

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



DISSERTAÇÃO DE MESTRADO

**Isolamento e caracterização de marcadores  
moleculares para análise genética de  
população de *Prochilodus costatus* (Pisces,  
Characiformes) do rio Pará**

**ORIENTADA:** Tatiana Moura Barroca

**ORIENTADOR:** Prof. Dr. Evanguedes Kalapothakis

**CO-ORIENTADORA:** Dra. Cibeles Soares de Castro

**BELO HORIZONTE**

**Outubro - 2009**

**Tatiana Moura Barroca**

Isolamento e caracterização de marcadores  
moleculares para análise genética de população de  
*Prochilodus costatus* (Pisces, Characiformes) do rio  
Pará

Dissertação apresentada ao curso de mestrado em Genética do  
Instituto de Ciências Biológicas da Universidade Federal de Minas  
Gerais com requisito parcial para obtenção de título de Mestre em  
Genética

Orientador: Prof.Dr. Evanguedes Kalapothakis  
Co-orientadora: Dra. Cibeles Soares de Castro

Universidade Federal de Minas Gerais  
Instituto de Ciências Biológicas  
Departamento de Biologia Geral  
Belo Horizonte  
Outubro de 2009

À minha família, especialmente ao meu marido e filha,  
Cláudio e Gabriela.

## AGRADECIMENTOS

Ao meu querido orientador e professor Dr. Evanguedes Kalapothakis pela oportunidade, confiança e ensinamentos. Você é minha referência profissional.

À minha co-orientadora Dra. Cibele Castro pelas dicas e correções que foram de grande valia para a elaboração deste trabalho.

Ao Programa de Pós-Graduação em genética do Departamento de Biologia Geral pela presença de professores e funcionários que contribuíram de alguma forma para a minha formação.

Ao professor Dr. Jorge Dergam por ser tão prestativo e me ensinar a trabalhar com os programas de genética de populações e pela ajuda na interpretação dos resultados. Obrigada por estar sempre tão disponível.

Ao professor Dr. Gilmar Santos pelo auxílio na ecologia e zoologia do Curimatá-pioa, pelas lindas fotos e pela coleta de material. Obrigada por me ajudar a solucionar minhas dúvidas.

Ao Dr. Yoshimi Sato por contribuir com este trabalho nos fornecendo o histórico de repovoamento do rio Pará.

Ao Cacá e ao comitê de bacias hidrográficas pelas fotos.

Aos meus amigos do Laboratório de Biotecnologia e Marcadores Moleculares: Anderson, Arthur, Bárbara, Caio, Camila, Carol, Egiancarla, Érika, Isabella, Flávia, Flavinha, Kelly, Luiz, Marcelle, Tatiani, Valéria. De alguma forma, todos vocês contribuíram para a elaboração deste trabalho e tornam o nosso laboratório um ambiente agradável para trabalhar. E aos meus amigos tão especiais Carol, André, Thaís, Gabriel e Ana Luiza, que já não estão mais no laboratório mas também sempre me auxiliaram. Adoro todos vocês. Aos vizinhos de laboratório pela agradável convivência e em especial à Lu, aluna da professora Cleusa, por ajudar de maneira especial.

Aos meus amigos pelos momentos de descontração principalmente nestes meses tão cansativos.

À minha família. Meus pais e irmão pelo exemplo de vida, pelo amor, pelo suporte e por terem proporcionado uma educação de boa qualidade. Vocês são meus pilares. Aos meus sogros pela constante ajuda. Ao meu amado Claudio, pelo apoio incondicional sempre, por estar sempre presente na minha vida, pelo amor e cuidado. Sem você eu não conseguiria. E à minha razão de viver, minha querida Gabriela. Você é meu anjinho que transborda os meus dias de alegria. Amo muito todos vocês!

## SUMÁRIO

LISTA DE FIGURAS .....	vii
LISTA DE TABELAS.....	ix
LISTA DE ABREVIATURAS.....	x
RESUMO .....	12
ABSTRACT .....	13
1) INTRODUÇÃO: .....	14
1.1) Área de estudo.....	14
1.2) <i>Prochilodus costatus</i> .....	18
1.3) As hidrelétricas e o impacto na população de peixes .....	20
1.4) Marcadores Moleculares.....	24
1.5) Importância da conservação da variabilidade genética.....	26
1.6) Análise Genética de populações via sequenciamento automático de DNA ..	27
2) OBJETIVOS .....	29
2.1) Objetivo geral:.....	29
2.2) Objetivos específicos: .....	29
3) METODOLOGIA:.....	30
3.1) Material Coletado: .....	30
3.2) Biblioteca Genômica: .....	30
3.3) Transformação dos plasmídeos selecionados na triagem com sonda radioativa em bactérias <i>Escherichia coli</i> eletrocompetentes: .....	31
3.4) Extração de plasmídeos por lise alcalina: .....	32
3.5) Sequenciamento: .....	33
3.6) Análise do sequenciamento e desenho dos iniciadores: .....	34
3.7) Extração de DNA genômico:.....	35
3.8) Padronização de PCR: .....	35
3.9) Padronização de <i>Nested</i> PCR: .....	36
3.10) Subclonagem do produto de <i>Nested</i> PCR em plasmídeo PCR 2.1 TOPO <i>Invitrogen</i> :.....	36
3.11) Transformação dos plasmídeos PCR 2.1 TOPO ( <i>Invitrogen</i> ) em bactérias <i>Escherichia coli</i> eletrocompetentes:.....	37
3.12) Extração dos plasmídeos PCR 2.1 TOPO ( <i>Invitrogen</i> ) por lise alcalina: ...	38
3.13) Análise do sequenciamento:.....	38
3.14) Análise estatística dos dados:.....	38
4) RESULTADOS:.....	40
4.1) Busca por novos marcadores para <i>Prochilodus</i> : .....	40

4.1.1) Transformação dos plasmídeos previamente selecionados pelo laboratório de Biotecnologia e Marcadores Moleculares na triagem com sonda radioativa em bactérias <i>Escherichia coli</i> eletrocompetentes e extração por lise alcalina: .....	40
<b>4.2) Sequenciamento e análise de clones positivos:</b> .....	42
<b>4.3) Padronização de PCR:</b> .....	46
<b>4.4) Padronização da <i>Nested</i> PCR:</b> .....	48
<b>4.5) Subclonagem do produto da <i>Nested</i> PCR em plasmídeo pCR 2.1 TOPO <i>Invitrogen</i>:</b> .....	50
<b>4.6) Análises de sequenciamentos:</b> .....	52
<b>4.7) Análises Estatísticas:</b> .....	55
4.7.1) Índices de diversidade: .....	55
4.7.2) <i>Fst</i> e teste exato da cadeia de Markov: .....	55
4.7.3) AMOVA: .....	56
4.7.4) Análise de subestruturação populacional: .....	57
4.7.5) Árvores de Haplótipos: .....	59
4.7.6) Biometria e Fator de Condição: .....	61
<b>5) DISCUSSÃO E CONCLUSÃO:</b> .....	64
<b>6) REFERÊNCIAS BIBLIOGRÁFICAS:</b> .....	69
<b>ANEXOS</b> .....	80

## LISTA DE FIGURAS

Figura 1: Hidrografia regional da bacia do rio Pará. ....	16
Figura 2: Usina Hidrelétrica (UHE) de Gafanhoto .....	17
Figura 3: Usina Hidrelétrica (UHE) de Cajuru.....	17
Figura 4: Indivíduos de <i>Prochilodus costatus</i> - Curimatá-pioa.....	19
Figura 5: Migração.....	23
Figura 6: Mapa do vetor pUC 18.....	32
Figura 7: Mapa do vetor pCR 2.1Topo .....	37
Figura 8: Gel de agarose 0,8% mostrando plasmídeos contendo insertos de microsatélites de DNA de <i>Prochilodus</i> .....	42
Figura 9: Gráfico de proporção entre as seqüências de clones positivos do peixe <i>Prochilodus lineatus</i> .....	42
Figura 10: Marcadores moleculares encontrados no seqüenciamento de clones do peixe <i>Prochilodus lineatus</i> .....	44
Figura 11: Gel de agarose 1,5%. Exemplo de amplificação utilizando o clone 2V35 com variação de temperaturas de anelamento.....	47
Figura 12: Gel de agarose 1,5%. Exemplo de amplificação utilizando o clone 2V35 obtidos com tampão IVB e diferentes pares de iniciadores. ....	48
Figura 13: Gel de agarose 1,5% contendo o produto de amplificação da primeira etapa da <i>Nested</i> PCR .....	49
Figura 14: Gel de agarose 1,5% contendo o produto de amplificação da segunda etapa da <i>Nested</i> PCR .....	49
Figura 15: Amplificação da região do marcador molecular HPL de um exemplar de <i>P.</i> <i>costatus</i> feita por <i>Nested</i> PCR em gel de agarose 1,5%.....	50
Figura 16: Placa de Petri contendo meio de cultura LB-Agar com ampicilina mostrando a alta eficiência da subclonagem. ....	51
Figura 17: PCR de colônia do exemplar de <i>P. costatus</i> A em gel de agarose 1,5% .....	51
Figura 18: Exemplo de quatro eletroferogramas resultantes do seqüenciamento da região de microsatélite HPL do exemplar 1278 de <i>P. costatus</i> .....	53
Figura 19: Alinhamento dos 15 exemplares de <i>P. costatus</i> coletados à jusante da UHE de Gafanhoto .....	54
Figura 20: Alinhamento feito no Clustal W implementado no programa MEGA 4.0 (Tamura <i>et al.</i> 2007) mostrando um polimorfismo detectado no marcador molecular HPL .....	54
Figura 21: O exemplo extraído do manual do <i>Structure</i> (Pritchard <i>et al.</i> , 2007).....	58
Figura 22: Diagrama de agrupamento dos genótipos dos 45 exemplares de <i>P. costatus</i> coletados em três diferentes localidades do rio Pará .....	58
Figura 23: Árvore de haplótipos evidenciando as conexões entre os 25 haplótipos evidenciados em <i>P. costatus</i> .....	60

Figura 24: Comprimento padrão das amostras coletadas à jusante da UHE de Gafanhoto, no trecho entre as UHE de Gafanhoto e Cajurú e na UHE de Cajurú .....	61
Figura 25: Peso corporal das amostras coletadas à jusante da UHE de Gafanhoto, no trecho entre as UHE de Gafanhoto e Cajurú e na UHE de Cajurú .....	62
Figura 26: Fator de condição (K) das amostras coletadas à jusante da UHE de Gafanhoto, no trecho entre as UHE de Gafanhoto e Cajurú e na UHE de Cajurú .....	62



**LISTA DE TABELAS**

Tabela 1: Clones de <i>Prochilodus</i> utilizados no trabalho .....	41
Tabela 2: Seqüenciamentos realizados na identificação dos marcadores moleculares...	45
Tabela 3: Seqüência dos iniciadores específicos para cada marcador .....	46
Tabela 4: Índices de diversidade de <i>Prochilodus costatus</i> observados para as três amostragens .....	55
Tabela 5: Matriz com os valores de $F_{st}$ (acima) e o valor de probabilidade P (abaixo) das amostragens analisadas. ....	56
Tabela 6: Teste global de não diferenciação da cadeia de Markov .....	56
Tabela 7: Resultados do programa AMOVA .....	57
Tabela 8: Resultado do programa <i>GraphPad Instat</i> : ANOVA .....	63

## LISTA DE ABREVIATURAS

**AFLP** – *Amplified Fragment Length Polymorphism*, polimorfismo do comprimento de fragmentos amplificados

**AL** – Alagoas

**AMOVA** – *Analysis of Molecular Variance*, Análise de Variância Molecular

**ANOVA** – *Analysis of Variance*, Análise de Variância

**BH** – Belo Horizonte

**CBH - PARÁ** – Comitê da Bacia Hidrográfica do rio Pará

**CEMIG** – Companhia Energética de Minas Gerais

**CODEVASF** – Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba

**CP** – Comprimento Padrão

**DNA** – Ácido Desoxirribonucleico

**DNTP** – Desoxinucleotídeo Tri-fosfato

**EDTA** – Ácido etilenodiamino tetra-acético

**F** – Índice de Fixação

**FAO/DVWK** – *Food and Agriculture Organization/German Association for Water Management and Improvement*, Organização para a Alimentação e Agricultura/Associação Alemã para Gestão da Água

**F<sub>ST</sub>** – Índice de Fixação

**g** – Gravidade

**gl** – Grau de Liberdade

**H** – Diversidade Gênica

**HCl** – Ácido Clorídrico

**HPL** – *Highly Polymorphic Locus*, Locus Hipervariável

**INDEL** – Inserção/Deleção

**IQA** – Índice de Qualidade da Água

**K** – *Clusters*

**K** – Fator de Condição

**Kb** – Kilobases

**Km** – Kilometros

**kVa** – Quilovolt-amperes

**LB** – Luria-Bertani

**MG** – Minas Gerais

**mg** – Miligrama

**MgCl<sub>2</sub>** – Cloreto de Magnésio

**MgSO<sub>4</sub>** – Sulfato de Magnésio

**min** – Minutos

**mL** – Mililitros

**mM** – Milimolar

**NaCl** – Cloreto de Sódio

**NaOH** – Hidróxido de Sódio

**NCBI** – *National Center for Biotechnology Information*, Centro Nacional para Informação em Biotecnologia

**P.A.** – Pró-análise

**pb** – Pares de Bases

**PC** – Peso Corporal

**PCR** – *Polymerase chain reaction*. Reação em cadeia da polimerase

**pH** – Potencial Hidrogeniônico

**pmol** – picomol

**RAPD**– *Random amplification of polymorphic DNA*, amplificação aleatória de DNA polimórfico

**RFLP**– *Restriction Fragment Length Polymorphism*, polimorfismo de comprimento de fragmentos de restrição

**RNA** – Ácido Ribonucleico

**SDS** – Dodecil Sulfato de Sódio

**seg** – Segundos

**T<sub>m</sub>** – Temperatura Média

**UFMG** – Universidade Federal de Minas Gerais

**UHE** – Usina Hidrelétrica

**°C** – Graus Celsius

**µg** – Micrograma

**µL** – Microlitros

**π** – Diversidade Nucleotídica

**%** – Porcento

## RESUMO

Nas últimas décadas, muitos rios vêm sendo represados para a construção de usinas hidrelétricas. As represas causam impacto negativo significativo sobre as populações de peixes locais. Diante desse contexto, é de grande relevância a realização de estudos populacionais de espécies cujos habitats sofreram fragmentação devido à construção de usinas hidrelétricas. Os marcadores moleculares vêm sendo utilizados neste tipo de estudo pois apresentam alto grau de polimorfismo, permitindo avaliar a estrutura e variabilidade genética de populações de peixes bem como o impacto que as barragens e o repovoamento causam sobre elas. O objetivo do presente trabalho foi o isolar marcadores moleculares de *Prochilodus costatus*, e caracterizar um marcador para análises populacionais dessa espécie na bacia do rio Pará. O isolamento dos marcadores foi feito a partir do seqüenciamento dos clones positivos encontrados na varredura da biblioteca genômica de *Prochilodus lineatus* existente no laboratório e confirmação de regiões microssatélites. Foram seqüenciados 71 clones, sendo que desses, 19 (26%) apresentaram pelo menos uma seqüência repetitiva. Foram determinadas as melhores condições de amplificação para um *locus*, posteriormente chamado de HPL ( *Highly Polymorphic Locus*), que foi testado em 45 indivíduos coletados em três diferentes regiões do rio Pará. Sua utilidade em estudos de genética de populações foi verificada através de análises de índices de diversidade, além da análise de partição da variabilidade genética (AMOVA) das amostras coletadas ao longo do rio, análise de agrupamento de genótipos para verificar a existência de subestrutura na amostragem. Uma árvore haplotípica foi construída utilizando-se o programa TCS. Também foram realizados testes estatísticos submentendo dados biométricos das amostras à análise de variância (ANOVA). O seqüenciamento do marcador HPL permitiu que o caracterizássemos como uma região polimórfica, sendo constatados 25 haplótipos diferentes. A utilização desse marcador como ferramenta molecular irá contribuir de forma determinante para o desenvolvimento de sistemas mais adequados para a manutenção da diversidade das comunidades de peixes além de acelerar e aumentar a exatidão das decisões tomadas para a conservação e manejo das espécies em questão.

## ABSTRACT

In the last decades, many rivers have been dammed for the construction of reservoirs. Dams cause significant adverse impact on the population of local fish. Thus, it is of huge relevance the accomplishment of populational studies in species whose habitat has been fragmented as a result of dam building. Because of their high level of polymorphism, molecular markers have been largely used in populational analysis of many species of fish, in order to evaluate the genetic structure and variability of populations and the impact of the dams and the hatchery fish have on wild populations. The objective of this study was to isolate molecular markers of *Prochilodus costatus* and to characterize its populations in the Pará River. Isolation of the molecular markers was achieved by sequencing positive clones from a *Prochilodus lineatus* genomic library. Of 71 analyzed clones, 19 (26%) contained at least one repetitive sequence. Optimized reactions were carried out for a locus called HPL (Highly Polymorphic Locus), which was tested in 45 individuals collected in three different regions of the Pará River. Its usefulness in population genetic studies was evaluated by analyzing indices of molecular diversity, molecular variance (AMOVA), a test on the structure of the samples, and a haplotypic tree constructed using the TCS program. Biometric data was also subjected to statistical using ANOVA. Sequencing of the marker HPL revealed 25 haplotypes.

## 1) INTRODUÇÃO:

### 1.1) Área de estudo:

A Bacia do São Francisco é a terceira maior do país, cobrindo uma área de 631.133Km<sup>2</sup>, ou seja, 7,5% do território nacional. Com nascentes na Serra da Canastra, na parte sul do estado de Minas Gerais estende-se para o leste correndo 400 quilômetros entre os estados da Bahia, Pernambuco, Sergipe e Alagoas para desaguar no oceano Atlântico (Sato e Godinho, 2004).

O São Francisco é um rio federal, o maior totalmente nacional, que percorre uma extensão de 2.700 km. Ele deságua no Oceano Atlântico, com a foz entre os estados de Sergipe e Alagoas, próxima a Piaçabuçu (AL). O São Francisco tem 36 afluentes de porte significativo, sendo 19 perenes. Os principais são, pela margem esquerda, os rios Abaeté, Paracatu, Urucuia, Carinhanha, Corrente e Grande, e, pela margem direita, os rios Paraopeba, das Velhas, Jequitaí, Pacuí, Verde Grande e Pará (Portal Peixe Vivo. Rios de Minas. Disponível *online* em: <<http://www.portalpeixevivo.com.br>>, acesso: Janeiro de 2009).

A Bacia do rio Pará (Figura 1) está situada no estado de Minas Gerais e compreende uma área de aproximadamente 12.500 km<sup>2</sup>. Ela alonga-se na direção norte-sul, tendo sua origem na região serrana de Entre Rios, Piracema e Desterro de Entre Rios e sua foz no Rio São Francisco, à montante do reservatório de Três Marias, no município de Pompeu (Pellegrin, 2001). É uma das bacias mais importantes do sistema do Rio São Francisco, abrangendo 38 municípios nos quais vivem cerca de 650 mil habitantes.

O rio Pará nasce no município de Resende Costa e compreende uma área de drenagem de 12.262 km<sup>2</sup>, atingindo uma extensão de 310,6 km. Seus principais afluentes são: os rios Peixe, Itapecerica, São João, Lambari e do Picão, antes de desembocar próximo à represa de Três Marias no município de Pompéu, MG (Instituto Mineiro de Gestão das águas. Disponível em: <<http://www.igam.mg.gov.br>>. Acesso em: Janeiro de 2009).

Por não cortar nenhuma cidade em seu percurso regional, o rio Pará é um dos poucos cuja água ainda apresenta um Índice de Qualidade das Águas (IQA) satisfatória, classe 1 e 2, segundo Deliberação Normativa COPAM nº 010/86 (CBH-PARÁ. Projeto

Água e Vida. Comitê da Bacia Hidrográfica do Rio Pará. Disponível em: <<http://www.cbhpara.org.br/projetos.htm>>. Acesso em: janeiro 2009).

Em seu curso estão instaladas as Usinas Hidrelétricas de Gafanhoto (Figura 2), localizada no município de Divinópolis (Minas Gerais), cuja operação iniciou-se em 1946, tornando-a a usina mais antiga da CEMIG; e a de Cajuru (Figura 3), localizada nos municípios de Divinópolis, Carmo do Cajuru e Cláudio (Minas Gerais), cujo início de operação foi em 1959 (Portal Peixe Vivo. Rios de Minas. Disponível *online* em: <<http://www.portalpeixe vivo.com.br>>, acesso: Janeiro de 2009).

## BACIA HIDROGRÁFICA DO RIO PARÁ

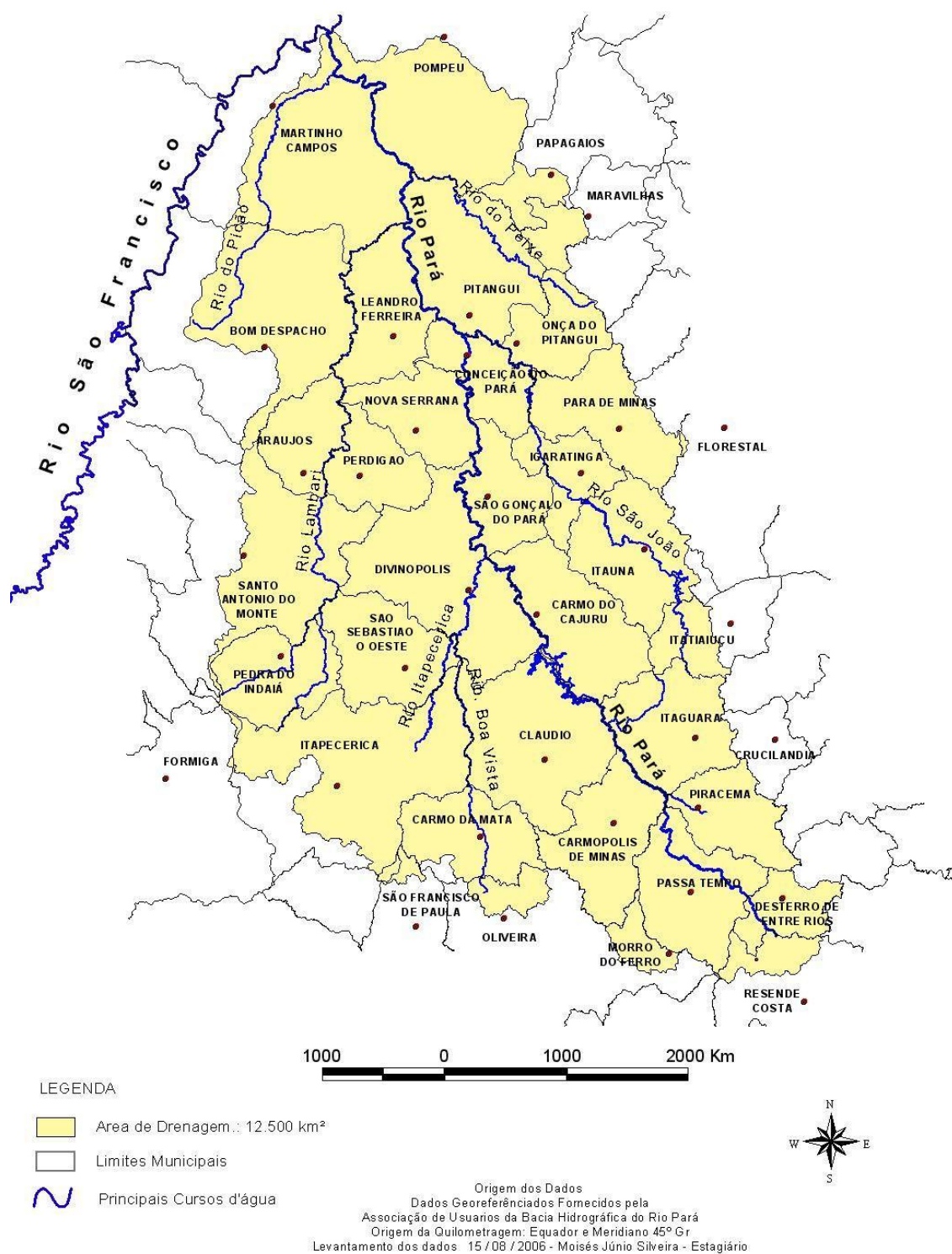


Figura 1: Hidrografia regional da bacia do rio Pará. A malha hidrográfica está representada em azul escuro. Escala em quilômetros. Fonte: Comitê da Bacia Hidrográfica do rio Pará.





Figura 2: Usina Hidrelétrica (UHE) de Gafanhoto, Divinópolis, Minas Gerais. Instalada no rio Pará com início de operação em 1946. Tem potência de 14,00MW e volume útil de 3,95hm<sup>3</sup>.



Figura 3: Usina Hidrelétrica (UHE) de Cajuru, Carmo do Cajuru, Minas Gerais. Instalada no rio Pará com início de operação em 1959. Tem potência de 7,2MW e volume útil de 132,00hm<sup>3</sup>.

## 1.2) *Prochilodus costatus*

O gênero *Prochilodus* possui grande relevância, pois está entre os peixes de água doce mais abundantes e amplamente distribuídos da América do sul. É constituído de espécies detritívoras e iliófagas, ou seja, que se alimentam de lodo e detritos encontrados no fundo lamacento de lagos e rios. Esses peixes são em geral grandes migradores e parte importante do volume de pesca continental no Brasil (Castro e Vari, 2004; Sivasundar *et al.*, 2000; Petrere, 1989; Welcome, 1979). Relatos destas longas migrações, até 1200 km por ano, já haviam sido feitos por Bonetto, 1963. Durante o ciclo anual, eles nadam rio abaixo, para as planícies alagadas onde se alimentam ou regiões mais baixas para a desova, perto das cabeceiras (Sivasundar *et al.*, 2000).

*Prochilodus costatus* (Figura 4) pertence à família Prochilodontidae, ordem Characiforme, é uma espécie inicialmente endêmica da Bacia do rio São Francisco mas que foi posteriormente introduzida no rio Jequitinhonha (Castro e Vari, 2003). É conhecido comumente como Curimatá-pioa (*FishBase - a global information system of fish*. Disponível online em: <<http://www.fishbase.org>>. Acesso: fevereiro de 2009).

Existem espécies de peixes que realizam grandes migrações para fins alimentícios ou reprodutivos em quase todos os rios da América Latina (Agostinho, 2003). O *Prochilodus costatus* é um peixe de piracema, ou seja, os adultos realizam deslocamentos em cardumes, chamadas de migrações reprodutivas, rio acima na estação chuvosa, que vai de novembro a março, e os ovos e larvas derivam passivamente junto com a correnteza, por dezenas de quilômetros rio abaixo. A espécie possui um importante papel ecológico considerando seus hábitos bentopelágicos e detritívoros, ou seja, ela vive e se alimenta nas camadas médias e mais profundas de água, se alimentando de detritos orgânicos, desenvolvendo um papel importantíssimo na ciclagem de matéria orgânica em ecossistemas límnicos (Flecker, 1996; Darnell, 1964), sendo uma espécie chave para programas de conservação.

Existe um grande interesse comercial no Curimatá-pioa por servir de alimento para a população de baixa renda, pela sua abundância e por suas ótimas características para a piscicultura como, por exemplo, grande prolificidade, ou seja, uma única fêmea pode desovar mais de um milhão de óvulos por temporada, desenvolve-se com relativa facilidade, sua reprodução em cativeiro já é bem controlada por profissionais especializados e seu alimento é barato (FAO/DVWK, 2002).



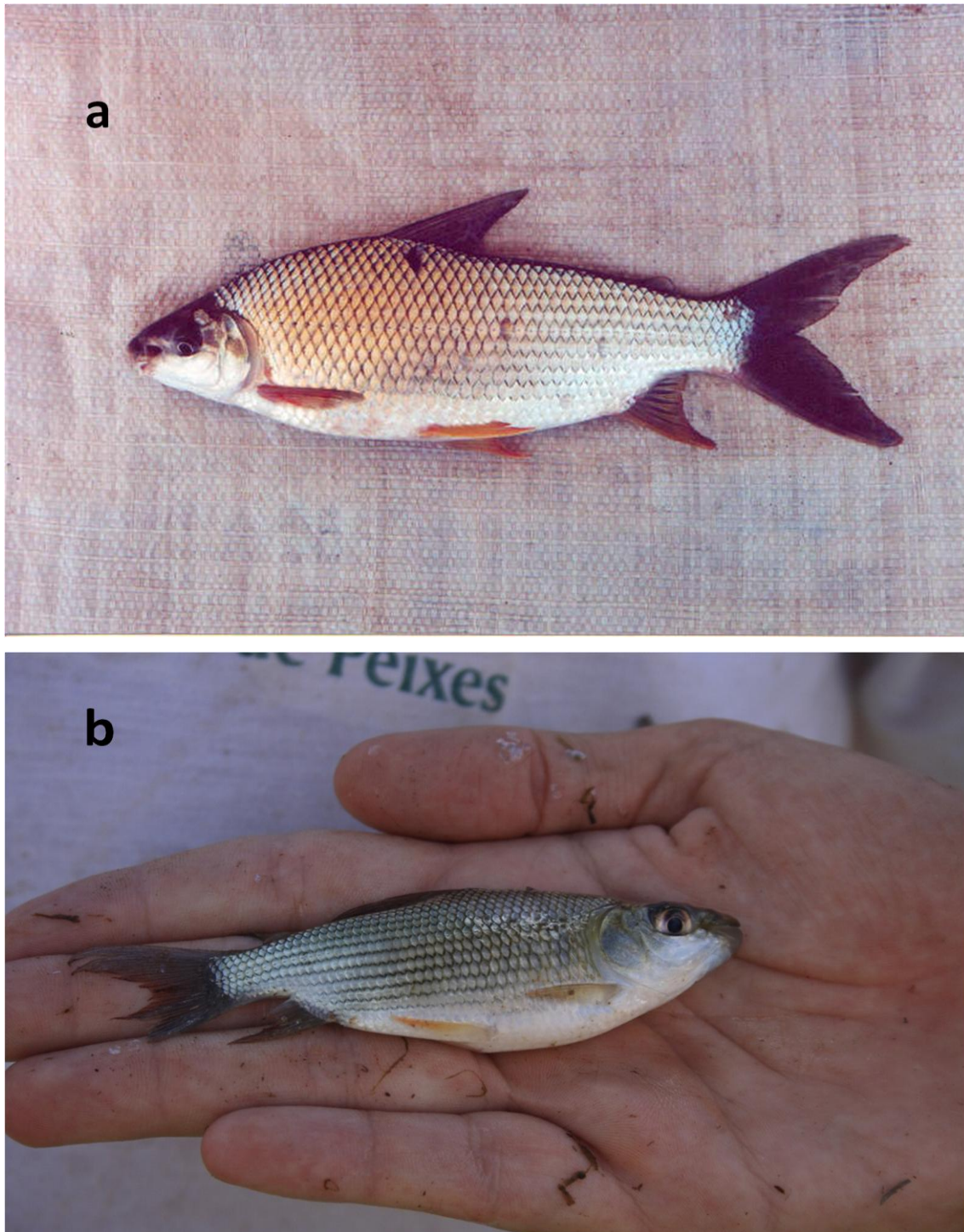


Figura 4: Indivíduos de *Prochilodus costatus* - Curimatá-pioa: a) indivíduo adulto. Foto: Carlos Bernardo Mascarenhas; b) jovem capturado em lagoa marginal do rio Itapecirica, situada a montante da cidade de Divinópolis (MG). Foto: Tiago Cerasim Pessali.

### 1.3) As hidrelétricas e o impacto na população de peixes

Os rios e suas planícies de inundação são considerados pelos hidrologistas como um sistema inseparável, com relação ao balanço de água, sedimento e matéria orgânica. Nestes sistemas, estão incluídos habitats lóticos, habitats permanentemente lênticos e a planície alagável; cada um possuindo propriedades únicas, caracterizando-se em um ecossistema específico (Junk *et al.*, 1989). Dentre esses ambientes, estão as lagoas marginais, que ocupam as depressões laterais do rio e apresentam comunicação constante ou intermitente com o mesmo. Durante as cheias as lagoas formam áreas inundadas e canais que enriquecem o habitat, pois são abastecidas com nutrientes e matéria orgânica vindas do rio. Isso permite que muitas espécies de peixes migratórios as utilizem como locais de acasalamento e desova, berçário, proteção e abrigo e área de descanso para adultos ou crescimento de juvenis (Vazzoler *et al.*, 1997; Agostinho *et al.*, 1993; Medeiros e Maltchik, 2001; Welcomme, 1979).

Nas últimas décadas, muitos rios vêm sendo represados no mundo inteiro para a construção de usinas hidrelétricas, devido à crescente demanda de energia elétrica, consequência do aumento da população mundial e do desenvolvimento industrial.

O Brasil contém cerca de 15% de toda a água doce do planeta, sendo dotado de uma vasta e densa rede hidrográfica. Seus rios possuem um alto potencial para a geração de energia elétrica, devido a suas características como grandes extensões, largas e profundidades, além de possuírem em seu leito rupturas de declive e vales encaixados, já que são em sua grande maioria rios de planalto. Cerca de 90% da energia consumida no país provém de Usinas Hidrelétricas e por esse motivo, muitos rios brasileiros têm sido transformados numa sucessão de barragens, provocando uma fragmentação de habitats, inundações de grandes áreas, alterações no regime hidrológico natural, na qualidade da água e na reprodução da ictiofauna local (Martins e Tamada, 2000; Petts, 1989).

As represas causam impacto negativo significativo sobre as populações de peixes locais. Nos primeiros anos após o barramento, observa-se uma tendência no aumento da produtividade de peixes, sendo que seu nível de equilíbrio da produtividade é igual ou até mesmo superior ao encontrado anteriormente (Kimmel e Groeger, 1986; Welcomme, 1989). Com o passar dos anos, ocorre uma depressão trófica dos reservatórios. Muitas vezes é observado um declínio de algumas populações de peixes devido a alterações nas características físico-químicas da água e na oferta alimentar para

os mesmos, que juntamente com o isolamento dos mesmos de seus sítios de desova (que interrompe seu processo reprodutivo), levam a uma diminuição na produtividade pesqueira. Esta ruptura pode levar à extinção local se sítios críticos forem perdidos ou não mais acessíveis (Nerras *et al.*, 2001; Godinho e Godinho 1994; Welcomme, 1989).

A migração desempenha um papel fundamental no sucesso reprodutivo de peixes reofílicos (Figura 5), pois além de ser um estímulo para o desenvolvimento das gônadas, também promove o encontro e uma alta concentração de ambos os sexos em local favorável à fertilização e ao desenvolvimento dos filhotes devido à alta oxigenação; e com pequena taxa de predação já que ocorre uma baixa na transparência da água. O ciclo biológico dessas espécies está diretamente relacionado aos ciclos hidrológicos dos rios, ou seja, os ciclos de cheias e secas, que também são alterados com a construção de barragens. Se lagoas marginais, antes alagáveis, deixam de receber água, há um comprometimento no seu papel de berçário de jovens peixes migradores. Assim, as espécies reofílicas são as mais afetadas com a transformação súbita do ambiente lótico para lântico, pois as adaptações e estratégias reprodutivas estavam associadas ao primeiro (Carosfeld *et al.*, 2003; Braga, 2001).

O declínio e a extinção de algumas espécies de peixes têm chamado a atenção de agências governamentais e especialistas, no intuito de restaurar ambientes degradados, conduzindo pesquisas e programas de desenvolvimento com o objetivo de reintroduzir e propagar espécies em questão.

O repovoamento, que consiste na reprodução de peixes em cativeiro e reintrodução nas represas, é uma das estratégias mais utilizadas para mitigar esse impacto (Agostinho *et al.*, 2007). No Brasil, a estratégia de repovoamento vem sendo empregada desde as últimas décadas na tentativa de restaurar as populações de peixes, mantendo-as em equilíbrio, embora envolva riscos relativos à eficiência do programa quanto aos seus resultados, dúvidas relativas à preservação do *pool* gênico e à possibilidade de introdução de doenças, além de outros aspectos ecológicos e econômicos (Hickley, 1994).

A reprodução induzida, em laboratório, das espécies cultivadas em tanques ou capturadas próximo às barragens, na época da piracema, está sendo amplamente utilizada, visando à conservação da biodiversidade e à manutenção da pesca (CEMIG- Companhia Energética de Minas Gerais. Sustentabilidade. Disponível *online* em <<http://www.cemig.com.br>>. Acesso: fevereiro de 2009). Entretanto, a influência da mesma sobre populações de peixes é pouco conhecida quando utilizada isoladamente

(Madeira *et al.*, 2005; Salvanes e Braithwait, 2006, Fraser, 2008). Quando realizados de forma não sistemática e planejada, os peixamentos podem ter efeitos adversos nas populações naturais (Miller e Kapuscinski, 2003). O manejo racional de espécies de peixes deve prevenir a perda da diversidade genética. Para isso, é necessário o conhecimento científico da composição genética dos estoques e das populações naturais para permitir o embasamento objetivo das decisões a serem tomadas.

Houve um aumento no impacto da genética de populações na conservação da biodiversidade e concomitantemente o desenvolvimento de diversas técnicas em biologia molecular. Assim, análises populacionais que utilizam marcadores moleculares como os microssatélites podem orientar medidas de manejo e conservação de espécies de peixes (Araki *et al.*, 2007).



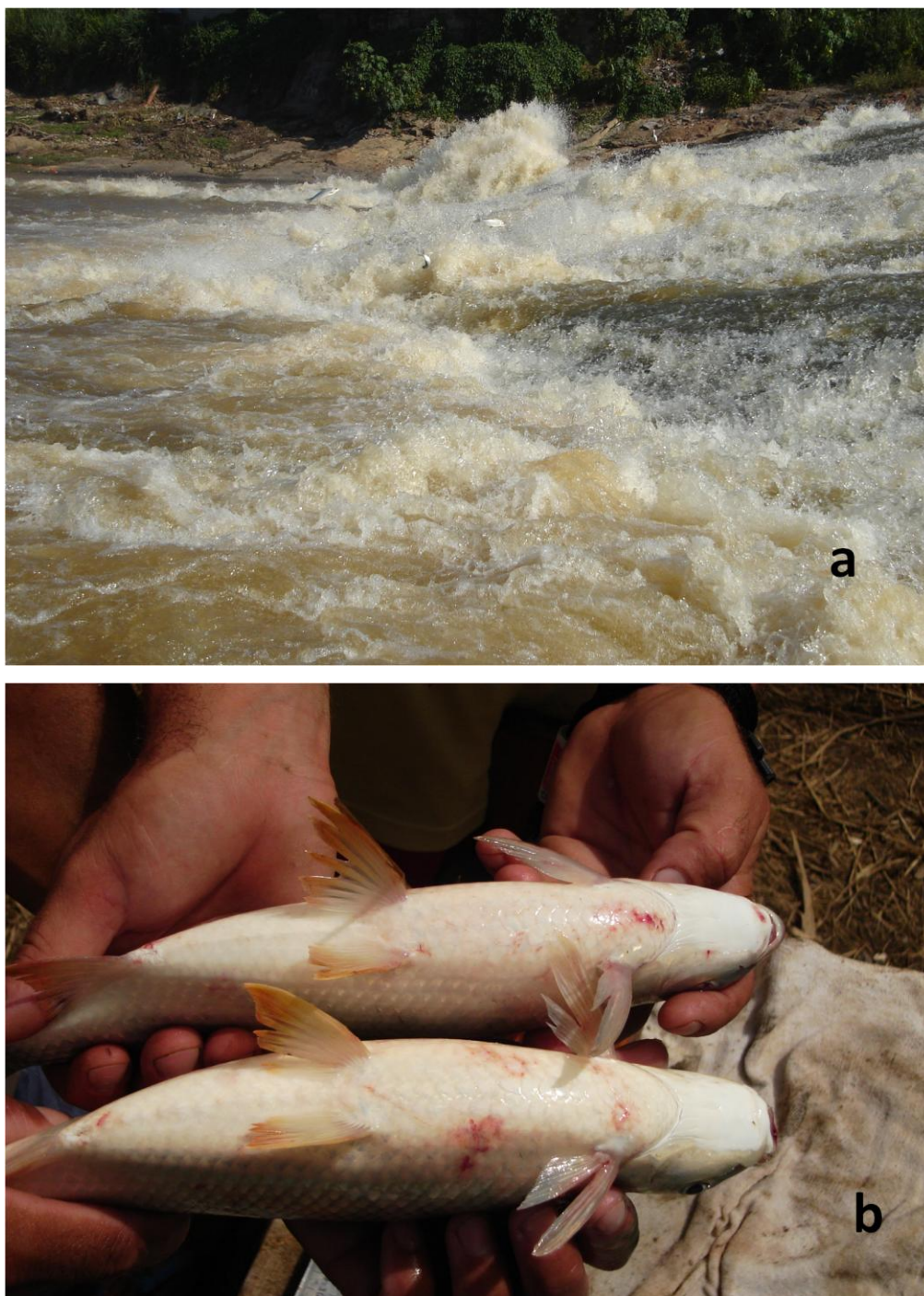


Figura 5: Migração: a) migração ascendente de Curimbas no rio Itapeçirica, área metropolitana de Divinópolis, MG, durante a piracema de 2006/2007; b) fêmeas de Curimatá-pioa capturadas no rio Itapeçirica, região metropolitana de Divinópolis, MG, durante a cheia de 2006/2007, mostrando o ventre volumoso devido ao desenvolvimento dos ovários. Fotos: Tiago Cerasim Pessali.

#### 1.4) Marcadores Moleculares

As ferramentas moleculares são os mais modernos e efetivos recursos que têm possibilitado análises genéticas em larga escala, se tornando de uso rotineiro em muitos laboratórios do mundo (Poke *et al.*, 2005).

Os marcadores moleculares são regiões do DNA que contém polimorfismos (de sequência ou de comprimento) e que apresentam grande variação entre os indivíduos, sendo assim, facilmente utilizadas em estudos e análises de diferenciação entre os mesmos. De acordo com Ferreira *et al.*, 1998, ao se verificar o padrão de herança do marcador, de acordo com as leis de herança mendeliana, um determinado marcador é definido como marcador genético.

Os marcadores moleculares genéticos estão sendo amplamente utilizados entre os biólogos e a escolha do marcador a ser utilizado depende da natureza da informação (dominância, co-dominância, ploidia), dos equipamentos disponíveis em laboratórios e do nível de variabilidade do marcador, ou seja, se ele é muito ou pouco polimórfico (Anne, 2006). Este último fator é muito importante, já que existem marcadores que evoluem rapidamente, ou seja, apresentam uma taxa mutacional bastante alta, como por exemplo, alguns tipos de marcadores nucleares. Esse tipo de marcador é útil para o estudo de indivíduos, famílias e populações, podendo-se comparar grupos com divergências recentes. Por outro lado, existem marcadores que evoluem de forma mais lenta e são fixados rapidamente em populações, como por exemplo, os marcadores mitocondriais os quais são freqüentemente utilizados em estudos evolutivos (Avisé *et al.*, 1987).

A grande vantagem do uso de marcadores nucleares que se encontram em regiões não codificantes, é seu alto grau de polimorfismo, ou seja, os *loci* são geralmente caracterizados por uma alta heterozigosidade (Ellegren, 2004, Schlötterer, 2000, Zane *et al.*, 2002). Praticamente todas as mutações irão contribuir para o polimorfismo (a não ser que o *locus* esteja próximo a uma região codificante), enquanto que a maioria das mutações é eliminada por seleção em marcadores que estão em regiões codificantes (Chenuil, 2006).

Hoje, estão disponíveis várias técnicas de biologia molecular para a detecção de variabilidade genética ao nível de sequência de DNA, como por exemplo, RFLP, AFLP, RAPD, minissatélites, microsatélites, dentre outros (Parker *et al.*, 1998) e com o desenvolvimento das técnicas de biologia molecular, novos marcadores vêm sendo



localizados em regiões do genoma nuclear e mitocondrial. Esses marcadores são aplicados a um grande número de questões biológicas como: reconstruções filogenéticas, testes de paternidade, aplicações na área forense, mapeamento e estudos genéticos de populações. Dentro da genética da conservação, eles têm sido úteis no planejamento e monitoramento de programas de reprodução em cativeiro, no entendimento de sistemas de acasalamento, na detecção da estrutura geográfica da variabilidade genética, no manejo de fluxo gênico e no entendimento de fatores que contribuem para o aumento ou declínio do *fitness* (Vrijenhoek, 1998).

Desde a última década, os microssatélites tornaram-se um dos marcadores moleculares mais utilizados em análise populacional (Schlötterer, 2004). Os marcadores microssatélites caracterizam-se por blocos repetitivos de seqüências idênticas ou similares dispostas em *tandem*, encontrados em todos os genomas eucariotos estudados até o momento (Zane *et al.*, 2002). Tais marcadores moleculares são constituídos por seqüências nucleotídicas de 2 a 6 pb repetidas muitas vezes, tendo uma disposição estabelecida de modo bem aleatório (Hearne *et al.*, 1992). A variação entre diferentes alelos nesses marcadores se dá pela diferença de comprimento de segmentos de DNA, em função de diferentes números de temas repetidos lado a lado. As repetições podem ser perfeitas ou imperfeitas; simples ou compostas. A distribuição e o tamanho de *loci* de microssatélites são altamente variáveis entre populações e até mesmo entre indivíduos (Ellegren, 2004).

Os marcadores moleculares são ferramentas importantes no estudo e avaliação genética de populações de peixes selvagens e em cativeiro (Hiltsdorf e Krieger, 1998). No Brasil, estudos genéticos que utilizam essas ferramentas para o estudo populacional de peixes vêm crescendo (Dergam *et al.*, 2002; Yazbeck e Kalapothakis, 2007; Paiva e Kalapothakis, 2008). O isolamento e caracterização de marcadores moleculares constituem etapas fundamentais, pois permitem avaliar a estrutura e a variabilidade genética de populações de peixes nativos e de cativeiro, bem como o impacto que as barragens e o repovoamento causam sobre elas (Yazbeck e Kalapothakis 2007; Araki *et al.*, 2007; Paiva e Kalapothakis 2008; Barbosa *et al.*, 2008; Costa *et al.*, 2008) e propor medidas mais indicadas para a manutenção da diversidade das comunidades de peixes além de acelerar e aumentar a exatidão das decisões tomadas para a conservação e manejo das espécies em questão.

### 1.5) Importância da conservação da variabilidade genética

A diversidade biológica do planeta está sendo rapidamente esgotada devido às consequências diretas e indiretas das atividades humanas como, por exemplo, a destruição, exploração e fragmentação de habitats reduzindo o fluxo gênico, e a movimentação de espécies para novas localidades. A preocupação genética dentro da conservação surge desses eventos estocásticos que podem levar à perda da diversidade genética, reduzindo as habilidades das populações de se adaptarem às mudanças ambientais (Frankham *et al*, 2004).

A variação genética dentro da população consiste da variabilidade de alelos presentes e de suas frequências em todos os membros da população. Essa diversidade pode ser influenciada pela seleção natural, deriva genética e fluxo gênico (Suzuki *et al.*, 1998).

A variação genética entre populações se dá pelas diferenças encontradas entre as frequências gênicas e pela presença de determinados alelos existentes em populações separadas geograficamente. Em geral, essas populações apresentam adaptações locais para sua sobrevivência, tornando-se mais especializadas (Templeton, 1980).

A heterozigose é um fator importante na variabilidade genética das espécies. Um indivíduo é considerado heterozigoto quando possui dois alelos diferentes para o mesmo *locus*. A heterozigose ou elevada variação genética entre indivíduos de uma mesma espécie está diretamente relacionada com o *fitness* ou valor adaptativo de uma determinada população (Araki *et al.*, 2007; Carroll *et al.*, 1997; Hartl *et al.*, 1989).

A diversidade genética dentro e entre populações exerce um importante papel na sobrevivência das mesmas através dos tempos. Populações geralmente são adaptadas localmente, devido às interações durante a evolução do seu genoma com os fatores bióticos e abióticos presentes no ambiente em que vivem e devido também, a fatores seletivos do local (Frankham *et al*, 2004; Templeton, 1979). A perda da diversidade genética ou a introdução de novos genes a populações altamente especializadas podem levar a uma perda da diversidade genética, reduzindo futuras opções evolutivas (Dobzhansky, 1955).

### 1.6) Análise Genética de populações via seqüenciamento automático de DNA

A análise genética de populações tomou uma importância muito grande nas últimas décadas. Uma das razões para isso é podermos analisar as diferenças genéticas através do seqüenciamento automático de DNA, ou seja, não é mais necessário focalizar em organismos que sofreram mutações e manifestam suas diferenças visivelmente, através de seu fenótipo.

O isolamento das primeiras enzimas de restrição, ainda no início dos anos 70 e o seqüenciamento do genoma completo do bacteriófago MS2, foram tidos como um marco da engenharia genética (Sanger e Coulson, 1975; Fiers *et al.*, 1976). O seqüenciamento automático é hoje um dos grandes avanços no conhecimento da genealogia dos alelos quando comparado com sistemas baseados na modificação química do DNA, como o proposto na década de 70 por Allan Maxam e Walter Gilbert e até mesmo ao sistema de terminação de cadeia proposto por Frederick Sanger na mesma época, utilizando métodos enzimáticos (Braslavsky *et al.*, 2003; Sanger e Coulson, 1977; Gilbert e Maxam, 1973).

A partir da década de 90, o seqüenciamento se tornou totalmente automatizado e hoje é possível obter bons resultados com praticamente qualquer tipo de material: RNA, DNA genômico, plasmídeos, cromossomos artificiais, produtos da reação em cadeia da polimerase (PCR); com as vantagens de um método rápido, prático e confiável devido à precisão da técnica (Braslavsky *et al.*, 2003; Shendure, 2005; Hall, 2007).

O seqüenciamento automático possui hoje várias aplicações como, por exemplo, o estudo sistemático do genoma, a biologia forense, a análise filogenética e o estudo genético populacional.

Os marcadores moleculares que não se encontram em regiões codificantes, podem acumular mutações, o que irá gerar um grande polimorfismo do *locus*. O desenvolvimento de técnicas que permitem o seqüenciamento de produtos de PCR que amplificam esses marcadores simplificou a geração de dados para análise da população (Hillis, 1996).

No caso do rio Pará, com a construção da Usina Hidrelétrica de Gafanhoto na década de 40 e da Usina Hidrelétrica de Cajuru no final da década de 50, parte do rio ficou fragmentada, havendo um bloqueio do movimento migratório de algumas espécies de peixes, inclusive da Curimatá-pioa. A condição de fragmentação pode levar à perda

da variabilidade genética e à diferenciação entre as populações devido à subdivisão populacional através da interrupção do fluxo gênico entre grupos (Frankham, 2003). Os trechos de rio entre as barragens de Cajuru e Gafanhoto e à montante de Gafanhoto não permitem a reprodução da população de *P. costatus* (Sato, com pess.), determinando uma condição de perda da população na ausência de medidas mitigadoras. Nestas condições, mesmo a realização de peixamentos (*i.e.* introdução de peixes reproduzidos em cativeiro nos reservatórios) podem ter efeitos adversos nas populações naturais quando realizados de forma não sistemática e planejada (Utter, 2003).

## 2) OBJETIVOS

### 2.1) Objetivo geral:

Esse projeto tem como objetivo isolar e caracterizar marcadores moleculares apropriados para estudos genéticos de populações na espécie *Prochilodus costatus*, peixe popularmente conhecido como Curimatá-pioa.

### 2.2) Objetivos específicos:

- Realizar o sequenciamento de DNA dos clones positivos encontrados na varredura da biblioteca genômica do peixe *Prochilodus lineatus*;
- Padronizar a reação de PCR para um marcador escolhido;
- Utilizar o marcador desenvolvido do peixe *P. lineatus* na análise genético-populacional de amostras de *Prochilodus costatus* coletadas nos diferentes pontos geográficos do rio Pará, utilizando o método de sequenciamento;
- Analisar estatisticamente os dados obtidos;
- Realizar uma análise preliminar das populações isoladas pelas barragens de Cajuru e Gafanhoto;
- Tentar correlacionar os dados moleculares com os dados fenotípicos;
- Sugerir novos rumos de pesquisa para teste de hipóteses de alguns dos padrões observados.

### **3) METODOLOGIA:**

#### **3.1) Material Coletado:**

Foram coletados tecidos de nadadeiras de exemplares da espécie *Prochilodus costatus* e o respectivo material fixado em Etanol 70% (P.A.). A coleta foi realizada pela equipe de biólogos do Professor Doutor Gilmar Bastos Santos, da Pontifícia Universidade Católica de Minas Gerais, através de projetos colaborativos.

Os animais foram coletados entre os municípios de Carmo do Cajurú e Divinópolis, MG, à jusante da Usina Hidrelétrica de Gafanhoto, no trecho entre as Usinas Hidrelétricas de Gafanhoto e Cajuru e na Barragem de Cajurú, rio Pará.

#### **3.2) Biblioteca Genômica:**

A biblioteca de DNA genômico do peixe *Prochilodus lineatus*, construída no Laboratório de Biotecnologia e Marcadores Moleculares/UFMG, utilizou o vetor pUC 18 no sítio da enzima de restrição BamH1 (cliva GGATCC, presente apenas no Sítio Múltiplo de Clonagem). O DNA genômico deste peixe foi digerido com a enzima de restrição Sal3A1 (cliva GATC) e submetido à eletroforese em gel de agarose, onde foram selecionados fragmentos de DNA maiores que 400pb. Esse material foi ligado ao vetor pUC 18, digerido previamente com BamH1 (Yazbeck, 2007).

Foi realizada a transformação de bactérias eletrocompetentes *Escherichia coli* da linhagem Topo 10 com o vetor pUC 18 acima mencionado e, em seguida, a seleção de colônias que continham uma inserção de DNA genômico da espécie alvo no vetor como descrito em Paiva, 2007; Yazbeck, 2007.

A varredura da biblioteca foi feita com o uso de sondas de oligonucleotídeos marcadas com fósforo radioativo. Foram utilizados os seguintes temas de repetição: GATA, TATC, GACA, GGAT, TATT, TACT, CGGA, GTCA, CCGT, GAGG, GATT, CATG, CCAG, CATG, CCAG, CTTG, CAGC, TATG, TAGT, AAGT, AAGT, CAT, CAC, GTA, GAA, AAT, CTT, CAC, CCG, CAG, CA, CT e TA (Yazbeck, 2007).

Setenta e um clones positivos encontrados na varredura da biblioteca genômica foram isolados para o desenvolvimento do presente trabalho.

### 3.3) Transformação dos plasmídeos selecionados na triagem com sonda radioativa em bactérias *Escherichia coli* eletrocompetentes:

Os clones positivos de *Prochilodus lineatus* foram precipitados com 10% do volume de acetato de sódio 3 Molar e 2 vezes o volume de etanol 70% para a retirada do excesso de sal existente nos mesmos. Depois foram centrifugados por 10 minutos, a 16000 g, a 4 °C. Retirou-se o sobrenadante e acrescentou-se 200 µL de etanol 70%. Centrifugou-se novamente durante 5 min., a 16000 g, a 4 °C. O sobrenadante foi retirado e acrescentou-se novamente 200 µL de etanol 70%. Foi realizada mais uma centrifugação durante 4 minutos, a 16000 g, a 4 °C. O sobrenadante foi retirado do tubo contendo o clone e o mesmo secou a temperatura ambiente.

Os plasmídeos foram ressuspensos em 6 µL de água bidestilada (Phoneutria, BH, MG) e então preparados para a eletroporação.

As bactérias eletrocompetentes utilizadas na eletroporação foram as *Escherichia coli* da linhagem Topo 10, lavadas em glicerol 10% e armazenadas a -80 °C (Phoneutria, BH, MG).

Incubou-se o plasmídeo com 100 µL de *Escherichia coli* e a solução de plasmídeo e bactéria foi então transferida para uma cubeta e submetida a um choque de 2 kVa (quilovolt-ampères) no eletroporador de pulso Micropulser da Bio-Rad. Após o choque foi adicionado 1 mL de solução SOC (Phoneutria, BH, MG) (20 mM de MgSO<sub>4</sub>, 20 mM de glicose) e após a homogeneização incubou-se a 37 °C por 1 hora. A solução foi transferida para outro tubo de 15 mL contendo meio 2XYT (5 g de NaCl, 10 g de peptona, 5 g de extrato de levedura, água destilada q.s.p. 1 L, pH 7.5 e ampicilina (1 µL de ampicilina/ mL de meio), ficando a temperatura ambiente por cerca de 18 horas. O plasmídeo pUC 18 (Figura 6) confere resistência à ampicilina, então utilizou-se esse antibiótico para selecionar apenas as bactérias que tivessem o plasmídeo.

Após 18 horas, 100 a 200 µL de meio foram plaqueados em meio LB ágar (10 g de NaCl, 10 g de bactotripton, 5 g de extrato de levedura, 20 g de ágar, água destilada q.s.p. 1 L, pH 7,5) com ampicilina (1 µL de ampicilina/mL de meio). Alguns dos clones selecionados pela sonda radioativa não estavam em tubos individualizados, tinha-se tubos com 2 ou 3 diferentes clones, então, na tentativa de que todos os clones fossem selecionados, coletou-se um grande número de colônias. As colônias foram então incubadas em 10 ml de meio 2XYT com ampicilina a 37 °C com agitação por 18 horas para garantir o crescimento das bactérias.

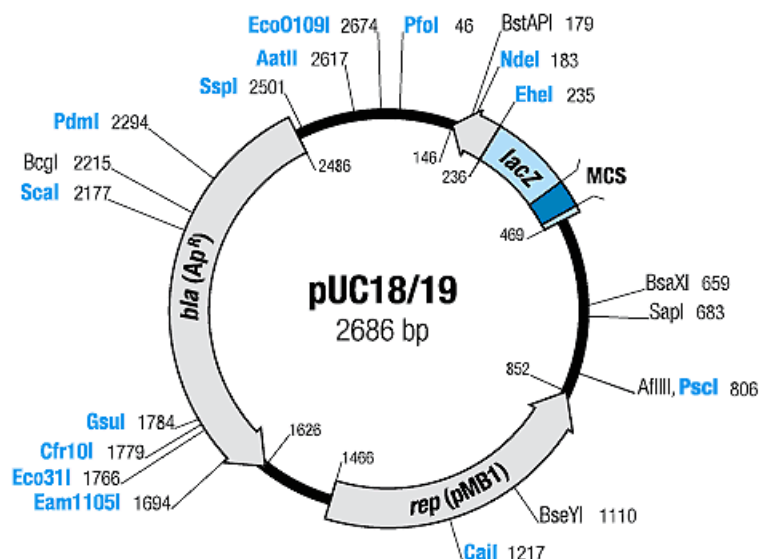


Figura 6: Mapa do vetor pUC 18, evidenciando o gene de resistência à ampicilina (*bla* ( $Ap^r$ )).

### 3.4) Extração de plasmídeos por lise alcalina:

Os tubos que continham as bactérias que estavam sendo incubadas a 37°C por 18 horas em meio 2XYT com ampicilina foram centrifugados por 5 minutos a 10000 g e o sobrenadante descartado. Foram então adicionados 300 µL de Solução I (glicose 50 mM, Tris-HCl 25 mM pH 8,0, EDTA 10 mM pH 8,0), ressuspensando o pellet, e então, adicionou-se 300 µL de Solução II (NaOH 0,2 N, SDS 1%). Logo em seguida, foram adicionados 300 µL de Solução III (acetato de potássio 3 M, 11,5% ácido acético glacial) e após incubação em gelo por 5 minutos, os tubos foram novamente centrifugados por 10 minutos a 16000 g. Transferiu-se o sobrenadante para um novo tubo e um volume de 600 µL de isopropanol foi adicionado, agitando-os esporadicamente por 15 minutos. Descartou-se o sobrenadante após centrifugação a 16000 g por 10 minutos. O precipitado foi lavado em 1 mL de etanol 70%, centrifugando-o a 16000 g por 5 minutos, e seco a temperatura ambiente, após a eliminação do excesso de etanol.

Após a extração por lise alcalina, o plasmídeo foi ressuspensado em 50 µL de RNase 10 ug/µL, deixando-a agir por 1 hora a 37 °C para ocorrer a eliminação do RNA. Aplicou-se 1 µL da solução em um gel de agarose 1% onde a coloração foi realizada com Brometo de Etídeo para verificar se as bactérias que haviam sido coletadas possuíam plasmídeos diferentes, selecionar quais clones seriam seqüenciados



visualizando a qualidade das bandas no gel e quantificá-los através da intensidade de bandas vista no gel de agarose.

### 3.5) Seqüenciamento:

Para a realização da reação de seqüenciamento utilizou-se 200 ng de plasmídeo, 0,3 µL de BigDye Terminator v3.1 Cycle da Applied Biosystems, 1,6 pmol de um iniciador do plasmídeo pUC18 (Figura 6): M13 direto 5' GTAAAACGACGGCCAGT 3' ou reverso 5' CAGGAAACAGCTATGACC 3' (para fazer o consenso), 1,77 µL de Tampão de seqüenciamento BigDye Terminator v1.1, v3.1 da *Applied Biosystems* e água suficiente para 10 µL. Essa reação foi então levada para o termociclador onde foi utilizado o programa padronizado para o seqüenciador 3130 *Genetic Analyzer* da *Applied Biosystems*:

96 °C por 2 minutos	}	44 ciclos
96 °C por 30 segundos		
Rampa 1 °C/seg.		
50 °C por 15 segundos		
Rampa 1 °C/seg.		
60 °C por 4 minutos		
Mantém 4 °C		

O produto de amplificação foi então precipitado com 40 µL de isopropanol 65%, deixado 30 minutos a temperatura ambiente e centrifugado a 16000 g por 25 minutos a 25 °C. O isopropanol foi então removido e lavou-se o pellet com 300 µL de etanol 60%. Em seguida, os tubos foram centrifugados a 16000 g por 5 minutos e deixados secar a temperatura ambiente, após a eliminação do excesso de etanol para então serem guardados no freezer -20 °C. Esse precipitado era então encaminhado para o seqüenciamento no *ABI 3130 Genetic Analyzer* do Laboratório de Biotecnologia e Marcadores Moleculares do ICB/UFMG.

### 3.6) Análise do seqüenciamento e desenho dos iniciadores:

Os plasmídeos foram seqüenciados em ambas as direções. As seqüências resultantes foram analisadas por meio dos programas Phred/Phrap/Consed (Ewin e Green, 1998; Gordon *et al.*, 1998) para obtenção de seqüências consenso. Para os seqüenciamentos realizados com o iniciador M13 reverso a seqüência era apresentada no sentido 3'-5'. Para se conseguir a fita complementar no sentido 5'-3', o programa *Reverse Complement* disponível em [http://bioinformatics.org/sms/rev\\_comp.html](http://bioinformatics.org/sms/rev_comp.html) foi utilizado para que a análise da fita fosse feita. A região da seqüência que correspondia à do plasmídeo de clonagem pUC18 era excluída das análises. Para isso, o programa *VecScreen* (NCBI- *National Center for Biotechnology Information*, disponível online em <<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>>, acesso: maio de 2008) foi utilizado.

Utilizou-se o programa *Microsatellite Repeats Finder* (BioPHP- PHP for bioinformatics, disponível *on line* [http://biophp.org/minitools/microsatellite\\_repeats\\_finder/demo.php](http://biophp.org/minitools/microsatellite_repeats_finder/demo.php). Acesso em: junho, 2008) para a localização dos microssatélites. Três sequencias que possuíam marcadores moleculares (clone 2M3, clone 2V35 e clone 2X7) foram selecionadas para o desenho dos iniciadores.

O *locus* do marcador é margeado por uma seqüência única de bases. Se a seqüência flanqueadora se torna conhecida, iniciadores podem ser sintetizados de modo a complementar essas sequencias, de tal forma que o *locus* do marcador pode ser amplificado pela Reação da Polimerase em Cadeia (PCR).

Para desenho dos iniciadores, foi utilizado o programa *Primer3* v. 0.4.0 (Rozen e Skaletsky, 2000) de onde foram selecionados aqueles que mais se enquadravam com os nossos objetivos que eram aproximadamente 60% de bases C e G e temperatura média de anelamento entre 54 °C e 60 °C, para que não ocorresse a formação de bandas inespecíficas.

Para todos os programas utilizados, os resultados obtidos foram criteriosamente analisados para a verificação de possíveis erros dos mesmos.

### 3.7) Extração de DNA genômico do peixe *P. costatus*:

Foi extraído o DNA de 45 indivíduos de *Prochilodus costatus* coletados ao longo do Rio Pará em 3 trechos distintos: Jusante UHE Gafanhoto (15 amostras), Trecho UHE Gafanhoto/Cajuru (15 amostras), Montante UHE Cajuru (15 amostras); pelo método de proteinase K e fenol (Sambrook *et al.*, 2001).

Um pequeno pedaço de nadadeira foi macerado em gelo seco, utilizando-se graal e pistilo. Adicionou-se no próprio graal 500 µL de TEN 9 (50 mM Tris-HCl, pH 9, 100 mM EDTA, 200 mM NaCl) e 2,5 µL de RNase (100 mg/ml) e em seguida, transferiu-se a solução para um tubo de 15 mL. O tubo foi agitado suavemente por inversão durante 10-30 minutos, até formar uma mistura homogênea. Adicionou-se 30 µL de SDS 20% e o tubo foi agitado novamente. Um volume de 30 µL de proteinase K (10mg/ml) foi adicionado, homogeneizando o conteúdo do tubo que, em seguida, foi deixado por aproximadamente 14 horas a 37 °C.

No outro dia, adicionou-se 1 volume de fenol/clorofórmio, agitando suavemente por 1 hora. A mistura foi centrifugada por 10 minutos a 3000 g e a interface foi retirada e lavada com clorofórmio por 20 minutos. Quando necessário, a lavagem com clorofórmio foi repetida para eliminar completamente o fenol da amostra. Foi acrescentado então, 1/5 do volume de acetato de amônio 10 M e 1 volume de isopropanol. O tubo foi invertido gentilmente. O DNA foi retirado e transferido para outro tubo utilizando-se uma pipeta *Pasteur* selada. Em seguida, ele foi lavado com 1 mL de etanol 70% por 1 hora. O DNA foi tirado do álcool e deixado secar a temperatura ambiente. Em seguida, o DNA foi dissolvido com 20 µL de TE (10 mM Tris-HCl, pH 8, 1 mM de EDTA).

### 3.8) Padronização de PCR:

Após os iniciadores serem desenhados (Item 3.6), testou-se um conjunto que margeava o *locus* de um marcador selecionado. As reações de PCR foram padronizadas de tal forma que se obteve uma quantidade suficiente de DNA de qualidade para o sequenciamento. Para a padronização, variou-se a temperatura de anelamento dos iniciadores desenhados, iniciando sempre pela  $T_m$  (*melting temperature*) que corresponde à temperatura na qual metade dos potenciais sítios de ligação possivelmente terá iniciadores ligados a eles. Para calcular a Temperatura Média a

fórmula:  $T_m = 4(G+C) + 2(A+T)$  onde G, C, A e T são as bases nitrogenadas (HILLIS, 1996), foi utilizada. Utilizando-se uma máquina de PCR de gradiente (Eppendorf Master Cycler Gradiente), permitindo 12 diferentes temperaturas de anelamento em um único programa de PCR, foi realizada a variação de temperatura de anelamento com incremento de 2 °C.

Variou-se também o número de repetições do ciclo de desnaturação, tipo de tampão, além das quantidades de DNA molde e iniciadores.

### **3.9) Padronização de *Nested* PCR:**

A reação de *Nested* PCR consiste em realizar duas reações de PCR consecutivas, onde a segunda PCR é realizada com o produto de outra PCR, sendo a primeira realizada com iniciadores mais externos do que a segunda, otimizando assim os resultados obtidos. Foi realizada a *Nested* PCR dos 45 indivíduos de Curimatá-pioa que tiveram seu DNA extraído.

O tamanho do fragmento de amplificação esperado foi calculado contando-se o número de pares de base de um iniciador interno ao outro, incluindo as seqüências dos iniciadores direto e reverso.

### **3.10) Subclonagem do produto de *Nested* PCR em plasmídeo PCR 2.1 TOPO *Invitrogen*:**

Para a subclonagem, os produtos da *Nested* PCR foram quantificados em gel de agarose 1,5%.

Para a ligação ao vetor utilizou-se 10 ng de produto de *Nested* PCR, 50 ng do vetor pCR 2.1 (Figura 7), 1 µL de tampão de ligação 10x, 4 unidades de T4 ligase e água estéril suficiente para 10 µL de reação e incubação por 16 horas a 14 °C. O plasmídeo foi precipitado utilizando-se acetato de sódio 3 M pH 5,2 e etanol 100%, deixando-o a uma temperatura de -80 °C por 20 minutos. Em seguida, centrifugou-se o plasmídeo a 16000 g por 10 minutos e o sobrenadante foi retirado. Lavou-se o pellet acrescentando 200 µL de etanol 70%, centrifugando por mais 5 minutos a 16000 g, retirando o sobrenadante, acrescentando novamente 200 µL de etanol 70%, centrifugando por mais 4 minutos a 16000 g, retirando o sobrenadante e deixando a



### 3.12) Extração dos plasmídeos PCR 2.1 TOPO (*Invitrogen*) por lise alcalina:

Os tubos que continham as bactérias que estavam sendo incubadas tiveram seus plasmídeos extraídos conforme o descrito na página 32.

Após a extração do plasmídeo por lise alcalina, o mesmo foi ressuspenso em 50 µL de RNase (10 mg/ml-estoque), 2 µL/mL de TE (10 mM pH8,0 Tris-HCl/0.1 mM pH 8,0 EDTA), deixando agir por 1 dia a 37 °C, e meio dia a temperatura ambiente para a eliminação do RNA. Logo após, fez-se a purificação com fenol/clorofórmio (Sambrook *et al.*, 2001) e em seguida o plasmídeo foi ressuspenso em 20 µL de TE. Foi aplicado 1 µL da solução em um gel de agarose 0,8% com coloração em Brometo de Etídeo para quantificar o plasmídeo que foi então seqüenciado como descrito na página 33.

### 3.13) Análise do seqüenciamento:

Após o seqüenciamento foram feitas análises para confirmar a existência do inserto contendo o marcador molecular no vetor PCR 2.1 TOPO.

Para os seqüenciamentos realizados com o iniciador M13 reverso a seqüência se apresentava no sentido 3'-5'. Para se obter a fita complementar no sentido 5'-3' foi utilizado o programa *Reverse Complement* (*The Sequence Manipulation Suite. Reverse Complement*. Disponível online em: <<http://www.bioinformatics.org>>. Acesso: julho de 2008).

A região da seqüência que correspondia à do vetor de clonagem PCR 2.1 TOPO era excluída das análises. Para isso, o programa *VecScreen* (NCBI- *National Center for Biotechnology Information*, disponível online em <<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>>. Acesso: julho de 2008) foi utilizado.

### 3.14) Análise estatística dos dados:

As seqüências consenso foram obtidas a partir de 4 seqüências de alta qualidade geradas com os iniciadores M13 direto e reverso, utilizando os programas Phred/Phrap/Consed (Ewin e Green, 1998; Gordon *et al.*, 1998). Os alinhamentos das

seqüências obtidas foram construídos utilizando-se o Clustal W (Higgins e Sharp, 1998), implementados no programa MEGA 4 (Tamura *et al.* 2007).

Para análise de partição da variabilidade genética das amostras coletadas ao longo do rio, foi empregada a metodologia da Análise de Variância Molecular (AMOVA) (Excoffier *et al.*, 1992), incluída no pacote para análises de genética de populações Arlequin v. 3.0 (Excoffier *et al.* 2005). Essa análise da variância (AMOVA) é baseada na comparação de duas estimativas diferentes da variância comum às diferentes populações: a variância entre amostras, que se baseia na variabilidade entre as médias amostrais e a variância dentro das amostras, baseado nas variâncias amostrais.

No método de AMOVA a significância do índice de fixação é testada utilizando-se permutação não paramétrica (Excoffier *et al.* 2005). Utiliza-se a distribuição F não paramétrica, sendo que esse tipo de distribuição é assimétrica à direita, onde se tem uma distribuição F diferente para cada par de graus de liberdade. Os valores de F podem variar de zero a positivos, mas nunca poderão ser negativos.

No programa Arlequin foram também calculadas estimativas de diversidade gênica (H), diversidade nucleotídica ( $\pi$ ), Fst (Índice de Fixação) e o teste exato da cadeia de Markov, para avaliação do potencial discriminatório do marcador.

Uma análise de agrupamento de genótipos foi feita para os 45 indivíduos coletados nas três regiões do rio Pará com o programa *Structure* (Pritchard *et al.*, 2000). Essa análise avalia, por meio de inferência Bayesiana, a designação de indivíduos à populações hipotéticas (K), na tentativa de encontrar subestrutura em uma amostragem.

Para estimar as relações genealógicas entre as amostras coletadas, através do critério da máxima parsimônia (95%), o programa TCS 1.21 (Clement *et al.*, 2000) foi utilizado. O programa insere a frequência dos haplótipos na população, gerando uma rede haplotípica.

Após realizar o teste de normalidade de Kolmogorov-Smirnoff e verificar a presença de *outliers* através do teste de Grubbs (Barnett e Lewis, 1994), os dados biométricos (peso, comprimento e fator de condição) foram submetidos à análise de variância (ANOVA), implementado no programa de estatística *GraphPad InStat* (GraphPad InStat Software versão 3.00, disponível *online* em: [www.graphpad.com](http://www.graphpad.com). Acesso: agosto 2009), seguido pelo teste de Duncan para ver se a variância era ou não significativa. Utilizando esses testes estatísticos, foi feita a comparação entre os dados biométricos dos peixes coletados em diferentes trechos do rio Pará. O nível de significância considerado no presente trabalho foi de  $p < 0,05$ .

#### **4) RESULTADOS:**

##### **4.1) Busca por novos marcadores para *Prochilodus*:**

4.1.1) Transformação dos plasmídeos previamente selecionados pelo laboratório de Biotecnologia e Marcadores Moleculares na triagem com sonda radioativa em bactérias *Escherichia coli* eletrocompetentes e extração por lise alcalina:

A Tabela 1 mostra os 71 clones selecionados a partir da biblioteca genômica de *P. lineatus* utilizados para eletro transformação, a identificação de cada um deles e a indicação do número de seqüenciamentos realizados por clone. A Figura 8 mostra exemplos de clones positivos, ou seja, que cresceram em meio contendo ampicilina, cujos plasmídeos foram extraídos por lise alcalina em gel de agarose 0,8%.



Tabela 1: Clones de *Prochilodus* utilizados no trabalho. Primeira coluna: clones utilizados para eletrotransformação, segunda coluna: identificação do clone, terceira coluna: indicação do número de seqüenciamentos realizados por clone.

Clone	Identificação	Seqüenciados
1	2M3	2
2	2O14	2
3	2O20	2
4	2O30	2
5	2P19	2
6	2P22	2
7	2R25	2
8	2R31	2
9	2S3	2
10	2S20	2
11	2S35	2
12	2T10	2
13	2T14	2
14	2T21	2
15	2U12	2
16	2U14	2
17	2u32	2
18	2V26	2
19	2V35	2
20	2X7	2
21	2x27	2
22	2Z2	2
23	2Z11	2
24	3A5	3
25	3A19	3
26	3A26	3
27	3B18	3
28	3B26	3
29	3C4	3
30	3C10	3
31	3C18	3
32	3C27	3
33	3D10	3
34	3D22	3
35	3E9	3

Clone	Identificação	Seqüenciados
36	3E27	3
37	3E32	3
38	3F5	3
39	3M29	3
40	5N34	3
41	5O10	3
42	5O32	3
43	5P6	3
44	5Q11	3
45	5Q18	3
46	5R13	3
47	5S7	3
48	5S18	3
49	5S31	3
50	5T22	3
51	5T25	3
52	5T28	3
53	5U29	3
54	5V34	3
55	5W11	3
56	5W18	3
57	5W27	3
58	5Y9	3
59	5Y22	3
60	5Y30	3
61	5Y33	3
62	5Z3	3
63	6A10	3
64	6A20	3
65	6A33	3
66	6B9	3
67	6B30	3
68	6C30	3
69	6D11	3
70	6D16	3
71	6E6	3

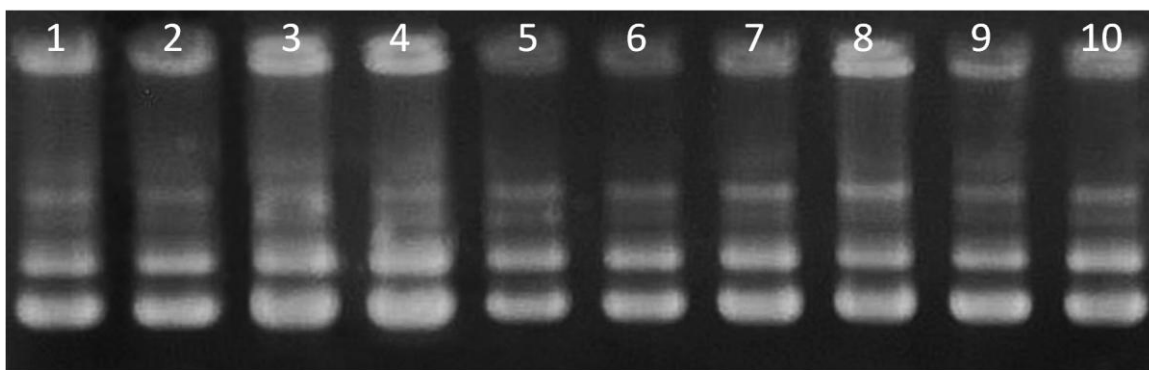


Figura 8: Gel de agarose 0,8% mostrando plasmídeos contendo insertos de microssatélites de DNA de *P. lineatus*. Visualização de rotina objetivando a análise de qualidade dos plasmídeos e quantificação para sequenciamento.

#### 4.2) Sequenciamento e análise de clones positivos:

Um total de 71 clones foi sequenciado, sendo que 19 (26%) apresentaram pelo menos uma sequência repetitiva (Figura 9).

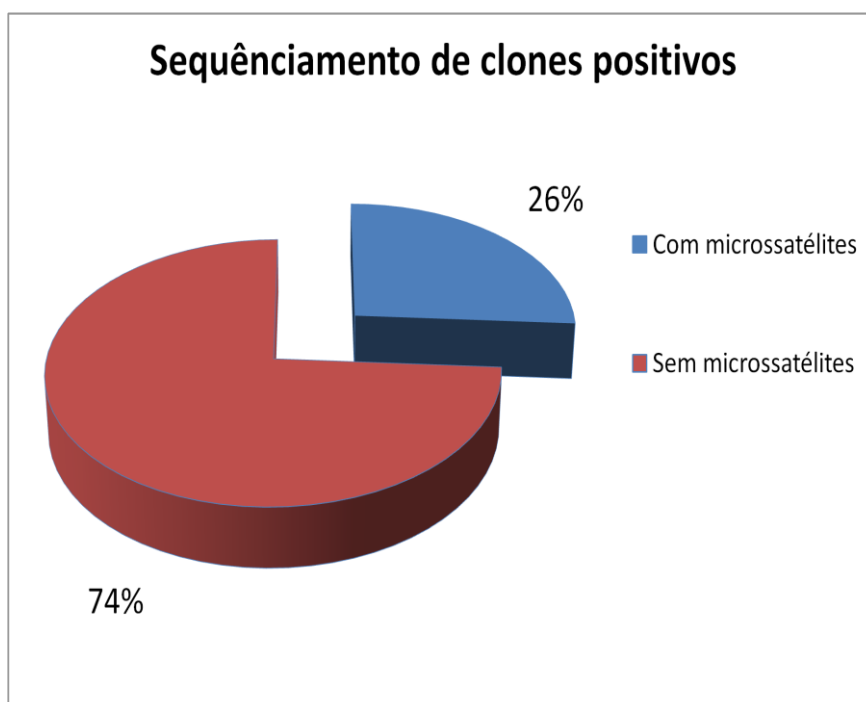


Figura 9: Gráfico de proporção entre as sequências de clones do peixe *P. lineatus* que possuem microssatélites e as que não possuem.

Cerca de 66% das seqüências repetitivas encontradas (18) foram da classe dos dinucleotídeos, seguidos pelas classes de seis trinucleotídeos, de dois tetranucleotídeos e de um pentanucleotídeo.

Analisando-se essas 27 regiões de microssatélites, três clones (2M3, 2V35 e 2X7) do peixe *Prochilodus lineatus* foram selecionados. Em seguida, analisaram-se as extremidades flanqueadoras das regiões consenso dos microssatélites, chamadas a partir de agora de Marcadores Moleculares, para que os iniciadores internos e externos a serem utilizados na *Nested PCR* fossem desenhados utilizando o programa *Primer3* v. 0.4.0 (Rozen e Skaletsky, 2000) (Figura 10).

#### A) Clone 2M3/1

TTAACCTAAAAGCGGCTCAGTCGATCTAGAGGTCATCTGTGTGCAGACCAACAACAGGT  
CAGATTTACCAAGCTCACCAAGTCTCCAGCTTCAGGTCTGAGAGAAGCACAAGGTCTCA  
GCTTTACAGCCCTCCGGCACCCTCTCAAAGTGGAAGCACCATCATTAGGCATCATCAGA  
GGCTGGTTTTTTATAAAAAAGAGGCCATGCCACCCACGTAATGACAGGCAGTCCCCACA  
GCACTGAGCTCCAGCTCTAATTACCGTTCTCTCCCTGAAAGCGGCTCTTTTTGACACAG  
CTCTGTGTGTGTGTGATTGACTGGAGTGGGGGAGAGCGCCACACACACCCCGTATTTAC  
CGCACC CGCTGGTTAAAAGTTAATGCTCAAAACATTAGTGAACTG **GACAGTGGAGCGA**  
**AGAGG**ATAA **CACAACAAGGATGGGGAATC** TAGATAAACAGAGACGCACACACACACA  
CACACACACACACACACACGCACACGCACACTCACACTCACACTCACACACGCACAC  
GCACACGCACACGCACATGCACACACACACACGCACACACACACACACGCACACACGCG  
CACACAAACGTACACGCACACATACACACACACGCACACGCACACGACACTCACACAC  
GCAAACACACACGCACACGCACACGCACATGCACACACACACACACACACACACACA  
CACACGCGCACACGCACACATACACACACGCACACTCACACACGCACATGCACACACAC  
GCGCGCACGCACACACGCGCACACGCACACACATAC **GTCTGTTTCTGTACCTTG**ACAG  
**AACATATGGTGTT**CTATATGTTAATTATGTATATGTAGGGATCCCCGGGTACCGAGCT  
CGAAATTCGTAATCATGTATAGCTGTTTTCCTGTGTGAAATTGTTTATCCGCTTCAAAT  
TCCACACCAACATAACGAGCCGGAAGCAAAAAGTGTAAGGCTTGGGGGGGCTTAAGGA  
GTGAACTAACTCACATTAATTGGCGTTGCGCCTCCTGCCCGCTTCAGTCGAAACTTG  
CCCTGGCCACCTTCATTAATGAATCGGCCAAGCGCCCGGGGAAAGGGCCGTATGGCTT  
ATGGGGCCCTCCTCATTTTCGCTTTCTCTCTCTGGCTACATGAATCTGCATCGCGTTCTGG  
GCTCGTTTCGCTGGGGACGTTAACGCTCACTTCAAGCGTTATATCGG

#### B) Clone 2v35/5

AGCCGTCCTGCAGTCGACTCTAGAGGTC **TAATGATTCTCTTTGCTTGTGTC**TTGCTCT  
CTCTCTCTTTCTCTTTGTCTCTCCCTTTGTCCCACCTTCCCTTCTCACTCTATCATTTT  
TCTGCCTTTTCATCTTTGTCACTTTTTCTCCAACGCTTCTCTCTCTCTTTTTTTCTCTT  
CTTTTTTTAACTCCCTCTGTGCCATTTTCTGATGTCACCTTTTTTCTCTCTCTCCATG  
TCTCTACCTCTTTTTTCTCTTCCCTCTCTCCCCCATATTGCCTCTTTCACTTTCTCCCTCT  
GTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTACCGTCTTTT  
TTTTCTGTCTCTTTCTCCATGTCTCTCCCTCTTTCTCTCTTGTGTCTTTTTTTCTGTCT  
AGCCTTTGTTTCATCACCTTGCAAACCTCTCATGCTGTCTCTGCGGTTTCTGAG **TGCAC**  
**TGTATTTACAGGCCTT**TTGTAAAGTGACATGAACTCCTGAATATTTAATGTGTGTGTGT

GGGTATTGGGGGGAAGATAGACTACGGTTCGCTCTAGAGGTTCAAGGTTAGGATAAAGGTCGTC  
TGTTTCAGGTTAAGGAAAATTAAGTCTGCTGAAAATGAATGGAAATCTCACAATACGTAC  
AAACAAGTGTGTGTGTGTGTGGGTGCGTGTGTGTGTTCTTACAGACTGAGA **GTC**  
**CTGTGCTCGAGTGCAGCACATA** AAGAACCTCAACCTCCGA **CTCCA**ACTACTCGCCTGA  
**CC**CTCCTCCAGTCTGTGTCTGAGAGAGAGAGAGGAGAGAGAGAGAGAAAGAAAAAGAGAG  
AAAGAGAGAGAGAGAAAGAGAAAGAAGGGGGAGGAGGGGAAGGAAGACAAAAAGGAAGAGG  
GCAGAAAGAAAGACAGAAGGAAAGACAGAAAGAAAGAAAAAACAAAAACACAGAAAG  
AACACACAAAAGAAGAAATAAGAGACAGAAAGAGAGAAAAAAATGAAAAATAACAAC  
CCAGAAAGAACACTTAAAGAAGCAAGGAGGCCAGAGAGAAAAACAGACATGAAGAAAAGAC  
AAATGAAATAAAGAAAAATGAATTGAAAGAAAAAAGACAAAAAGAAGAAACAGAAAG  
AAAGAACGAAAGAAAGAACGAAAGAACGAAAGAAAGAACGAAAGAAAGAAAGAACAAAA  
GAAAGAACGAAAGAAAGAAAGAAAGAACGAAAGAAAGAAAGAAAGAACGAAAGAAAGAA  
AGAAAGAAAGAAAGAAAGAAAGGAAGAAAGAACAAAAGAACGAAAGAAAGAAAGAAAA  
GGAAGAAAGAACAAAAGAACGAAAGAAAGAACAAAAGAACGAAAGTAAATGAAGAAAAA  
GAAAGAGAAAAGGTTAGAAGAAGGAAATGAAAAAGAAAAGAAAGAAAGTAAAGTAAAGG  
GAGATAAAGAAAGAGAAAGAGGTAGAAGAAGGAAATTTAAAAAAGAAAGAAAGAA  
AGAAAGTAAAGGAAGAAAAAGAAAGAGAAAGAGGTAGAAGAAGGAAATAGAAAAGGAAA  
AGAAAAGAAAAGCATTAAGAAAGAAACAAAGAAAAATGAAGGACGAAAGAAACAAACA  
GAGAAAGAAAAATGAAGAAAGGAAGAAAGAAAGAAAGGAACAAACAAAAAAGAAACGT  
TAAAGACTTTCTCA **CCAAAATCAGACAGGGGAGA** TGGGTGATAAAACGTGTCC **CCGG**  
**TACCGAGCTCC**AATATA

Figura 10: Marcadores moleculares encontrados no sequenciamento de clones do peixe *Prochilodus lineatus* cujas seqüências flangeadoras foram selecionadas para desenho dos iniciadores internos e externos a serem utilizados em *Nested* PCR. Todas as seqüências apresentadas estão no sentido 5'-3'. As regiões correspondentes ao plasmídeo PUC 18 foram indicadas por letras maiúsculas e em cor cinza claro. As seqüências correspondentes ao DNA do peixe *Prochilodus* estão escritas com letras maiúsculas e em preto. As regiões dos marcadores moleculares estão grifadas e as seqüências dos iniciadores escolhidos foram marcadas em negrito, sendo os iniciadores externos destacados em caixas de bordas e sombreamento verdes e os iniciadores internos destacados em caixas de bordas e sombreamento vermelhos. A) clone 2M3, B) clone 2V35, C) clone 2X7.

Na Tabela 2, tem-se todos os seqüenciamentos realizados para cada região polimórfica selecionada.

Tabela 2: Seqüenciamentos realizados na identificação dos marcadores moleculares. Na primeira coluna encontram-se as siglas de identificação dos clones; na segunda coluna os números dos seqüenciamentos realizados para identificar com precisão toda a seqüência dos marcadores descobertos; na terceira coluna, tem-se o nome do iniciador utilizado em cada reação de seqüenciamento.

Identificação do Clone	Seq. N°	Iniciador
2M3	AM 618	M13 Direto
	AM 619	M13 Direto
	AM 620	M13 Direto
	AM 621	M13 Direto
	AM 622	M13 Direto
2V35	AM 894	M13 Direto
	AM 895	M13 Direto
	AM 896	M13 Direto
	AM 897	M13 Direto
	AM 953	M13 Reverso
2X7	AM 998	M13 Direto
	AM 999	M13 Direto

### 4.3) Padronização de PCR:

A temperatura de anelamento de cada iniciador foi calculada pelo programa *Primer3* v.0.4.0 (Rozen e Skaletsky, 2000) e posteriormente testada na PCR para ajuste preciso a ser utilizado na amplificação do DNA genômico.

A Tabela 3 mostra além das seqüências dos iniciadores internos e externos desenhados para cada clone, as temperaturas de anelamento padronizadas no laboratório para a reação de *Nested* PCR do clone 2V35.

Tabela 3: Seqüência dos iniciadores específicos para cada marcador. Na primeira coluna encontram-se as siglas de identificação dos clones; na segunda coluna as seqüências dos iniciadores internos e externos; na terceira a temperatura de anelamento calculada pelo programa *Primer3* v.0.4.0 (Rozen e Skaletsky, 2000); na quarta coluna a temperatura de anelamento padronizada no laboratório; na quinta coluna, o peso molecular esperado para cada fragmento amplificado como produto de PCR.

Identificação do Clone	Iniciadores	Tm Programa Primer3 v.0.4.0 (°C)	Tm padronizada (°C)	Fragmento amplificado (pb)
2M3	Direto Externo 5'GACAGTGGAGCGAAGAGG 3'	55,2	*	403
	Direto Interno 5'CACAACAAGGATGGGGAATC 3'	60,17		
	Reverso Externo 5'AACACCATAGTTC 3'	58,5		
	Reverso Interno 5'CAAGGTACAGAAACAGAC3'	54,3		
2V35	Direto Externo 5'TAATGATTCTCTTGCTTGTGTC 3'	53,3	51	455
	Reverso Externo 5'AAGGCCTGAAATACAGTGCA3'	56,0		
	Reverso Interno 5'GCAGACCCCTCAGCAC3'	54,5		
2X7	Direto Externo 5'GTGCTCGAGTGCAGCACATA 3'	57,7	*	937
	Direto Interno 5'CTCCAATACTCGCCTGACC3'	59,87		
	Reverso Externo 5'GGAGCTCGGTACCGG 3'	55,0		
	Reverso Interno 5'TCTCCCCTGTCTGATTTGG3'	60,04		

O asterisco (\*) indica as temperaturas que ainda não foram calibradas na padronização da PCR.

A padronização das condições ideais de amplificação para o *locus* do marcador do clone 2V35 começou com a utilização de um *mix* com volume final de 25  $\mu$ l, contendo: 5 pmols de cada iniciador externo (direto e reverso), 0,2 unidades de Taq Polimerase (Phoneutria, BH, MG) e tampão IVB especial (Phoneutria, BH, MG). A partir desse ponto, foram testadas as melhores condições de anelamento, testando as amostras em um gradiente de temperatura de anelamento, entre 47  $^{\circ}$ C e 67  $^{\circ}$ C, com incremento de 2  $^{\circ}$ C. O programa do termociclador foi ajustado para os seguintes passos: 1) 2 min. a 94  $^{\circ}$ C; 2) 30s a 94  $^{\circ}$ C 3) 30s a X  $^{\circ}$ C; 4) 40s a 72  $^{\circ}$ C; 40 vezes de volta ao passo 2 e 5' a 72  $^{\circ}$ C, onde X é a temperatura de anelamento que foi variada.

A Figura 11 mostra um exemplo de amplificações feitas com os iniciadores internos utilizando-se o plasmídeo do clone 2V35. O par de iniciadores resultou em amplificações consideradas satisfatórias para a tipagem de alelos em um padrão diplóide. As bandas de alto peso molecular são do plasmídeo e não amplificações inespecíficas, podendo ser comprovado quando comparado com a amplificação de DNA genômico de *Prochilodus costatus* (figura 13). O *locus* escolhido e então caracterizado neste trabalho foi o HPL (*Highly Polymorphic Locus*), referente ao clone 2V35.

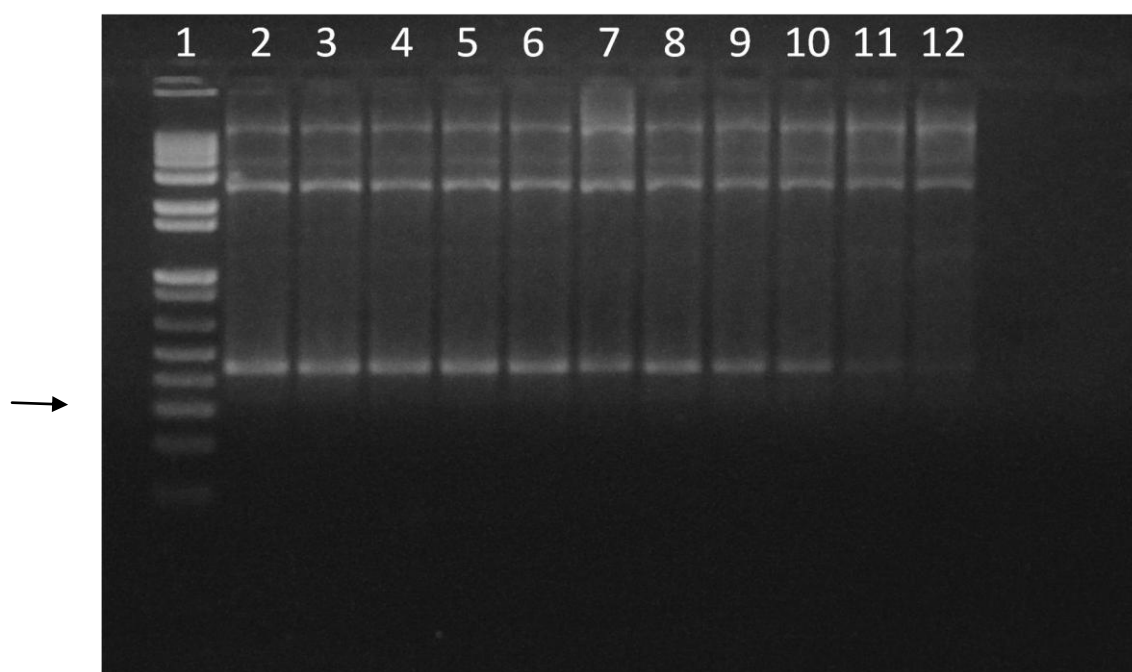


Figura 11: Gel de agarose 1,5%. Exemplo de amplificação utilizando o clone 2V35 com variação de temperatura entre 47  $^{\circ}$ C e 67  $^{\circ}$ C (canaletas 2 a 12). A primeira canaleta contém o padrão de peso molecular de 1 Kb plus (*Invitrogen*). A seta compreende o tamanho de 400 pb.

A temperatura de anelamento escolhida foi a de 51 °C; 3,5 °C abaixo da média das temperaturas dos iniciadores direto e reverso (54,65 °C). O programa utilizado foi ajustado para os seguintes passos: 1) 2 min. a 94 °C; 2) 30s a 94 °C 3) 30s a 51 °C; 4) 40s a 72 °C; 35 vezes de volta ao passo 2 e 6) 5' a 72 °C. O número de ciclos foi diminuído para obtenção de bandas mais específicas.

Inicialmente, o tampão IVB especial (Phoneutria, BH, MG) foi o escolhido para as reações de PCR com os iniciadores do marcador do clone 2V35, como mostra a Figura 12, mas após a extração de DNA dos indivíduos da espécie *P. costatus*, foram realizados vários testes com o DNA genômico e decidiu-se utilizar o Pré-mix IC (Phoneutria, BH, MG), que já contém DNTPs, Taq Polimerase e Tampão IC (500 mM KCl, 100 mM Tris-HCl pH 8,4, 1% Triton X-100, 200 mM MgCl<sub>2</sub>), sendo necessário acrescentar apenas o DNA, os iniciadores e a água.

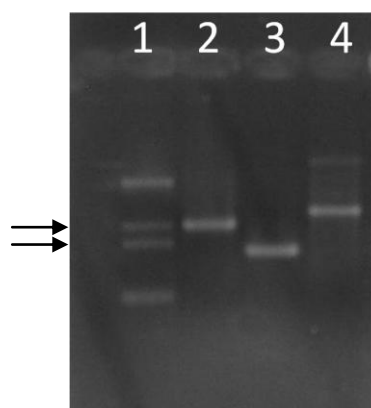


Figura 12: Gel de agarose 1,5%. Exemplo de amplificação utilizando o clone 2V35 e tampão IVB. Canaleta 1: padrão de peso molecular 1.200pb. Canaletas 2 a 4: produto de PCR do plasmídeo 2V35 com iniciadores externos, internos e M13, respectivamente. As setas indicam os tamanhos de 500 pb e 700 pb.

#### 4.4) Padronização da *Nested* PCR:

Para a padronização da primeira etapa de amplificação da *Nested* PCR foi constatado que o uso do Pré-mix IC e 5 pmols de iniciadores externos (direto e reverso) e aproximadamente 100 ng de DNA genômico, geraram os melhores resultados (Figura 13). Para a segunda etapa de amplificação, a utilização de 5 pmols de iniciadores internos (direto e reverso), 02 unidades de Taq Polimerase (Phoneutria) e Tampão IVB



(Phoneutria), apresentaram os melhores resultados. O volume de 0,2  $\mu$ l de *amplicon* foi o selecionado para todas as amplificações (Figura 14).

A temperatura de anelamento escolhida foi a de 51<sup>0</sup>C, praticamente igual à média das temperaturas dos iniciadores direto e reverso internos (50,8<sup>0</sup>C). O programa utilizado foi ajustado para os seguintes passos: 1) 2 min. a 94<sup>0</sup>C; 2) 30s a 94<sup>0</sup>C 3) 30s a 51<sup>0</sup>C; 4) 40s a 72<sup>0</sup>C; 25 vezes de volta ao passo 2 e 6) 5' a 72<sup>0</sup>C.

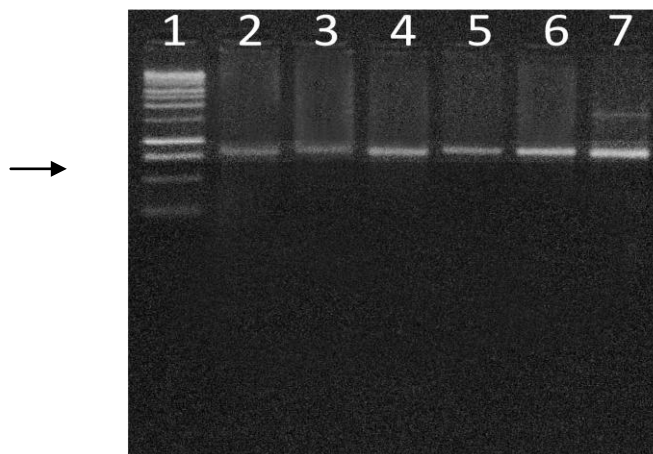


Figura 13: Gel de agarose 1,5% contendo o produto de amplificação da primeira etapa da *Nested* PCR, utilizando-se diferentes concentrações de DNA extraído de um exemplar de *P. costatus* e Pré-mix 1C. Canaleta 1: padrão de peso molecular de 1 Kb (Promega). A seta indica o tamanho de 750 pb. Nas canaletas 2, 3, 4, 5 e 6 foram usadas diferentes concentrações de DNA genômico de *P. costatus*. Canaleta 7: controle positivo da reação.

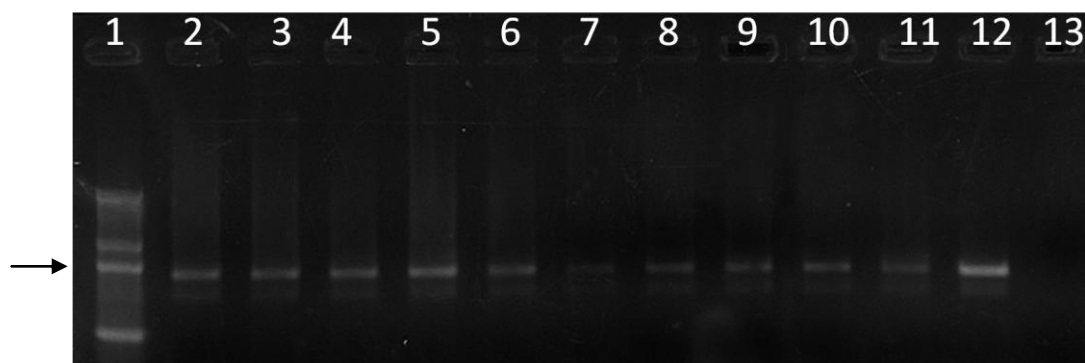


Figura 14: Gel de agarose 1,5% contendo o produto de amplificação da segunda etapa da *Nested* PCR, utilizando-se o tampão IV B. Canaleta 1: padrão de peso molecular 1.200 pb. A seta indica o tamanho de 500 pb. Canaletas 2 a 11: produtos da *Nested* PCR. Canaleta 12: controle positivo da reação. Canaleta 13: controle negativo da reação.

#### 4.5) Subclonagem do produto da Nested PCR em plasmídeo pCR 2.1 TOPO *Invitrogen*:

A subclonagem em plasmídeo pCR 2.1 TOPO foi realizada com os produtos da *Nested* PCR de cada um dos 45 indivíduos de *P. costatus* que apresentaram os melhores resultados, ou seja, bandas mais definidas e nítidas e foram purificados em coluna de Sephacril 400, com o intuito de excluir fragmentos de ampliações inespecíficas (Figura 15). Após a reação de subclonagem, os plasmídeos foram eletrotransformados em *E. coli* XL1Blue. Foram selecionadas 10 colônias de cada indivíduo (Figura 16) e os plasmídeos foram extraídos e então a PCR de colônia foi realizada para a verificação de inserto de DNA exógeno, ou seja, o marcador HPL, no vetor de clonagem pCR 2.1 TOPO, como está demonstrada na Figura 17. Em seguida, os quatro melhores *amplicons* de cada indivíduo foram selecionados para o sequenciamento, como se pode observar também na Figura 17 onde os *amplicons* do exemplar A de *P. costatus* das canaletas 2, 9, 10 e 11 foram selecionados por apresentarem bandas com melhor definição e menos “rastros”.

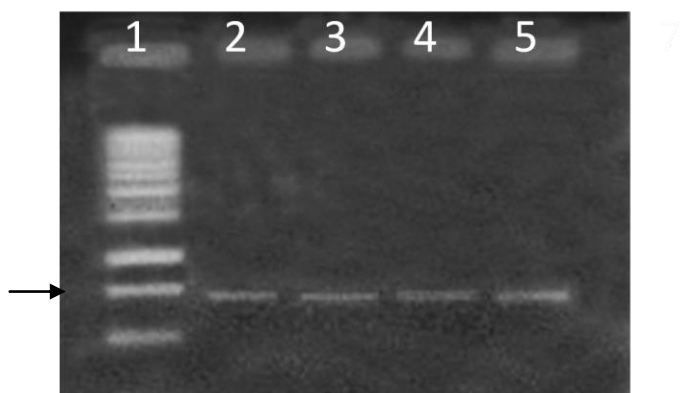


Figura 15: Amplificação da região do marcador molecular HPL de um exemplar de *P. costatus* feita por *Nested* PCR em gel de agarose 1,5%, mostrando bandas livres de rastros. Canaleta 1: padrão de peso molecular de 1 Kb (Promega). A seta indica o tamanho de 500 pb. Canaletas 2 a 5: indivíduo A.

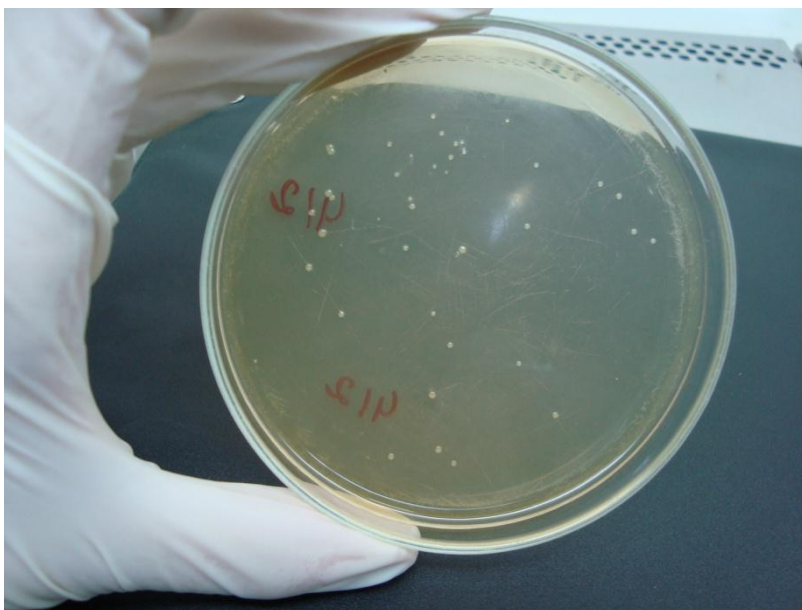


Figura 16: Placa de Petri contendo meio de cultura LB-Agar com ampicilina mostrando a alta eficiência da subclonagem. As colônias brancas possuem o plasmídeo pCR 2.1 TOPO, resistente à ampicilina.

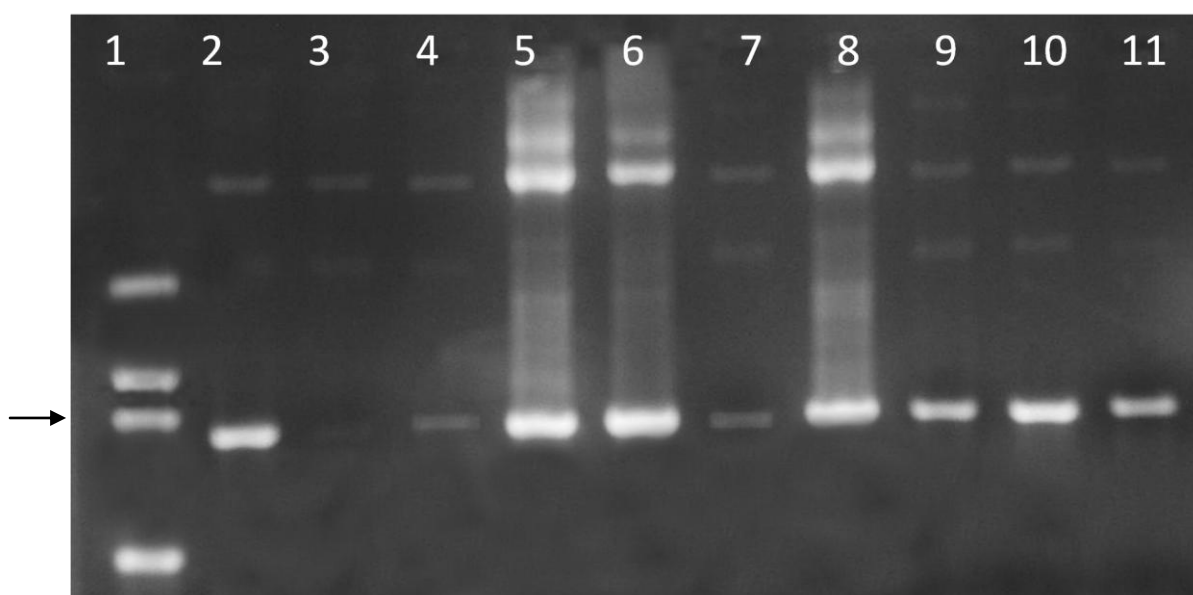


Figura 17: PCR de colônia do exemplar de *P. costatus* A em gel de agarose 1,5% para verificação de inserto de DNA exógeno no vetor de clonagem pCR 2.1 TOPO. Canaleta 1: padrão de peso molecular 1.200pb. A seta indica o tamanho de 500 pb. Canaletas 2 a 11: produto de PCR de colônia do Indivíduo A. Canaletas 2, 9, 10, 11: *amplicons* selecionados para seqüenciamento.

#### **4.6) Análises de seqüenciamentos:**

Para cada indivíduo dos 45 selecionados para esse estudo foram obtidas 4 seqüências de alta qualidade (Figura 18), totalizando 180 seqüências. Das 180, foram obtidas 45 seqüências consenso do marcador molecular HPL.

As seqüências foram então alinhadas pelo Clustal W (Figura19), implementado no programa MEGA 4.0 (Tamura *et al.* 2007) e todo polimorfismo detectado era então conferido manualmente, voltando-se no eletroferograma que continha a seqüência em questão (Figura 20).

Após o alinhamento, foi constatado que o marcador molecular HPL não é apenas um microssatélite imperfeito, contendo polimorfismos de comprimento (diferentes números de temas repetidos lado a lado), mas que também apresenta mutações por transversão e transição, e “indels” (inserção/deleção de pares de bases).

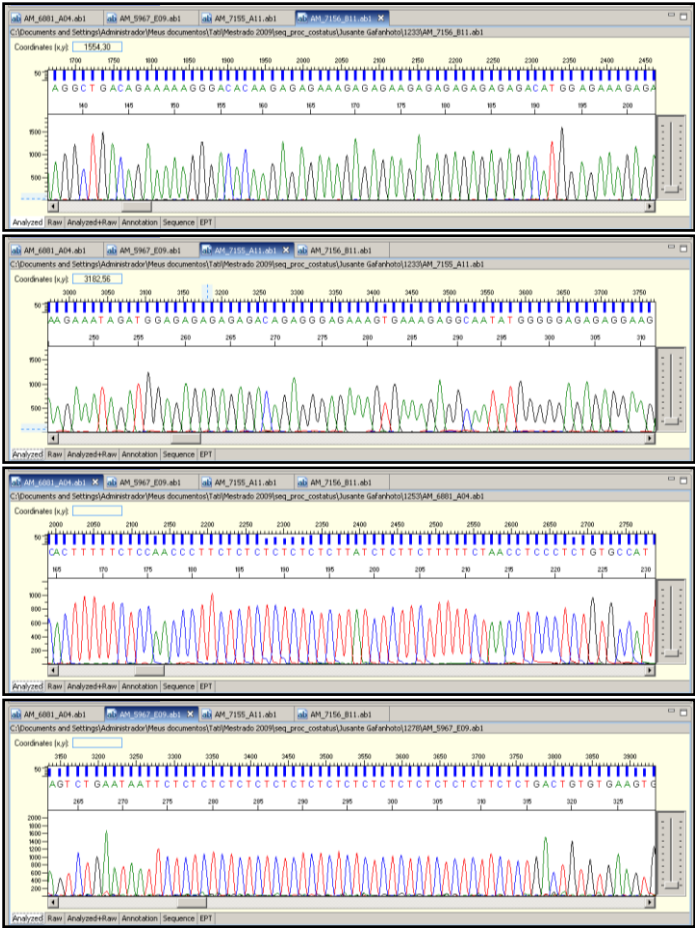


Figura 18: Exemplo de quatro eletroferogramas resultantes do sequenciamento da região de microssatélite HPL do exemplar 1278 de *P. costatus*, mostrando alta qualidade dos picos.

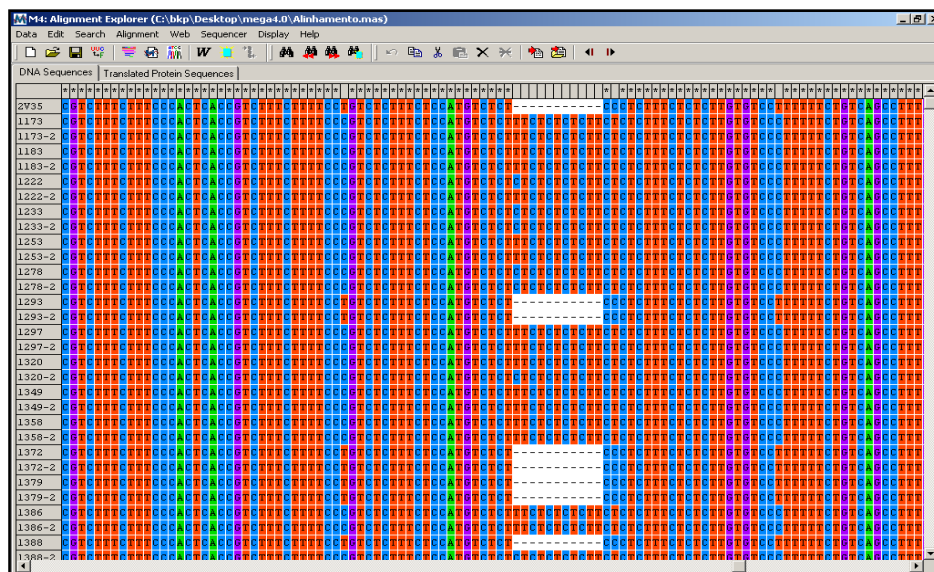


Figura 19: Alinhamento dos 15 exemplares de *P. costatus* coletados à jusante da UHE de Gafanhoto realizado pelo Clustal W implementado no programa MEGA 4.0 (Tamura *et al.* 2007). Pode-se observar “indels” e mutações por substituição.



Figura 20: Alinhamento feito no Clustal W implementado no programa MEGA 4.0 (Tamura *et al.* 2007) mostrando um polimorfismo detectado no marcador molecular HPL no indivíduo 1173 de *P. costatus* e dois de seus eletroferogramas onde os picos de alta qualidade confirmam esse polimorfismo.

#### 4.7) Análises Estatísticas:

##### 4.7.1) Índices de diversidade:

Foram encontrados 25 haplótipos para a espécie analisada, onde foram observadas 96 substituições, sendo 63 transições e 33 transversões. Foram observados 44 “indels”. O polimorfismo de comprimento variou entre 448 pb e 458 pb. As amostragens apresentaram haplótipos em comum sendo que as coletadas à jusante de Gafanhoto e na barragem de Cajurú possuem os haplótipos 1 e 4 em comum e as coletadas no trecho Gafanhoto/Cajurú e na barragem de Cajurú possuem em comum o haplótipo 16 e o 19. Na Tabela 4 tem-se um resumo dos índices de diversidade observados para essas amostragens.

Tabela 4: Índices de diversidade de *Prochilodus costatus* observados para as três localidades.

	Jusante Gafanhoto	Trecho Gafanhoto/ Cajurú	UHE Cajurú
N	15	15	15
h	12	8	9
H	0,9264 +/- 0,0242	0,8690 +/- 0,0384	0,8874 +/- 0,0324
$\pi$	0,042471 +/- 0,021511	0,017420 +/- 0,009268	0,035805 +/- 0,018257

Onde N= número de indivíduos, h= número de haplótipos, H= diversidade gênica e  $\pi$  = diversidade nucleotídica.

##### 4.7.2) Fst e teste exato da cadeia de Markov:

Na Tabela 5, tem-se o resultado dos cálculos realizados pelo programa ARLEQUIN 3.1.1 para os valores do índice de fixação (Fst) (Wright, 1951), que é a medida clássica de diferenciação das populações, usada para avaliar os níveis de conectividade entre elas. A amostra localizada no trecho entre as duas barragens foi a que apresentou a maior diferença genética entre as três amostragens analisadas, comparando-se o índice de fixação: amostras coletadas à jusante da UHE de Gafanhoto e Trecho (0.23); amostras coletadas à jusante da UHE de Gafanhoto e na UHE de Cajurú (0.09); amostras coletadas na UHE de Cajurú e Trecho (0.19).

Os valores de probabilidade de P foram significativos ( $P < 0,05$ ) para o Fst calculado para as três amostragens. Sendo assim, os valores encontrados foram considerados.

O teste da cadeia de Markov indicou uma enorme diferenciação genética entre as três amostragens ( $P < 0,05$ ), como mostra a Tabela 6.

Tabela 5: Matriz com os valores de Fst (acima) e o valor de probabilidade P (abaixo) das amostragens analisadas. Foram realizadas 1023 permutações para o cálculo do Fst e para  $P(0,050)$ .

	Jusante Gafanhoto	Trecho Cajurú/Gafanhoto	UHE Cajurú
Jusante Gafanhoto	0.00001		
Trecho Cajurú/Gafanhoto	0.23024 0.00001+-0.0001	0.00001	
UHE Cajurú	0.09031 0.02703+-0.0139	0.19351 0.00001+-0.0001	0.00001

Nota: Quando P for maior que 0,05, o valor é considerado não significante, o que significa que o valor calculado para Fst é plausível de ter sido encontrado ao acaso. Neste caso, desconsidera-se o valor.

Tabela 6: Teste global de não diferenciação da cadeia de Markov. Foram realizados 6000 passos da cadeia de Markov com o P exato igual a 0.00001 +- 0.00001. O valor P indica a diferenciação entre as populações.

	Jusante Gafanhoto	Trecho Cajurú/Gafanhoto
Trecho Cajurú/Gafanhoto	0.00001+-0.0001	
UHE Cajurú	0.00001+-0.0001	0.00001+-0.0001

#### 4.7.3) AMOVA:

O programa AMOVA, implementado no ARLEQUIN 3.1.1, foi utilizado para determinar a diferença entre a variância molecular dentro e entre as amostragens. Para esse cálculo, as três amostragens foram agrupadas em apenas um grupo.



O resultado do AMOVA mostrou uma maior variação entre os indivíduos de cada uma das três amostragens (83,27%), do que entre as três (16,73%). O índice de fixação  $F_{st}$  (16,6%), aponta uma variação pouco significativa entre as três amostragens analisadas (Tabela 7). O nível de significância considerado foi de  $P < 0,05$ .

Tabela 7: Resultados do programa AMOVA sugerindo ausência de subestruturação, com poucas diferenças genéticas, sem fluxo gênico elevado. Onde g.l.= grau de liberdade,  $F_{st}$  =índice de fixação.

Fonte de variação	g.l.	Soma dos quadrados	Componente de variância	Porcentagem de variação
Entre populações	2	102.989	1.47218 $V_a$	16.73
Dentro de populações	87	637.633	7.32912 $V_b$	83.27
Total	89	740.622	8.80130	
Índice de Fixação		$F_{ST} :$	0.16727	

#### 4.7.4) Análise de subestruturação populacional:

Para a avaliação da presença de subestruturação dentre os 45 indivíduos analisados, foi empregada a técnica de agrupamento (*clustering*) de genótipos através do método apresentado por Pritchard *et al.* (2007); no programa de análise de estrutura genética *Structure* 2.3.1. Um exemplo de *output* da análise está representado na Figura 21, sendo que  $K=3$  (três *clusters*). A opção de utilizar três *clusters* foi escolhida porque permite que todos os *clusters* construídos sejam representados de modo direto, ou seja, não há um ângulo que represente mais do que um *cluster*. Os resultados para os valores de  $K$  (*Clusters*) são representados por um triângulo equilátero onde cada ângulo representa um *cluster*. Cada população presumida é representada por uma cor (Jusante Gafanhoto, Trecho Cajurú/Gafanhoto e UHE Cajurú) e, na presença de estruturação, é possível perceber a distribuição dos indivíduos de cada uma delas pelos diferentes *clusters*, o que permite a associação entre os grupos presumidos e os grupos genéticos construídos pelo programa, como está exemplificado na Figura 21. O diagrama da Figura 22 representa um perfil que retrata a ausência de estruturação com pontos de todas as cores preenchendo as áreas mais centrais do triângulo.

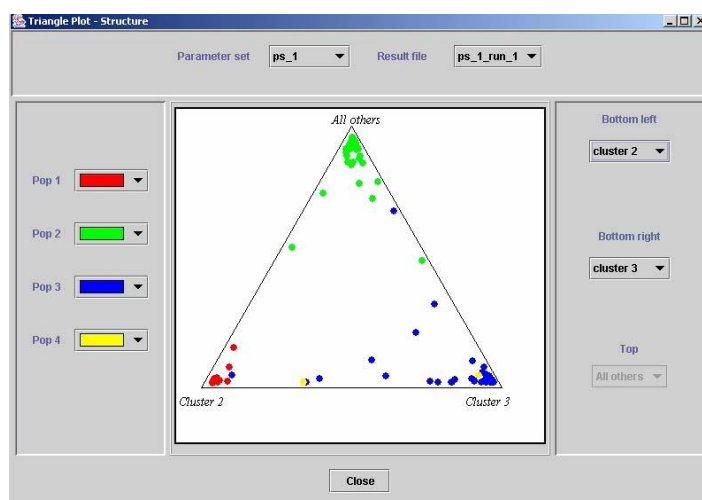


Figura 21: O exemplo extraído do manual do *Structure* (Pritchard *et al.*, 2007) mostra um perfil característico de estruturação genética. No caso são 4 as subpopulações presumidas e  $K=3$ .

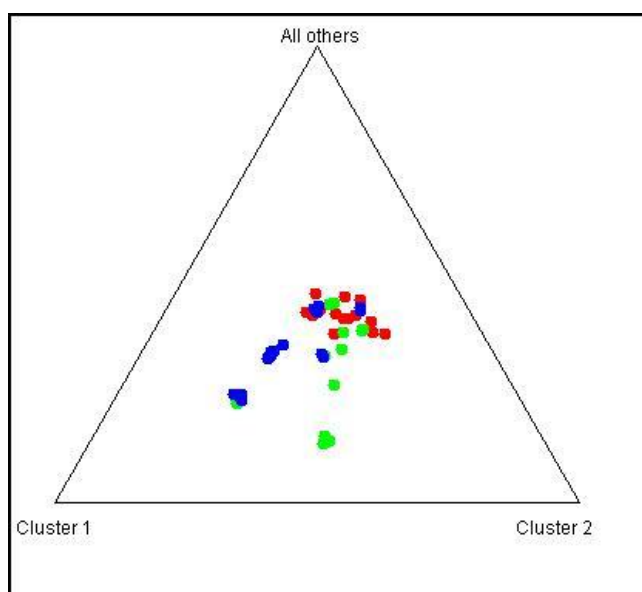


Figura 22: Diagrama de agrupamento dos genótipos dos 45 exemplares de *P. costatus* coletados em três diferentes localidades do rio Pará: no rio à jusante da represa Gafanhoto, no trecho entre as barragens de Cajurú e Gafanhoto e na barragem de Cajurú. Os pontos agrupados representam os indivíduos das amostras, sendo que a cor vermelha representa exemplares coletados à jusante da represa de Gafanhoto, a verde representa exemplares coletados entre as represas de Cajurú e de Gafanhoto e a azul, exemplares coletados na represa de Cajurú. Observa-se um perfil que retrata a ausência

de estruturação com pontos de todas as cores preenchendo as áreas mais centrais do triângulo.

#### 4.7.5) Árvores de Haplótipos:

Para a construção do cladograma foi utilizado o princípio de máxima parcimônia, de forma que os haplótipos tivessem probabilidade igual ou maior que 95%, de se conectar parcimoniosamente, obtendo-se assim, dois haplótipos isolados e duas sub-árvores, sendo esses correspondentes às unidades evolutivas (Figura 23).

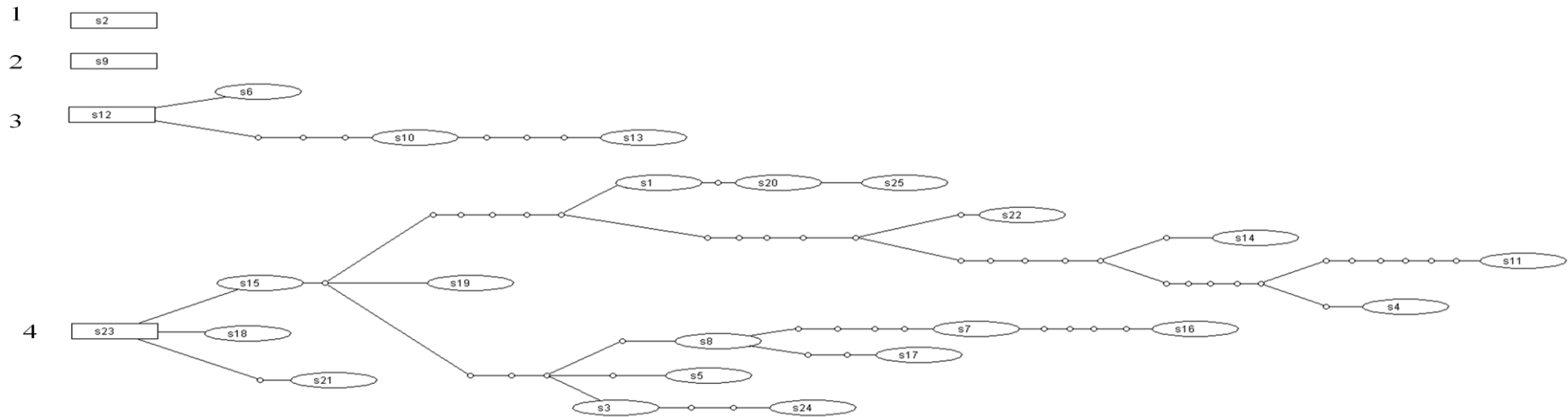


Figura 23: Árvore de haplótipos evidenciando as conexões entre os 25 haplótipos evidenciados em *P. costatus*. Os círculos pequenos significam haplótipos inferidos, mas não detectados. Os haplótipos estão distribuídos da seguinte forma: Unidade evolutiva 1: S2= rio Pará (jusante UHE Gafanhoto); unidade evolutiva 2: S9= rio Pará (jusante UHE Gafanhoto); unidade evolutiva 3: S6, S10, S12 = rio Pará (jusante Gafanhoto) e S13= UHE Cajuru; unidade evolutiva 4: S1, S3, S4, S5, S7, S8 e S11= rio Pará (jusante Gafanhoto); S14, S15 e S18= UHE Cajuru; S16, S17 e S19= trecho entre UHE de Cajuru e Gafanhoto e UHE Cajuru; S20 a S25= trecho entre UHE de Cajuru e Gafanhoto.

#### 4.7.6) Biometria e Fator de Condição:

As análises das variações de comprimento padrão e peso corporal das amostras, bem como o fator de condição (K) foram feitas para avaliar o bem-estar dos peixes, ou seja, a adequação deles ao meio ambiente em que vivem. O teste estatístico ANOVA foi utilizado para medir a variância entre as amostras das três regiões do rio Pará.

Os valores do comprimento padrão e peso corporal das amostras da barragem de Cajurú foram maiores do que as do rio (jusante da UHE Gafanhoto) e do trecho entre UHE de Cajurú e Gafanhoto (Figuras 24 e 25). O fator de condição (K) foi maior para as amostras da barragem de Cajurú e do rio (jusante da UHE Gafanhoto) em comparação com as do trecho entre UHE de Cajurú e Gafanhoto (Figura 26). O teste estatístico ANOVA mostrou uma variação significativa entre as amostras da barragem de Cajurú e do rio (jusante da UHE Gafanhoto) em relação às do trecho entre as duas barragens, como mostra a Tabela 8. O nível de significância considerado foi de  $P < 0,05$ .

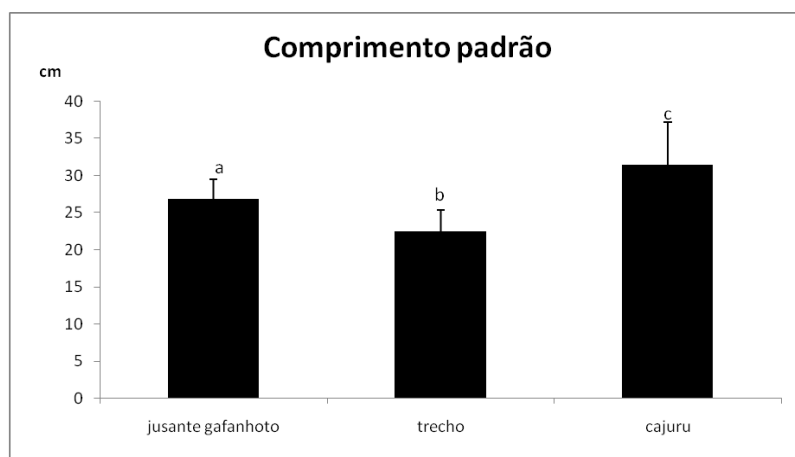


Figura 24: Comprimento padrão das amostras coletadas à jusante da UHE de Gafanhoto, no trecho entre as UHE de Gafanhoto e Cajurú e na UHE de Cajurú. A média do comprimento das amostras à jusante da barragem de Gafanhoto foi de 26,73 cm com desvio-padrão de 2,75 cm; média de 22,35 cm para as amostras do trecho entre as duas barragens com desvio-padrão de 3 cm e para as amostras da barragem de Cajurú, média de 31,43 cm com desvio-padrão de 5,74 cm ( $P < 0,05$ ).

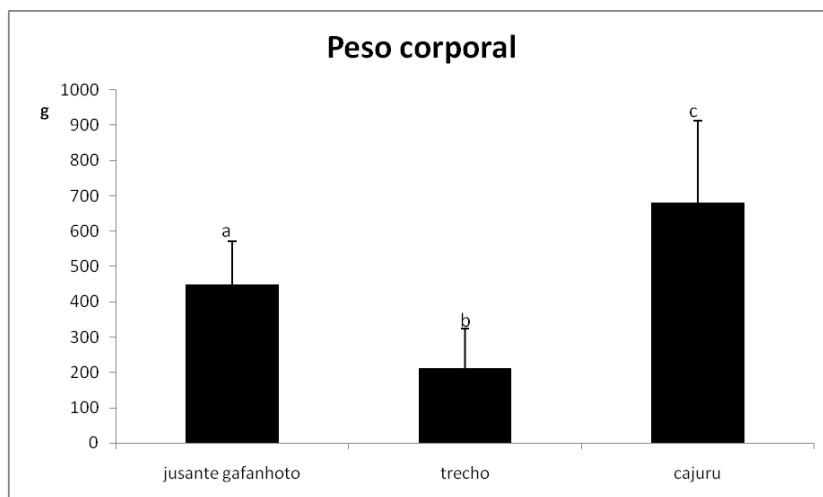


Figura 25: Peso corporal das amostras coletadas à jusante da UHE de Gafanhoto, no trecho entre as UHE de Gafanhoto e Cajurú e na UHE de Cajurú. A média do peso das amostras à jusante da barragem de Gafanhoto foi de 449,66 g com desvio-padrão de 122,3 g; média de 211,42 g para as amostras do trecho entre as duas barragens com desvio-padrão de 112,58 g e para as amostras da barragem de Cajurú, média de 681,66 g com desvio-padrão de 230,56 g ( $P < 0,05$ ).

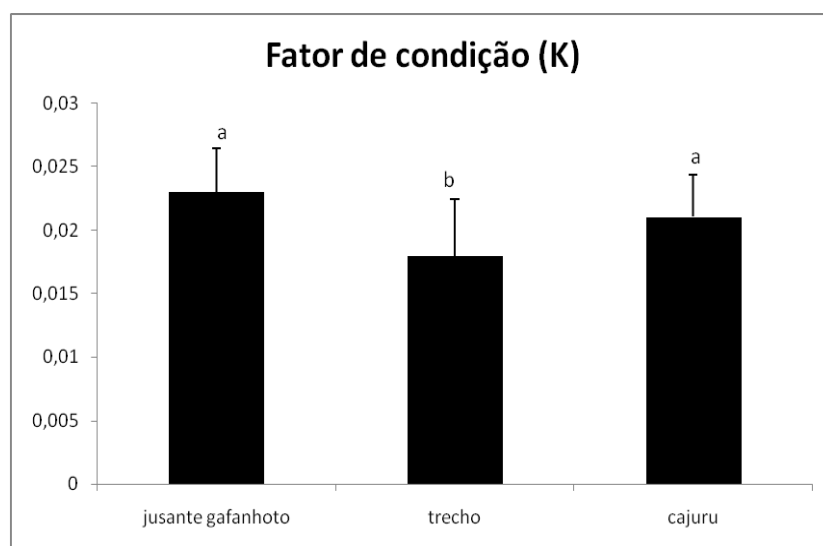


Figura 26: Fator de condição (K) das amostras coletadas à jusante da UHE de Gafanhoto, no trecho entre as UHE de Gafanhoto e Cajurú e na UHE de Cajurú, onde a média das amostras à jusante da barragem de Gafanhoto foi de 0,023 com desvio-padrão de 0,003; média de 0,017 para as amostras do trecho entre as duas barragens com desvio-padrão de 0,004 e para as amostras da barragem de Cajurú, média de 0,021 com desvio-padrão de 0,003 ( $P < 0,05$ ).

Tabela 8: Resultado do programa *GraphPad Instat*: ANOVA mostrando a variância entre as três amostragens. Há uma variação significativa entre as amostras da barragem de Cajurú e do rio (jusante da UHE Gafanhoto) em relação às do trecho entre as duas barragens.

	Jusante Gafanhoto	Trecho	Cajurú
<b>PC</b>	449,6 ± 122,3 <sup>a</sup>	211,4 ± 112,6 <sup>b</sup>	681,7 ± 230,6 <sup>c</sup>
<b>CP</b>	26,7 ± 2,7 <sup>a</sup>	22,35 ± 3,002 <sup>b</sup>	31,4 ± 5,7 <sup>c</sup>
<b>K</b>	0,02 ± 0,003 <sup>a</sup>	0,017 ± 0,004 <sup>b</sup>	0,02 ± 0,003 <sup>a</sup>

Nota: PC= Peso Corporal, CP= comprimento Padrão, K= fator de condição e P>0,05.

## 5) DISCUSSÃO E CONCLUSÃO:

No presente trabalho, foram sequenciados ao todo, 71 clones positivos selecionados a partir da biblioteca genômica do peixe *Prochilodus lineatus*, sendo que 19 (26%) possuíam pelo menos uma sequência repetitiva com potencial para se tornar um marcador molecular. Foram encontradas ao todo 27 sequências repetitivas e a maior parte das sequências encontradas (18) foram da classe dos dinucleotídeos, seguido de seis trinucleotídeos, dois tetranucleotídeos e um pentanucleotídeo.

A frequência do tipo e do motivo de repetições variou muito entre os grupos taxonômicos, mas de maneira geral têm-se notado que dinucleotídeos são mais abundantes em muitos genomas (Katti *et al.*, 2001), o que corrobora a proporção encontrada neste trabalho (Carreras *et al.*, 2005; Hatanaka *et al.*, 2006, Paiva e Kalapothakis, 2008). Destacando também que apesar de terem sido utilizadas sondas com vários motivos diferentes de repetições na hibridização da biblioteca, poucas sequências com tri, tetra ou pentanucleotídeos foram encontradas. O que, novamente, sugere que as repetições de dinucleotídeos sejam muito mais abundantes que as demais no genoma de *P. lineatus*.

De acordo com Estoup e Angers (1998), iniciadores de regiões de marcadores microsatélites desenvolvidos para uma espécie, algumas vezes amplificam o *locus* em outra espécie proximamente relacionada. O marcador encontrado na biblioteca genômica de *P. lineatus* se mostra presente em *P. costatus*, levando a concluir que as regiões flangeadoras onde os iniciadores foram desenhados permanecem conservadas dentro do gênero.

Uma condição importantíssima para a obtenção de *loci* analisáveis é a região flangeadora da sequência que deve possuir uma composição de bases que permitam o desenho de iniciadores eficientes na amplificação via PCR. Esta é uma etapa limitante para escolha dos marcadores, pois muitos apesar de possuírem boas repetições, não possuem sequências flangeadoras suficientes ou com boa composição de bases. Dos vinte e sete marcadores encontrados, três foram escolhidos para o desenho dos iniciadores, pois possuíam bordas 3' e 5' com regiões flangeadoras ideais, sendo que um *locus* chamado posteriormente de HPL- *Highly Polymorphic Locus* foi analisado nesta Dissertação.

A padronização da amplificação da região de microsatélite é outra etapa crucial para que seja realizada uma análise correta dos resultados. Essa etapa envolve um



esforço considerável e demanda certo tempo até que a reação esteja corretamente calibrada. A padronização da reação de *Nested* PCR para o *locus* HPL foi realizada variando-se temperatura de anelamento, tipo de tampão e quantidade de DNA genômico. A reação de *Nested* PCR apresenta a vantagem de permitir, através de seu perfil de restrição do produto amplificado um melhoramento da especificidade e eficiência da reação.

O método de extração de DNA escolhido para este trabalho foi o de proteinase K e fenol (Sambrook *et al.*, 2001). O método já foi testado com sucesso pelo nosso grupo (Paiva, 2007) e proporcionou ótima eficiência também em passos subsequentes como a amplificação via *Nested* PCR.

A análise de variabilidade genética via seqüenciamento vem sendo utilizada em estudos populacionais de peixes (Dergam *et al.*, 2002; Hatanaka *et al.*, 2006; Barbosa *et al.*, 2008; Carvalho-Costa *et al.*, 2008). A grande vantagem da utilização desse método em relação à análise via gel de poliacrilamida é a obtenção de resultados rápidos e altamente sensíveis. Além de seu tamanho em pares de bases, vários fatores influenciam a mobilidade eletroforética de um *amplicon* na análise por gel de poliacrilamida (Rosenblum *et al.*, 1997), como por exemplo, fatores aleatórios de variações na velocidade da corrida e temperatura do gel no momento da análise (Amos *et al.*, 2007). Esses fatores não estão presentes na análise de seqüências.

O seqüenciamento do marcador HPL permitiu que o caracterizássemos como uma região polimórfica presente no genoma do peixe *P. costatus*, sendo constatados 25 haplótipos diferentes (ver anexo).

Para as amostras coletadas à jusante da barragem de Gafanhoto, foram constatados 12 haplótipos, sendo que para as coletadas no trecho entre as duas barragens constatou-se 8 e nas amostras coletadas na barragem de Cajurú, 9 haplótipos, sendo que as amostras coletadas à jusante de Gafanhoto e na barragem de Cajurú possuem 2 haplótipos em comum (1 e 4) e as coletadas no trecho Gafanhoto/Cajurú e na barragem de Cajurú também possuem 2 haplótipos em comum (16 e 19). A diversidade gênica entre as três amostras variou de 0,86 a 0,92, considerada alta quando comparada a outros trabalhos do gênero *Prochilodus* (Hatanaka *et al.*, 2006; Barbosa *et al.*, 2008; Carvalho-Costa *et al.*, 2008). As amostras coletadas à jusante da barragem de Gafanhoto apresentam uma diversidade gênica um pouco maior que as demais, além de 6 haplótipos exclusivos que podem ser constatados através da árvore haplotípica (Figura 23). A presença de haplótipos exclusivos pode ser explicado pelo alto grau de

polimorfismo encontrado no *locus* estudado (Hatanaka e Galetti, 2003; Hatanaka *et al.*, 2006). No rio Pará, um intenso peixamento vem sendo realizado desde 1986 (Dr. Yoshimi Sato, com. pess.) como pode ser constatado na Tabela 10 em anexo. Isso reforça a idéia de que unidades reprodutivas distintas, ou seja, peixes nativos e peixes provenientes do repovoamento podem estar se misturando reprodutivamente (Dr. Gilmar B. Santos, com. pess.).

De acordo com os critérios de Wright, 1951, os valores de  $F_{st}$  com P menor que 0,05 indicam que as populações estudadas são geneticamente distintas, mas com diferenciação pequena. No presente trabalho, o índice de fixação  $F_{st}$ , quando calculado para as três amostras, apresentou variação de 0,16 (Tabela 8) podendo-se considerar uma diferenciação genética pequena entre as três amostras (Li *et al.*, 2009). Resultados semelhantes foram encontrados por Revaldaves *et al.* (1997) e Sivasundar *et al.* (2000), utilizando-se diferentes marcadores; onde foi verificado baixo nível de diferenciação genética nas subpopulações do peixe *Prochilodus lineatus*.

O baixo nível de diferenciação genética entre as três amostragens é reforçada diante da análise dos resultados do programa *Structure*, que apontam para a falta de agrupamento entre as amostras e da árvore haplotípica, resultado gerado pelo programa TCS 2.1., podendo-se observar que a unidade evolutiva 4 possui haplótipos compartilhados pelas três amostragens. Assim, infere-se que os 45 indivíduos analisados pertencem de fato à mesma população, com uma pequena diferenciação genética das amostras coletadas no rio em relação às coletadas no trecho entre as duas barragens e às coletadas na represa de Cajurú. Isso se deve ao fato de que as amostras do rio que foram coletadas à jusante de Gafanhoto são as únicas que ainda possuem população nativa e que fazem a migração reprodutiva anual no rio Itapecerica, retornando para o Pará, ou seja, provavelmente não existe recrutamento à montante da UHE de Gafanhoto (Dr. Yoshimi Sato, com. pess.). O cálculo da diferenciação genética utilizando o programa AMOVA determinou o grau de variabilidade genética como sendo de apenas 16,73% entre todas as amostragens de *P. costatus* utilizadas neste trabalho, sendo que 83,27% da variação foi encontrada dentro de cada amostragem.

A origem de matrizes do peixe *P. costatus* que deram origem ao plantel das estações de piscicultura da CODEVASF pode explicar essa pequena diferenciação genética entre os peixes coletados nas três regiões do rio Pará, que vem recebendo alevinos provenientes da estação há mais de 20 anos (ver tabela fornecida pelo Dr. Yoshimi Sato em anexo).

O fato de que as amostras coletadas no trecho entre UHE Gafanhoto e UHE Cajurú possuem uma divergência genética maior quando comparado seu índice de fixação ( $F_{st}$ ) com as demais (Trecho Gafanhoto/ Cajurú X Jusante Gafanhoto:  $F_{st}$  de 23% e Trecho Gafanhoto/ Cajurú X UHE Cajurú:  $F_{st}$  de 19,35%) pode ser explicado devido à fragmentação do hábitat já que não existem mecanismos de transposição em nenhuma das duas Usinas. Essas amostras estão isoladas em um pequeno trecho de aproximadamente 32 Km e não têm como completar seu ciclo de vida pois não existem locais para a desova. As condições para reprodução são particularmente desfavoráveis devido à influência diária de nível causado pela operação da usina de Cajurú (é um trecho regulado pela operação deste reservatório) (Dr. Gilmar B. Santos, com. pess.).

Outros dados que reforçam este indício são as análises de biometria e fator de condição feitas com todas as amostras analisadas. Os resultados observados neste trabalho estão dentro do padrão do gênero *Prochilodus* quando comparados com outros estudos (Rocha *et al.*, 1997; Vicentin *et al.*, 2004). Análises das variações de comprimento e peso corporal dos peixes permitem avaliar como a espécie obtém recursos do ambiente em que vive e indicam a condição do peixe através do bem-estar geral, desenvolvimento gonadal, adequação ao meio-ambiente e do acúmulo de gordura; sendo que a variação no teor de gordura dos peixes pode estar relacionada com o ciclo hidrológico do rio ou do reservatório (Le Cren, 1951; Bennemann *et al.*, 1996; Orsi *et al.*, 2000). Quando esses dados são associados ao fator de condição  $K$ , que tem sido amplamente incorporado a modelos estatísticos aplicados à biologia, como medida do estado fisiológico de peixes (Barbieri *et al.*, 1996), pode-se constatar que os maiores valores de  $K$  e da biometria realizada são das amostras do rio e da Usina hidrelétrica de Cajurú e que a variância do comprimento padrão (CP) e peso corporal (PC), mas não totalmente de  $K$  das amostras do trecho localizado entre as barragens, difere significativamente das outras, como mostra a Tabela 8 (teste estatístico ANOVA).

A utilização de novos marcadores moleculares do peixe *P. costatus* pode ser realizada para aumentar a confiabilidade dos resultados obtidos neste trabalho. Contudo, deve-se lembrar que o marcador molecular HPL não é apenas um microsatélite imperfeito, contendo polimorfismos de comprimento (diferentes números de temas repetidos lado a lado), mas que também apresenta mutações por transversão e transição, e “indels” (inserção/deleção de pares de bases) a serem analisados. Esse conjunto de polimorfismos tornou esse marcador uma ferramenta molecular eficiente na indicação da estrutura genética dos peixes *P. costatus* estudados nesse trabalho.

A utilização de outros exemplares de Curimatá-pioa (*P. costatus*) que não foram alvo de reprodução em cativeiro para repovoamentos também seria de grande valia para a determinação da eficiência do repovoamento. Tal estudo, em conjunto com os resultados obtidos até o momento, irá contribuir de forma determinante para o desenvolvimento de sistemas mais adequados para a manutenção da diversidade das comunidades de peixes além de acelerar e aumentar a exatidão das decisões tomadas para a conservação e manejo das espécies em questão, pois trarão subsídios técnicos capazes de minimizar o impacto gerado pela construção de hidrelétricas.

## 6) REFERÊNCIAS BIBLIOGRÁFICAS:

Agostinho, A. A., Vazzoler, A. E. E. M., Gomes, L. C.; Okada, E. K, 1993. Estratificación y comportamiento de *Prochilodus scrofa* em distintas fases del ciclo de vida, em la planície de inundação del alto rio Paraná y embalse de Itaipu, Paraná, Brasil. Rev. Hydrobio. Trop. 26(1): 79-90.

Agostinho, A. A.; Gomes, L. C.; Suzuki, H. I.; Júlio Jr., H. F., 2003. Migratory fish from the upper Parana river basin, Brazil. Em: J. Carolsfeld, B. Harvey, C. Ross, A. Baer. Migratory fishes of South America: biology, fisheries and conservation status. World Fisheries Trust, British Columbia, Canadá.

Agostinho AA, Gomes LC, Pelicice FM. 2007. Ecologia e manejo de recursos pesqueiros em reservatórios do Brasil. Maringá: Eduem.

Amos, W.; Hoffman, J.I.; Frodsham, A. et al., 2007 Automated binning of microsatellite alleles: problems and solutions. Molecular Ecology Notes 7: 10-14.

Anne, C., 2006. Choosing the right molecular genetic markers for studying biodiversity: from molecular evolution to practical aspects. Genetica, 127(1-3):101-120

Araki H, Cooper B, Blouin S., 2007. Genetic effects of captive breeding cause rapid, cumulative fitness decline in the wild. Science, 318: 100-103.

Avise, J. C.; Arnold, J.; Ball, R. M.; Bermingham, E.; Lamb, T.; Neigel, J. E.; Reeb, C. A.; Saunders, N. C., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematic. Annu. Rev. Ecol. Syst. 18: 489-522.

Barbieri, G.; Hartz, S.M.; Verani, J.R., 1996. O fator de condição e índice hepatossomático como indicadores do período de desova de *Astyanax fasciatus* da represa do lobo, São Paulo (Osteichthyes, Characidae). Iheringia 81:97-100.

Barbosa, A.C.D.R.; Galzerani, F.; Corrêa, T.C.; Galetti, P.M.; Hatanaka, T., 2008. Description of novel microsatellite loci in the Neotropical fish *Prochilodus argenteus* and cross-amplification in *P. costatus* and *P. lineatus*. *Genetics and Molecular Biology* 31(1): 357-360.

Barnett, V. e Lewis, T., 1994. *Outliers in Statistical Data* (3rd edition), Wiley.

Bennemann, S.T.; Orsi, M.L.; Shibatta, O., 1996. Atividade alimentar de species de peixes do rio Tibagi, relacionada com o desenvolvimento de gorduras e gônadas. *Revista Brasileira de Zoologia*. 13(2): 501-512.

BioPHP- PHP for bioinformatics, 2008. Disponível *on line* em [http://biophp.org/minitools/microsatellite\\_repeats\\_finder/demo.php](http://biophp.org/minitools/microsatellite_repeats_finder/demo.php). Acesso em: junho, 2008.

Bonetto, A. A., 1963. Investigaciones sobre migraciones de peces em los rio de La cuenca del Plata. *Ciencia e Investigación* 19 (12): 12-26.

Braga, F. M. S., 2001. Reprodução de peixes (Osteichthyes) em afluentes do reservatório de Volta Grande, Rio Grande, sudeste do Brasil. *Iheringia, Ser. Zool*, 91: 67-74.

Braslavsky, I.; Hebert, B.; Kartalov, E.; Quake, S. R., 2003. Sequence information can be obtained from single DNA molecules. *Proc. Natl. Acad. Sci. USA* 100 (7): 3960–4. doi : 10.1073/pnas.0230489100 : PMID 12651960 .

Carosfeld, J., Harvey, B., Ross, C., Baer, A., 2003. *Migratory fishes of South America: biology, fisheries and conservation status*. Victoria: World Fisheries Trust.

Carroll, R. C., Meffe, G. K., 1997. *Principals of Conservation Biology*, 2nd ed, pp.161-201. Sinauer Associates, Sunder Land, Mass.

Castro, R.M.C.; Vari, R.P., 2004. Detritives of the South American Fish Family Prochilodontidae (Teleostei: Ostariophysi: Characiformes): A Phylogenetic and Revisionary Study. 187p. Washington: Smithsonian.

Castro, R.M.C.; Vari, R.P. 2003. Family Prochilodontidae. In: Reis R.E.; Kullander S.O.; Ferraris Jr., C.J. Check List of the Freshwater Fishes of South and Central America, EDIPUCRS 1: Porto Alegre, 65-70.

CBH-PARÁ. Projeto Água e Vida. Comitê da Bacia Hidrográfica do Rio Pará, 2001. Disponível em <<http://www.cbhpara.org.br/projetos.htm>>. Acesso em: janeiro 2009.

CEMIG- Companhia Energética de Minas Gerais, 2009. Sustentabilidade. Disponível online em <<http://www.cemig.com.br>>, acesso: fevereiro de 2009.

Costa, L.F.C.; Hatanaka, T.; Galetti, P.M., 2008. Evidence of lack of population substructuring in the Brazilian freshwater fish *Prochilodus costatus*. Genetics and Molecular Biology 31(1):377-380.

Chenuil, A., 2006. Choosing the right molecular markers for studying biodiversity: from molecular evolution to practical aspects. Genética, 127: 101-120.

Clement, M.; Posada, D.; Crandall K., 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9(10): 1657-1660 DOI 10.1046/j. 1365-294x.2000.01020.x

Darnell, R. M., 1961. Trophic spectrum of an estuarine community based upon studies of Lake Pontchartrain, Louisiana. Ecology, Temple, 42:552-568.

Dergam, J. A., Paiva, S.R., Schaeffer, C. E., Godinho, A. L., Vieira, F., 2002. Phylogeography and RAPD-PCR variation in *Hoplias malabaricus* (Bloch, 1794) (Pisces, Teleostei) in southeastern Brazil. *Gen Mol Biol*, 25 (4): 379-387.

Dobzhansky, T., 1955. Evolution: genetics and man. Wiley. New York.

Ellegren, H., 2004. Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, 5: 435 – 445.

Estoup, A., and B. Angers., 1998. Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. Pp. 55–86 in G. Carvalho, ed. *Advances in molecular ecology*. IOS Press, Amsterdam

Excoffier, L.; Smouse, P.E.; Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Excoffier, LG. Laval, and S. Schneider .2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

Ewing, B.; Green, P.,1998. Basecalling of automated sequencer traces using phred. II. Error probabilities. *Genome research* 8: 186-194.

Ferreira, M. E.; Grattaglia, D., 1998. Introdução ao uso de marcadores moleculares em análise genética. 3<sup>a</sup> Ed. Brasília: EMBRAPA-CENARGEN, pp.220

Fiers, W.; Contreras, R.; Duerinck, F., 1976. Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature* 260 (5551): 500–7. doi : 10.1038/260500a0 . PMID 1264203.

FishBase- A Global Information System on Fishes. 2004. FishBase Book 2000. Disponível *online* em <<http://www.fishbase.org>>, acesso: fevereiro de 2009.

Flecker, A. S., 1996. Ecosystem engineering by a dominant detritivore in a diverse tropical stream. *Ecology* 77: 1845-1854.

FAO/DVWK (Food and Agriculture Organization/German Association for Water Management and Improvement). 2002. Fish passes: design, dimensions and monitoring. FAO, Rome, Italy.



Frankham, R., 2003. Genetics and conservation biology. *Comptes Rendes Biologies*, 326: S23 -S29.

Frankham, R.; Ballou, J. D.; Briscoe, D. A., 2004. A primer of conservation genetics. Cambridge University Press, Cambridge.

Fraser, DJ., 2008. How well can captive breeding programs conserve biodiversity? A review of salmonids. *Evolutionary Applications*, DOI:10.1111/j.1752-4571.2008.00036.x.

Gilbert, W.; Maxam, A., 1973. The nucleotide sequence of the lac operator. *Proc. Natl. Acad. Sci USA* **70** (12): 3581-4. Doi: 10.1073/pnas.70.12.3581. PMID 4587255.

Godinho, H. P.; Godinho, A. L., 1994. Ecology and conservation of fish in southeastern brazilian river basins submitted to hydroelectric impoudments. *Acta Limnologica Brasiliensa*, vol V: 187-197. Workshop: Brazilian programme on conservation and management on inland water.

Gordon, D.; Abajian, C., Green, P., 1998. Consed: A graphical tool for sequence finishing.

GraphPad InStat version 3.00 for Windows 95, GraphPad Software, 1995, San Diego California USA, disponível *online* em: [www.graphpad.com](http://www.graphpad.com), acesso: agosto 2009.

Hall, N., 2007. Advanced sequencing technologies and their wider impact in microbiology. *J. Exp. Biol.* 210 (Pt 9): 1518–25. Doi: 10.1242/jeb.001370. PMID 17449817.

Hartl, D. L., Clark, A. G., 1989. Principles of population genetics, 2nd ed, Sinauer Associates, Sunder Land, Mass.

Hatanaka, T., Henrique-Silva, F., Galetti, Jr. P.M., 2006. Population substructutring in a migratory freshwater fish *Prochilodus argenteus* (Characiformes, Prochilodontidae) from the São Francisco River. *Genetica*, 126: 1-7.

- Hatanaka, T., Galetti, Jr. P.M., 2003. RAPD markers indicate the occurrence of structured populations of a migratory freshwater fish species. *Genetics and Molecular Biology*, 26: 19-25.
- Hearne CM, Ghosh S, Todd JA. 1992. Microsatellite for linkage analysis of genetic traits. *T.I.G.*, 8: 288-294.
- Hickley P., 1994. Stocking and introduction of fish: a synthesis. In: COWX, I.G. (ed.). *Rehabilitation of freshwater fisheries*. Oxford: Fishing News Books, p. 247-254.
- Higgins DH, Sharp PM. 1998. Clustal: A package for performing multiple sequence alignment on a microcomputer. *Gene*, 73, p. 237-244.
- Hillis, D. M.; Moritz, C.; Mable, B. K., 1996. *Molecular Systematics*. Sinauer Associates, Inc. E.U.A.
- Hilsdorf, A.; Krieger, J. E., 1998. *Biologia molecular na conservação de peixes*. Biotecnologia, ciência e desenvolvimento, 1(5): 10 - 12.
- IGAM- Instituto Mineiro de Gestão das Águas, 2009. Disponível em: <<http://www.igam.mg.gov.br>>. Acesso em: Janeiro de 2009.
- Junk, W.J.; Bayley, P.B.; Sparks, R.E., 1989. The flood pulse concept in river-floodplain systems. *Canadian Journal of Fishers and Aquatic*, 106: 110-127.
- Katti, M. V., Ranjekar. P. K., Gupta. v. s., 2001. Differential Distribution of Simple Sequence Repeats in Eukaryotic Genome Sequences. *Molecular. Biological Evolution*. 18(7):1161–1167.
- Kimmel, B. L.; Groeger, A. W., 1986. Limnological and ecological changes associated with reservoir aging. In: Hall, G. E.; Van Den Avyle, M. J. (Eds), 1986. *Reservoir*

fisheries management: strategies for the 80's. American fisheries Society. Bethesda. p.103-109.

Le Cren, E. D., 1951. The length-weight relationship and seasonal cycle in godad weight and condition in the perch (*Perca fluviatilis*) Journal of Animal Ecology, v.20,n.2, p. 201-209.

Li, C.; Bessert, M.L.; Macrander, J.; Ortí, G., 2009. Low variation but strong population structure in mitochondrial control region of plains topminnow, *Fundulus sciadicus*. Journal of Fish Biology 74: 1037-1048.

Madeira, M.J.; Gómez-Moliner, B.J.; Barbe, A.M., 2005. Genetic introgression on freshwater fish populations caused by restocking programmes. Biol Invasions, 7: 117–125.

Martins, S. L., Tamada, K., 2000. Sistemas para a transposição de peixes (Boletim Técnico da Escola Politécnica da USP, Departamento de Engenharia Hidráulica e Sanitária, BT/ PHD/72). São Paulo: EPUSP.

Medeiros, E. S. F.; Maltchick, L., 2001. Fish assemblage stability in an intermittently stream from the Brazilian semiarid region. Austral Ecology 26, 156–164.

Miller, L. M. e Kapuscinski, A. R., 2003. Genetic guidelines for hatchery supplementation programs. In E. M. Hallerman, ed Population Genetics: Principles and Practices for Fisheries Scientists. American Fisheries Society, Bethesda: 329-355.

NCBI- National Center for Biotechnology Information, 2008. VecScreen. Disponível online em <http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>. Acesso: maio de 2008.

Neraas, L. P.; Spruell, P., 2001. Fragmentation of riverine systems: the genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. Molecular Ecology, 10: 1153-1164.

Orsi, M.L.; Foresti, F.; Carvalho, E.D., 2000. Relação peso/comprimento e composição em comprimento de *Astyanax bimaculatus* em quatro trechos do rio Tibagi, PR, p.363. In: Congresso Brasileiro de Zoologia, 23, Cuiabá. 746p.

Paiva, A. L. B., Kalapothakis, E., 2008. Isolation and characterization of microsatellite loci in *Pimelodus maculatus* (Siluriformes: Pimelodidae). Molecular Ecology Resources 07/487 (in press). doi: 10.1111/j.1755-0998.2008.02160.x

Paiva, A. L. B., 2007. Isolamento e caracterização de microssatélites de *Pimelodus maculatus* (Pisces, Siluriformes). Dissertação de mestrado, Programa de pós-graduação em genética, Universidade Federal de Minas Gerais, Belo Horizonte, 75 p.

Parker, P. G., Snow, A. A., Schug, M. D., Booton, G. C., Furerst, P. A., 1998. What molecules can tell us about populations: choosing and using a molecular marker. Ecology 79(2): 361-382.

Pellegrin, L. A. Técnicas de sensoriamento remoto e geoprocessamento aplicadas ao mapeamento do uso do solo: a bacia do rio Pará como exemplo, 2001, 109 f. Dissertação de Mestrado – PUCMinas, Belo Horizonte, Minas Gerais, 2001.

Petts, G. E., 1989. Perspectives for ecological management of regulated rivers. In: Gore, A.; Petts, G. E. (ed). Alternatives in regulated river management. Boca Raton: CRC Press: 3-24.

Petrere, M., 1989. River Fisheries in Brazil: a review. In: Regulated Rivers: Research and Management. Vol. 4: 1-16.

Poke, F. S. *et al.*, 2005. Genomic research in Eucalyptus. Genética, vol.125, p.79-101.

Portal Peixe Vivo, 2007. Rios de Minas. Disponível *online* em <<http://www.portalpeixevivo.com.br>>, acesso: janeiro de 2009

Pritchard, J. K.; Stephens, M.; Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Rocha, M. A.; Ribeiro, E.L.A.; Mizubuti, I.Y., 1997. Comparação entre os fatores de condição de Fulton e alométrico em curimatá (*Prochilodus lineatus*) criados em dois ambientes. *Arch. Latinoam. Prod. Anim.* 5(Supl. 1): 459-460.

Rozen, S. e Skaletsky, H. J., 2000. Primer3 on the www for general users and for biologist programmers. In: Krawetz, S.; Misener, S. (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.

Rosenblum, B.B., Oaks, F., Menchen S. et al., 1997. Improved single-strand DNA sizing accuracy in capillary electrophoresis *Nucleic Acids Research* 25: 3925-3929.

Salvanes AGV, Braithwaite V., 2006. The need to understand the behaviour of fish reared for mariculture or restocking. *J Mar Sci*, 63: 346-354.

Sambrook, J. and Russell, D., 2001. *Molecular cloning: A laboratory manual*, 3rd Ed. Cold Spring harbor Laboratory Press, Cold Spring Harbor, NY, USA, 682p.

Sanger, F., Coulson, A. R., 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Biology* 94 (3): 441–8. doi : 10.1016/0022-2836(75)90213-2 . PMID 1100841.

Sanger, F.; Nicklen, S.; Coulson, A. R., 1977. DNA sequencing with chain-terminating inhibitors . *Proc. Natl. Acad. Sci. USA* 74 (12): 5463–7. Doi: 10.1073/pnas.74.12.5463. PMID 271,968.

Sato, Y. e Godinho, H.P. Migratory Fishes of the São Francisco River. In:Carolsfeld, J. *et al.* *Migratory Fishes of South America: biology, fisheries and conservation status*. Victoria: IDRC, 2004. 380 p.

Schlötterer, C., 2000. Evolutionary dynamics of microsatellite DNA. *Chromosoma*, 109(6): 365-371.

Schlötterer C. 2004. The evolution of molecular markers – just a matter of fashion? *Nat Rev Gen*, 5: 63-69.

Shendure, J., 2005. Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome. *Science* 309 : 1728. Doi : 10.1126/science.1117389 .

Sivasundar, A.; Bermingham, E.; Ortí, G, 2000. Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Charasiformes) in major south American rivers. *Molecular Ecology*, Lincoln, vol.10 (2), p. 407-417.

Suzuki, D. T., Griffiths, A. J. F., Miller, J. H., Lewontin, R. C., Gelbart, W. M., 1998. *Introdução à Genética*, 2a edição, Rio de Janeiro: Guanabara Koogan.

The Sequence Manipulation Suite, 2000. Reverse Complement. Disponível *online* em<<http://www.bioinformatics.org>>, acesso: julho de 2008).

Tamura K, Dudley J, Nei M, Kumar S.2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*24: 1596-1599.

Templeton, A. R., 1979. Once again why, 300 species of Hawaiian *Drosophila*? *Evolution* 33: 513-517.

Templeton, A. R., 1980. The theory of speciation via the founder principle. *Genetics*: 94: 1011-1038.

Utter, F. M., 2003. Genetics impacts of fish introductions. In: EM Hallerman ed, *Population genetics: principles and applications for fisheries scientists*. American Fisheries Society, 357-378.

Vazzoler, A. E. A. M.; Lizama, M. A. P.; Inada, P., 1997. Influências ambientais sobre a sazonalidade reprodutiva. p. 267-280. In: Vazzoler, A. E. A. M.; Agortinho, A. A.;

Hahn, E. N. S. (ed). A planície de inundação do alto rio Paraná: aspectos físicos, biológicos e socioeconômicos . Maringá: EDUEM, p.460.

Vincentin, W.; Costa, F.E.S.; Marques, S.P.; Zuntini, D.; Barbosa, E.G., 2004. Fator de condição e relação peso X comprimento de *Prochilodus lineatus*, capturados na cabeceira do Rio Miranda, MS. IV Simpósio sobre Recursos Naturais e Sócio Econômicos do Pantanal. Corumbá, MS.

Vrijenhoek, R. C., 1998. Conservation genetics of freshwater fish. Journal of Fish Biology 53 (Supplement A), 394 - 412.

Welcomme, R.L., 1979. Fisheries ecology of floodplain rivers. London: Logman Group Limited. 316p.

Welcomme, R.L., 1989. Floodplain fisheries management. In: Gore, J. A.; Petts, G.E. Alternatives in regulated river management. Boca Raton: CRC Press, p.210-233.

Yazbeck, G. M., 1997. Microsatélites em estudos populacionais de peixes migratórios, 139 f. Tese de Doutorado – Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais.

Yazbeck, G. M. e Kalapothakis, E., 2007. Isolation and characterization of microsatellite DNA in the piracema fish *Prochilodus lineatus* (Characiformes). Genetics and Molecular Research 6 (4): 1026-1034.

Zane, L., Barceloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. Molecular ecology 11: 1-16.

**ANEXOS**



**Peixamento com curimatã piao (*Prochilodus costatus*) – a montante da barragem de Três Marias (alto São Francisco)**

<b>Data</b>	<b>Local</b>	<b>Quantidade (UN)</b>
04/1983	Represa Três Marias / Três Marias	30.000
09/1984	Represa Três Marias / Três Marias	35.000
03/1985	Represa Três Marias / Três Marias	50.000
03/1986	Represa Três Marias / Três Marias	50.000
03/1986	Rio Paraopeba / Betim	15.000
03/1986	Rio Pará / Martinho Campos	12.000
03/1986	Represa Gafanhoto / Divinópolis	10.000
03/1986	Represa Cajuru / C do Cajuru	10.000
09/1986	Rio Pará / Martinho Campos	15.000
09/1986	Represa Três Marias / Morada Nova de Minas	44.000
09/1986	Rio São Francisco / Bom Despacho	25.000
04/1987	Rio Pará / Martinho Campos	22.000
04/1987	Represa Gafanhoto / Divinópolis	7.500
04/1987	Represa Cajuru / Carmo do Cajuru	7.500
04/1987	Represa Três Marias / Três Marias	35.000
04/1987	Rio Paraopeba / Betim	15.000
04/1987	Represa Três Marias / Morada Nova de Minas	13.500
04/1987	Rio São Francisco / Bom Despacho	15.000
04/1987	Represa Cajuru / C do Cajuru	7.000
04/1987	Rio Pará / Martinho Campos	14.500
04/1988	Rio Pará / São S do Oeste	9.500
04/1988	Rio Pará / S G do Pará	8.000
04/1988	Rio Pará / Divinópolis	12.000

04/1988	Represa Gafanhoto / Divinópolis	8.500
04/1988	Represa Cajuru / C do Cajuru	8.000
04/1988	Represa Três Marias / Três Marias	25.000
04/1988	Rio S Fco / Pompéu	11.000
04/1988	Rio São Fco / Bom Despacho	13.000
03/1989	Rio S Fco / Pompéu	30.000
03/1989	Rio São Fco / Bom Despacho	18.000
03/1989	Represa Três Marias / Três Marias	35.000
03/1989	Represa Cajuru / C do Cajuru	12.000
04/1989	Rio Pará / São S do Oeste	10.000
04/1989	Rio Pará / S G do Pará	10.000
04/1989	Rio Pará / Divinópolis	12.000
04/1989	Rio Pará / Martinho Campos	15.000
04/1989	Represa Gafanhoto / Divinópolis	10.000
04/1989	Represa Três Marias / Felixlandia	14.000
04/1990	Rio S Fco / Pompéu	10.000
04/1990	Rio São Fco / Bom Despacho	20.000
04/1990	Represa Três Marias / Morada Nova de Minas	30.000
09/1990	Rio Paraopeba / Betim	15.000
09/1990	Represa Três Marias / Três Marias	59.500
09/1990	Represa Cajuru / C do Cajuru	10.000
09/1990	Rio Pará / Martinho Campos	15.000
03/1992	Represa Três Marias / Três Marias	36.500
03/1992	Rio Pará / Martinho Campos	12.000
04/1993	Represa Gafanhoto / Divinópolis	15.000
04/1993	Represa Cajuru / C do Cajuru	15.000

09/1993	Rio Pará / Martinho Campos	15.000
09/1993	Represa Três Marias / Três Marias	18.800
04/1994	Rio Pará / Martinho Campos	24.500
04/1994	Rio Paraopeba / Betim	15.000
04/1994	Represa Cajuru / C do Cajuru	21.000
04/1994	Rio S Fco / Pompéu	23.100
09/1995	Rio Pará / Martinho Campos	16.500
09/1994	Rio Paraopeba / Betim	12.000
09/1995	Represa Três Marias / Três Marias	15.500
03/1996	Represa Três Marias / Três Marias	43.000
03/1996	Represa Três Marias / Morada Nova de Minas	20.000
04/1999	Rio Paraopeba / Betim	15.000
04/1999	Rio São Fco / Bom Despacho	20.000
09/1999	Represa Gafanhoto / Divinópolis	25.000
09/1999	Represa Cajuru / C do Cajuru	25.000
10/1999	Represa Três Marias / Três Marias	36.680
10/1999	Represa Três Marias / Morada Nova de Minas	30.000
10/1999	Rio S Fco / Pompéu	25.000
10/1999	Rio Pará / Martinho Campos	17.000
11/1999	Represa Três Marias / Três Marias	3.675
02/2000	Represa Três Marias / Morada Nova de Minas	14.300
02/2000	Represa Três Marias / Três Marias	4.620
05/2000	Represa Três Marias / Três Marias	15.800
04/2002	Rio São Fco / Moema	17.500
04/2002	Represa Cajuru / C do Cajuru	12.000
04/2002	Rio Pará / São S do Oeste	12.500

04/2002	Rio Pará / S G do Pará	12.600
04/2002	Rio Pará / Divinópolis	12.000
10/2002	Represa Três Marias / Três Marias	16.740
10/2002	Represa Cajuru / C do Cajuru	5.500
10/2002	Represa Gafanhoto / Divinópolis	5.960
11/2002	Rio Pará / São S do Oeste	1.600
11/2002	Represa Três Marias / Três Marias	10.880
10/2003	Represa Gafanhoto / Divinópolis	15.300
10/2004	Represa Cajuru / C do Cajuru	13.300
11/2004	Represa Três Marias / Três Marias	9.500
09/2005	Represa Três Marias / Três Marias	5.800
09/2005	Represa Cajuru / C do Cajuru	3.000
09/2005	Represa Gafanhoto / Divinópolis	4.860
10/2005	Represa Três Marias / Três Marias	15.980
10/2005	Rio São Fco / Iguatama	9.720

## Sequências dos 25 haplótipos encontrados para o marcador Molecular HPL

Hap1 TCTTTGCTTGTGCTTGCTCTCTC--TTTCTCTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTTCTC-  
CAACCCTTCTCTCTCTCTCTTATCTCTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTCTTCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTCCATGTCTCTTTCTCTCTCT  
TCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCCAGCAAACCTCTCATGCTGTCTCTGCCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap2 TCTTTGCTTGTGCTTGCTCTCTCTCTTTCTCTTGTCTCTCCCTTGTCCC-ACCTTCCGAATTGATTATCATCTCTCTGCCTTTCATCTTTGTCAACC----  
CTCCAACCCTTCTCTCTCTCTCTTATCTCTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCACATTGCCTCCTTCACTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTCCATGTCTCTTTCTCTC  
TCTTCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap3 TCTTTGCTTGTGCTTGCTCTCTCTC--TTTCTCTTGTCCCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTT-  
CTTCAACTCTTCTCTCTCTCTTTTTCTCTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCACATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTCCATGTCTCTCTCTC  
TCTTCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap4 TCTTTGCTTGTGCTTGCTCTCTCTCTTTCTCT-TTTGTCTCTCCCTCTGTCCCCACCTTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCAACC----  
CTCCAACCTTCTCTCTCTCTCTTATCTCTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCACATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTCCATGTCTCTCTCTC  
TCTTCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCCAGCAAACCTCTCATGCTGTCTCTGCCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap5 TCTTTGCTTGTGCTTGCTCTCTCTC--TTTCTCTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTATGCCTTTCATCTTTGTCACTTTTT-  
CTTCAACTCTTCTCTCTCTCTTTTTCTCTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCACATTGCCCCCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTCCATGTCTCTCTCTC  
TCTTCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap6 TCTTTGCTTGTGCTTGCTCTCTCTCTTTCTCTTGTCTCTCCCTTGTCCC-ACCTTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCACTTTTT-  
CTCCAACGCTTCTCTCTCTCCCTTTTTCTCTTCTTTTTTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCATATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCTGTCTCTTTCTCCATGTCTCT-----  
CCCTCTTTCTCTCTTGTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap7 TCTTTGCTTGTGTCTTGCTCTCTC--TTTCTCTGTCTCTCCCTTTGTCCC-ACCTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCACTTTTT-  
CTCCAACGCTTCTCTCTCTCTCTTTTTTCTCTCTTTTTTAACTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCATATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTCCATGTCTCTTTCTCTC  
TCTTCTCTCTTTCTCTCTTGCTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCCTCTGCCGTTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap8 TCTTTGCTTGTGTCTTGCTCTCTC--TTTCTCTTGTCTCTCCCTCTGTCCC-ACCTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTT-  
CTTCAACTCTTCTCTCTCTCTCTTTTTTCTCTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTCTTCTCTCTCCCC  
CATATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCT  
CTTCTCTCTTTCTCTCTTGCTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCCTCTGCCGTTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap9 TCTTTGCTTGTGTCTTGCTCTCTC--TTTCTCTTGTCTCTCCCTTTGTCCC-  
AGGGTCCCTTCTCCCTCTATCATTTCTCTGCCTTTCATCTTTGTCACTTTTAAAAACAACGCTTCTCTCTCTCTTTTTTCTCTCTTTTTTAACTCCCTCAGTGCCATTTTCTGAT  
GTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTCTTCTCTCTCTCCCCATATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTT  
CTTTCCCACTCACCGTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTTTCTCTCTCTTTCTCTCTTTCTCTCTTTGTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCA  
TGCTGTCTCTGCCGTTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap10 TCTTTGCTTGTGTCTTGCTCTCTCTCTTTCTCTTGTCTCTCCCTTTGTCCC-ACCTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCACTTTTTCTC-  
CAACGCTTCTCTCTCTCTCTTTTTTCTCTCTTTTTTAACTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTCTTCTCTCTCCCCA  
TATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCGTCTCTTTCTCCATGTCTCT-----  
CCCTCTTTCTCTCTTGCTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCCTCTGCCGTTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap11 TCTTTGCTTGTGTCTTGCTCTCTCTCTTTCTCTTGTCTCTCCCTCTGTCGC-ACGTTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCAACC----  
CTCCAACCCTTCTCTCTCTCTCTTATCTCTTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCACATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTTTCTCTC  
TCTTCTCTCTTTCTCTCTTGCTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCCTCTGCCGTTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap12 TCTTTGCTTGTGTCTTGCTCTCTCTCTTTCTCTTGTCTCTCCCTTTGTCCC-ACCTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCACTTTTT-  
CTCCAACGCTTCTCTCTCTCTCTTTTTTCTCTCTTTTTTAACTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTCTTCTCTCTCCC  
CCATATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCGTCTCTTTCTCCATGTCTCT-----  
CCCTCTTTCTCTCTTGCTGTCCCTTTTTCTGTCAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCCTCTGCCGTTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap13 TCTTTGCTTGTGTCTTGCTCTCTCTCTTTCTTTGTCTCTCCCTTTGTCCC-ACCTTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCACTTTTTCTC-  
CAACGCTTCTCTCACTCTCTTTTTTCTCTCTTTTT-  
AACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCCCTCCATGTCTCTACCTCTTTTTCTTTCTCTCTCCCCATATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTC  
TCCATCTATTTCTTATTCTGCGCTTTCTTTCCCACTCACCGTCTTTCTTTCTGTCTCTTTCTCCATGTCTCT-----  
CCCTCTTTCTCTCTTGTGTCTTTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap14 TCTTTGCTTGTGTCTTGCTCTCTCTCTTTCTCT-TT-GTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCAACC---CTC-  
CAACCTTCTCTCTCTCTCTCTTTATCTCTTTCTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTTTCTCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCTCT  
TCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap15 TCTTTGCTTGTGTCTTGCTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTTCTT-  
CAACTCTTCTCTCTCTCTCTTTTTTCTTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTTTCTCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCTCT  
TCTCTCTTTCTCTCTTTGTGCCCTTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap16 TCTTTGCTTGTGTCTTGCTCTCTCTC--TTTCTCTGTCTCTCCCTTTGTCCC-ACCTTCCCTTCTCACTCTATCATTTCTCTGCCTTTCATCTTTGTCACTTTTTCTC-  
CAACGCTTCTCTCTCTCTCTTTTTTCTTTCTTTTTTAACTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTTTCTCTCTCTCCCCA  
TATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCATCTATTTCTTATTCTACGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTTTCTCTCTCT  
TCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap17 TCTTTGCTTGTGTCTTGCTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTTCTT-  
CAACTCTTCTCTCTCTCTCTTTTTTCTTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGTCTCTACCTCTTTTTCTTTCTCTCTCTCCCCA  
TATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCTCT  
TCTCTCTTTCTCTCTTTGTGCCCTTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap18 TCTTTGCTTGTGTCTTGCTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTTCTT-  
CAACTCTTCTCTCTCTCTCTTTTTTCTTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTTCTCTCTCTCCATGCCTCTACCTCTTTTTCTTTCTCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCCCATCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCTCT  
TCTCTCTTTCTCTCTTGTGTCCCTTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTTTCTGAGTGCACTGTATTTAGGCCTT

Hap19 TCTTTGCTTGTGCTTGTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTCTT-  
CAACTCTTCTCTCTCTCTCTCTTTTCTATTCTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTCTCTCTCTCCATGCCTCTACCTCTTTTCTCTTCCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCATCTATTTCTATTCTGCGTCTTTCTTTCCCACTCACCCTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCTCT  
TCTCTCTTCTCTCTTGTGTCCCTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTCTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap20 TCTTTGCTTGTGCTTGTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTCTC-  
CAACCTTCTCTCTCTCTCTCTTATCTCTTCTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTCTCTCTCTCCATGCCTCTACCTCTTTTCTCTTCCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCACCTATTTCTATTCTGCGTCTTTCTTTCCCACTCACCCTCTTTCTTTTCCCGTCTCTTTCTTCATGTCTCTTTCTCTCTCT  
TCTCTCTTCTCTCTTGTGTCCCTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTCTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap21 TCTTTGCTTGTGCTTGTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTCTT-  
CAACTCTTCTCTCTCTCTCTTTTCTCTTCTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTCTCTCTCTCCATGCCTCTACCTCTTTTCTCTTCCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCCCATCTATTTCTATTCTGCGTCTTTCTTTCCCACTCACCCTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCT--  
TCTCTCTTCTCTCTTTTGTCCCTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTCTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap22 TCTTTGCTTGTGCTTGTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCAACC---CTC-  
CAACCTTCTCTCTCTCTCTCTTATCTCTTCTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTCTCTCTCTCCATGCCTCTACCTCTTTTCTCTTCCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCATCTATTTCTATTCTGCGTCTTTCTTTCCCACTCACCCTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTTTCTCTCTCT  
TCTCTCTTCTCTCTTGTGTCCCTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTCTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap23 TCTTTGCTTGTGCTTGTCTCTCTC--TTTCTCTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTCTT-  
CAACTCTTCTCTCTCTCTCTTTTCTCTTCTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTCTCTCTCTCCATGCCTCTACCTCTTTTCTCTTCCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCCCATCTATTTCTATTCTGCGTCTTTCTTTCCCACTCACCCTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCTCT  
TCTCTCTTCTCTCTTTTGTCCCTTTTCTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTCTTCTGAGTGCACTGTATTTCAAGGCCTT

Hap24 TCTTTGCTTGTGCTTGTCTCTCTC--TTTCTCTTTGTCCCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACTTTTCTT-  
CAACTCTTCTCTCTCTCTCTTTTCTCTTCTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTCTCTCTCTCCATGCCTCTACCTCTTTTCTCTTCCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCCATCTATTTCTATTCTGCGTCTTTCTTTCCCACTCACCCTCTTTCTTTTCCCGTCTCTTTCTCCATGTCTCTCTCTCTCT  
TCTCTCTTCTCTCTTTTGTCCCTTTTGTGAGCCTTTGTTTCATCACCTGCAAACCTCTCATGCTGTCTCTGCGTCTTCTGAGTGCACTGTATTTCAAGGCCTT



Hap25            TCTTTGCTTGTGTCTTGCTCTCTC--TTTCTTTTGTCTCTCCCTCTGTCCC-ACCTTCCCTTCTCACTCTATCATCTCTCTGCCTTTCATCTTTGTCACCTTTTCTC-  
CAACCCCTTCTCTCTCTCTCTCTTATCTCTTCTTTTTCTAACCTCCCTCTGTGCCATTTTCTGATGTCACTCTTTTCTCTCTCTCCATGCCTCTACCTCTTTTCTCTTCCTCTCTCCCCA  
CATTGCCTCTTTCACCTTCTCCCTCTGTCTCTCTCTCTCCACCTATTTCTTATTCTGCGTCTTTCTTTCCCACTCACCGTCTTTCTTTCCCGTCTCTTTCTTCATGTCTCTTCTCTCT  
TCTCTCTTCTCTCTTGTGTCCCTTTTTTGTGAGCCTTTGTTTCATCACCCAGCAAACCTCTCATGCTGTCCTTGCCGTTTCTGAGTGCACTGTATTCAGGCCTT