-214.
- 18 Mirande, M. (2010). 'Phylogeny of the family Characidae (Teleostei: Characiformes):
19 from characters to taxonomy', *Neotropical Ichthyology*, 8(3):385-568.
20 <http://dx.doi.org/10.1590/S1679-62252010000300001>
- 21 Miot, H. A. et al. (2011) 'Qual o tamanho da amostra ideal para se realizar um ensaio
22 clínico?', *Food Chemistry*, 54(4), pp. 275–278. doi: 10.1590/S1677-54492011000400001.
- 23 Morgan, J. A. T. et al. (2005) 'Origin and diversification of the human parasite
24 *Schistosoma mansoni*', *Molecular Ecology*, 14, pp. 3889–3902. doi: 10.1111/j.1365-
25 294X.2005.02709.x.
- 26 Moritz, C. and Cicero, C. (2004) 'DNA barcoding: Promise and pitfalls', *PLoS Biology*,
27 2(10). doi: 10.1371/journal.pbio.0020354.
- 28 Nishimura, Y., Yoshinari, T., Naruse, K., Yamada, T., Sumi, K., Mitani, H.,
29 Higashiyama, T., and Kuroiwa, T. (2006). *Proc. Natl. Acad. Sci.* 103, 1382–1387.
- 30 Núñez-Rodríguez, D. et al. (2016) 'Complete mitochondrial DNA of *Hoplias*
31 *intermedius* (Günther, 1864) (*Ostariophysi*: *Characiformes*: *Erythrinidae*)', *Mitochondrial*
32 *DNA Part B: Resources*, 1(1), pp. 742–743. doi: 10.1080/23802359.2016.1186521.
- 33 Ojala D., Montoya J., Attardi G. (1981). tRNA punctuation model of RNA processing
34 in human mitochondria. *Nature*. 290:470–474.
- 35 Oliveira, E. A. De et al. (2015) 'Comparative cytogenetics in the genus *Hoplias*
36 (*Characiformes*, *Erythrinidae*) highlights contrasting karyotype evolution among
37 congeneric species', *Molecular Cytogenetics*. *Molecular Cytogenetics*, pp. 1–10. doi:
38 10.1186/s13039-015-0161-4.

- 1 Oyakawa, O. T. and Mattox, G. M. T. (2009) 'Revision of the Neotropical trahiras of
2 the *Hoplias lacerdae* species-group (*Ostariophysi* : *Characiformes* : *Erythrinidae*) with
3 descriptions of two new species', 7(2), pp. 117–140.
- 4 Palomares, F. et al. (2002) 'Faecal genetic analysis to determine the presence and
5 distribution of elusive carnivores: Design and feasibility for the Iberian lynx', *Molecular*
6 *Ecology*, 11(10), pp. 2171–2182. doi: 10.1046/j.1365-294X.2002.01608.x.
- 7 Paredes, J. F., Paim, A. J., da Costa-Doria, E. M. & Rocha, W. L. (1983). Sao
8 Francisco River:Hydrological studies in the dammed lake of Sobradinho. In *Transport of*
9 *Carbon and Minerals in Major World Rivers, Part 2* (Degens, E. T., Kempe, S. & Soliman,
10 H., eds),pp. 193–202. Hamburg: UNEP/SCOPE.
- 11 Parson, W. et al. (2000) 'Species identification by means of the cytochrome b gene',
12 *Canadian Journal of Fisheries and Aquatic Sciences*, 114(1), pp. 23–28. Available at:
13 <http://www.springerlink.com/index/J2M7DU0A9PCLUC1H.pdf>.
- 14 Pereira, A. H. et al. (2019) 'Complete mitochondrial genome sequence of *Hypostomus*
15 *francisci* (*Siluriformes* : *Loricariidae*) Complete mitochondrial genome sequence of
16 *Hypostomus francisci*', *Mitochondrial DNA Part B: Resources*. Taylor & Francis, 4(1), pp.
17 155–157. doi: 10.1080/23802359.2018.1544860.
- 18 Pereira, L. H. G. et al. (2013) 'Can DNA barcoding accurately discriminate
19 megadiverse Neotropical freshwater fish fauna?', pp. 1–14.
- 20 Pimentel, J. da S. M. et al. (2014) 'Complete mitochondrial genome sequence of
21 *Piaractus mesopotamicus*', *Mitochondrial DNA*, 1736, pp. 1–2. doi:
22 10.3109/19401736.2014.971297.
- 23 Potter, P. E. (1997). The Mesozoic and Cenozoic paleodrainage of South America: a
24 natural history. *Journal of South American Earth Sciences*, 10(5-6): 331-344.
- 25 Raza, K. and Ahmad, S. (2016) 'Recent advancement in Next Generation Sequencing
26 techniques and its computational analysis', *bioRxiv*, 13, p. 341. doi:
27 10.1016/j.apcata.2006.04.017.
- 28 Reis, R. E. et al. (2016) 'Fish biodiversity and conservation in South America', *Journal*
29 *of Fish Biology**Fish Biology*, 89, pp. 12–47. doi: 10.1111/jfb.13016.
- 30 Reis, R. E., Kullander, S. O. and Ferraris, C. J. J. (2003) *Check List of the Freshwater*
31 *Fishes of South and Central America* Edited. Edited by R. E. Reis, S. O. Kullander, and
32 C. J. J. Ferraris. Porto Alegre: EDIPUCRS.
- 33 Resende, L. C. et al. (2016) '*Pimelodus maculatus* (*Siluriformes* , *Pimelodidae*):
34 complete mtDNA sequence of an economically important fish from the São Francisco river
35 basin', *Mitochondrial DNA Part B: Resources*. Informa UK Ltd., 0(0), p. 000. doi:
36 10.1080/23802359.2016.1219646.
- 37 Ribeiro, A. C. (2006). Tectonic history and the biogeography of the freshwater fishes
38 from the coastal drainages of eastern Brazil: an example of faunal evolution associated
39 with a divergent continental margin. *Neotropical Ichthyology*, 4(2): 225-246.

- 1 Rojas, M. et al. (2011) 'Development of a real-time PCR assay to control the illegal
2 trade of meat from protected capercaillie species (*Tetrao urogallus*)', *Forensic Science*
3 *International*, 210, pp. 133–138. doi: 10.1016/j.forsciint.2011.02.021.
- 4 Saitou, N. and Nei, M. (1987) 'The neighbor-joining method: a new method for
5 reconstructing phylogenetic trees.', *Molecular Biology and Evolution*, 4(4), pp. 406–425.
6 doi: 10.1093/oxfordjournals.molbev.a040454.
- 7 Sanger, F., Nicklen, S., Coulson A.R. 1977. DNA sequencing with chain-terminating
8 inhibitors. *Proc Natl Acad Sci*, 74:5463–5467.
- 9 Satoh, T. P. et al. (2016) 'Structure and variation of the mitochondrial genome of
10 fishes', *BMC Genomics*. *BMC Genomics*, 17(719), pp. 1–20. doi: 10.1186/s12864-016-
11 3054-y.
- 12 Schmieder, R. and Edwards, R. (2015) 'Quality control and preprocessing of
13 metagenomic datasets', *Kukila*, 18(2), pp. 46–59. doi: 10.1093/bioinformatics/btr026.
- 14 Siqueira, F. de F. et al. (2014) 'Complete mitochondrial genome sequence of *Brycon*
15 *orbignyanus*', *Mitochondrial DNA*, 1736, pp. 1–2. doi: 10.3109/19401736.2014.971298.
- 16 Spencer, P. B. S., Schmidt, D., Hummel, S. (2010) ' Identification of historical
17 specimens and wildlife seizures originating from highly degraded sources of kangaroos
18 and other macropods', *Forensic Sci Med Pathol* (2010) 6:225–232, doi: 10.1007/s12024-
19 009-9119-3
- 20 Stein, L. (2001). *Reviews genome annotation : from sequence to biology*, 2(July).
- 21 Stothard, P. (2000) 'The Sequence Manipulation Suite: JavaScript Programs for
22 Analyzing and Formatting Protein and DNA Sequences', *Biotechniques*, 28(6). doi:
23 10.2144/00286ir01.
- 24 Tamura, K. et al. (2013) 'MEGA6: Molecular evolutionary genetics analysis version
25 6.0', *Molecular Biology and Evolution*, 30(12), pp. 2725–2729. doi:
26 10.1093/molbev/mst197.
- 27 Teixeira, L. V. et al. (2012) 'Extração de DNA e avaliação da composição espécie-
28 específica de queijos', *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 64(3), pp.
29 721–726. doi: 10.1590/S0102-09352012000300025.
- 30 Tobe, S. S., Kitchener, A. and Linacre, A. (2009) 'Cytochrome b or cytochrome c
31 oxidase subunit I for mammalian species identification-An answer to the debate', *Forensic*
32 *Science International: Genetics Supplement Series*, 2(1), pp. 306–307. doi:
33 10.1016/j.fsigss.2009.08.053.
- 34 Untergasser, A. et al. (2007) 'Primer3Plus, an enhanced web interface to Primer3',
35 *Nucleic Acids Research*, 35(SUPPL.2), pp. 71–74. doi: 10.1093/nar/gkm306.
- 36 Vari, R. P. & A. S. Harold. (2001). Phylogenetic study of the neotropical fish genera
37 *Creagrutus* Günther and *Piabina* Reinhardt (Teleostei: Ostariophysi: Characiformes), with
38 a revision of the cis-andean species. *Smithsonian Contributions to Zoology*, 613, 239p.

- 1 Vieira, F., C. Alves, P. Santos. 2009. Peixes. 81-121 p. Em: Gláucia Moreira
2 Drumond; Cássio Soares Martins; Magda Barcelos Greco; Fábio Vieira. Biota Minas.
3 Fundação Biodiversitas. Belo Horizonte. Yang, D.Y. & Speller, C.F., 2006. Co-
4 amplification of cytochrome b and D-loop mtDNA fragments for the identification of
5 degraded DNA samples. *Molecular Ecology Notes*, 6, pp.605–608.
- 6 Vink, C. J., Paquin, P. and Cruickshank, R. H. (2012) 'Taxonomy and Irreproducible
7 Biological Science', 62(5), pp. 451–452. doi: 10.1525/bio.2012.62.5.3.
- 8 Yang, D. Y. and Speller, C. F. (2006) 'Co-amplification of cytochrome b and D-loop
9 mtDNA fragments for the identification of degraded DNA samples', *Molecular Ecology*
10 *Notes*, 6(3), pp. 605–608. doi: 10.1111/j.1471-8286.2006.01370.x.
- 11 Yang, L. et al. (2014) 'Species identification through mitochondrial rRNA genetic
12 analysis', *Scientific Reports*, 4, pp. 1–11. doi: 10.1038/srep04089.
- 13 Yang, Y., Xie, B. and Yan, J. (2014) 'Application of next-generation sequencing
14 technology in forensic science', *Genomics, Proteomics and Bioinformatics*. Beijing
15 Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China,
16 12(5), pp. 190–197. doi: 10.1016/j.gpb.2014.09.001.
- 17 Ye, J. et al. (2012) 'Primer-BLAST: A tool to design target-specific primers for
18 polymerase chain reaction'.
- 19 Xiang, J. M. and Hickey D. (2007). 'Assessing the effect of varying sequence length
20 on DNA barcoding of fungi', *Molecular Ecology Notes* (2007) 7, 365–373. doi:
21 10.1111/j.1471-8286.2007.01698.x
- 22

1

2

3

4

5

6

7 **8. ANEXOS**

8

9

Tabela 13. Lista de tecidos analisados a partir de coleções científicas.

Coleção	Espécie	Número Amostra	Bacia	Latitude	Longitude
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC609	São Francisco	16°10' 38.18" S	45° 41' 03.69" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC639	São Francisco	16°10' 38.18" S	45° 41' 03.69" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC650	São Francisco	16°10' 38.18" S	45° 41' 03.69" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC657	São Francisco	16°10' 38.18" S	45° 41' 03.69" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC1184	São Francisco	18° 8' 9.76" S	45° 14' 41.78" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC1188	São Francisco	18° 8' 9.76" S	45° 14' 41.78" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC1228	São Francisco	17° 36' 58.1" S	44° 40' 32.9" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC1234	São Francisco	17° 36' 58.1" S	44° 40' 32.9" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC1273	São Francisco	17° 35' 24.9" S	44° 39' 33.9" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC1999	São Francisco	20°26' 44.89" S	44° 1' 7.04" O

Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC3551	Rio Doce	19° 01' 19.13" S	42°07' 20.07" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC3552	Rio Doce	19° 01' 19.13" S	42°07' 20.07" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC3553	Rio Doce	19° 01' 19.13" S	42° 07' 20.07" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC3573	Rio Doce	20° 06' 22.54" S	43° 24' 12.03" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC3589	Rio Doce	19° 01' 55.07" S	42° 07' 34.23" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC 3707	Rio Doce	18° 57' 13" S	43° 26' 21" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC 3714	Rio Doce	18° 55' 01" S	43° 27' 42" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC 4003	Rio Mucuri	17° 50' 97.5" S	40° 22' 60.2" O
Laboratório de Genética da Conservação PucMG	<i>Hoplias intermedius</i>	LGC 4045	Rio Mucuri	17° 44' 08.1" S	40° 35' 82.0" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1095	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1556	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O

Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1557	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1558	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1645	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1646	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1647	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1648	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1650	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	1651	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	748_C8	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	576_C8	Rio Corrente Grande tributário do rio Doce	19° 01' 05.14"S	42° 07' 20.49" O

Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	517_C8	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	07_S	Rio Santo Antônio tributário do Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	860_C35	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	785_C34	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	742_74	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	741_74	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	738_74	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	702_74	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	675_C34	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	597	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O

Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	434	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	180	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	156	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	155	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	154	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	126	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	112	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Coleção de Ictiologia do Museu de Ciências Naturais da PUC Minas	<i>Hoplias intermedius</i>	111	Rio Doce	19° 01' 05.14"S	42° 07' 20.49" O
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias aimara</i>	INPA 40347	Rio Bacajái	03° 35' 30" S	051° 45' 56" O
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias curupira</i>	INPA 43428	Rio Xingu	03° 33' 11" S	051° 51' 22" O
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050268	Rio Trombetas	-1,4258333	-56,8552778

Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050333	Rio Trombetas	-1,4558333	-56,7791667
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050051	Rio Trombetas	-1,4308333	-56,7644444
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050322	Rio Trombetas	-1,4644444	-56,7291667
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050458	Baixo rio Madeira	-3.9823100000	-59.1194900000
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050512	Baixo rio Madeira	-4.0335700000	-59.1202500000
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050512	Baixo rio Madeira	-4.0335700000	-59.1202500000
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 050776	Baixo rio Madeira	-3.9275700000	-58.7679100000
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 052883	Rio Negro	-2,71791	-60,74293
Instituto Nacional de Pesquisas da Amazônia	<i>Hoplias malabaricus</i>	INPA-ICT 052937	Rio Negro	-2,82201	-60,87086
Museo de Ciência e Tecnologia da PUCRS	<i>Hoplias australis</i>	MCP-Peixes 000051218	Rio Forquilha	- 276.766.666.667	-51.7360277778

Tabela 14. Localização de possíveis marcadores específicos para *H. intermedius*. É indicado o tamanho de janela, a quantidade de *mismatches* a posição de início do marcador ou janela analisada e a região gênica a que pertence o possível marcador.

Tamanho da janela	Quantidade de <i>mismatches</i>	Posição de início da janela	Região do genoma
100	3	397 a 402 1.899 1.951 1.978 a 1.999 4.490 a 4.546	12S rRNA 16S rRNA 16S rRNA 16S rRNA ND2
	4	1.952 a 1977	16S rRNA
200	3	297 a 402 1.799 a 1.850 1.978 a 1.999 4.390 a 4.546 12.179 a 12.214 12.223 a 12.276 12.355 a 12.379 12.415	12S rRNA 16S rRNA 16S rRNA ND2 ND5 ND5 ND5 ND5
	4	1.851 1.900 a 1.977 12.215 a 12.222	16S rRNA 16S rRNA ND5
	5	1.852 a 1.899	16S rRNA
		183 a 201 278 a 402 1.780 a 1.831 1.978 a 1.998 4.371 a 4.546 12.160 a 12.195 12.223 a 12.276 12.336 a 12.379 12.396 a 12.415 12.541 a 12.555	12S rRNA 12S rRNA 16S rRNA 16S rRNA ND2 ND5 ND5 ND5 ND5 ND5
	4	1.832 1.900 a 1.977 12.196 a 12.222	16S rRNA 16S rRNA ND5
	5	1.833 a 1.899	16S rRNA
300	3	105 a 127 208 a 405 916 a 924 1.602 a 1.606 1.702 a 1.753 1.981 a 2.002 4.293 a 4.549 11.155 a 11.173 12.085 a 12.117	12S rRNA 12S rRNA 12S rRNA a 16SrRNA 16S rRNA 16S rRNA 16S rRNA ND2 ND4 ND5

		12.226 a 12.257 12.280 a 12.317 12.382 a 12.418 12.463 a 12.558 12.566 a 12.618	ND5 ND5 ND5 ND5 ND5
	4	128 a 199 205 a 207 1.754 1.903 a 1.980 12.118 a 12.225 12.258 a 12.279 12.318 a 12.382	12S rRNA 12S rRNA 16S rRNA 16S rRNA ND5 ND5 ND5
	5	200 a 204 1.755 a 1.902	12S rRNA 16S rRNA
400	4	28 a 99 205 a 207 816 a 861 1.580 a 1.601 1.607 a 1.653 1.903 a 1.980 12.018 a 12.157 12.280 a 12.362 12.383 a 12.418 12.466 a 12.558	tRNA-Phe a 1SrRNA 12S rRNA 12S rRNA a 16SrRNA 16S rRNA 16S rRNA 16S rRNA ND5 ND5 ND5 ND5
	5	100 a 204 1.602 a 1.606 1.654 1.661 a 1.902 12.158 a 12.217 12.226 a 12.279 12.363 a 12.382	12S rRNA 16S rRNA 16S rRNA ND5 ND5 ND5 ND5
	6	1.655 a 1.660 12.218 a 12.225	16S rRNA ND5
500	5	1 a 201 1.499 a 1.550 1.658 a 1.899 12.055 a 12.114 12.223 a 12.259 12.277 a 12.362 12.380 a 12.415	tRNA-Phe a 1SrRNA 16S rRNA 16S rRNA ND5 ND5 ND5 ND5
	6	1.551 1.604 a 1.657 12.115 a 12.222 12.260 a 12.276 12.363 a 12.379	16S rRNA 16S rRNA ND5 ND5 ND5
	7	1.552 a 1.603	16S rRNA

511	5	1 a 201 1.488 a 1.539 1.658 a 1.899 1.973 a 1.977 12.044 a 12.103 12.223 a 12.248 12.277 a 12.351 12.380 a 12.415	tRNA-Phe a 1SrRNA 16S rRNA 16S rRNA 16S rRNA ND5 ND5 ND5 ND5
	6	1.540 1.604 a 1.657 12.104 a 12.222 12.249 a 12.276 12.352 a 12.379	16S rRNA 16S rRNA ND5 ND5 ND5
	7	1.541 a 1.603	16S rRNA
583	5	1 a 201 338 a 402 1.416 a 1.467 1.658 a 1.977 11.972 a 12.031 12.277 a 12.279 12.380 a 12.415	tRNA-Phe a 1SrRNA 12S rRNA a 16SrRNA 16S rRNA 16S rRNA ND5 ND5 ND5
	6	1.468 1.604 a 1.657 12.032 a 12.176 12.223 a 12.276 12.288 a 12.379	16S rRNA 16S rRNA ND5 ND5 ND5
	7	1.469 a 1.603 12.177 a 12.222	16S rRNA ND5
600	5	1 a 201 321 a 402 1.399 a 1.450 1.659 a 1.883 1.900 a 1.977 11.955 a 12.014 12.380 a 12.415	tRNA-Phe a 1SrRNA 12S rRNA 16S rRNA 16S rRNA 16S rRNA ND5 ND5
	6	1.451 1.604 a 1.657 1.884 a 1.899 12.015 a 12.159 12.223 a 12.262 12.277 a 12.379	16S rRNA 16S rRNA 16S rRNA ND5 ND5 ND5
	7	1.452 a 1.603 12.160 a 12.222 12.263 a 12.276	16S rRNA ND5 ND5

Tabela 15. Localização de possíveis marcadores específicos para *H. malabaricus*. É indicado o tamanho de janela, a quantidade de *mismatches* a posição de início do marcador ou janela analisada e a região gênica a que pertence o possível marcador.

Tamanho da janela	Quantidade de <i>mismatches</i>	Posição de início da janela	Região do genoma
100	3	2456 a 2473	16S rRNA
		2488 a 2541	16S rRNA
		2556 a 2574	16S rRNA
100	4	5240 a 5268	tRNA-Asn
		8726 a 8800	ATPase 6 a COIII
		10775 a 10831	ND4
100	5	10833 a 10873	ND4
		13104 a 13163	ND5
		13188 a 13202	ND5
100	5	13722 a 13791	ND5 a ND6
		2474 a 2478	16S rRNA
		2485 a 2487	16S rRNA
100	5	2542 a 2556	16S rRNA
		13164 a 13187	ND5
		2479 a 2484	16S rRNA
200	4	2374 a 2378	16S rRNA
		2488 a 2556	16S rRNA
		1733 a 1831	16S rRNA
200	5	10872 a 10873	ND4
		13064 a 13187	ND5
		2379 a 2441	16S rRNA
200	6	2485 a 2487	16S rRNA
		2442 a 2484	16S rRNA
		2355 a 2359	16S rRNA
219	4	2488 a 2556	16S rRNA
		10714 a 10831	ND4
		10583 a 10873	ND4
219	5	13045 a 13187	ND5
		13603 a 13618	ND5 a ND6
		2360 a 2422	16S rRNA
219	6	2485 a 2487	16S rRNA
		2423 a 2484	16S rRNA
		2256 a 2273	16S rRNA
300	4	2488 a 2556	16S rRNA a tRNA-Leu
		10633 a 10771	ND4
		10832 a 10873	ND4
300	5	12964 a 13187	ND5
		13493 a 13521	ND5 a ND6
		13586 a 13618	ND5 a ND6
300	5	2274 a 2278	16S rRNA
		2314 a 2341	16S rRNA
		2485 a 2487	16S rRNA a tRNA-Leu

		10772 a 10831 13522 a 13585	ND4 ND5 a ND6
	6	2279 a 2313 2342 a 2484	16S rRNA 16S rRNA a tRNA-Leu
400	5	2174 a 2178 2485 a 2487 10672 a 10831 13185 a 13187 13422 a 13585	16S rRNA 16S rRNA a tRNA-Leu ND4 ND5 ND5 a ND6
	6	2179 a 2241 2314 a 2484	16S rRNA 16S rRNA a tRNA-Leu
	7	2242 a 2313	16S rRNA
500	5	1987 a 2049 2056 a 2073 2485 a 2487 10433 a 10505 10572 a 10831 13085 a 13117 13118 a 13202 13322 a 13585 13689 a 13791	16S rRNA 16S rRNA 16S a ND1 ND4 ND4 ND5 ND5 ND5 a ND6 ND5 a ND6
	6	2074 a 2078 2086 a 2141 2314 a 2484 13118 a 13187	16S rRNA 16S rRNA 16S rRNA a ND1 ND5
	7	2079 a 2085 2142 a 2313	16S rRNA 16S rRNA a tRNA-Leu
		10422 a 10505 10561 a 10831 13074 a 13106 13188 a 13202 13311 a 13585 13678 a 13791 13107 a 13187	ND4 ND4 ND5 ND5 ND5 a ND6 ND5 a ND6 ND5 ND4
		10489 a 10505 13035 a 13187 13239 a 13264 13583 a 13585 13606 a 13618	ND5 a ND6 ND5 a ND6 ND5 a ND6 ND5 a ND6
600	7	1974 a 1978 2086 a 2313 13193 a 13202	16S rRNA 16S rRNA a ND1 ND5 a ND6
	8	1979 a 2041 2050 a 2085	16S rRNA 16S rRNA
	9	2042 a 2049	16S rRNA

1.1. Sequencias dos amplicons obtidos a partir das amostras do rio doce usando o marcador específico para *H. intermedius* Hin5F-Hin3R

Amplicon esperado Hin5F-Hin3R:

GCCAGTCGTA CTTGAGCAAAGAGAACTTTAGTTCAAAGCCCCGAAACCAAGCGAGC
TACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGGGAAG
ATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTGTCTAGGA
AATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCACCACACC

1. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAA
AGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGC
2. CTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGC
CCCCGCACTCCTTAACTCAYYKWSRAMCACCACACC
3. GCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAA
GAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATA
GCTGGTTGTCTAGGAAATGGATAGAAGTTCA
4. CAAGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGC
AAAAGAGTGGGAAGATNNCTTTGGGTAGAGGTGACAAGCCTACCNGAGCT
-TGGTGATAGC-TGGTTGTC-TAGGAAATGGATAGAAGTTCAGC
5. AAGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAA
AAGAGTGGGAAGATCTTTGGGTAGANGGTGACAAGCCTACCNGAGCTTGGT
GATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAA
CTCACCARGRCCACCACACC
6. ACWGAGCAAAGAGAACTTTAGTTCAAAGCCCCGAAACCAAGCGAGCTACCC
CAGGACAGCCTANTTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGGGAA
GATCTNTTGGGTAGAGGTGACAAGCCTNACCGAGCNTTGGTGATAGCTGGT
TGTCTNAGGAAATGGATAGATAGTTCAGCCCCGCACTCCTTAACTCACCWK
RGCCACCACACCYR
7. AACTTACCCYAAGGAWAGCCTATTWAGGGCCAACCCGTCTCTGTGGCAWA
TAGWAGTGGGAWGATCTTTGGGTWGWGGTGACAWGCCTACCGWGCTTG
GTGATAGCTGGTTGTCTAGGAAATGGATAGAWGTTTCAGCCCCGCACTCCT
AACTCACCWGGRCCACCACACC
8. AGCGAGCTNACCCAGGACAGCCNTATTAGGGCCAACCCGTCTCTGTGGC
AAAAGAGTGGGAAGATCT-TTGGGTAGAGGTGACA-AGCCT-ACCGNAGC-
TTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTC
CTTAACTCACCARGRCCACCACACCT
9. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAA
AGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGAT
AGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTY
WCWARGGCCACCACACCA

10. GCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAA
GAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATA
GCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCA
CCAGGRMCACCACACCTG
11. CAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCA
AAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTG
ATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAMC
TCMSCAGGGCCACCACACC
12. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAA
AGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGAT
AGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCA
CCAGGRCCACCACACC
13. AAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCA
AAAAGAGTGGGAAGGATCTTTGGGTAGANGGTGACAAGCCTACCG-AGCT-
TGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCC
TTAACTCACCARGRCCACCACACC
14. GCCCCGAAACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCG
TCT-CTGTGGCAAAAGGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCT-
ACCNGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCC
CCGCACTCCTTAACTCACCARARCCACCACACC
15. CAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAA
AAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGA
TAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTC
ACCAGGRCCACCACACC
16. AAAGCGAGCTACCCCAGNGACAGCCTNATTAGGGCCCAAACCCGTCTCTGT
GGCAAAAAGGAGTGGGAAN-GATNCTTTGGGTAGAGGTGNACAAGCCTACC-
GAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGC
A
17. CAAAGCCCCGAAACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAA
CCCGTCTCTGTGGCAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGC
CTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGC
CC
18. TTAGTTCAAAGCCCCGAAACCAAGCGAGCTACCCCAGGACAGCCTATTAGG
GCCAACCCGTCTCTGTGGCAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGA
CAAGCCTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGT
TCAGCCCCGCACTC
19. TCAAAGCCCCGAAACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAA
CCCGTCTCTGTGGCAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGC
CTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGC

C—CC

20. CAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCA
AAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTG
ATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCC
21. AAAGCGAGCNTACCCCAGGACAGCCTNATTAGGGCCAACCCGTCTCTGTGG
CAAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTG
GTGAT-AGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACT
22. CCCCAGAAACCAAGCGAGCNTACCCCAGGACAGCCTATTAGGGCCAACCCG
TCTCTGTGGCAAAAAGAGTGGGAAGNATCTTNTGGGTAGAGGTGACAAGCC
TACCGNAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGC
CCCGCA
23. CCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGG
CAAAAGAGTGGGAAGATCTTTGGGTANGAGGTGACAAGCCTACCGAGCTTG
GTNGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACT
24. ACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGG
CAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGG
TGATAGCTGGTTGTCTAGGAAATGGATAGAAGTT—CAGCC
25. GCTACCCCAGGACAGCCTATTAAGGGCCAACCCGTCTCTGTGGCAAAAGAG
TGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACCTACCA
RGRCCACACACC
26. TCAAAGCCCCGAAACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAA
CCCGTCTCTGTGGCAAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGC
CTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGC
CC
27. ACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGG
CAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGG
TGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCC
28. AAGCCCCGAAACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACC
CGTCTCTGTGGCAAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCT
ACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGC
29. AAACCAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGT
GGCAAAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTT
GGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGT—TCAGCCC

1.2. Sequencias dos amplicons obtidos a partir das amostras do rio doce usando o marcador específico para *H. intermedius* Hin5F-Hin4R

Amplicon esperado Hin5F-Hin4R:

1. GCCAGTCGTA CTTGAGCAAAGAGAACTTTAGTTCAAAGCCCCGAAACCAAGCGA
GCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGG
GAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTG
TCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCAC
CACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCC
TTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAGATC
TTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCA
GGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTACTA
GACCGTTCCATGCCAACATGGAAGTGACCCTGCTAAAATGAGTAATAAGAAGGT
GTACCCTTCTCCTATAGCCCGT
2. CTGTGGCAAAGAGTGGGAAGAKCTTTGGGTAGAGRTGACAAGCCTACCGAGC
TTGGTGATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTT
AACTCACCAGGGCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCA
AAGGGGGTACAGCCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTTAGGA
TCATACTAAACAAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATA
GAAAGCGTTAAAGCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCA
CACCCCTTCACTTACTAGACCGTT-CCATGCCAACATG
3. CGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGAG
TGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGG
TTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCC
ACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCC
CCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAGA
TCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCT
CAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTAC
TAGAC
4. GTCTCTGTGGCAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACC
GAGCTTGGTGATAGCTGGTTGTCTAGGAAAATGGATAGNAAGTTCAGCCCCGC
ACTCCTTAACTCACCAGGGCCACCACACCACGAGACAAAGAAAGACGCGAGAG
TTAGTCAAAGGGGGTACAGCCCCCTTTGAATCAGGACACAACCTTACCAGGAGG
TTAAGGATCATACTAAACAAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCAT
CTGCATAGAAAGCGTTAAAGCTCAGGCAGTTTATAAATCTATTATTCTGATACCC
CATCCCACACCCCTTCACTTACTNAGACCGTTCCATGCC
5. GCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT

TACTAGACCGTTCCATGCCAACATG

6. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGA
7. AGCGAGCTACCCCAGGACAGCCTATTAGGGNCCCAACCCCGTCTCTGTGGCA
AAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACNAAGCCNTACCGAGCTTGGT
GATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGNCCCCGCACTCCCTTTAA
CTCACCAGGGCCACCACACCACGAGACAANAGAAAGACGCGAGAGTTAGTCAA
AGGGGGTACAGCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGAT
CATACTAAACAAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAG
AAAGCGTTAAAGCTCAGGCAGTTTTATAAATCTATTATTCTGATACCCCATCCA
CACCCNNTTCACTTACTAANGACCGTTCCATGCCAACATG
8. AGCGAGCTACCCCAGGACAGCCTATTAGGGNCCCAACCCCGTCTCTGTGGCA
AAAAGAGTGGGAAGATCTTTGGGTAGAGGTGACNAAGCCNTACCGAGCTTGGT
GATAGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGNCCCCGCACTCCCTTTAA
CTCACCAGGGCCACCACACCACGAGACAANAGAAAGACGCGAGAGTTAGTCAA
AGGGGGTACAGCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGAT
CATACTAAACAAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAG
AAAGCGTTAAAGCTCAGGCAGTTTTATAAATCTATTATTCTGATACCCCATCCA
CACCCNNTTCACTTACTAANGACCGTTCCATGCCAACAT
9. CAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAA
AGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAG
CTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAG
GGCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTAC
AGCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAAC
AAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
AGCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTAC
TACTAGACCGTTCCATGCCAACATG
10. TACAAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCA
AAAGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGAT
AGCTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACC
AGGGCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGG
TACAGCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTA
ACAAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGT
TAAAGCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTT
CACTTACTAGACCGTTCCATG
11. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGBCAAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT

GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGACCCTGCT

12. CCCCAGGACAGCCTATTAGNGGCCAACCCGTCTCTGTGGCAAAAAGAGTGGGA
AGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTGTCT
AGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCACCAC
ACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCCTTT
GAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAGATCTTC
TGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCAGG
CAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTACTAGA
CCGTTCCATGCCAACATGGAAGTGA

13. CAAGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAA
AGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAG
CTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAG
GGCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTAC
AGCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAAC
AAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
AGCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTAC
TACTAGACCGTTCCATGCCAACATGGAAGTGA

14. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGC
TGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGACCCTG

15. CAAGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAA
AGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAG
CTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAG
GGCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTAC
AGCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAAC
AAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
AGCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTAC
TACT-AGACCGTTCCATGCCAACATGGAA

16. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA

- GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTT-
CACTTACTA--GACCGTTCCATGCCAACATGGAAGTGACC
17. GCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGA
GTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTG
GTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGC
CACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGC
CCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAG
ATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGC
TCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTA
CTAGACCGTTCCATGCCAACATGGAAGTGACC
 18. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAG-ACCGTTCCATGCCAACATGGAAGTGACC
 19. GCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGG
GAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTG
TCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCAC
CACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCC
TTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAGATC
TTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCA
GGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTACTA
GACCGTTCCATGCCAACATGGAAGTGAC
 20. GCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGG
GAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTG
TCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCAC
CACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCC
TTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAGATC
TTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCA
GGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTACTA
GACCGTTCCATGCCAACATGGAAGTGA
 21. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAAGCGTTAA
AGCTCAGGCAGTTTATAAATCTNATTATNTCTGATACCCCATCCCACACCCCTT
ACTTACTANGACCGTTCCATGCCAACATGGAAGTGACCCTGCTAAAATGAGT
 22. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT

GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTANTTCTGATACCCCATCCCACACCCCTNTCA
CTTACTAN-GACCGTTCCNATG

23. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTNTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAAC
AAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
AGCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTAC
TT-ACTA—GACCGTTCCATGCCAACATGGAAGTGACC

24. GCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGA
GTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTG
GTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGC
CACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGC
CCCTTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAA
GATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAG
CTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTT
ACTAGACCGTTCCATGCCAACAT

25. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGAC

26. GCGAGCTACCCCAGGACAGCCTATTAGGGCCNAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAA

27. GACAGCCTATTAGGGCCAACCCGTCTCTGTGGCRAAAGAGTGGGAAGATCTTT
GGGTAGAGGTGASAAGCCTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAAT
GGATAGAAGTTCAGCCCMGCWCTCCTTAACTCACCAGGGCCACCACACCACGA
GACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCCTTTGAATCAG
GACACAWCCTTACCAGGAGGTTAAGGATCATACTAAACAAGATCTTCTGCTTCA
GTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCAGGCAGTTTAT

AAATCTATTATTCTGATACCCCATCCCACACMCCTTCACTTACTAGACCGTACCA
TGCSAACA

28. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGA

29. CAACGAGCTACCCAGACACCTATTAGGGCCACCCGTCTCTGTGGCAAAGAGT
GGGAAGACTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTT
GTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCA
CCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCC
CTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAGAT
CTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTC
AGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTACT
AGACCGTACCATGCCAACA

30. CAAGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAA
AGAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAG
CTGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAG
GGCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTAC
AGCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAAC
AAGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
AGCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCC-
TTCAC-TTACTA—GACCGTTCCATGCCAACATGGAAGTGA

31. GCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGA
GTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTG
GTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGC
CACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGC
CCCTTTGAATCAGGACACAACCTTACCAGGAGGNTTAAGGATCATACTAAACAA
GATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAG
CTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTT
ACTAGACCGTTCCATGCCAACATGGAAGTGACCCTGCT

32. AGCGAGCTACCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAANTGACCCTGCTAAAATGAGT

33. CCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGGGAAGAT
CTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTGTCTAGG
AAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCACCACACC
ACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCCTTTGAA
TCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAGATCTTCTGC
TTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCAGGCAGT
TTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTACTAGACCGT
TCCATGCCAACATGGAAGTGACCCTGCTAACATGAG
34. GCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGA
GTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTG
GTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGC
CACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGC
CCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAG
ATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGC
TCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTA
CTAGACCGTTCCATGCCAACATGG-AAGTGAC
35. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGACC
36. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGA
37. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGAC
38. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA

GCCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGA

39. GCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGA
GTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTG
GTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGC
CACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGC
CCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAG
ATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGC
TCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTA
CTAGACCGTTCCATGCCAACATGGAAG

40. AGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAG
AGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCT
GGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACATGGAAGTGAC

41. AAGCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAA
GAGTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGC
TGGTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGG
GCCACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACA
GCCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACA
AGATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAA
GCTCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACT
TACTAGACCGTTCCATGCCAACAT

42. GCGAGCTACCCCAGGACAGCCTATTAGGGCCAACCCGTCTCTGTGGCAAAGA
GTGGGAAGATCTTTGGGTAGAGGTGACAAGCCTACCGAGCTTGGTGATAGCTG
GTTGTCTAGGAAATGGATAGAAGTTCAGCCCCGCACTCCTTAACTCACCAGGGC
CACCACACCACGAGACAAAGAAAGACGCGAGAGTTAGTCAAAGGGGGTACAGC
CCCTTTGAATCAGGACACAACCTTACCAGGAGGTTAAGGATCATACTAAACAAG
ATCTTCTGCTTCAGTGGGCCTAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGC
TCAGGCAGTTTATAAATCTATTATTCTGATACCCCATCCCACACCCCTTCACTTA
CTAGACCGTTCCATGCCAACATGGAAGTGACC

43. TATTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGGGAAGATCTTTGGGTAGA
GGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGA
AGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCACCACACCACGAGACAAAG
AAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCCTTTGAATCAGGACACAA
CCTTACCAGGAGGTTAAGGATCATACTAAACAAGATCTTCTGCTTCAGTGGGCC
TAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCAGGCAGTTTATAAATCTAT
TATTCTGATACCCCATCCCACACCCCTTCACTTACTAGACCG

44. TATTAGGGCCAACCCGTCTCTGTGGCAAAGAGTGGGAAGATCTTTGGGTAGA
GGTGACAAGCCTACCGAGCTTGGTGATAGCTGGTTGTCTAGGAAATGGATAGA
AGTTCAGCCCCGCACTCCTTAACTCACCAGGGCCACCACACCACGAGACAAAG
AAAGACGCGAGAGTTAGTCAAAGGGGGTACAGCCCCTTTGAATCAGGACACAA
CCTTACCAGGAGGTTAAGGATCATACTAAACAAGATCTTCTGCTTCAGTGGGCC
TAAAAGCAGCCATCTGCATAGAAAGCGTTAAAGCTCAGGCAGTTTATAAATCTAT
TATTCTGATACCCCATCCCACACCCCTTCACTTACTAGACCGTTCCATGCCAACA
TGGAAGT