

**SUBPROJETO III A. DINÂMICA POPULACIONAL E
EVOLUTIVA DA ICTIOFAUNA NA BACIA DO RIO DOCE
- ANÁLISE POPULACIONAL COM USO DE
MARCADORES RADSEQ**

**Projeto: Estudo da ictiofauna da bacia do Rio
Doce após o rompimento da barragem de rejeito
da Samarco**

Dra. Natália Maria de Freitas Vicente – UFV

Dra. Jorge Dergam – UFV

Dr. Pedro Vidigal – UFV

Isabela Malo – UFV

Janeiro, 2023

SUMÁRIO

SUMARIO EXECUTIVO	4
1. OBJETIVOS	5
2.....	INTRODUÇÃO 5
2.1 Caracterização do problema: o desastre de Mariana	5
2.2 Histórico da degradação ambiental da bacia do Rio Doce	6
2.3 Impactos do rompimento da barragem de “Fundão” sobre os habitats e sua biota	7
3.MATERIAIS	E MÉTODOS 8
3.1 Espécies estudadas	8
3.2 Desenho Amostral	9
3.3 Extração e diluição do DNA.....	13
3.4 Preparação da biblioteca de RADSeq	14
3.5 Controle de qualidade, trimagem e filtragem	15
3.6 Mapeamento genômico	15
3.7 Análise dos polimorfismos e das populações	16
4.....	RESULTADOS 17
4.1 <i>Astyanax lacustris</i>	17
4.2 <i>Loricariichthys castaneus</i>	23
4.3 <i>Geophagus brasiliensis</i>	28
5.....	DISCUSSÃO 33
5.1 Variabilidade e estruturação genética das populações de peixes dos tributários e da calha principal da bacia do rio Doce	33
5.2 Conhecimento sobre a diversidade e a distribuição da ictiofauna da bacia do rio Doce 35	
5.3 Conectividade entre os tributários e a calha principal do rio Doce	36
6..	CONCLUSÃO 37
7.....	REFERÊNCIAS 38

LISTA DE FIGURAS

Figura 1. O Acará <i>Geophagus brasiliensis</i> (Quoy & Gaimard 1824)	8
Figura 2. O Lambari <i>Astyanax lacustris</i> (Lütken 1875).....	9
Figura 3.O Cascudo, <i>Loricariichthys castaneus</i> (Castelnaud, 1855).....	9
Figura 4. Mapa da bacia hidrográfica do Rio Doce. Pontos em verde são pontos-referência, representados por afluentes que não tiveram contato com o rejeito. As áreas afetadas pelo rejeito estão representadas por pontos vermelhos.....	10

Figura 5. Valores do F_{ST} par a par entre populações para todos os indivíduos de <i>Astyanax lacustris</i> . Segundo Wright (1978) F_{ST} entre 0 e 0.05 indica pouca diferenciação genética; F_{ST} entre 0.05 e 0.15 indica diferenciação genética moderada; F_{ST} entre 0.15 e 0.25 indica alta diferenciação genética e F_{ST} acima de 0,25 diferenciação genética muito alta.	20
Figura 6. Análise Discriminante de Componentes Principais das variações de SNPs gerados pela ferramenta RAD-seq, obtidas em populações do Lambari-de-rabo amarelo <i>Astyanax lacustris</i>	21
Figura 7. Dendograma hierárquico para os indivíduos analisados de <i>A. lacustris</i> (cada um deles é representado por letras e números únicos), gerado pela matriz de distâncias pareadas calculada partir da tabela de genotipagem usando a distância Euclidiana usando o método Ward.D2 implementado na função hclust do R version. Os três grupamentos da Figura 6 (A, B e C) estão representados por três tons de cinza no anel mais externo. O anel intermediário representa duas cores (calha e tributário). O anel mais interno, representa as três seções da bacia (Alto, Médio e Baixo Rio Doce) e cada tributário com uma cor única.	22
Figura 8. Gráfico de estrutura populacional mostrando o agrupamento das populações em dois componentes genéticos distintos para <i>A. lacustris</i> . Cada barra vertical representa um indivíduo. Os componentes genéticos identificados são; população 1 (verde) e população 2 (vermelho). 23	
Figura 9. Valores do F_{ST} par a par entre populações para todos os indivíduos de <i>L. castaneus</i> . Segundo Wright (1978) F_{ST} entre 0 e 0.05 indica pouca diferenciação genética; F_{ST} entre 0.05 e 0.15 indica diferenciação genética moderada; F_{ST} entre 0.15 e 0.25 indica alta diferenciação genética e F_{ST} acima de 0,25 diferenciação genética muito alta.	25
Figura 10. Análise Discriminante de Componentes Principais das variações de SNPs gerados pela ferramenta RAD-seq, obtidas em populações de <i>Loricariichthys castaneus</i>	26
Figura 11. Dendograma hierárquico gerado para os indivíduos analisados de <i>L. castaneus</i> (cada um deles é representado por letras e números únicos) pela matriz de distâncias pareadas calculada partir da tabela de genotipagem usando a distância Euclidiana usando o método Ward.D2 implementado na função hclust do R. Os três grupamentos da Figura 10 (A, B e C) estão representados por três tons de cinza no anel mais externo. O anel intermediário representa duas cores (calha e tributário). O anel mais interno, representa as três seções da bacia (Alto, Médio e Baixo Rio Doce) e cada tributário com uma cor única.	27
Figura 12. Gráfico de estrutura populacional mostrando o agrupamento das populações em três componentes genéticos distintos para <i>L. castaneus</i> . Cada barra vertical representa um indivíduo. Os componentes genéticos identificados são; população 1 (verde), população 2 (vermelho), população 3 (azul).	28
Figura 13. Valores do F_{ST} par a par entre populações para todos os indivíduos de <i>G. brasiliensis</i> . Segundo Wright (1978) F_{ST} entre 0 e 0.05 indica pouca diferenciação genética; F_{ST} entre 0.05 e 0.15 indica diferenciação genética moderada; F_{ST} entre 0.15 e 0.25 indica alta diferenciação genética e F_{ST} acima de 0,25 diferenciação genética muito alta.	30
Figura 14. Análise Discriminante de Componentes Principais das variações de SNPs gerados pela ferramenta RADSeq, obtidas em populações de <i>G. brasiliensis</i> . Os filtros utilizados para a seleção dos alelos foi: eliminação dos alelos que não se ajustam às expectativas de equilíbrio de Hardy-Weinberg (HWE), dos alelos que ocorrem em baixa frequência (MAF) e loci que não foram amplificados (MD).	31
Figura 15. Dendograma hierárquico gerado para os indivíduos analisados de <i>G. brasiliensis</i> (cada um deles é representado por letras e números únicos) pela matriz de distâncias pareadas calculada partir da tabela de genotipagem usando a distância Euclidiana usando o método Ward.D2 implementado na função hclust do R. Os três grupamentos da Figura 14 (A, B, C e D) estão representados por quatro tons de cinza no anel mais externo. O anel intermediário representa duas cores (calha e tributário). O anel mais interno, representa as três seções da bacia (Alto, Médio e Baixo Rio Doce) e cada tributário com uma cor única.	32
Figura 16. Gráfico de estrutura populacional mostrando o agrupamento das populações em três componentes genéticos distintos para <i>G. brasiliensis</i> . Cada barra vertical representa um indivíduo. Os componentes genéticos identificados são; população 1 (verde), população 2 (vermelho), população 3 (azul).	33

LISTA DE TABELAS

Tabela 1. Localização, coordenadas geográficas, data da coleta no período chuvoso e seco. Área afetada se refere a trechos que tiveram contato direto com a passagem do rejeito	10
Tabela 2. Filtros aplicados aos dados de sequenciamento de <i>A. lacustris</i> . HWE (Equilíbrio Hardy Weinberg); MAF (menor frequência alélica)	18
Tabela 3. Heterozigosidade observada (H_o), Heterozigosidade esperada (H_e) e sua aplicação no cálculo do Coeficiente de endocruzamento (F_{IS}) nas populações estudadas de <i>Astyanax lacustris</i>	19
Tabela 4. Filtros aplicados aos dados de sequenciamento do Cascudo-viola <i>L. castaneus</i> . HWE (Equilíbrio Hardy Weinberg); MAF (menor frequência alélica)	23
Tabela 5. Heterozigosidade observada (H_o), Heterozigosidade esperada (H_e) e sua aplicação no cálculo do Coeficiente de endocruzamento (F_{IS}) nas populações estudadas de <i>Loricariichthys castaneus</i>	24
Tabela 6. Filtros aplicados aos dados de sequenciamento de <i>G. brasiliensis</i> . HWE (Equilíbrio Hardy Weinberg); MAF (menor frequência alélica)	28
Tabela 7. Heterozigosidade observada (H_o), Heterozigosidade esperada (H_e) e sua aplicação no cálculo do Coeficiente de endocruzamento (F_{IS}) nas populações estudadas de <i>Geophagus brasiliensis</i>	29
Tabela 8. Grau de correlação entre a abundância de espécies não-nativas e o Acará, no Baixo Rio Doce. O teste foi significativo no nível $p > 0,05$	Erro! Indicador não definido.

SUMARIO EXECUTIVO

A análise molecular utilizando RADSeq em três espécies de peixes: o lambari *Astyanax lacustris*, o acará *Geophagus brasiliensis* e o cascudo *Loricariichthys castaneus* da bacia do Rio Doce, permitiu definir aspectos gerais dos padrões de variação que são relevantes para o estabelecimento de uma linha histórica de referência e como subsídio para medidas mitigadoras dos impactos no ecossistema aquático na bacia. Os padrões gerais mostram uma baixa variabilidade genética e alto grau de endocruzamento para as três espécies estudadas ao longo da calha do Rio Doce e dos tributários. Nossos resultados sugerem também que *A. lacustris* e *L. castaneus* não apresentam estruturação profunda de suas populações, apresentando um alto grau de panmixia. O mesmo vemos para as populações da porção alta do Rio Doce e do tributário Piracicaba em *A. lacustris*. Já *G. brasiliensis* apresentou resultados que sugere diferenciação genética alta, principalmente entre a população do baixo Rio Doce e as restantes da bacia. A análise bayesiana de agrupamento mostrou a diversidade genética para as três espécies estudadas, com *L. castaneus* e *G. brasiliensis* apresentando três componentes genéticos distintos e *A. lacustris*, dois. Para *G. brasiliensis*, os diferentes componentes genéticos aparentam estar relacionados com maior diferenciação das populações da espécie na bacia. O agrupamento hierárquico demonstrou que cada espécie possui tributários ou porções da calha que são relevantes como fonte de espécimes na dinâmica de recolonização do Rio Doce. Para o acará *G. brasiliensis*, os rios Piracicaba e Caratinga foram os tributários que apareceram mais agrupados com as porções da calha principal do Rio Doce, porém somente nas porções média e alta. Já o Rio Guandu foi encontrado agrupado somente com a porção média do Rio Doce. A

população do baixo Rio Doce permaneceu isolada dos tributários e das demais porções da calha. *Astyanax lacustris* apresentou uma lista mais diversificada de rios que aportam peixes para a calha principal do Rio Doce, com os rios Santo Antônio, Manhuaçu, Suaçuí e Guandu, os quais aparecem agrupados com espécimes do médio Rio Doce. Encontramos o baixo Rio Doce agrupado com espécimes do médio Rio Doce, sugerindo que, para essa porção da bacia, o aporte de peixes para a recolonização pode partir de porções menos degradadas ou menos afetadas pelo rejeito da porção média da calha. A porção alta não apareceu agrupada com espécimes de tributário e seguiu o padrão da porção baixa, agrupando-se com espécimes da porção média do Rio Doce. Para *L. castaneus*, os tributários demonstraram relação genética próxima com outros tributários, agrupando entre si, enquanto a porção alta da calha agrupa com a porção baixa e os espécimes da porção média formou um cluster isolado. Os resultados de *A. lacustris* e *L. castaneus* sugerem que um contingente de cada espécie sobreviveu a passagem do rejeito de minério, mas esse padrão é menos evidente para o acará *G. brasiliensis*.

1) OBJETIVOS

a. Objetivo Geral

Investigar o papel dos tributários no processo de recolonização da ictiofauna na calha principal do rio Doce após o rompimento da barragem de rejeitos “Fundão”.

b. Objetivos Específicos

- (i) Avaliar temporalmente a variabilidade e estruturação genética das populações de peixe dos tributários e da calha principal da bacia do rio Doce;
- (ii) Ampliar o conhecimento sobre a diversidade e a distribuição da ictiofauna da bacia do rio Doce;
- (iii) Produzir um ranking da importância ecológica e conservacionista de cada tributário baseado na contribuição potencial ou real do seu aporte de peixes para a calha principal.
- (iv) investigar a conectividade (expresso pelo fluxo gênico entre as populações de espécies de peixes) entre os tributários e a calha principal do rio Doce e
- (v) produzir um estudo da história evolutiva recente da ictiofauna da bacia do rio Doce.

2) INTRODUÇÃO

a. Caracterização do problema: o desastre de Mariana

Em 5 de Novembro de 2015 o Brasil presenciou seu maior desastre ambiental quando uma barragem das mineradoras Samarco/Vale/BHP (barragem de “Fundão”) no município de Mariana em Minas Gerais se rompeu, descarregando aproximadamente 60 milhões de toneladas de rejeito de minério de ferro. A enxurrada percorreu 663 km de rios e córregos da bacia do Rio Doce, levando à impactos na biodiversidade, perda de recursos pesqueiros, comprometimento da qualidade da água para o consumo e irrigação, perda de área cultivável, entre outros tantos danos.

O rompimento da barragem da Samarco/Vale/BHP causou um aumento instantâneo da turbidez do Rio Doce e queda do nível de oxigênio (IGAM, 2015), causando a morte imediata de 11 toneladas de peixes (IBAMA, 2015) e de muitos outros organismos. A devastação atingiu aproximadamente 1500 ha de cobertura florestal, incluindo áreas de preservação permanente (IBAMA, 2015; Governo de Minas, 2016). Os rejeitos de minério alcançaram o Oceano Atlântico 17 dias após o rompimento, estendendo os impactos aos ecossistemas costeiros e estuarinos (Miranda & Marques, 2016).

Sete anos após o desastre, a potencial extensão, magnitude e reversibilidade dos impactos causados pelo rejeito depositado na bacia ainda não são conhecidos em sua totalidade, o que ressalta a importância de se realizar uma análise da composição genética de peixes dos afluentes e da calha principal do rio Doce.

b. Histórico da degradação ambiental da bacia do Rio Doce

Todo o processo histórico de ocupação extrativista, aliado à densidade populacional e intensa atividade econômica da região tornaram a bacia do Rio Doce uma das drenagens mais comprometidas em termos de impactos antrópicos. A poluição por esgotos é a principal causa de degradação da qualidade da água da bacia, como mostra os dados do IGAM (2005) em resultados obtidos entre 1997 e 2003: para as estações de amostragem da parte mineira da bacia, os parâmetros que apresentaram maior número de violações foram fosfato total, coliformes fecais e coliformes totais. Essas ocorrências, acima dos limites legais, têm como principal causa a falta de tratamento de esgoto na bacia: das 230 cidades que formam a bacia, 191 (90%) lançam seus dejetos *in natura* nos corpos receptores (Consórcio EcoPLAN-Lume, 2010). A contaminação por fosfato também pode se dar por carreamento para os rios após o uso de técnicas de adubação nas plantações de eucalipto para elevar os níveis de fósforo no solo. Segundo ainda o IGAM (2005), a presença da suinocultura do Alto curso do Rio Doce é responsável pela ocorrência de elevados níveis de zinco.

Os impactos antrópicos advindos da extração de ouro na bacia teriam disponibilizado ou liberado uma potente carga de sedimentos enriquecidos com elementos tóxicos que foram transportados como carga de canal e em suspensão desde o passado, até os dias atuais. Foram encontradas elevadas concentrações de metais como Mercúrio (Hg) e Arsênio (As) em sedimentos do ribeirão do Carmo e no Rio Gualaxo, atribuídas às atividades extrativo-minerais (Eleutério, 1997; Costa, 2006). Diferentes estudos têm apresentado evidências variadas sobre a presença de metais pesados no Rio Doce após o desastre de Mariana (níveis normais: Prefeitura de Governador Valadares, 2015; CPRM, 2015 - níveis elevados: Prefeitura de Baixo Guandu, 2015, IGAM 2015, GIAIA 2015; Carvalho et al 2017, e provavelmente o revolvimento do leito do rio pela avalanche do rejeito disponibilizou novamente esses elementos provenientes de atividades antigas na bacia.

c. Impactos do rompimento da barragem de “Fundão” sobre os habitats e sua biota

O rompimento da barragem de “Fundão” configura-se como o maior impacto ambiental ocorrido no Brasil e soma-se ao longo histórico de degradação da bacia do Rio Doce e da Mata Atlântica. Os impactos foram maiores ao longo dos primeiros 100 km percorridos pela lama de rejeito, por extrapolar a calha dos Rios Gualaxo do Norte, Carmo e Doce (Governo de Minas, 2016; EMBRAPA, 2015). Parte dos rejeitos de minério foi retido pela barragem da Usina Hidrelétrica de Candonga e, após essa barreira, o material seguiu pela calha do Rio Doce.

Em seu trajeto, cerca de 60 milhões de toneladas de rejeito removeram um total de 1587 ha de vegetação, sendo 511,08 ha de Mata Atlântica (Governo de Minas, 2016). Foram devastados remanescentes de matas ciliares e de matas de galeria, essenciais para o equilíbrio ecológico na interface entre os habitats aquáticos e terrestres. A biota aquática depende de fontes de alimentos externas ao meio aquático além de uma diversidade de substratos para que possam se alimentar, reproduzir e evitar predadores (Bunn et al., 1997; Roque et al., 2003; Rios & Bailey, 2006). Além disso, as florestas ripárias evitam mudanças estruturais do canal, como mudanças na profundidade e na estabilidade das margens (Ferreira et al., 2011). Essa integridade estrutural dos habitats aquáticos é de fundamental importância para a preservação da biodiversidade aquática. A vegetação ripária funciona ainda como uma barreira natural para a carga de partículas de solo trazida pelas enxurradas, impedindo que sedimentos e poluentes afetem a dinâmica da biota aquática (Dosskey et al., 2010).

A enxurrada de rejeito atingiu 663 km de rios da bacia, alterando a qualidade dos cursos d'água e causando imediata perda de biodiversidade. As margens dos Rios Gualaxo do Norte, Carmo e parte do Doce foram soterradas, extirpando a biota desses locais. A elevada turbidez causada pela quantidade de sólidos em suspensão levou a mortandade de plantas aquáticas e da ictiofauna, causadas respectivamente pela inviabilidade da fotossíntese e pelo colapso das brânquias e asfixia (IBAMA, 2015). Foram retiradas 11 toneladas de peixes mortos (IBAMA, 2015) da bacia do Rio Doce após a passagem do rejeito. Essa perda de espécies de peixes pode ser ainda maior, pois as estimativas não consideram em seus cálculos os efeitos de longo prazo sobre as comunidades da ictiofauna durante seus ciclos de vida (Vieira, 2009).

No seu trajeto, a onda de rejeitos de minério ressuspendeu os sedimentos estabilizados nas partes mais profundas no leito desses rios, aumentando para níveis tóxicos a concentração de metais pesados (Governo de Minas, 2016). O grande aporte de sedimento acumulado nos leitos alterou completamente o habitat de espécies aquáticas de fundo. O assoreamento, principalmente nos leitos dos Rios Gualaxo do Norte, Carmo e parte do Doce causou fragmentação de habitat e interrupção do fluxo gênico da biota aquática.

A diversidade genética das populações de peixes sobreviventes a grandes impactos pode ser reduzida devido à baixa qualidade da água, à menor integridade de habitats e às alterações nas teias alimentares (Piorski et al., 2008). Esta diversidade genética reduzida é um desafio para

a recolonização de locais afetados, uma vez que reduz a possibilidade das espécies em lidar com o estresse a longo prazo imposto pela perturbação do habitat (Piorski et al., 2008).

Nesse estudo utilizamos a avaliação genômica com a técnica RADSeq para avaliar as dinâmicas populacionais e a estrutura genética da ictiofauna da bacia do Rio Doce, após o rompimento da barragem do “Fundão”. A técnica de RADSeq se baseia na quebra (por ação enzimática) do genoma, permitindo a varredura ampla do DNA de cada indivíduo, o qual resulta na produção de um grande número de fragmentos que permitem a comparação de cada indivíduo dentro de sua população e a comparação entre as populações da mesma espécie. Os parâmetros obtidos permitem determinar se as populações são ricas ou pobres em termos da variação genética. As populações com baixa riqueza genética, podem ter ficado assim após uma redução drástica do seu tamanho. Além disso, é possível definir se as populações tem altos níveis de troca de espécimes, o qual permite inferir se os locais de coleta apresentam baixo grau de isolamento. Os casos contrários, quando há uma grande diferenciação genética entre as populações, pode ser inferida a existência de barreiras físicas ou ecológicas entre elas.

3) MATERIAIS E MÉTODOS

a. Espécies estudadas

As três espécies que foram analisadas com a abordagem RADSeq são nativas da bacia do Rio Doce e são de natureza tolerante às alterações ambientais (ICMBIO, 2018), em contraposição a outras espécies, que são intolerantes e que já eram de ocorrência rara antes do evento (Figuras 1 a 3). Dentro da ordem Characiformes, o Lambari-de-rabo-amarelo *Astyanax lacustris* é provavelmente a espécie de maior distribuição na bacia. Como representante da ordem Cichliformes, o acará *Geophagus brasiliensis* acompanha em termos de distribuição, o lambari *A. lacustris*. No grupo dos Siluriformes, o cascudo-viola *Loricariichthys castaneus* é uma das espécies de ampla distribuição, junto ao cascudo-manteiga *Hypostomus affinis*; esta última espécie poderá ser objeto de estudos futuros.

Figura 1. O acará *Geophagus brasiliensis* (Quoy & Gaimard 1824)



Figura 2. O lambari *Astyanax lacustris* (Lütken 1875)



Figura 3. O cascudo-viola, *Loricariichthys castaneus* (Castelnau, 1855)



b. Desenho Amostral

As amostras foram coletadas dentro da malha amostral do projeto (Figura 4, Tabela 1). As coletas foram iniciadas no dia 14 de fevereiro de 2020 e se encerraram em 30 de setembro de 2020. Os peixes foram coletados com redes de espera de 10 metros de comprimento e malhas de 15, 20, 25, 30, 35, 40, 50, 60, 70 e 80 mm (distância entre nós), perfazendo assim um conjunto de 10 redes de 10 metros. Em cada ponto de coleta, foram utilizados dois conjuntos de redes totalizando 20 redes e 200 metros em cada ponto. As redes foram armadas no entardecer do dia (às 18:00 horas) e retiradas na manhã do dia seguinte (às 6:00 horas), perfazendo 12 horas na coluna d'água. Além dos pontos da malha amostral, foram incluídas 41 amostras da calha do baixo Rio Doce coletadas pelo projeto "Rede Rio Doce Mar" (RRDM) para complementar nosso conjunto de dados.

Amostras de tecido muscular foram obtidas de espécimes frescos ou de indivíduos coletados vivos e eutanasiados em solução contendo eugenol a 4%. O eugenol é um composto ativo do óleo de cravo, de acordo com a PORTARIA CFBio Nº 148/2012, que regulamenta os procedimentos de captura, contenção, marcação e coleta de animais vertebrados previstos nos artigos 4º, 5º, 6º e 8º da Resolução CFBio nº 301/2012 e pelo CONCEA em diretrizes para prática de eutanásia em peixes. As amostras foram fixadas imediatamente em etanol 90% e

condicionadas em tubos tipo Eppendorf. Posteriormente, os tubos foram transferidos e estocados em freezer -20 °C até seu processamento.

Figura 4. Mapa da bacia hidrográfica do Rio Doce. Pontos em verde são pontos-referência, representados por afluentes que não tiveram contato com o rejeito. As áreas afetadas pelo rejeito estão representadas por pontos vermelhos.

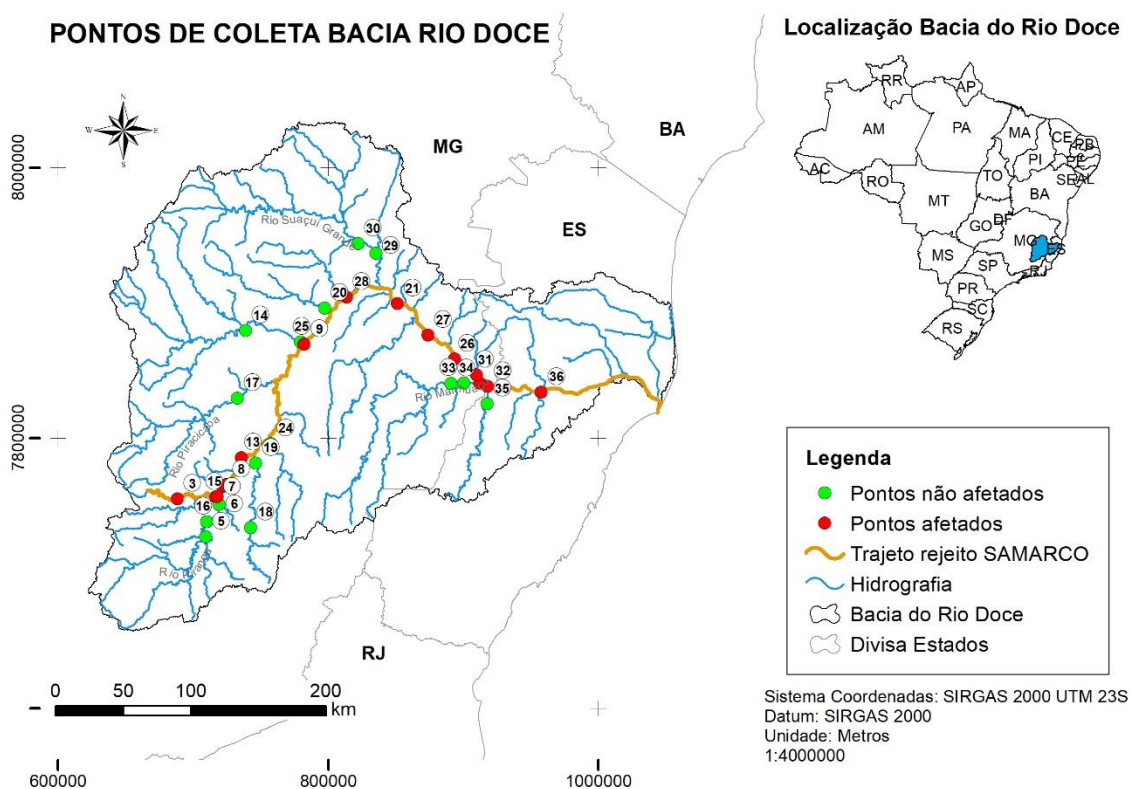


Tabela 1. Localização, coordenadas geográficas, data da coleta no período chuvoso e seco. Área afetada se refere a trechos que tiveram contato direto com a passagem do rejeito

Código do ponto	Nome da drenagem	Localidade	Coordenadas	Data coleta (chuvoso)	Data coleta (seco)	Área afetada
Ponto 3	Rio Gualaxo do Norte	Paracatu de baixo, subdistrito de Monsenhor Horta, Mariana, MG	23k 0688284 7754717	19/02/2020	11/08/2020	Sim

Código do ponto	Nome da drenagem	Localidade	Coordenadas	Data coleta (chuvoso)	Data coleta (seco)	Área afetada
Ponto 4	Rio do Carmo	Barra Longa, MG	23k 0704609 7756231	03/03/2020	25/09/2020	Sim
Ponto 5	Rio Piranga	Guaraciaba, MG	23k 0709518 7726784	06/03/2020	18/08/2020	Não
Ponto 6	Rio Piranga	Rio Doce, MG	23k 0719392 7750689	29/02/2020	10/09/2020	Não
Ponto 7	Rio Doce	Rio Doce, MG	23k 0717886 7756690	16/02/2020	05/09/2020	Sim
Ponto 8	Rio Doce	Rio Doce, MG	23k 0724451 7765328	18/02/2020	03/09/2020	Sim
Ponto 9	Rio Doce	Naque, MG	23k 0782354 7869156	19/03/2020	28/08/2020	Sim
Ponto 10	Rio Doce	Rio Doce, UHE Risoleta neves, MG	23k 0721890 7764082	15/02/2020	02/09/2020	Sim
Ponto 13	Rio Doce	São José do Goiabal, MG	23k 0735965 7785306	20/03/2020	21/08/2020	Sim
Ponto 15	Rio Carmo	Rio Doce, MG	23k 0716721 7756192	27/02/2020	04/09/2020	Sim
Ponto 16	Rio Piranga	UHE Brito, MG	23k 0710222 7737948	05/03/2020	11/09/2020	Não
Ponto 17	Rio Piracicaba	Jaguaraçu, MG	23k 0732869 7829391	13/03/2020	18/09/2020	Não
Ponto 18	Rio Casca	Jequiri, MG	23k 0742458 7733441	19/03/2020	09/09/2020	Não

Código do ponto	Nome da drenagem	Localidade	Coordenadas	Data coleta (chuvoso)	Data coleta (seco)	Área afetada
Ponto 19	Rio Casca	São José do Goiabal, MG	23k 0746357 7781360	16/03/2020	12/09/2020	Não
Ponto 20	Rio Corrente	Periquito, MG	23k 0797357 7896137	10/03/2020	18/09/2020	Não
Ponto 21	Rio Doce	Tumiritinga, MG	24k 0219512 7900508	10/03/2020	02/09/2020	Sim
Ponto 22	Rio Caratinga	Barra do Cuieté, MG	24k 0233687 7890184	05/03/2020	20/08/2020	Não
Ponto 23	Rio Doce	Aimorés, MG	24k 0282205 7844032	15/02/2020	08/09/2020	Sim
Ponto 24	Rio Matipó	Raul Soares	23k 0757449 7795917	18/03/2020	22/09/2020	Não
Ponto 25	Rio Santo Antônio	Naque, MG	23k 0779627 7871089	18/03/2020	17/09/2020	Não
Ponto 26	Rio Doce	Resplendor, MG	24k 0263294 7861088	04/03/2020	19/08/2020	Sim
Ponto 27	Rio Doce	Conselheiro Pena, MG	24k 0242767 7877983	06/03/2020	21/08/2020	Sim
Ponto 28	Rio Doce	Governador Valadares, MG	23k 0813646 7904202	11/03/2020	15/09/2020	Sim
Ponto 29	Rio Suaçuí Grande	Matias Lobato, MG	24k 0202153 7937309	11/03/2020	19/09/2020	Não

Código do ponto	Nome da drenagem	Localidade	Coordenadas	Data coleta (chuvoso)	Data coleta (seco)	Área afetada
Ponto 30	Rio Suaçuí Grande	Matias Lobato, MG	24k 0188844 7943845	13/03/2020	20/09/2020	Não
Ponto 31	Rio Doce	Aimorés, MG	24k 0279689 7849766	27/02/2020	12/08/2020	Sim
Ponto 32	Rio Doce	Baixo Guandu, ES	24k 0288311 7841895	16/02/2020	14/08/2020	Sim
Ponto 33	Rio Manhuaçu	Aimorés, MG	24k 0270455 7843334	17/02/2020	15/08/2020	Não
Ponto 34	Rio Manhuaçu	Aimorés, MG	24k 0261172 7842938	18/02/2020	10/09/2020	Não
Ponto 35	Rio Guandu	Baixo Guandu, ES	24k 0288291 7828762	28/02/2020	13/08/2020	Não
Ponto 36	Rio Doce	Colatina, ES	24k 0327949 7838898	29/02/2020	03/09/2020	Sim

c. Extração e diluição do DNA

O DNA genômico foi extraído utilizando o protocolo de CTAB (brometo de cetil-trimetil amônio) a 2%. Após a extração medimos a quantidade e a qualidade do DNA por espectrofotometria no equipamento NanoDrop™. O NanoDrop™ é um espectrofotômetro que estima a concentração de amostras de DNA, RNA e de proteínas, usando o princípio da absorvância do espectro ultravioleta-visível (UV-Vis). Além de medir a concentração, o NanoDrop™ indica a presença de contaminantes nas amostras a partir das taxas de absorção 260/280 e 260/230. Um valor abaixo de 1,8 na taxa 260/280 indica contaminação da amostra por proteínas indesejadas, enquanto um valor menor que 1,8 na taxa 260/230 indica contaminação por compostos orgânicos ou outros agentes contaminantes. Após as medições, as amostras foram diluídas para adequarem-se às exigências do protocolo de sequenciamento. As diluições foram realizadas em duas etapas, após a primeira medição no NanoDrop™ o DNA genômico total foi diluído a uma concentração de 50 ng/µL. Todas as amostras passaram por uma segunda

medição e foram novamente diluídas para a concentração de 25 ng/uL. Uma terceira medição foi realizada para confirmar esse valor final que é um dos parâmetros que garantem a boa qualidade do sequenciamento. Após as diluições, as amostras foram conferidas em gel de agarose em busca de bandas indesejadas, que indicariam fragmentação de DNA, contaminantes ou RNA. Em seguida as amostras foram enviadas para a empresa, Floragenex (Corvallis, Oregon, Estados Unidos), onde foram sequenciadas.

d. Preparação da biblioteca de RADSeq

O DNA genômico foi digerido e processado em bibliotecas RAD semelhantes ao método de Baird et al. (2008). Iniciamos a construção da biblioteca com a digestão do DNA genômico com a enzimas de restrição *SbfI* por 60 min a 37°C em 15 uL de reação com 20 unidades (U) de enzima (New England Biolabs [NEB]).

Após a digestão, as amostras foram inativadas por calor por 20 min a 80°C. Adaptadores P1 de 1 uM (adaptador Solexa© modificado Illumina, Inc.) foram adicionados a cada amostra digerida juntamente com buffer 10x de T4 DNA Ligase (Enzymatics, Inc) e 0,8 uL H₂O. Cada adaptador P1 contém uma sequência curta de nucleotídeos multiplex exclusiva (código de barras), que é lida durante a corrida e identifica a origem dos fragmentos após o sequenciamento. A amostras foram então deixadas à temperatura ambiente por 20 minutos para a ligação dos adaptadores pela enzima T4 DNA ligase. Em seguida, as amostras foram inativadas pelo calor por 20 min a 65°C, agrupadas e cortadas aleatoriamente com um Bioruptor (Diagenode) para um tamanho médio de 300pb - 500pb.

Uma reação do módulo End Repair / dA-Tailing (NEB) foi usada para polir as extremidades do DNA. Após purificação subsequente, 1 µL de 1 µM de adaptador P2, um adaptador Solexa© modificado divergente (Illumina, Inc.) foi ligado, em temperatura ambiente, aos fragmentos de DNA obtidos. As amostras foram novamente purificadas e eluídas em 17 µL. A reação foi quantificada usando um Qubit Fluorometer com o kit de ensaio dsDNA HS (Invitrogen) e 10 ng deste produto foi usado em uma amplificação por PCR com 50 µL 2x Phusion Master Mix (NEB), 5 µL de 10 µM modificado Solexa© Amplification primer mix (Illumina, Inc.) e até 45 µL H₂O.

As bibliotecas amplificadas foram purificadas e quantificadas com um fluorômetro Qubit e analisada em um Agilent Bioanalyzer com o kit de alta sensibilidade para determinar a distribuição de tamanho e o tamanho médio do fragmento, que foi de 491 pb. A biblioteca foi sequenciada no NovaSeq 6000 na Universidade de Oregon GC3F com leituras de 2x150 bp. A etapa de seleção de tamanho foi realizada em um gel de 1,5% de agarose (Sigma), 0,5X TBE. O DNA foi isolado usando um kit de extração de gel MinElute (Qiagen).

e. Controle de qualidade, trimagem e filtragem

Durante o sequenciamento, milhões de sequências de DNA foram geradas contendo um código que indica a amostra a qual ela pertence. Nessa primeira etapa da análise, essas sequências foram separadas e atribuídas às diferentes amostras enviadas, além de passarem por um controle de qualidade. Os dados do RADSeq foram separados de acordo com as suas respectivas amostras usando o CLC Genomics Workbench versão 8.5.4 (QIAGEN Aarhus) e a tabela de indexadores (pequeno fragmento de nucleotídeos que identifica a origem da sequência) fornecida pela empresa Floragenex Inc. A qualidade dos dados foi inspecionada usando os relatórios gerados pelo FASTQC versão 0.11.9 (<https://github.com/s-andrews/FastQC>) e MultiQC versão 1.11, com o objetivo de identificar os erros de sequenciamento e outros artefatos da técnica do RADSeq que poderiam interferir na identificação dos polimorfismos. As sequências dos adaptadores (pequeno fragmento de nucleotídeos usado pela empresa para fazer a corrida de sequenciamento) foram detectadas e removidas dos dados do sequenciamento usando a configuração de auto-deteção do TrimGalore versão 0.6.7. Em seguida, as sequências passaram pelo processo de trimagem, que consiste em remover os nucleotídeos sequenciados com baixa qualidade ou que são indesejados por causarem incidência de erros nas análises dos dados. As sequências foram então filtradas pelo tamanho usando o Fastp versão 0.23.2 e selecionando os seguintes parâmetros: `--trim_poly_g --poly_g_min_len 10; --trim_poly_x --poly_x_min_len 10; --dedup --dup_calc_accuracy 6; --trim_front 15 --trim_tail 5; and --length_required 50.`

f. Mapeamento genômico

Após a separação e conferência da qualidade das sequências, é necessário saber a posição das mesmas no genoma para podermos compará-las e assim identificarmos os polimorfismos entre elas. Isso é realizado de duas formas: alinhando-se as sequências a um genoma já conhecido (que pode ser da mesma espécie, ou de espécies próximas) ou, quando não temos um genoma de referência adequado, fazemos a montagem *de novo*. Na montagem *de novo* é criado um genoma de referência a partir da sobreposição das sequências geradas. Para identificação dos polimorfismos em *Astyanax lacustris* e *Geophagus brasiliensis* foram selecionados no banco de dados NCBI Genome (<https://www.ncbi.nlm.nih.gov/genome/>) como referência, os genomas de *Astyanax mexicanus* (montagem: *Astyanax_mexicanus-2.0*; código de acesso: GCF_000372685.2) e *Archocentrus centrarchus* (montagem: *fArcCen*; código de acesso: GCF_007364275.1), respectivamente. As sequências foram mapeadas nos genomas de referência usando o algoritmo BWA-MEM do BWA versão 0.7.15. Em seguida, os arquivos Sequence Alignment Map (SAM), contendo as informações do mapeamento, foram convertidos em arquivos ordenados no formato Binary Alignment Map (BAM) usando o Samtools versão 1.12. O genoma do cascudo-viola *Loricariichthys castaneus* foi analisado com a abordagem *de novo*.

g. Análise dos polimorfismos e das populações

Neste trabalho, os pontos de coleta de cada tributário foram colapsados. O mesmo foi feito, unindo os pontos de coleta de acordo com as porções alta, média e baixa da calha principal. Dessa forma, consideramos as diferentes populações de acordo com o tributário e porção da calha principal da qual ela pertence.

Após o mapeamento das sequências, realizou-se a identificação de nucleotídeos que variam entre elas (ou polimorfismos). A análise dos polimorfismos foi feita usando o Stacks versão 2.6.0, seguindo o manual de referência: <https://catchenlab.life.illinois.edu/stacks/manual>. Os polimorfismos de nucleotídeo único (single nucleotide polymorphisms, SNPs) foram identificados usando a ferramenta gstacks para os dados do lambari *Astyanax lacustris* e o grupo do acará *Geophagus brasiliensis* a partir dos arquivos BAM e selecionando a opção para remover as reads duplicadas, que são artefatos da técnica da reação em cadeia da polimerase (PCR). Os SNPs para *Loricariichthys castaneus* foram identificados pela ferramenta ustacks usando uma predição *de novo* diretamente a partir dos dados do RADSeq. A ferramenta “population” do Stacks foi usada para gerar arquivos Variant Call Format (VCF) e calcular a divergência do equilíbrio de Hardy-Weinberg (HWE) para cada locus. Os loci que desviaram do HWE ($P < 0.05$) e apresentaram menor frequência alélica (*minor allele frequency*, MAF) menor ou igual que 5% foram selecionados usando o vcftools versão 0.1.16 e removidos. Essa filtragem dos dados é importante para garantir a qualidade da identificação e análise dos polimorfismos entre as sequências.

Os arquivos VCF obtidos contendo os dados filtrados foram convertidos em arquivos do formato GENEPOP e STRUCTURE usando o PGDSpider versão 2.1.1.5 para as análises populacionais. Os arquivos GENEPOP foram analisados usando diferentes pacotes do R versão 3.6.3. Primeiro, os dados ausentes (*missing data*, MD) foram removidos usando a função “missing” do pacote Poppr versão 2.9.3 aplicando os seguintes filtros: (i) remoção dos loci contendo frequências de MD maior que 75%; (ii) remoção das amostras/indivíduos com mais de 75% dos loci contendo MD.

Em seguida, foram calculados os seguintes parâmetros populacionais:

1. Coeficiente de endocruzamento (F_{IS}): expressa a porção da variância populacional que ocorre em cada indivíduo (altos valores deste parâmetro sugerem muito endocruzamento, enquanto valores negativos indicam excesso de heterozigotos).
2. Índice de fixação (F_{ST}): As medidas F_{ST} podem ser interpretadas como a variância de alelos entre populações. Em condições de pequenas populações ou de autofecundação, o valor do F_{ST} é de 1 e em populações muito grandes e panmíticas (onde todos os indivíduos têm a mesma chance de cruzar com outros), é próximo a 0.

Tanto o F_{IS} quanto o F_{ST} podem ser interpretados como a probabilidade, dentro de cada indivíduo, da herança do mesmo gene, via paterna e via materna (Weir e Cockerham, 1984). Quando as probabilidades disso acontecer são altas (como por exemplo, no caso de reprodução assexuada), o valor desses parâmetros é 1.

As estimativas das frequências alélicas, as heterozigosidades observadas e as diversidades genéticas foram calculadas usando a função `basic.stats` do pacote `hierfstat` versão 0.5-11 (<https://github.com/jgx65/hierfstat>). O índice de fixação (F_{ST}) foi calculado usando a função Weir e Cockerham 1984 (WC84).

A estruturação da população e a identificação de grupos de indivíduos geneticamente relacionados foram avaliadas usando três abordagens diferentes. Primeiro, uma matriz de distâncias pareadas foi calculada a partir da tabela de genotipagem usando a distância Euclidiana. Com base nessa tabela, buscou-se descobrir agrupamentos hierárquicos de indivíduos relacionados usando o método Ward.D2 implementado na função `hclust` do R version 3.6.3.

Na segunda abordagem, foi realizada uma análise discriminante de componentes principais (Discriminant Analysis of Principal Components, DAPC) usando o pacote `adegenet` versão 2.1.7 (<https://adegenet.r-forge.r-project.org/>). A DAPC identifica agrupamentos de indivíduos geneticamente relacionados para grupos definidos *a priori*.

Na terceira abordagem, foi utilizado o método de agrupamento bayesiano disponibilizado no Structure versão 2.3.4. Nesse método, o modelo utiliza uma distribuição de probabilidade *a priori* para inferir o número de populações (K) e então os indivíduos são agrupados para cada K, de acordo com a distância genética entre os grupos. A análise foi realizada testando todos os valores de K variando de 1 a n (n = número de localidades com amostras/indivíduos coletados). A inferência foi feita com 5 replicatas para cada hipótese de valor K, usando uma análise MCMC com 100,000 repetições e um período de burning de 100,00. O melhor valor de K foi selecionado usando o Structure harvester versão 0.6.94 (<https://taylor0.biology.ucla.edu/structureHarvester>).

4) RESULTADOS

a. *Astyanax lacustris*

A maioria dos loci foi removida por desviarem do equilíbrio de Hardy-Weinberg (HWE) (154.432 loci) e em seguida foram excluídos outros 13.997 loci que apresentaram menor frequência alélica (MAF). Em seguida foram retirados os loci e os indivíduos que possuíam mais que 75% de dados ausentes, que contabilizaram 1.885 loci e 32 indivíduos. Após essa filtragem dos dados um total de 131 indivíduos e 1.519 loci foram utilizados para as análises nesta espécie (Tabela 2).

Tabela 2. Filtros aplicados aos dados de sequenciamento de *A. lacustris*. HWE (Equilíbrio Hardy Weinberg); MAF (menor frequência alélica)

Filtro	Indivíduos	Loci
Stacks*	170	177.859
HWE ($P \leq 0.05$)	170	23.427
MAF ≤ 0.05	170	9.430
Missing Data (75%) – loci	170	3.404
Missing Data (75%) – indivíduos**	131	1.519

*Genoma de referência: *Astyanax mexicanus*;
 Assembly (NCBI ID): *Astyanax mexicanus*-2.0 (GCF_000372685.2)
 **indivíduos removidos:
 FB0897, FA168, FA169, FA170, FA171, FA172, FF1263, FF1264, FB0442, FF0439, FA648, FA650, FF0137, FF0373, FA234, FA235, FA238, FB0953, FA264, FA265, FA677, FA678, FA679, FA680, FA681, FA683, FF1392, FF1394, FB1008, FB1009, FF0085, FF0086, FF0312, FB0341, FA527, FB0067, FF0612, FB0108, FF1359

O parâmetro F_{IS} das populações do lambari de rabo amarelo (*A. lacustris*) apresentou uma estimativa média de 0,57 (tabela 3). As populações da calha do Alto, Médio e Baixo Rio Doce e dos Rios Piracicaba, Piranga e Santo Antônio apresentaram um valor de F_{IS} acima de 0,5, sugerindo isolamento reprodutivo e conseqüentemente acasalamentos não aleatórios entre os indivíduos (maior taxa de endocruzamento entre indivíduos aparentados). Comparando as três subdivisões da bacia, a região do Baixo Rio Doce apresentou valores mais baixos, sugerindo níveis maiores de cruzamento entre espécimes não aparentados. O Médio Rio Doce apresentou valores intermediários e o que apareceu com maior taxa de endocruzamento foi o Alto Rio Doce. O maior tributário do Alto Rio Doce, o Rio Piranga apresenta níveis mais baixos de endocruzamento em comparação com a calha do Alto Rio Doce e isso pode estar relacionado com os efeitos demográficos da passagem do rejeito nos rios Gualaxo do Norte e do Carmo. O Médio Rio Doce apresenta tributários com níveis menores ainda de endocruzamento, como os rios Caratinga e Casca, não afetados pela passagem do rejeito.

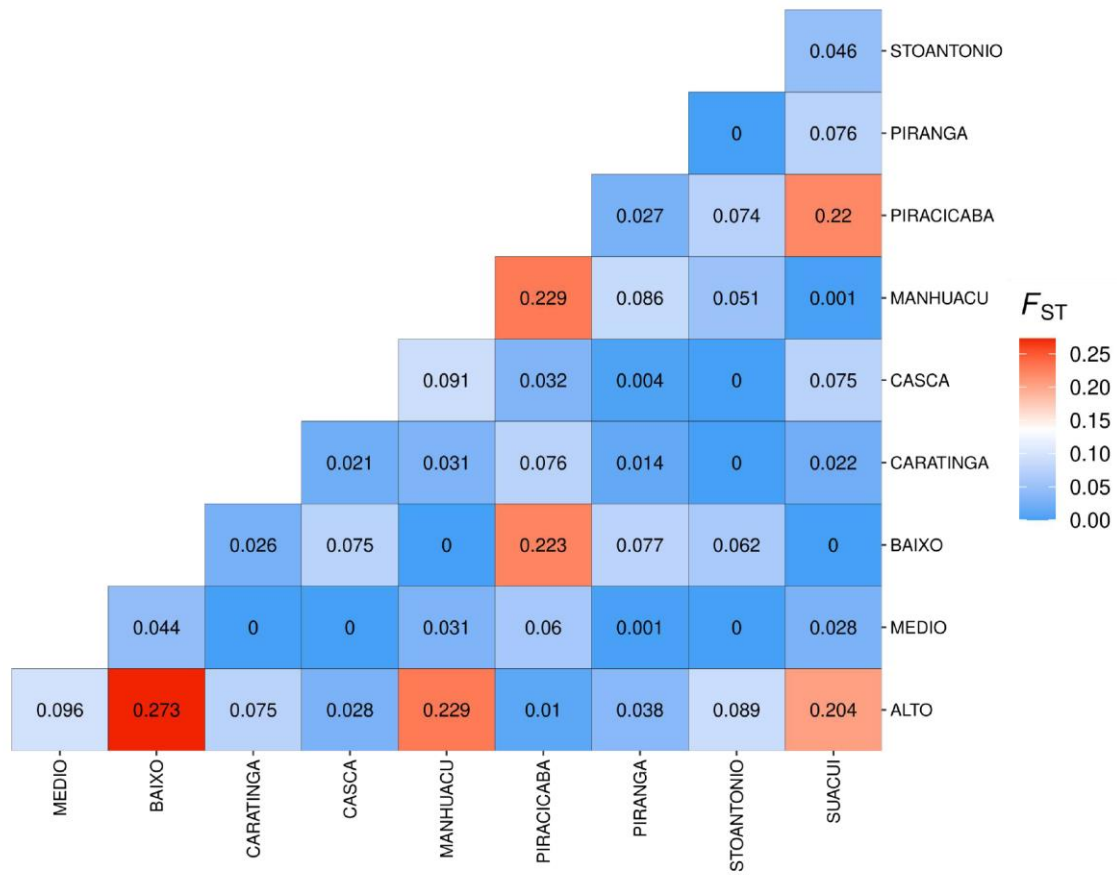
Ainda em relação ao F_{IS} , as populações do Manhuaçu e do Suaçuí apresentaram os valores 0,35 e 0,29 para F_{IS} , respectivamente, indicando uma menor taxa de endocruzamento ou populações grandes, com indivíduos que cruzam entre si. A população do Rio Santo Antônio, apesar de ser um hotspot de biodiversidade aquática, também apresentou um nível relativamente alto de endocruzamento, algo que já tinha sido evidenciado por Paiva et al. (2006) utilizando marcadores RAPD-PCR.

Tabela 3. Heterozigosidade observada (H_o), Heterozigosidade esperada (H_e) e sua aplicação no cálculo do Coeficiente de endocruzamento (F_{IS}) nas populações estudadas de *Astyanax lacustris*

População	H_o	H_e	F_{IS}
Alto Rio Doce	0,1042	0,2946	0,6464
Médio Rio Doce	0,0933	0,2482	0,6242
Baixo Rio Doce	0,0674	0,1371	0,5086
Caratinga	0,0944	0,1734	0,4556
Casca	0,1262	0,2506	0,4964
Manhuaçu	0,0727	0,1113	0,3562
Piracicaba	0,1206	0,2701	0,5535
Piranga	0,1299	0,2631	0,5061
Santo Antônio	0,0927	0,2498	0,6287
Suaçuí	0,0873	0,1237	0,2944
MEDIA	0,0882	0,2079	0,5756

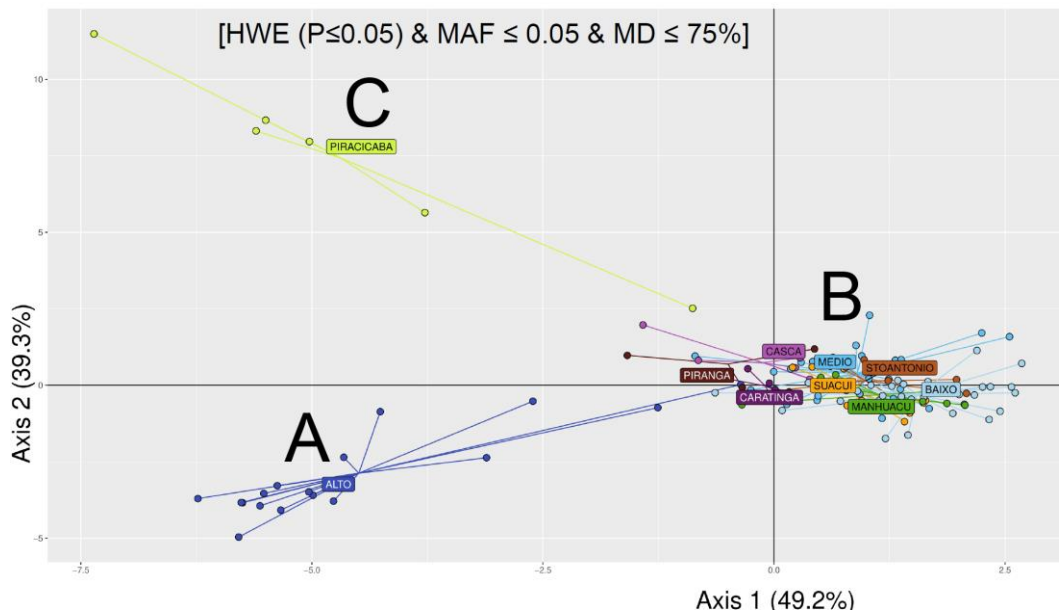
Os valores de F_{ST} entre as nove localidades variaram de 0 a 0,27 (Figura 5). Os maiores valores observados de divergência alélica indicada pelo F_{ST} foram entre o Baixo e o Alto Rio Doce (0,27), sugerindo maior grau de diferenciação entre esses dois segmentos da bacia. As comparações par-a-par entre os Rios Piracicaba x Suaçuí; Piracicaba x Manhuaçu; Piracicaba x Baixo Rio Doce e entre o Alto Rio Doce x Rio Manhuaçu (todos 0,22) indicaram uma estruturação moderada entre essas populações. Finalmente, os valores de F_{ST} iguais a zero foram entre as populações do Rio Santo Antônio x Piranga, Santo Antônio x Casca, Santo Antônio x Caratinga, Baixo x Manhuaçu, Baixo x Suaçuí, Médio x Caratinga, Médio x Casca, Médio x Santo Antônio. Os outros valores de F_{ST} observados variaram de 0,001 – 0,096, e, assim como os valores iguais a zero, indicam uma baixa diferenciação genética (ou ausência de populações diferenciadas) entre essas populações. Quando comparamos as seções da bacia com as amostras das localidades restantes, a que apresenta maior número de casos de estruturação populacional (ou maior grau de isolamento) é o Alto Rio Doce (em relação ao Baixo e aos rios Manhuaçu e Suaçuí). Em condição intermediária está o Baixo Rio Doce (em relação Rio Piracicaba) e finalmente, o Médio Rio Doce não apresentou estruturação em relação às outras localidades. Em geral, o padrão pode ser interpretado como uma interconectividade reduzida entre o Alto e o Baixo Rio Doce e uma população pouco diferenciada no Médio Rio Doce.

Figura 5. Valores do F_{ST} par a par entre populações para todos os indivíduos de *Astyanax lacustris*. Segundo Wright (1978) F_{ST} entre 0 e 0.05 indica pouca diferenciação genética; F_{ST} entre 0.05 e 0.15 indica diferenciação genética moderada; F_{ST} entre 0.15 e 0.25 indica alta diferenciação genética e F_{ST} acima de 0,25 diferenciação genética muito alta.



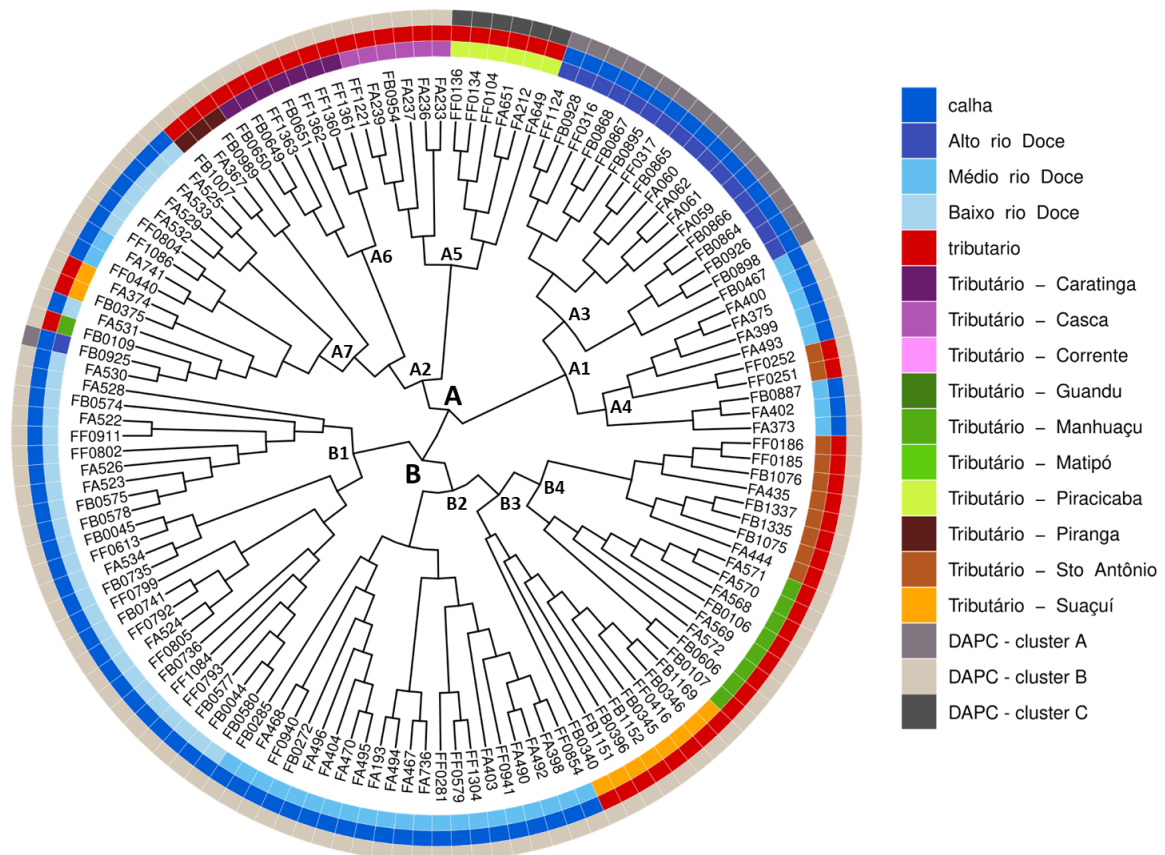
O agrupamento obtido pela DAPC é uma representação do padrão de estruturação indicado pelo F_{ST} e mostra três grupos genéticos: um formado por espécimes do Rio Piracicaba, um segundo, formado por espécimes da calha do Alto Rio Doce e um terceiro grupo que é uma sobreposição das outras localidades ou populações (Figura 6). Esse padrão explica 88,5% da variação genética total.

Figura 6. Análise Discriminante de Componentes Principais das variações de SNPs gerados pela ferramenta RADSeq, obtidas em populações do lambari-de-rabo amarelo *Astyanax lacustris*



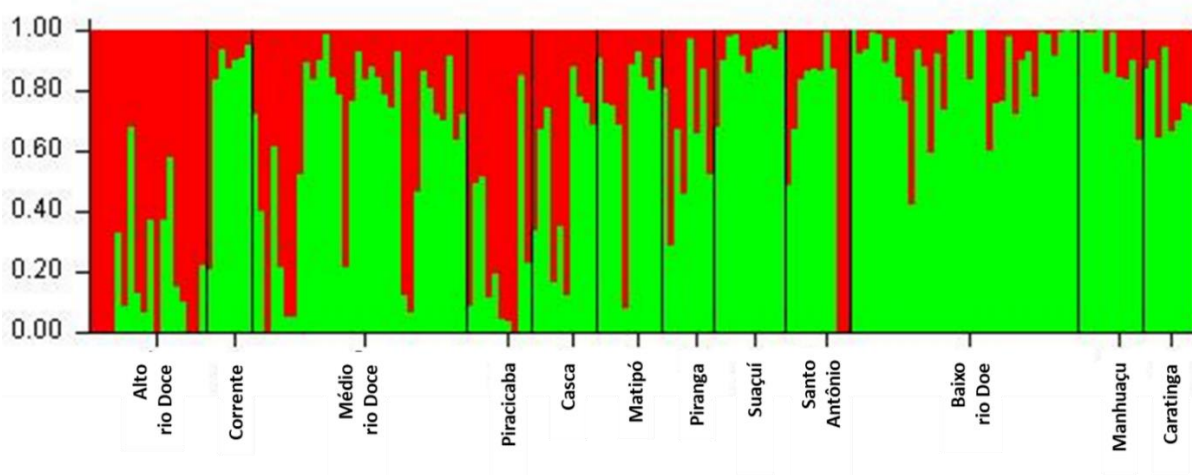
A análise de agrupamento hierárquico feita a partir do método de Ward classificou os espécimes de *A. lacustris* primeiramente em dois grupos: A e B (figura 7). Posteriormente o grupo A dividiu-se em dois outros grupos, A1 e A2. Em A1 temos um agrupamento (A3) formado por espécimes da calha do alto rio Doce + um único espécime da calha do médio rio Doce e um agrupamento (A4) que mistura espécimes da calha do alto e médio rio Doce, além de espécimes do tributário Santo Antônio. No agrupamento A2, encontramos um cluster (A5) formado por dois tributários: Piracicaba e Casca. Dentro os outros grupos formados em A2, um deles é composto por espécimes do tributário Caratinga (A6) e o outro (A7) é composto por espécimes tributário Piranga + espécimes da calha do médio e baixo rio Doce + espécimes do tributário Suaçuí + espécime do tributário Guandu. A primeira divisão do grupo B forma um grupo composto somente por espécimes da calha do baixo Rio Doce (B1) e outro composto por espécimes da calha do médio Rio Doce (B2), tributário Suaçuí (B3) e tributário Santo Antônio e Manhuaçu (B4).

Figura 7. Dendrograma hierárquico para os indivíduos analisados de *A. lacustris* (cada um deles é representado por letras e números únicos), gerado pela matriz de distâncias pareadas calculada partir da tabela de genotipagem usando a distância Euclidiana usando o método Ward.D2 implementado na função hclust do R version. Os três grupamentos da Figura 6 (A, B e C) estão representados por três tons de cinza no anel mais externo. O anel intermediário representa duas cores (calha e tributário). O anel interno, representa as três seções da bacia (Alto, Médio e Baixo Rio Doce) e cada tributário com uma cor única.



O resultado da análise bayesiana de agrupamento apresentou dois grupos genéticos distintos ($K=2$, vermelho e verde) para o lambari *A. lacustris*, com todas as populações apresentando uma mistura desses dois grupos em sua composição genética (Figura 8). A contribuição de cada grupo é variável por localidade, mas aparece de forma mais homogênea no Alto Rio Doce, onde predomina a cor vermelha e no Baixo Rio Doce, onde predomina a cor verde. Este resultado demonstra os dois extremos de diferenciação genética entre estas duas seções da bacia.

Figura 8. Gráfico de estrutura populacional mostrando o agrupamento das populações em dois componentes genéticos distintos para *A. lacustris*. Cada barra vertical representa um indivíduo. Os componentes genéticos identificados são; população 1 (verde) e população 2 (vermelho).



b. *Loricariichthys castaneus*

No caso do cascudo-viola, a análise envolveu a remoção inicial de 14.097 loci, por desviarem do equilíbrio de Hardy-Weinberg (HWE); em seguida, foram excluídos outros 418 loci que apresentaram menor frequência alélica (MAF). Em seguida foram retirados os loci e os indivíduos que possuíam mais que 75% de dados ausentes, que contabilizaram 3540 loci e 32 indivíduos. Após essa filtragem dos dados um total de 111 indivíduos e 1519 loci foram utilizados para as análises desta espécie (Tabela 4).

Tabela 4. Filtros aplicados aos dados de sequenciamento do cascudo-viola *L. castaneus*. HWE (Equilíbrio Hardy Weinberg); MAF (menor frequência alélica)

Filtro	Indivíduos	Loci
Stacks*	143	15.588
HWE ($P < 0.05$)	143	1.491
MAF < 0.05	143	1.073
Missing Data (75%) – loci	143	3540
Missing Data (75%) – indivíduos**	111	1.519
*predição <i>de novo</i>		
**indivíduos removidos:		
FA322, FA329, FB0851, FB0885, FB0886, FB0918, FB0929, FB0977, FF0026, FF0027, FA044, FA050, FB0973, FA694, FB0063, FB0073, FF0830, FF1082, FB0169, FB0534, FB170, FF1045, FB0400, FB0401, FB0404, FB0405, FB0406, FB1218, FF0352, FF0403, FF0404, FF0408		

Os resultados da análise de F_{IS} para *L. castaneus* indicaram uma variação de -0,0122 a 0,40 entre as populações da espécie nos pontos de coleta, com uma média estimada de 0,24. A população de *L. castaneus* do Rio Santo Antônio obteve o menor valor de F_{IS} (-0,0122) (Tabela 5), seguida pelos rios Manhuaçu (0,0657) e Guandu (0,125), indicando uma menor taxa de endocruzamento nesses locais e, possivelmente, populações maiores desta espécie nestes locais. No Rio Santo Antônio, esses valores são opostos aos encontrados para o lambari *A. lacustris*, que mostrou uma condição de maior

endocruzamento neste tributário. Os tributários Manhauçu e Guandu apresentaram também valores compatíveis com grandes populações, sendo estes três tributários, locais não expostos ao rejeito de minério. Outros tributários, como o Piranga, Casca e Piracicaba também apresentaram valores mais baixos de F_{IS} . Chama a atenção que dentre as três seções da calha, o Baixo Rio Doce apresentou valores superiores de F_{IS} , sugerindo um gargalo nessa seção. O panorama é complexo e pode representar questões históricas específicas, como a encontrada no Rio Matipó, o qual também apresentou altos valores de F_{IS} , superiores inclusive aos observados nas seções alta e média do Rio Doce. A heteroziguidade observada variou de 0,21 para o baixo Rio Doce a 0,30 para rio Piracicaba, obtendo uma média de 0,23. Já a heteroziguidade esperada variou de 0,24 para o Manhauçu e 0,36 para o Rio Piracicaba (tabela 5).

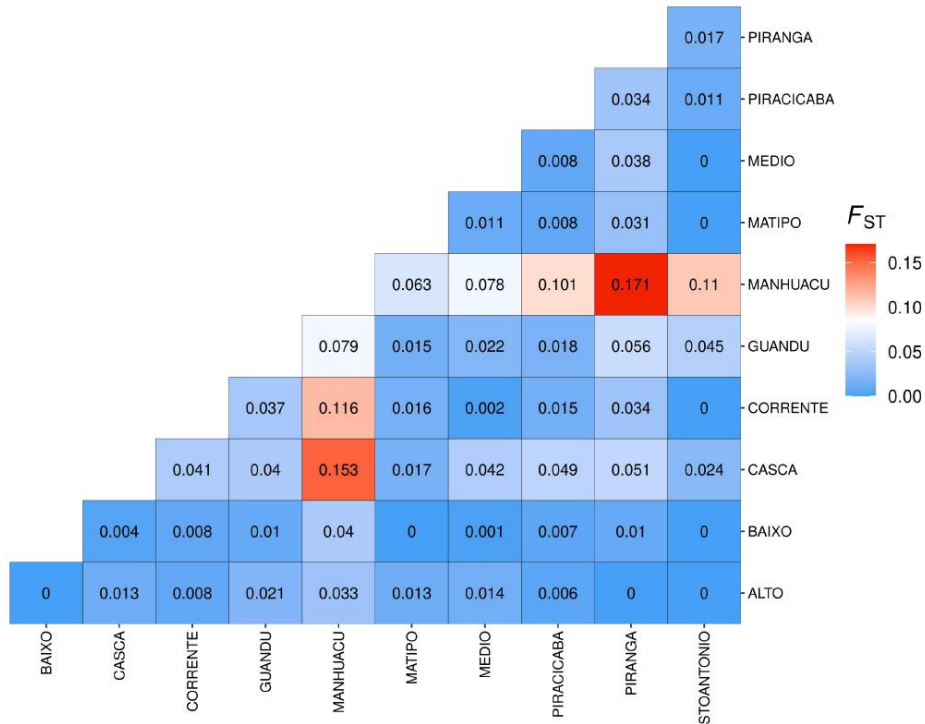
Tabela 5. Heteroziguidade observada (H_o), Heteroziguidade esperada (H_e) e sua aplicação no cálculo do Coeficiente de endocruzamento (F_{IS}) nas populações estudadas de *Loricariichthys castaneus*

	H_o	H_e	F_{IS}
Calha alto Rio Doce	0,2425	0,328	0,2606
Calha médio Rio Doce	0,2547	0,356	0,2844
Calha baixo Rio Doce	0,2119	0,3537	0,4008
Casca	0,2607	0,3185	0,1815
Corrente	0,2452	0,3354	0,2687
Guandu	0,2294	0,2622	0,125
Manhauçu	0,2271	0,2431	0,0657
Matipó	0,2295	0,3526	0,3492
Piracicaba	0,3031	0,3631	0,1654
Piranga	0,2317	0,2756	0,1593
Santo Antônio	0,2918	0,2883	-0,0122
Média geral	0,2348	0,3123	0,2482

No tocante à diferenciação genética entre as populações, as seções alta, média e baixa da bacia não apresentaram estruturação relevante (podem ser considerados uma população só) (Figura 9) e as maiores divergências foram observadas entre as localidades dos rios Piranga e Manhauçu (0,171) em relação ao restante da bacia, conforme indicado no DAPC (Figura 10). Os valores que indicaram maior diferenciação populacional foram observados entre as populações dos rios Manhauçu e Piranga (os valores mais altos de F_{ST} foram de 0,171, seguido pelo segundo valor mais alto (0,153). Em geral, observou-se que a população de cascudo-viola do Rio Manhauçu foi a mais diferenciada da bacia (Figura 9).

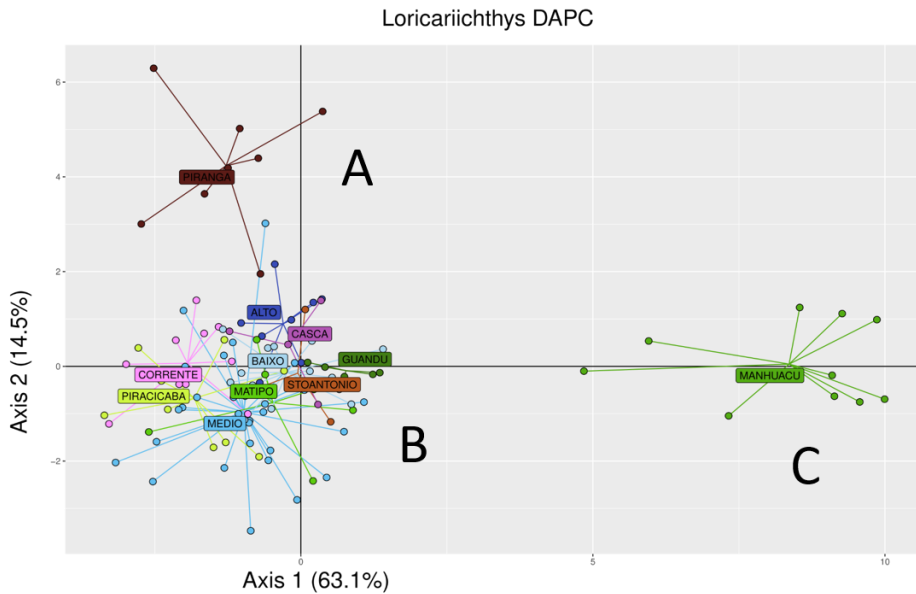
Em contrapartida, níveis mais altos de fluxo gênico estão expressos nos valores de F_{ST} iguais a zero entre as populações do Santo Antônio x calha do médio Rio Doce, Santo Antônio x Matipó, Santo Antônio x Corrente, Santo Antônio x calha do baixo Rio Doce, Santo Antônio x Calha do alto Rio Doce, Calha do alto Rio Doce x Piranga, Calha do alto Rio Doce x Calha do baixo Rio Doce, Calha do baixo Rio Doce x Matipó. Os outros valores de F_{ST} observados variaram de 0,001 – 0,079, e, assim como os valores iguais a zero, indicam uma baixa diferenciação genética para essas populações.

Figura 9. Valores do F_{ST} par a par entre populações para todos os indivíduos de *L. castaneus*. Segundo Wright (1978) F_{ST} entre 0 e 0.05 indica pouca diferenciação genética; F_{ST} entre 0.05 e 0.15 indica diferenciação genética moderada; F_{ST} entre 0.15 e 0.25 indica alta diferenciação genética e F_{ST} acima de 0,25 diferenciação genética muito alta.



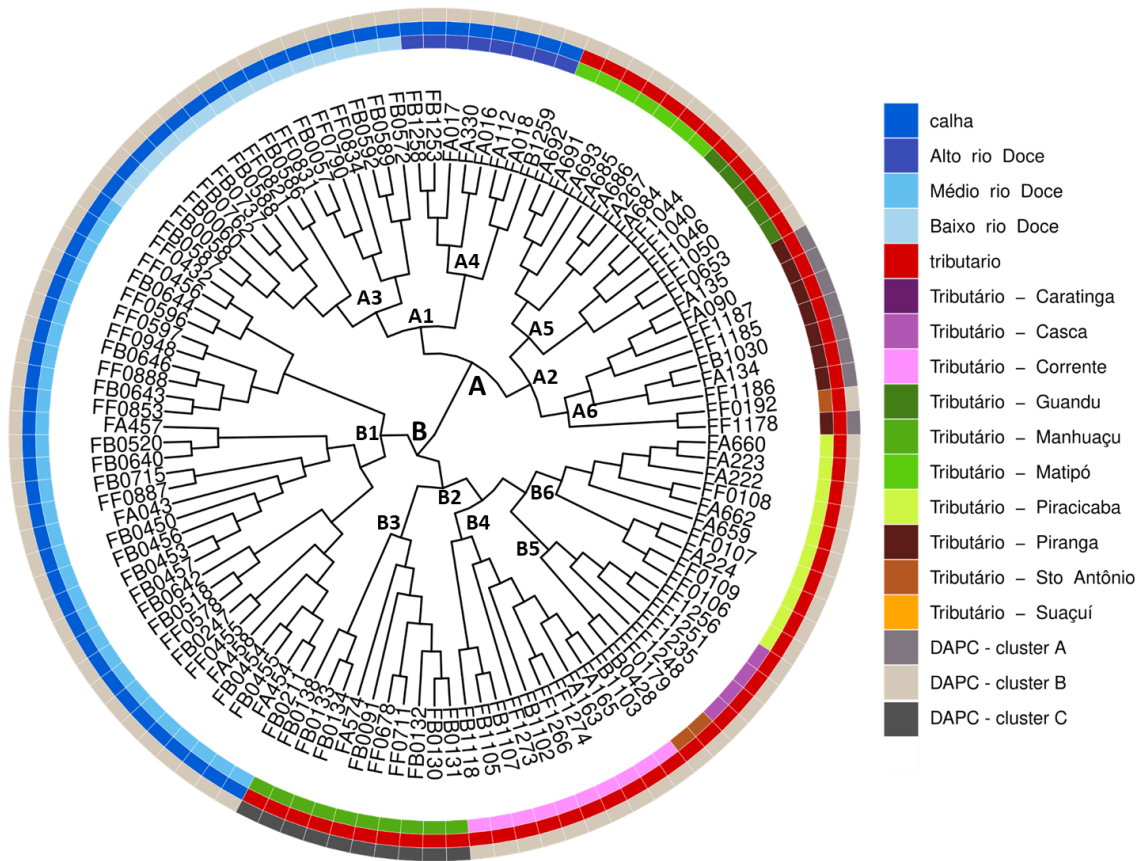
O modelo de agrupamento sugerido pela DAPC para *L. castaneus* mostra três grupos genéticos: um formado por espécimes do Rio Piranga, um segundo, formado por espécimes do Rio Manhuaçu e um terceiro grupo que é uma sobreposição das outras populações (Figura 10). Esse modelo explica 77,6% da variação genética total.

Figura 10. Análise Discriminante de Componentes Principais das variações de SNPs gerados pela ferramenta RAD-seq, obtidas em populações de *Loricariichthys castaneus*.



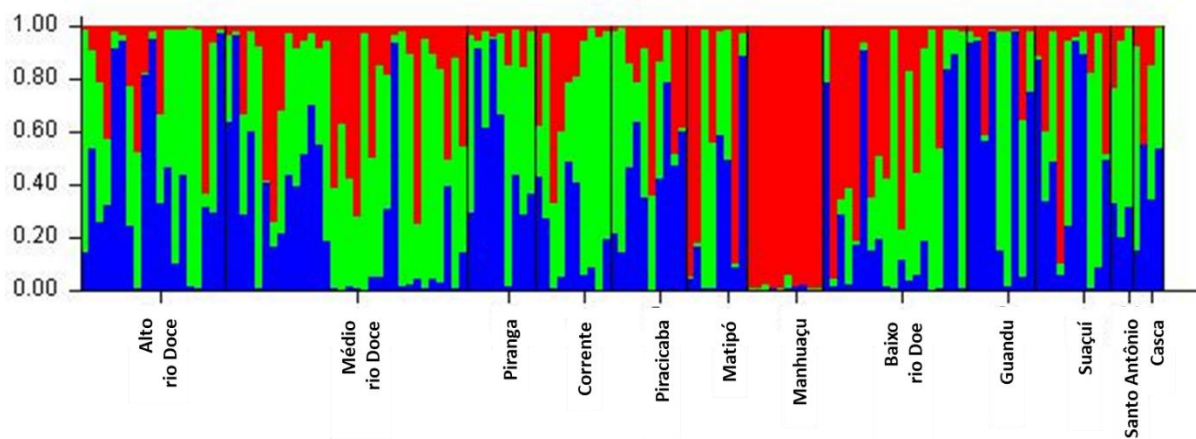
A análise de agrupamento hierárquico feita a partir do método de Ward classificou os espécimes de *L. castaneus* em dois grupos maiores: A e B (figura 11). O grupo A dividiu-se em dois outros grupos, A1 e A2. Em A1 temos um agrupamento (A3) formado por espécimes da calha do Baixo Rio Doce e um agrupamento (A4) que é composto por indivíduos da calha do Alto Rio Doce. No agrupamento A2, encontramos um cluster (A5) formado por dois tributários: Guandu e Matipó. O outro agrupamento em A2 é formado por espécimes do tributário Piranga e um espécime do tributário Santo Antônio. Em B temos dois grandes grupos, o primeiro formado somente por espécimes da calha do médio Rio Doce (B1) e um segundo, formado por tributários (B2). A primeira divisão do grupo B2 forma um grupo composto somente por espécimes do tributário Manhuaçu (B3). B2 então origina um cluster com três outros agrupamentos: o primeiro formado por espécimes do tributário Corrente (B4), outro composto por espécimes do Rio Casca + Rio Santo Antônio (B5) e o último formado por espécimes do Rio Piracicaba (B6).

Figura 11. Dendograma hierárquico gerado para os indivíduos analisados de *L. castaneus* (cada um deles é representado por letras e números únicos) pela matriz de distâncias pareadas calculada partir da tabela de genotipagem usando a distância Euclidiana usando o método Ward.D2 implementado na função hclust do R. Os três grupamentos da Figura 10 (A, B e C) estão representados por três tons de cinza no anel mais externo. O anel intermediário representa duas cores (calha e tributário). O anel mais interno, representa as três seções da bacia (Alto, Médio e Baixo Rio Doce) e cada tributário com uma cor única.



O resultado da análise bayesiana de agrupamento demonstrou três grupos genéticos distintos ($K=3$, vermelho, azul e verde) para o Cascudo *L. castaneus* (Figura 12), com os espécimes do Rio Manhuaçu apresentando majoritariamente, apenas um dos três componentes genéticos. As outras populações apresentaram uma mistura dos três grupos em sua composição genética.

Figura 12. Gráfico de estrutura populacional mostrando o agrupamento das populações em três componentes genéticos distintos para *L. castaneus*. Cada barra vertical representa um indivíduo. Os componentes genéticos identificados são; população 1 (verde), população 2 (vermelho), população 3 (azul).



c. *Geophagus brasiliensis*

Na análise do acará *Geophagus brasiliensis*, foi realizada a filtragem de dados 66.371 loci, por desviarem do equilíbrio de Hardy-Weinberg (HWE). Em seguida foram excluídos outros 67 loci que apresentaram menor frequência alélica (MAF). Em seguida foram retirados os loci e os indivíduos que possuíam mais que 75% de dados ausentes, que contabilizaram 259 loci e 07 indivíduos. Após essa filtragem dos dados um total de 60 indivíduos e 1.519 loci foram utilizados para as análises desta espécie (Tabela 6).

Tabela 6. Filtros aplicados aos dados de sequenciamento de *G. brasiliensis*. HWE (Equilíbrio Hardy Weinberg); MAF (menor frequência alélica)

Filtro	Indivíduos	Loci
Stacks*	67	68.216
HWE ($P < 0.05$)	67	1.845
MAF < 0.05	67	1.778
Missing Data (75%) – loci	67	1.519
Missing Data (75%) – indivíduos**	60	1.519

*Genoma de referência: *Archocentrus centrarchus*;
 Assembly (NCBI ID): fArcCen1 (GCF_007364275.1)
 **indivíduos removidos: FB0166, FB1217, FF0118, FF0221, FF0247, FF0668, JD11312

Os resultados da análise de F_{IS} para o acará *Geophagus brasiliensis* indicaram uma estimativa média de grau de endocruzamento das populações na bacia do Rio Doce bem alta, próxima de 1 (Tabela 7). Esse resultado sugere um alto nível de endogamia ao longo da calha principal e tributários para essa espécie, configurando um padrão muito diferente das duas espécies anteriores. As três seções da calha apresentaram valores semelhantes, embora foi observado um aumento gradual do endocruzamento (com valores aumentados de F_{IS}) desde a região alta à região baixa da bacia. Os

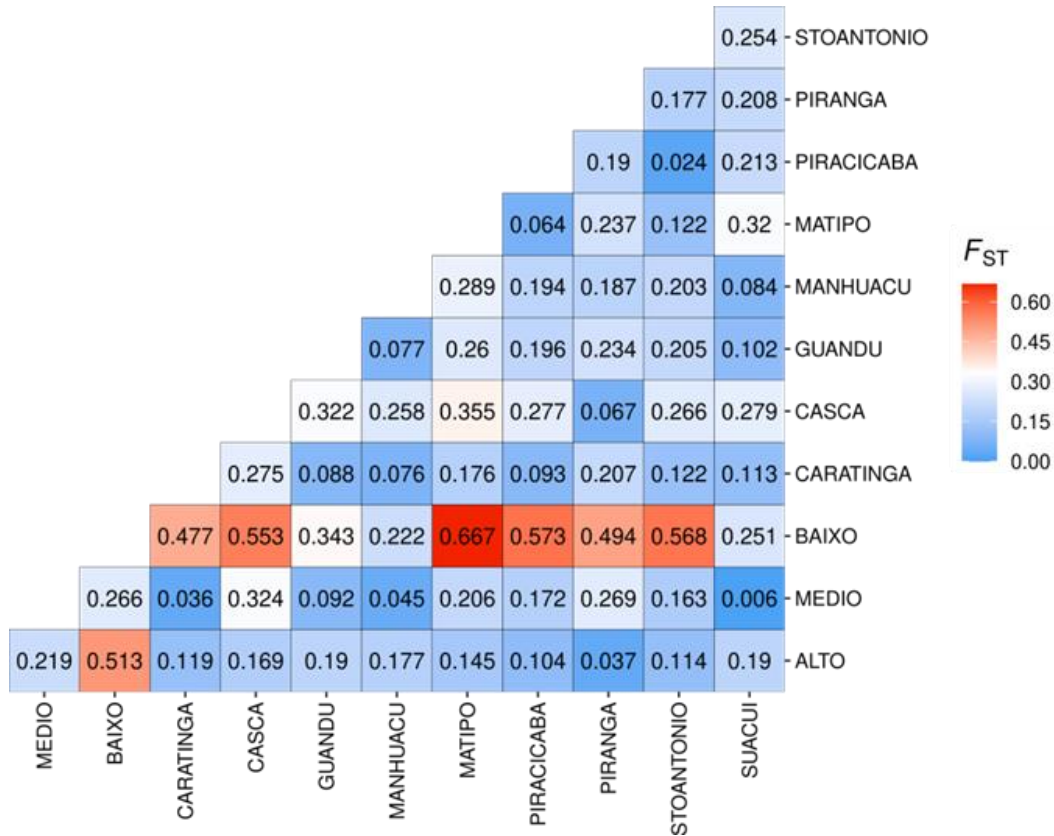
valores foram mais baixos em vários dos tributários, particularmente nos rios Caratinga, Piranga, Matipó e Manhuaçu. Esta espécie pode potencialmente ser uma boa indicadora da recuperação das condições do rio porque a pesar de ser abundante, seus parâmetros genéticos parecem ser mais sensíveis às alterações demográficas. A heterozigidade observada variou de 0,0044 para o Rio Matipó a 0,0421 para Rio Caratinga, obtendo uma média de 0,0154. Já a heterozigidade esperada variou de 0,0366 para o Matipó e 0,29 para o Rio Piranga (Tabela 7).

Tabela 7. Heterozigidade observada (H_o), Heterozigidade esperada (H_e) e sua aplicação no cálculo do Coeficiente de endocruzamento (F_{IS}) nas populações estudadas de *Geophagus brasiliensis*

	H_o	H_e	F_{IS}
Calha alto Rio Doce	0,0203	0,2317	0,9124
Calha médio Rio Doce	0,0178	0,2269	0,9216
Calha baixo Rio médio	0,0096	0,137	0,9298
Caratinga	0,0421	0,1646	0,7441
Casca	0,019	0,2293	0,9668
Guandu	0,0199	0,2084	0,9044
Manhuaçu	0,0236	0,2355	0,8997
Matipó	0,0044	0,0366	0,8806
Piracicaba	0,0162	0,131	0,876
Piranga	0,0216	0,2901	0,9254
Santo Antônio	0,0046	0,0615	0,9252
Suaçuí	0,0198	0,1647	0,8799
Média geral	0,0154	0,194	0,9207

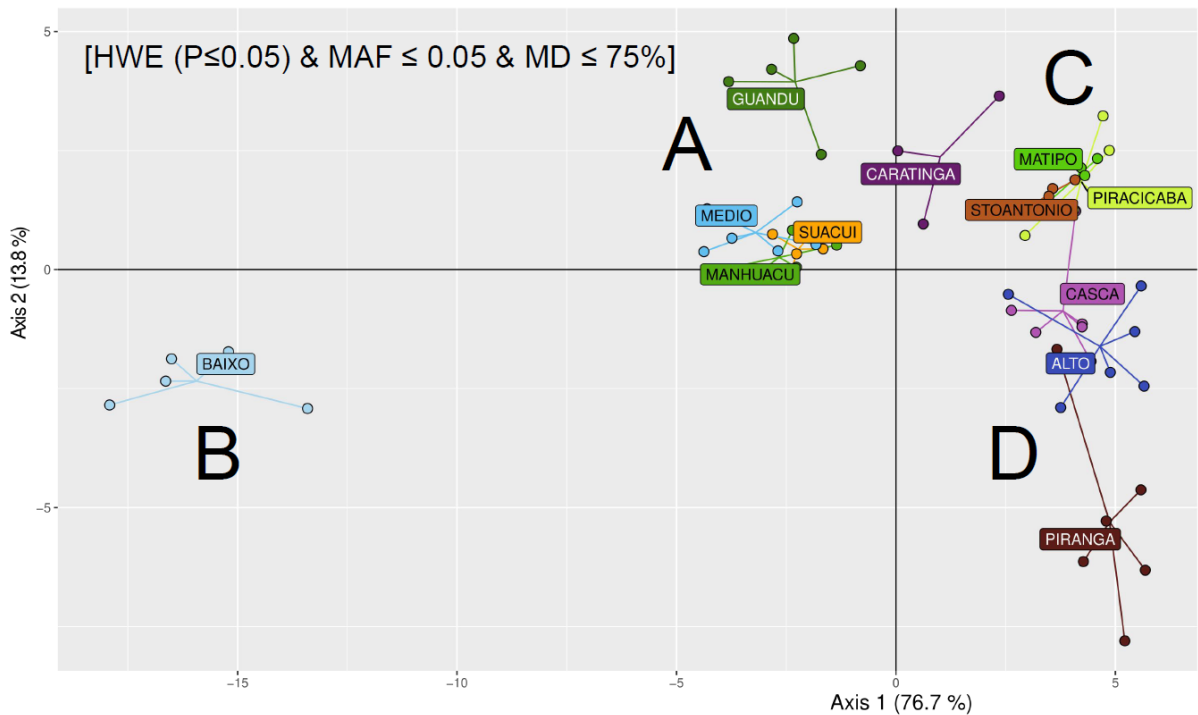
Em relação ao grau de estruturação (diferenciação genética) das populações, expressas pelo parâmetro F_{ST} na comparação par a par de *Geophagus brasiliensis* resultou em valores, em sua maioria, acima de 0,25 (Figura 13). Isso indica uma diferenciação genética média para essa espécie; chama a atenção a condição de forte estruturação do Baixo Rio Doce em relação a quase todos os outros locais, exceto os tributários Guandu, Manhuaçu e Suaçuí.

Figura 13. Valores do F_{ST} par a par entre populações para todos os indivíduos de *G. brasiliensis*. Segundo Wright (1978) F_{ST} entre 0 e 0.05 indica pouca diferenciação genética; F_{ST} entre 0.05 e 0.15 indica diferenciação genética moderada; F_{ST} entre 0.15 e 0.25 indica alta diferenciação genética e F_{ST} acima de 0,25 diferenciação genética muito alta.



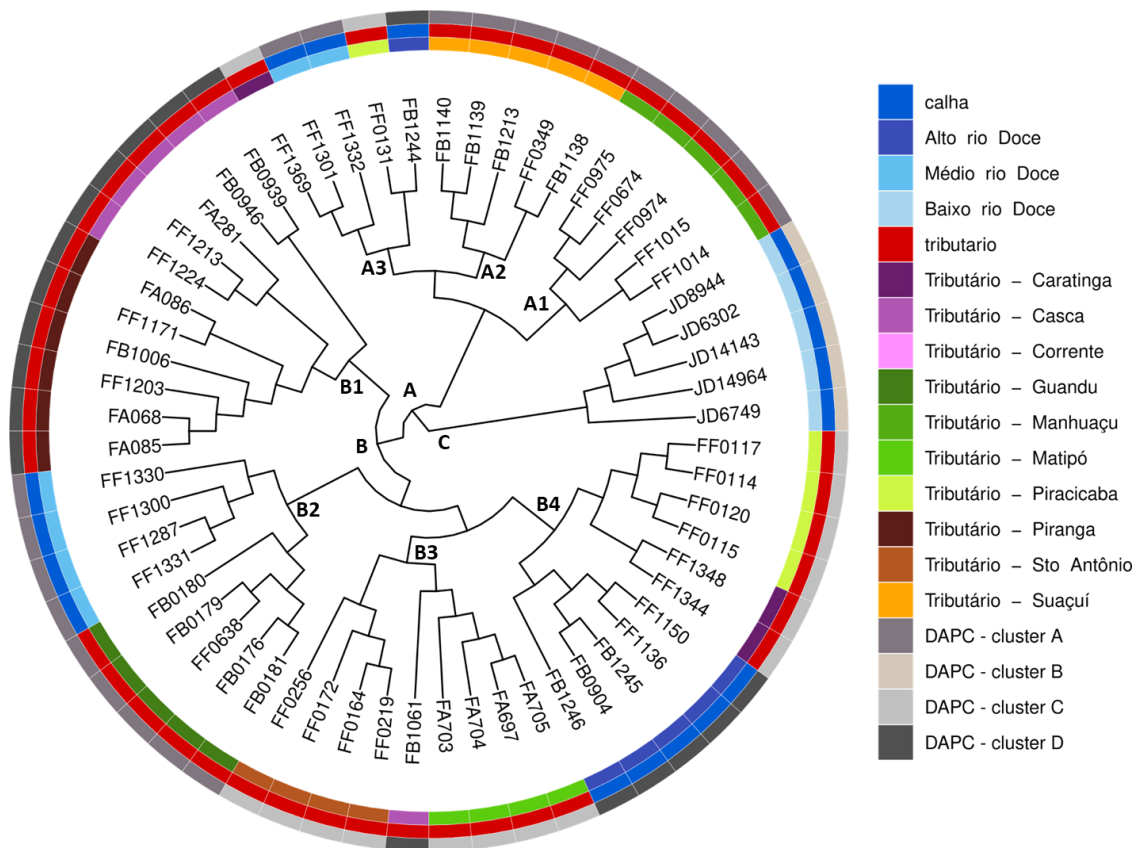
A análise discriminante de componente principal (DAPC) da espécie reflete a condição observada na comparação par-a-par das populações, reconhecendo 4 subgrupos, mas com uma diferenciação muito maior do Baixo Rio Doce (grupo B), com o eixo 1 explicando 76,7% da variação total (Figura 14). Embora as outras populações estejam mais agrupadas, elas não são tão próximas como foi observado nas duas espécies anteriores (Figuras 6 e 10), indicando uma maior diferenciação entre as populações.

Figura 14. Análise Discriminante de Componentes Principais das variações de SNPs gerados pela ferramenta RADSeq, obtidas em populações de *G. brasiliensis*. Os filtros utilizados para a seleção dos alelos foi: eliminação dos alelos que não se ajustam às expectativas de equilíbrio de Hardy-Weinberg (HWE), dos alelos que ocorrem em baixa frequência (MAF) e loci que não foram amplificados (MD).



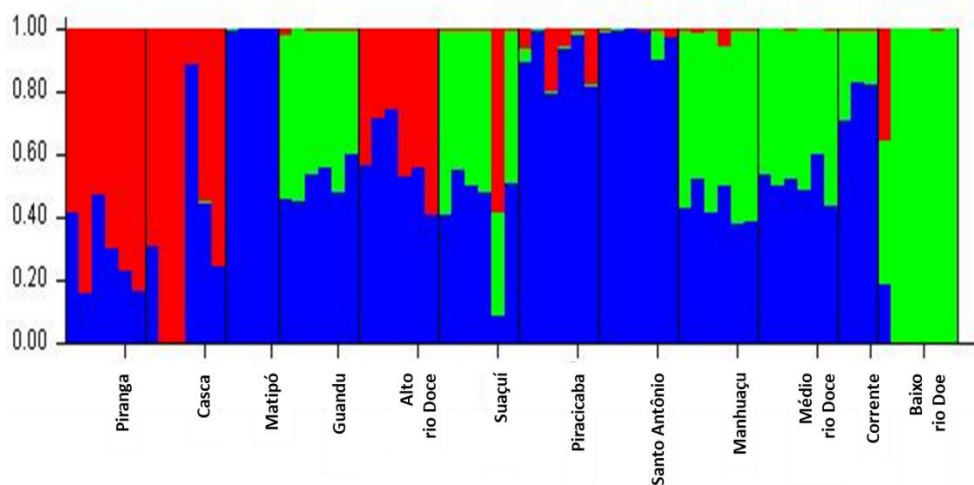
A análise de agrupamento hierárquico feita a partir do método de Ward classificou os espécimes de *Geophagus brasiliensis* em três grupos maiores: A, B e C (Figura 15). Posteriormente o grupo A dividiu-se em dois outros grupos, A1 composto por espécimes do tributário Manhuaçu, A2 formado por espécimes do tributário Suaçuí e A3, formado por uma mistura de espécimes da calha do alto Rio Doce, do tributário Piracicaba, da calha do médio rio Doce e do tributário Caratinga. Em B temos quatro grupos, B1 é formado por um grupo de espécimes do tributário Piranga e outro grupo por espécimes do tributário Casca. B2 também é formado por dois outros grupos, o primeiro por espécimes da calha do médio Rio Doce e o segundo, por espécimes do Rio Guandu. Em B3, temos um grupo formado por espécimes do Rio Santo Antônio e outro grupo composto por espécimes do rio Matipó, ao qual foi acrescentado um espécime do Rio Casca. B4, por sua vez, forma 3 outros agrupamentos, um deles formado somente por espécimes da calha do alto Rio Doce. Os outros dois são formados por espécimes Rio Caratinga e Piracicaba. E por fim, C é formado somente por espécimes da calha do baixo Rio Doce.

Figura 15. Dendograma hierárquico gerado para os indivíduos analisados de *G. brasiliensis* (cada um deles é representado por letras e números únicos) pela matriz de distâncias pareadas calculada partir da tabela de genotipagem usando a distância Euclidiana usando o método Ward.D2 implementado na função hclust do R. Os quatro grupamentos da Figura 14 (A, B, C e D) estão representados por quatro tons de cinza no anel externo. O anel intermediário representa duas cores (calha e tributário). O anel mais interno, representa as três seções da bacia (Alto, Médio e Baixo Rio Doce) e cada tributário com uma cor única.



A análise bayesiana de agrupamento gerou três grupos genéticos distintos ($K=3$, vermelho, azul e verde) (Figura 16) para o *Geophagus brasiliensis*, com Matipó, Santo Antônio e espécimes da calha do baixo Rio Doce formados majoritariamente por apenas um dos grupos genéticos. As outras populações apresentaram uma mistura desses três grupos em sua composição genética.

Figura 16. Gráfico de estrutura populacional mostrando o agrupamento das populações em três componentes genéticos distintos para *G. brasiliensis*. Cada barra vertical representa um indivíduo. Os componentes genéticos identificados são; população 1 (verde), população 2 (vermelho), população 3 (azul).



5) DISCUSSÃO

a. Variabilidade e estruturação genética das populações de peixes dos tributários e da calha principal da bacia do rio Doce

Das três espécies estudadas, o lambari *A. lacustris* e o cascudo-viola *L. castaneus* apresentaram espécimes coletados na calha, que são mais aparentados entre si que com espécimes coletados nos tributários (Figuras 7 e 11). Este padrão sugere que pelo menos estas espécies tiveram espécimes que sobreviveram a passagem do rejeito de minério e representam a primeira evidência que o desastre não eliminou totalmente a ictiofauna. A ampla distribuição destas espécies sugere que elas são extremamente tolerantes e são consistentes com os padrões observados nas figuras 5 e 9, as quais mostram um baixo grau de diferenciação genética dentro da bacia, mantida por grandes populações e altos níveis de fluxo gênico.

Astyanax lacustris apresentou valores baixos de heterozigosidade observada (entre 0,067 a 0,12), assim como para o Cascudo *L. castaneus* (entre 0,21 a 0,30) e o Acará *G. brasiliensis* (entre 0 e 0,04), demonstrando uma baixa variabilidade genética para as populações dessas espécies analisadas na calha principal e tributários da bacia do Rio Doce. Os mesmos resultados foram encontrados utilizando microssatélites (Dergam et al., 2022). A heterozigosidade média esperada aumenta para as espécies de 0,0882 para 0,20 para *A. lacustris*, de 0,23 para 0,31 para *L. castaneus* e de 0,01 para 0,19 para *G. brasiliensis*, porém ainda todas elas permaneceram em valores baixos, indicando a necessidade de monitoramento desse parâmetro, uma vez que a variabilidade genética é uma característica essencial para que a população enfrente as mudanças ambientais e desastres como por exemplo o rompimento da barragem da SAMARCO. Mesmo com baixos valores gerais de heterozigosidade, as populações dos rios Piracicaba e Piranga se repetem entre os maiores valores observados para as três

espécies. De uma maneira geral, não encontramos maior heteroziguidade para essas espécies nos tributários, como seria esperado devido ao impacto provocado pela passagem da lama na calha do rio Doce. Isso sugere que o longo histórico de degradação da bacia atinge também os tributários e medidas de recuperação devem estender-se a eles, aumentando o grau de conectividade dentro da bacia.

Exceto para *L. castaneus* no Rio Santo Antônio (-0,0122) e no Rio Manhuaçu (0,0657), cujos valores foram muito próximos de zero e negativos e indicam um excesso de heterozigotos, os valores de F_{IS} indicam uma deficiência de heterozigotos para a maioria das populações amostradas. Esse resultado sugere que esteja acontecendo endogamia nas populações. Devido à alta mortalidade de peixes ocasionada pelo rompimento da barragem, é de se esperar que essas populações estejam passando por um efeito *bottleneck* e tenham seu tamanho efetivo populacional reduzido, aumentando a probabilidade de acasalamentos não aleatórios, conforme indicado nos últimos três anos para a porção do Baixo Rio Doce (RRDM, 2022). Nesses casos é esperado também uma redução entre o número de alelos e consequentemente a redução da heteroziguidade (Allendorf et al., 2007).

Segundo Wright (1978) os valores de F_{ST} podem ser interpretados da seguinte maneira: F_{ST} entre 0 e 0,05 indica pouca diferenciação genética; F_{ST} entre 0,05 e 0,15 indica diferenciação genética moderada; F_{ST} entre 0,15 e 0,25 indica alta diferenciação genética e F_{ST} acima de 0,25 diferenciação genética muito alta. Num aspecto geral, *L. castaneus* teve uma diferenciação moderada, e somente a comparação entre o Rio Manhuaçu e o Rio Piranga apresentou alta diferenciação ($F_{ST}=1,71$). O cenário muda um pouco para *A. lacustris*, onde a comparação entre alto x baixo Rio Doce ($F_{ST} = 0,273$), indica que essas populações são bem definidas e divergentes entre si. Apesar dessas exceções, a grande maioria das comparações para essas duas espécies tiveram um valor de F_{ST} igual a zero. O longo histórico de degradação da bacia, agravado pela passagem do rejeito, determinou locais expostos ao rejeito de minério e com profundas alterações estruturais, os quais representam fatores potenciais de degradação e perda de habitat e de fragmentação de paisagem ao longo do rio, afetando a migração das espécies. Os resultados de F_{ST} contrariam essa interpretação, sugerindo a ocorrência de fluxo gênico atual entre as populações, mesmo entre as que são separadas por barreiras físicas, como hidrelétricas. Uma explicação possível para essa similaridade genética é explicada pela existência de altos níveis de fluxo gênico (Rousset, 1997) e a persistência de grandes populações que têm resiliência à perda de alelos (Allendorf, 1986).

Para *G. brasiliensis* a comparação par a par das populações do baixo Rio Doce com todas as outras populações indicam uma diferenciação genética muito alta. *G. brasiliensis* é um complexo de espécies crípticas (Alves-Silva & Dergam, 2015; Souza et al 2018; Argolo et al., 2020) e o que podemos estar vendo nos nossos resultados é um reflexo desse fenômeno. De acordo com Argolo et al. (2020) e com o relatório do PMBA (RRDM, 2022), a população do Baixo Rio Doce foi considerada como pertencente a *Geophagus santosi*. Por outro lado, os valores de F_{ST} mais baixos ocorreram em comparações com o Médio Rio Doce, sugerindo que *G.*

brasiliensis pode estar sofrendo os efeitos da baixa variabilidade genética causados pela degradação ambiental da porção mais populosa da bacia. É provável que esta espécie seja uma das que mais rapidamente desaparecem nos ambientes invadidos pela piranha-vermelha e pelos tucunarés, conforme observado nas lagoas do Médio Rio Doce (Fragoso-Moura, 2016; Dergam, obs. pess.) e no Baixo Rio Doce.

b. Conhecimento sobre a diversidade e a distribuição da ictiofauna da bacia do rio Doce

Nossos resultados demonstram que as populações de *G. brasiliensis* representam linhagens com alto grau de divergência molecular correlacionados à distribuição da espécie na bacia do Rio Doce. Os dados gerados pela análise de DAPC para *G. brasiliensis* complementam os resultados encontrados em microssatélites (Dergam, 2022), aumentando um agrupamento formado pelos tributários Santo Antônio, Matipó, Piracicaba e Casca à diversidade da espécie. A análise bayesiana de agrupamentos mostrou que esses quatro grupos da DAPC possuem componentes genéticos distintos, onde o agrupamento A (Guandu, Médio Rio Doce, Suaçuí e Manhuaçu) são formados por uma mistura dos componentes identificados pelas cores azul e verde; o agrupamento B (baixo Rio Doce) formado pelo componente genético identificado pela cor verde; o agrupamento C (Matipó, Piracicaba, Santo Antônio e Caratinga) formados pelos componentes genéticos identificados pela cor azul, vermelho e verde, e por fim o agrupamento D (Piranga, Casca, alto Rio Doce) formados por uma mistura dos componentes genéticos vermelho e azul (Figura 16). Os agrupamentos da DAPC também podem ser correlacionados com a distribuição da espécie na bacia, com o agrupamento A concentrando indivíduos encontrados nos tributários localizado no baixo Rio Doce ou mais próximos a essa porção da bacia, o agrupamento B por indivíduos exclusivamente da calha principal do baixo Rio Doce, o agrupamento C formado por espécimes dos tributários do médio, mais distantes da porção baixa da bacia e o agrupamento D, formado pela porção alta e tributários exclusivos do alto Rio Doce. Destacamos aqui a separação do rio Guandu do baixo Rio Doce, tanto nos agrupamentos sugeridos pela DAPC quanto na sua composição genética identificada pela análise bayesiana de agrupamentos. Esse rio encontra-se entre duas hidrelétricas, a de Aimorés a montante e a de Mascarenhas a Jusante. Nossos resultados indicam que a barragem de Mascarenhas pode representar uma forte barreira ao fluxo gênico de *G. brasiliensis*. Por outro lado, a barragem de Mascarenhas foi construída em 1974, e recentemente adotou a prática de transposição manual peixes (<https://www.canalenergia.com.br/noticias/53050435/edp-inicia-transposicao-manual-de-peixes-na-usina-de-mascarenhas>). O reservatório de Mascarenhas se caracteriza por uma grande abundância de espécies não-nativas que podem agir como filtro na contribuição do Rio Guandu com o Baixo Rio Doce.

Em contrapartida, os agrupamentos da DAPC encontrados em *A. lacustris* não são formados por componentes genéticos distintos e sim por uma homogeneização geral dos locais

amostrados, exceto a região do Alto Rio Doce e do Rio Piracicaba. O mesmo padrão de alto fluxo gênico foi observado no Cascudo *L. castaneus*. Esses resultados corroboram os baixos valores de F_{ST} que encontramos, confirmando *A. lacustris* e *L. castaneus* como espécies panmíticas (maior grau de cruzamento aleatório entre os indivíduos ou maiores populações) na bacia do Rio Doce.

c. Conectividade entre os tributários e a calha principal do rio Doce

A partir da análise de agrupamento hierárquico feita com o método de Ward analisamos quais tributários dividiam agrupamentos (*clusters*) com as porções da calha principal do Rio Doce, demonstrando assim as relações genéticas entre tributários e calha. Dessa forma, conseguimos relacionar a contribuição dos tributários para o aporte de peixes nas porções da calha principal do Rio Doce. Também utilizamos a análise de DAPC e as comparações par-a-par dos valores de F_{ST} para inferir sobre as relações genéticas a partir do agrupamento entre as populações nos quadrantes da análise.

Para o acará *G. brasiliensis*, os rios Piracicaba e Caratinga foram os tributários que apareceram mais agrupados com as porções da calha principal do Rio Doce. Isso indica que esses tributários representam uma importante fonte de peixes para a recolonização e recuperação da calha, porém somente nas porções média e alta. Já o Rio Guandu foi encontrado agrupado somente com a porção média do Rio Doce, sugerindo que, juntamente com o Rio Caratinga, esse tributário contribui para o aporte de peixes para a porção média da calha. Não encontramos nenhum cluster onde o Baixo Rio Doce fosse encontrado agrupado ou em agrupamentos próximos a tributários. Isso evidencia o isolamento da espécie no Baixo Rio Doce; um processo que pode refletir os efeitos da barragem da UHE Mascarenhas. Os parâmetros genéticos de população, como o Garza-Williamson (GW) analisados a partir de microssatélites, também sugerem forte endogamia de outras espécies nativas do Baixo Rio Doce (RRDM, 2022). A análise de DAPC mostra a população do Alto Rio Doce do acará *G. brasiliensis* é aparentada geneticamente com as populações dos tributários Casca e Piranga, enquanto a população do médio Rio Doce está geneticamente mais próxima com as populações dos tributários Suaçuí e Manhuaçu, além do Rio Guandu. O DAPC coloca as amostras do Baixo Rio Doce a uma grande distância de todas as restantes. Da mesma forma que foi observado no Médio Rio Doce (Fragoso-Moura *et al.*, 2016) os acarás são impactados pela presença das espécies não-nativas de tucunarés (*Cichla*) e possivelmente por piranhas.

Astyanax lacustris apresentou uma lista mais diversificada de tributários que mantém uma composição genética semelhante à da calha principal do Rio Doce. O Rio Guandu aparece mais uma vez agrupado com espécimes do médio Rio Doce. Destacamos também a importância dos rios Santo Antônio, Manhuaçu e Suaçuí como fonte de peixes para a recolonização dessa porção da calha. Encontramos o Baixo Rio Doce agrupado com espécimes do Médio Rio Doce, sugerindo que, exceto o Alto Rio Doce e o Rio Piracicaba, há uma única unidade demográfica.

Isso também é evidente no DAPC, onde o lambari *A. lacustris* apresenta um isolamento do alto Rio Doce.

Encontramos um cenário parecido para *L. castaneus*, com a diferença que nesta espécie, são as populações do Rio Manhuaçu e Piranga que se apresentam mais isoladas das populações restantes. A análise de DAPC mostra que o cascudo-viola *L. castaneus* representa em geral, uma única grande população na bacia do Rio Doce.

6) . CONCLUSÃO

As análises genéticas comparativas realizadas nesse trabalho para *A. lacustris*, *G. brasiliensis* e *L. castaneus* mostram que as três espécies analisadas apresentam padrões idiossincráticos de cada espécie. Além disso, para *A. lacustris* e *L. castaneus*, ocorreu um contingente de espécimes coletados na calha (alto, médio e baixo Rio Doce) que são mais aparentados entre si, sugerindo persistência de espécimes durante a passagem do rejeito de minério. Por outro lado, os níveis de variação detectados com RADSeq e a existência de populações diferenciadas também foram específicos de cada espécie. A diferenciação populacional foi mais baixa no lambari, intermediária no cascudo-viola e bem mais elevada no acará, sendo que neste caso, este resultado é esperado pela ocorrência de mais de uma espécie na bacia. No caso de redução da diversidade genética relacionada a seções da bacia, é possível que seja resultado do efeito conjunto de degradação ambiental e da pressão de espécies não-nativas.

Em relação aos objetivos propostos, os resultados permitiram:

- **Avaliar temporalmente a variabilidade e estruturação genética das populações de peixes dos tributários e da calha principal da bacia do Rio Doce.** A avaliação temporal ficou restrita ao esforço realizado nas duas campanhas realizadas em 2020, cinco anos após a passagem do rejeito de minério. *Loricariichthys castaneus* apresentou excesso de heterozigotos para o Rio Santo Antônio, mas para as outras populações desta e das outras espécies os resultados indicam um alto grau de endocruzamento. Os resultados de F_{ST} sugerem baixa estruturação com *L. castaneus* e *A. lacustris* apresentando um alto grau de panmixia. Já *G. brasiliensis* apresentou resultados que sugere diferenciação genética alta, principalmente para o baixo Rio Doce em relação ao restante da bacia.
- **Ampliar o conhecimento sobre a diversidade e a distribuição da ictiofauna da bacia do Rio Doce.** Nossos resultados mostram a diversidade gênica para as três espécies estudadas, com *L. castaneus* e *G. brasiliensis* apresentando três componentes genéticos distintos e *A. lacustris*, dois. Nosso trabalho corrobora a hipótese que *G. brasiliensis* da porção baixa do Rio Doce pode representar uma espécie diferente mais próxima a *Geophagus pintoí*, conforme os estudos de Argolo *et al.* (2020) e da RRDM (2022). Além disso, a análise de DAPC e valores de F_{ST} indicam que as populações do Rio Manhuaçu

e Piranga de *L. castaneus* possuem uma variação genética diferente das demais e aparentemente não participam do processo de repovoamento. O mesmo vemos para as populações da porção alta do Rio Doce e do tributário Piracicaba em *A. lacustris*.

- **Investigar a conectividade (fluxo gênico entre as populações de espécies de peixes) entre os tributários e a calha principal do Rio Doce.** O fluxo gênico foi avaliado conforme o grau de estruturação populacional de cada espécie a partir da análise do parâmetro F_{ST} e também pela relações encontradas na análise de DAPC. Nossos resultados indicam um alto grau de fluxo gênico em *L. castaneus* e *A. lacustris*, sendo interpretado que essas espécies apresentam um alto grau de panmixia na bacia do Rio Doce. Já *G. brasiliensis* demonstra ser mais estruturada e é provável que a falta de conectividade possa estar influenciando a variabilidade genética para essa espécie.
- **Produzir um estudo da história evolutiva recente da ictiofauna da bacia do Rio Doce.** Nesse trabalho geramos as bases genéticas para uma hipótese da história evolutiva dessas espécies na bacia do Rio Doce. Os padrões que observamos representam o marco-zero dos acontecimentos que já estavam presentes na bacia e que podem ter sido aprofundados pela passagem do rejeito de minério.
- **Produzir um ranking da importância ecológica e conservacionista de cada tributário baseado na contribuição potencial ou real do seu aporte de peixes para a calha principal.** Para o acará *G. brasiliensis*, os rios Piracicaba e Caratinga foram os tributários que apareceram mais agrupados com as porções da calha principal do Rio Doce, porém somente nas porções média e alta. Já o Rio Guandu foi encontrado agrupado somente com a porção média do Rio Doce. A população do Baixo Rio Doce permaneceu isolada dos tributários e das demais porções da calha. *Astyanax lacustris* apresentou uma lista maior de tributários que mantém altos níveis de fluxo gênico e que participam do processo de repovoamento, exceto a calha do alto Rio Doce e a população do Rio Piracicaba. A mesma condição foi observada no Cascudo-viola *L. castaneus*, os tributários demonstraram relação genética próxima com outros tributários, agrupando entre si, exceto a porção alta da calha do Rio Doce e do Rio Manhuaçu. Os resultados de *A. lacustris* e *L. castaneus* obtidos com RADSeq sugerem que as ações de conservação e recuperação devem abranger tanto os tributários quanto a calha, para manter os níveis de diversidade genética destas espécies. Para *G. brasiliensis*, estudos complementares são ainda necessários para determinar os riscos da espécie na calha do Rio Doce, considerando que a sua abundância está correlacionada com a presença de espécies não-nativas.

7) REFERÊNCIAS

- Allendorf, F.W. 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoobiology* 5(2): 181-190.
- Allendorf, F.W. 2007. Conservation and the genomics of populations. Oxford University Press.

Alves-Silva AP, Dergam JA (2015) Cryptic speciation within the Neotropical cichlid *Geophagus brasiliensis* (Quoy & Gaimard, 1824) (Teleostei Cichlidae): a new paradigm in karyotypical and molecular evolution. *Zebrafish*;12:91–101. <http://doi.org/10.1089/zeb.2014.0999>

Argolo, L.A., López-Fernández. H., Batalha-Filho. H., Affonso. P.R.A.M. 2020 Unraveling the systematics and evolution of the “*Geophagus*” *brasiliensis* (Cichliformes: Cichlidae) species complex. *Molecular Phylogenetics and Evolution*. 150(7):1–12. doi: 10.1016/j.ympev.2020.106855.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., et al. 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* 3(10): e3376. <https://doi.org/10.1371/journal.pone.0003376>

Bunn, S.E., Davies, P.M., Kellaway D.M. 1997. Contributions of sugar cane and invasive pasture grass to the aquatic food web of a tropical lowland stream. *Marine and Freshwater Research* 48: 173-179.

Carvalho, M.S., Moreira, R.M., Ribeiro, K.D., Almeida, A.M. 2017 Concentração de metais no rio Doce em Mariana, Minas Gerais, Brasil. *Acta Brasiliensis* 1(3): 37-41.

Consortio Ecoplan-Lume, 2010. Plano integrado de recursos hídricos da bacia hidrográfica do rio Doce e planos de ações para as unidades de planejamento e gestão de recursos hídricos no âmbito da bacia do rio Doce. Contrato Nº 043/2008 – IGAM, Vol. 1-Relatório Final, 2010. 472 p.In: www.cbhdoce.org.br, acessado em 15/01/2016.

CPRM/ANA, 2015. Monitoramento especial da bacia do rio Doce. Relatório 1-Acompanhamento da onda de cheia. CPRM, Belo Horizonte- MG, Dezembro 2015.In:www.cprm.org.br.

Costa A. T. 2006. Registro histórico de contaminação por metais pesados associados à exploração aurífera no alto e médio curso na bacia do ribeirão do Carmo. Departamento de geologia da Universidade Federal de Ouro Preto, Minas Gerais. Tese de doutorado, 258 p.

Dergam et al., 2022. Subprojeto III B. Dinâmica populacional e evolutiva da ictiofauna na bacia do Rio Doce. Análise populacional com uso de microssatélites. Relatório Técnico, UFV. 42 p.

Dosskey, M.G., Vidon, P., Gurwick, N.P., Allan, C.J., Duval, T.P., Lowrance, R. 2010 The role of riparian vegetation in protecting and improving chemical water quality in streams. *Journal of the American Water Resources Association* 46 (2): 261-277. doi:10.1111/j.1752.

Eleutério, L.1997. Diagnóstico da situação ambiental da cabeceira da bacia do rio Doce, MG, no âmbito das contaminações por metais pesados em sedimentos de fundo. Departamento de Geologia da Universidade Federal de Ouro Preto, Minas Gerais. Dissertação de mestrado, 154p.

EMBRAPA. 2015. Avaliação dos impactos causados ao solo pelo rompimento de barragem de rejeito de mineração em Mariana, MG: Apoio ao plano de recuperação agropecuária. Empresa Brasileira de Pesquisa Agropecuária. Dezembro de 2015.

Ferreira, A., Paula, F.R., Ferraz SFB, Gerhard, P., Kashiwaqui, E.A.L., Cyrino, J.E.P., Martinelli, L.A. 2012 Riparian coverage affects diets of characids in neotropical streams. *Ecology of Freshwater Fish* 2012 21: 12-22. <https://doi.org/10.1111/j.1600-0633.2011.00518.x>

Fragoso-Moura, E.N., Oporto, L.T., Maia-Barbosa, P.M., Barbosa, F.A.R. 2016. Loss of biodiversity in a conservation unit of the Brazilian Atlantic Forest: the effect of introducing non-native fish species. *Brazilian Journal of Biology* . 76: 18-27. <https://doi.org/10.1590/1519-6984.07914>

GIAIA. Relatórios técnicos. Novembro de 2015; Dezembro de 2015. In: <http://giaia.eco.br/>

Gonçalves, A.L. 2007: “As técnicas de mineração nas Minas Gerais do século XVIII”, In: RESENDE, Maria Efigênia Lage de; VILLALTA, Luiz Carlos (Orgs.). As Minas Setecentistas. Belo Horizonte: Autêntica; Companhia do Tempo, 2007. p. 159-186.

Governo de Minas Gerais 2016. Relatório: Avaliação dos efeitos e desdobramentos do rompimento da Barragem de Fundão em Mariana-MG. Secretaria de Estado de Desenvolvimento Regional, Política Urbana e Gestão Metropolitana, Minas Gerais, 287 p.

Instituto Brasileiro do Meio Ambiente e Recursos Naturais Renováveis – IBAMA.2015. Laudo Técnico Preliminar – Impactos ambientais decorrentes do desastre envolvendo o rompimento da barragem de Fundão, em Mariana, Minas Gerais. Diretoria de Proteção Ambiental - DIPRO & Coordenação Geral de Emergências Ambientais – CGEMA. Brasília, Novembro de 2015. 38 pp.

Instituto Chico Mendes de Conservação e Biodiversidade – ICMBIO 2018. Livro Vermelho da Fauna Brasileira Ameaçada de Extinção. Volume 4 – Peixes. Brasília – DF , p. 1-1232.

Instituto Mineiro de Gestão das Águas – IGAM, 2005 - DIAGNÓSTICO CONSOLIDADO DA BACIA. Disponível em: <http://www.riodoce.cbh.gov.br/Diagnostico2005/diagnostico.pdf>.

Instituto Mineiro de Gestão das Águas – IGAM. 2015. Acompanhamento da qualidade das águas do rio Doce após o rompimento da barragem da Samarco no distrito de Bento Rodrigues – Mariana/MG. Gerência de Monitoramento de Qualidade das Águas, Secretaria Estadual do Meio Ambiente, Governo do Estado de Minas Gerais, Belo Horizonte, 30 de Novembro (relatório 1) e 15 de Dezembro de 2015 (relatório 2). <http://www.igam.mg.gov.br/component/content/article/16/1632-monitoramento-da-qualidade-das-aguas-superficiais-do-rio-doce-no-estado-de-minas-gerais>.

Miranda, L.S., Marques, A.C., 2016. Hidden impacts of the Samarco mining waste dam collapse to Brazilian marine fauna an example from the staurozoans (Cnidaria). *Biota Neotropica* 16 (2), e20160169. <http://dx.doi.org/10.1590/1676-0611-BN-2016-0169>.

Paiva S.R., Dergam, J.A., Machado, F. 2006. Determining management units in southeastern Brazil: the case of *Astyanax bimaculatus* (Linnaeus, 1758) (Teleostei: Ostariophysi:Characidae) *Hydrobiologia* 560:393-404. <https://doi.org/10.1007/s10750-005-9415-1>

Piorski, N.M., Sanchez, A., Carvalho-Costa, L.F., Hatanaka, T., et al. 2008. Contribution of conservation genetics in assessing Neotropical freshwater fish biodiversity. *Brazilian Journal of Biology* 2008 68(4): 1039-1050. <https://doi.org/10.1590/S1519-69842008000500011>

Prefeitura de Baixo Guandu 2015. Relatório Analítico parcial SAAE. Disponível em: <http://www.pmbg.es.gov.br/adm/ckfinder/userfiles/files/RELATORIO%20002-63866-96.pdf>.

Prefeitura de Governador Valadares. 2015. Relatório SAAE. Disponível em: http://www.valadares.mg.gov.br/Materia_especifica/22642/SAAE-analisa-regularmente-agua-do-Rio-Doce

Ríos, S. L., Bailey, R.C. 2006. Relationship between riparian vegetation and stream benthic communities at three spatial scales. *Hydrobiologia* 553: 153-160. <https://doi.org/10.1007/s10750-005-0868-z>

Roque, F.O., Corbi, J.J., Trivinho-Strixino, S. 2003 Macroinvertebrates on different technosubstrates in a stream of an urban area of São Carlos - SP, Brazil. *Multiciência* 5: 172-177.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145(4): 1219-1228.

RRDM, Rede Rio Doce Mar. Relatório Anual do PMBA/Fest-RRDM – Evolução espaço-temporal do Ambiente Dulcícola. RT-39B, Programa de Monitoramento da Biodiversidade Aquática, Fundação Espírito-santense de Tecnologia. 613 pag. 2022.

Souza, C.R., Affonso, P.R.A.M.; Bitencourt, J.A., Sampaio, I. Carneiro, P.L.S. 2018. Species validation and cryptic diversity in the “*Geophagus*” *brasiliensis* Quoy & Gaimard, 1824 complex (Teleostei, Cichlidae) from Brazilian coastal basins as revealed by DNA analyses. *Hydrobiologia* 2018 ; 809(1): 309–321. doi: 10.1007/s10750-017-3482-y.

Vieira, F. 2009. Distribuição, impactos ambientais e conservação da fauna de peixes da bacia do rio Doce. *MG BIOTA*, 2: 5-22.

Weir, B.S., Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

Wright, S. 1978 *Evolution and the genetics of populations, volume 4: variability within and among natural populations*. University of Chicago Press.