

UNIVERSIDADE DO ESTADO DO AMAZONAS
ESCOLA NORMAL SUPERIOR
LICENCIATURA EM CIÊNCIAS BIOLÓGICAS

WALESKA ELIZANGELA DOS SANTOS BARBOSA

FILOGENIA DO COMPLEXO *Synallaxis rutilans* TEMMINCK, 1823 (AVES:
PASSERIFORMES: FURNARIIDAE), UTILIZANDO REPRESENTAÇÃO REDUZIDA
DO GENOMA POR CAPTURA DE SEQUÊNCIA

Manaus- AM

2017

WALESKA ELIZANGELA DOS SANTOS BARBOSA

FILOGENIA DO COMPLEXO *Synallaxis rutilans* TEMMINCK, 1823 (AVES:
PASSERIFORMES: FURNARIIDAE), UTILIZANDO REPRESENTAÇÃO REDUZIDA
DO GENOMA POR CAPTURA DE SEQUÊNCIA

Monografia apresentada como requisito
parcial para obtenção do título de
Licenciada em Ciências Biológicas pela
Universidade do Estado do Amazonas.

Orientador: Dr. Marcelo Salles Rocha

Coorientadora: Dra. Camila Cherem Ribas

Manaus- AM

2017

Barbosa, Waleska Eizangela dos Santos

B238f Filogenia do complexo *Synallaxis rutilans* Temminck, 1823 (Aves: Passeriformes: Furnariidae), utilizando representação reduzida do genoma por captura de sequência / Waleska Eizangela dos Santos Barbosa. 2017

29 f.: il. color; 31 cm.

Orientador: Marcelo Salles Rocha

Coorientadora: Camila Cherem Ribas

TCC de Graduação (Licenciatura em Ciências Biológicas) -
Universidade do Estado do Amazonas.

1. Filogenômica. 2. Uce. 3. Coalescência. 4. Furnariidae. I. Rocha, Marcelo Salles II. Universidade do Estado do Amazonas III. Título

À minha amada avó Cecília e ao meu avô José

AGRADECIMENTOS

Agradeço aos meus avós e aos meus pais pelo investimento, incentivo e apoio na minha educação e escolha acadêmica.

Aos meus orientadores e co-orientadores, Dra. Camila Ribas, Dr. Marcelo Rocha, Dr. Mateus Ferreira e Me. Érico Polo, por todo ensinamento, apoio e confiança para a realização deste estudo.

Aos amigos do EBBA por todas as conversas, discussões, ajuda, amizade e cafés: Romina Batista, Roberta Canton, Érik Choueri, Marina Maximiano, Leandro Moraes, Lídia Martins, Nelson Buainain, Mirna Amoedo, Glauco Kohler.

Às instituições que concederam as amostras de tecido para a realização deste trabalho: INPA, MPEG.

Ao pesquisador Michael Harvey pela concessão dos dados genômicos de algumas amostras.

Aos meus professores da UEA pela contribuição na minha formação acadêmica.

Aos meus colegas da graduação. Em especial: Power Rangers, Kaka, Mandinha e Gabi pela amizade, momentos descontraídos e desabafos.

Ao Kalebe Pinto por sempre me arrancar um sorriso do rosto mesmo quando estou num poço de depressão.

À minha irmã e melhor amiga Sabrina por todas as conversas, risadas, conselhos, confidências e desabafos.

Hakuna matata

-Timão e Pumba

RESUMO

A morfologia nem sempre representa toda a diversidade genética de um grupo, tal diversidade pode ser revelada por meio de estudos genéticos, que podem auxiliar na resolução de taxonomias ainda não resolvidas e revelar a história evolutiva dos organismos. *Synallaxis rutilans* é um complexo de espécies que pertence à família Furnariidae, uma das mais complexas e numerosas dentre os Passeriformes, e ainda não possui uma filogenia que represente suas unidades evolutivas. Este complexo foi alvo de uma revisão sistemática recente, utilizando dados morfológicos, e um estudo de sistemática molecular utilizando dois genes mitocondriais. No estudo morfológico, são reconhecidos três táxons: *S. r. rutilans*, *S. r. amazonica* e *S. r. omissa*, sendo que *S. r. omissa* se distingue por sua distribuição restrita à área de endemismo Belém e por possuir uma plumagem mais diferenciada quando comparada aos outros táxons do grupo, por esses fatores foi proposto que este táxon seja elevado à categoria de espécie. No estudo de sistemática molecular os resultados revelaram sete linhagens no complexo *S. rutilans*, tendo sua distribuição delimitada pelos principais afluentes da bacia Amazônica, e a linhagem de *omissa* encontra-se dentro da diversificação de *rutilans*, comparado aos dados morfológicos, as linhagens moleculares não formam clados monofiléticos. Neste estudo utilizando as regiões flangeadoras dos elementos ultra conservados do genoma, pelo método de captura de sequência, fizemos análises de estrutura populacional e de coalescência para gerar árvores de espécies e obter uma filogenia mais robusta. Obteve-se em média um valor de 2234 loci UCE por amostra, destes, ~1405 loci tem pelo menos 1 sítio informativo. A estrutura populacional revelou a existência de dois grandes grupos, um ao Norte, e outro ao Sul do rio Amazonas. A inferência Bayesiana recuperou uma filogenia com três clados principais, duas linhagens ao Norte do rio Amazonas, duas linhagens a Leste e três linhagens a Oeste do rio Xingu. As árvores de espécies geradas a partir de SNPs e de sequências, não são totalmente concordantes entre si, e com a árvore de inferência bayesiana. Revelando que as linhagens que ocorrem ao Sul do rio Amazonas possuem uma complexa história evolutiva que precisa ser investigada com maiores detalhes.

Palavras chave: Filogenômica, UCE, Coalescência, Furnariidae

ABSTRACT

Morphology does not always represent all the diversity of a group, such diversity can be revealed through genetic studies, which can help in solving unresolved taxonomies and reveal the evolutionary history of organisms. *Synallaxis rutilans* is a complex of species that belongs to the family Furnariidae, one of the most complex among the Passeriformes, and for which no phylogeny was proposed that represents its evolutionary units. This complex was the subject of a recent systematic review, using morphological data, and a systematic molecular study of two mitochondrial genes. In the morphological study, three taxa are recognized: *S. r. rutilans*, *S. r. amazonica* and *S. r. omissa*, being that *S. r. omissa* is distinguished by its distribution being restricted to the area of endemism Belém and its plumage is more differentiated from the others, by these factors it was proposed that this taxon be recognized as a species. The results of the molecular systematics study revealed seven lineages in the *S. rutilans* complex, with their distribution delimited by the main tributaries of the Amazon basin, and the *S. r. omissa* lineage embedded within the lineages of *S. r. rutilans*. In this study, we use the flanking regions of the ultra conserved elements of the genome by the sequence capture method, perform population structure and coalescence analyzes to generate species trees and obtain a more robust phylogeny. An average of 2234 UCE loci per sample was obtained, these 1405 loci have at least 1 informational site. A population structure revealed an existence of two large groups, one to the North and one to the South of the Amazon River. Bayesian inference recovered a phylogeny with three main clades, two Northern Amazonian lineages, two lineages to the east and three lineages to the west of the Xingu River. Revealing that the lineages that occur south of the Amazon River have a complex evolutionary history that needs to be investigated in greater detail.

Key words: Phylogeny, UCE, Coalescence, Furnariidae

LISTA DE FIGURAS

Figura 1. Áreas de endemismo identificadas para vertebrados terrestres na Amazônia.	12
Figura 2. Localidades de coleta dos espécimes de <i>S. rutilans</i> analisados por Stopiglia (2010). Círculos brancos com ponto no meio, representam <i>S. r. rutilans</i> , e círculos pretos, representam <i>S. r. amazonica</i> . Estrela, na área de endemismo Belém, corresponde a <i>S. r. omissa</i>	13
Figura 3. Cronograma, rede de haplótipos e distribuição das amostras no mapa do complexo <i>Synallaxis rutilans</i> . As cores representam cada linhagem, e estão relacionadas no mapa, árvore e rede de haplótipos.....	14
Figura 4. Mapa da distribuição geográfica dos indivíduos analisados neste trabalho.	16
Figura 5. BAPS utilizando SNPs, mostrando os agrupamentos populacionais, as cores representam os agrupamentos, e as barras representam cada indivíduo. Amarelo representa os indivíduos ao Norte do rio Amazonas, e vermelho os indivíduos do Sul.	19
Figura 6. Mapa de distribuição das linhagens do complexo e árvore bayesiana, a coloração dos clados está de acordo com a área de endemismo a qual ocupam descritas na legenda do mapa. A barra ao lado indica a qual táxon proposto por Stopiglia (2010) representa cada linhagem.....	20
Figura 7. Árvore de espécies utilizando SNPs.....	21
Figura 8. Árvore de espécies gerada no *BEAST 2.....	22
Figura 9. Comparação das árvores de espécies geradas a partir de SNPs e de sequências.	23

SUMÁRIO

1. INTRODUÇÃO.....	11
1.1 OBJETIVO GERAL	15
1.1.1 OBJETIVOS ESPECÍFICOS	15
2. MATERIAL E MÉTODOS.....	16
2.1 AMOSTRAGEM.....	16
PROCESSAMENTO DOS DADOS GENÔMICOS	17
2.4 ANÁLISES FILOGENÉTICAS	17
3. RESULTADOS.....	19
3.1 IDENTIFICAÇÃO DOS ELEMENTOS ULTRA CONSERVADOS	19
3.2 ESTRUTURA GENÉTICA POPULACIONAL.....	19
3.3 ANÁLISES FILOGENÉTICAS	19
4. DISCUSSÃO	22
5. CONSIDERAÇÕES FINAIS.....	25
6. REFERÊNCIAS	26

1. INTRODUÇÃO

A Amazônia é um dos biomas com a maior diversidade de aves do mundo. Estima-se que existem cerca de 1300 espécies, das quais 273 são endêmicas (MITTERMIER *et al.*, 2003). Tal biodiversidade tem instigado a curiosidade de pesquisadores a tentar entender os processos de sua origem e diversificação (e.g. RIBAS *et al.*, 2012; SMITH *et al.*, 2014; SCHULTZ *et al.*, 2016; FERREIRA *et al.*, 2017).

Diversas hipóteses foram propostas para explicar os padrões e processos que deram origem às espécies na Amazônia (SICK, 1967; HAFFER, 1969; ENDLER, 1982; AYRES, 1986; COLINVAUX, 1993), dentre as mais conhecidas estão a hipótese dos rios como barreiras geográficas (SICK, 1967) e a hipótese dos refúgios (HAFFER, 1969). A primeira propõe que a evolução da drenagem separou populações amplamente distribuídas em margens opostas dos rios e ao longo do tempo essas populações se tornaram espécies diferentes. A segunda está associada aos ciclos climáticos do Pleistoceno, durante o qual a floresta teria se retraído em períodos secos a fragmentos refúgios para as espécies, cercados por uma matriz de vegetação não florestal, enquanto em períodos úmidos a floresta se expandiria, reconectando os fragmentos e permitindo expansão e contato secundário entre populações diferenciadas pelo isolamento. Independente das hipóteses, o fato é que as espécies em geral não se distribuem uniformemente na região, e algumas possuem distribuição restrita a determinadas áreas conhecidas como áreas de endemismo (Figura 1) (CRACRAFT, 1985; SILVA *et al.*, 2005; BORGES & DA SILVA, 2012).

Figura 1. Áreas de endemismo identificadas para vertebrados terrestres na Amazônia.



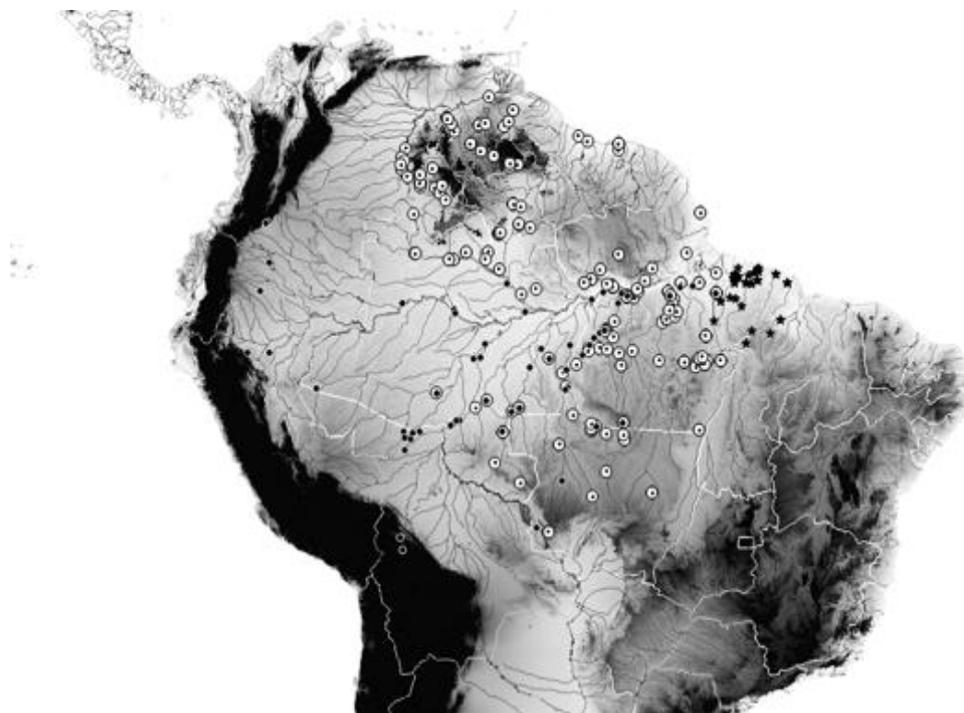
Fonte: Carvalho & Almeida. Biogeografia da América do Sul.

Apesar de as aves serem o grupo mais bem estudado e conhecido dentre os vertebrados, sua diversidade ainda é subestimada, tendo em vista que estudos apontam para a existência de uma diversidade críptica (D'HORTA *et al.*, 2013), pois, nem sempre a diversidade é representada pelas diferenças morfológicas, mas modificações genéticas são praticamente obrigatórias.

Entre os passeriformes mais complexos e diversos está a família Furnariidae, que tem despertado interesse para estudos filogenéticos e revisões taxonômicas (DERRYBERRY *et al.* 2011). A essa família pertence o complexo *Synallaxis rutilans* Temminck, 1823, que ocorre exclusivamente na bacia amazônica, e não possui uma resolução taxonômica satisfatória. Segundo a revisão de Remsen (2003), os táxons que compõem esse complexo são: *S. r. rutilans* Temminck, 1823, leste do rio Tapajós, ao rio Tocantins e noroeste do Mato Grosso; *S. r. omissa* Hartert, 1901, leste do rio Tocantins ao norte do Maranhão; *S. r. amazonica* Hellmayr, 1907, leste do Peru, sul do rio Amazonas até o leste do rio Tapajós; *S. r. tertia* Hellmayr, 1907, sudoeste do Mato Grosso e leste da Bolívia (leste de Santa Cruz); *S. r. caquetensis* Chapman, 1914, sudeste da Colômbia, leste do Equador e nordeste do Peru; *S. r. dissors* Zimmer, 1935, leste da Colômbia, sul da Venezuela, Guianas e norte do Brasil (leste do rio Negro); *S. r. confinis* Zimmer, 1935, oeste do rio Negro.

A revisão mais recente (STOPIGLIA, 2010), analisando caracteres morfológicos, reconhece três táxons: *S. r. rutilans*, distribuído no Escudo das Guianas e Escudo Brasileiro, *S. r. amazonica*, Oeste do rio Negro e Sul do rio Amazonas e *S. omissa* Leste do rio Tocantins, na área de endemismo Belém (Figura 2), e propõe a validação de *S. omissa* como espécie, devido à sua diferenciação na plumagem e distribuição geográfica, sendo o único representante do complexo *S. rutilans* a ocorrer na área de endemismo Belém.

Figura 2. Localidades de coleta dos espécimes de *S. rutilans* analisados por Stopiglia (2010). Círculos brancos com ponto no meio, representam *S. r. rutilans*, e círculos pretos, representam *S. r. amazonica*. Estrela, na área de endemismo Belém, corresponde a *S. r. omissa*.

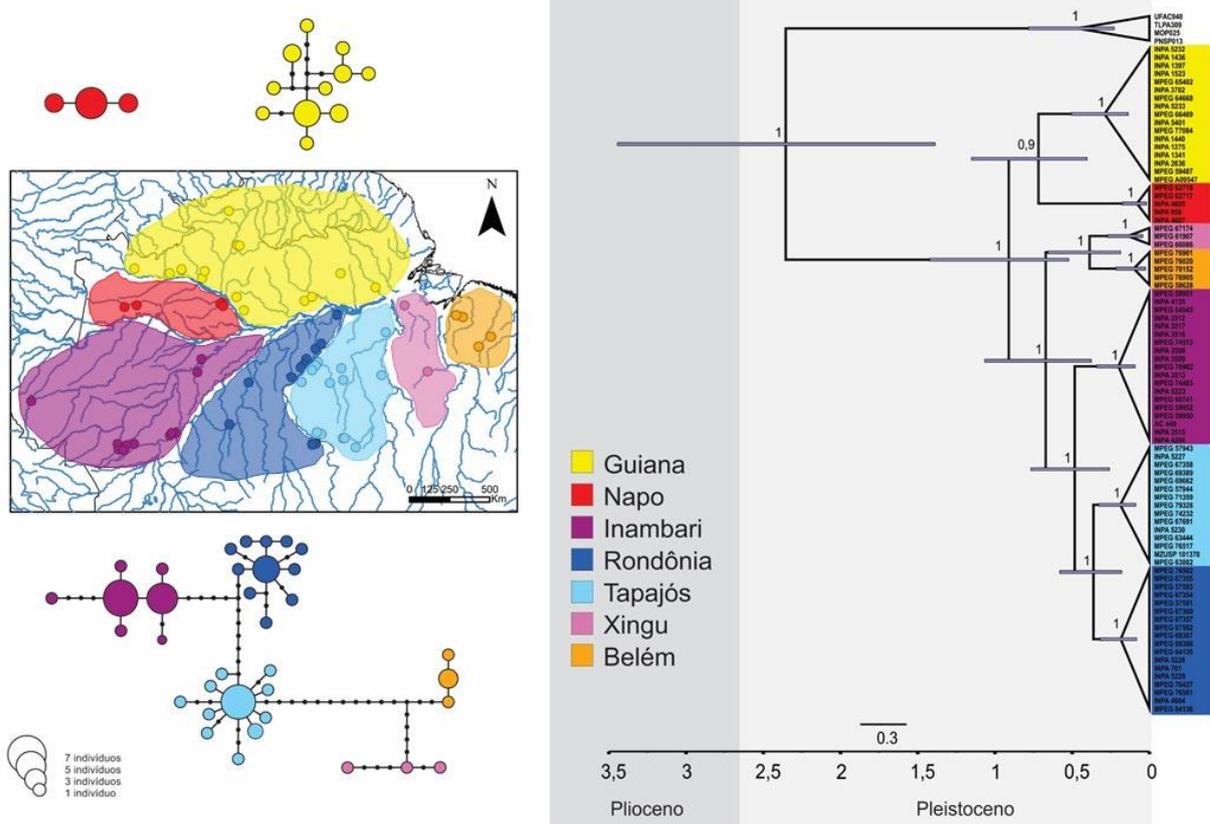


Fonte: Stopiglia, 2010.

Análises filogenéticas utilizando dois marcadores mitocondriais, COI- Cytocromo oxidase subunidade I, e ND2- Desidrogenase subunidade II (BARBOSA *et al.*, 2017) revelaram que este complexo é composto por sete linhagens tendo a distribuição delimitada pelos principais afluentes do rio Amazonas, sendo que cada linhagem ocupa uma área de endemismo reconhecida para aves e primatas na Amazônia (Figura 3), portanto, não corroboram os táxons propostos por Stopiglia (2010). Segundo Barbosa e colaboradores (2017) as sete linhagens para o complexo *S. rutilans* estão distribuídas assim: (a) linhagem 1, encontrada ao Norte do Amazonas e Leste do Rio Negro na área de endemismo Guiana (b) linhagem 2, oeste do rio

Negro, na área de endemismo Napo (c) linhagem 3, oeste do rio Madeira, na área de endemismo Inambari (d) linhagem 4, interflúvio Madeira-Tapajós, na área de endemismo Rondônia (e) linhagem 5, interflúvio Tapajós-Xingu, na área de endemismo Tapajós (f) linhagem 6, interflúvio Xingu-Tocantins, na área de endemismo Xingu e (g) linhagem 7, leste do rio Tocantins, na área de endemismo Belém.

Figura 3. Cronograma, rede de haplótipos e distribuição das amostras no mapa do complexo *Synallaxis rutilans*. As cores representam cada linhagem, e estão relacionadas no mapa, árvore e rede de haplótipos.



Atualmente, um campo de estudo que vem avançando e tem como proposta trazer maiores resoluções da árvore da vida bem como desvendar os mecanismos moleculares da evolução é a filogenômica (DELSUC *et al.*, 2005), que surgiu com os avanços das técnicas de sequenciamento de nova geração (NGS), que produz milhares de sequências de DNA em apenas uma corrida (BEHJATI & TARPEY, 2013). Diferentes métodos podem ser usados para se extrair a informação desejada do genoma (MARDIS, 2008), uma delas é a captura de sequência, utilizada para isolar partes específicas do genoma, tendo assim um agrupamento enriquecido de

sequências alvo, reduzindo o espaço de sequenciamento genômico, obtendo uma maior cobertura de sequência para cada região segmentada (GROVER *et al.*, 2012)

Uma classe de marcadores que pode ser acessada pelo método de captura de sequência são os Elementos Ultra Conservados do genoma (UCEs). Esses marcadores são curtos trechos de DNA (~500pb), que são compartilhados por diferentes táxons vertebrados, mesmo relativamente distantes entre si (BEJERANO *et al.*, 2004; CRAWFORD *et al.*, 2012). As regiões flangeadoras são bastante polimórficas, sendo utilizadas para reconstruir a história evolutiva em diferentes níveis taxonômicos (FAIRCLOTH *et al.*, 2012). Para isso, uma abordagem bastante utilizada para análises de árvore de espécies é a coalescência, um método matemático para estudar a história ancestral de genes ou sequência de DNA em uma população, rastreando os genes ou sequências entre os indivíduos de certa população até encontrar o ancestral comum mais recente (LIU *et al.*, 2009).

A proposta deste trabalho é utilizar as regiões flangeadoras dos elementos ultra conservados para realizar análises filogenéticas e comparar os resultados obtidos com os dados mitocondriais e morfológicos, para uma melhor compreensão dos processos evolutivos do complexo *Synallaxis rutilans*.

1.1 OBJETIVO GERAL

Investigar as relações filogenéticas entre as espécies do complexo *Synallaxis rutilans*, utilizando sondas para elementos ultra conservados do genoma, a fim de examinar se os dados genômicos corroboram os dados mitocondriais ou a variação fenotípica.

1.1.1 OBJETIVOS ESPECÍFICOS

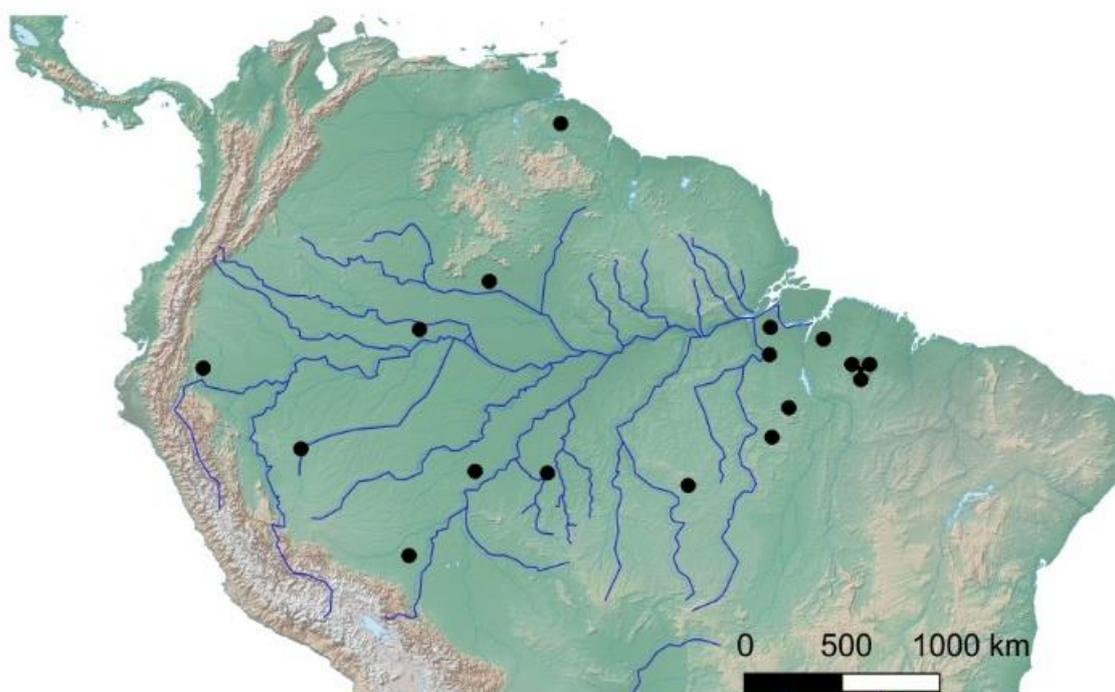
- Verificar a estrutura populacional do complexo *Synallaxis rutilans*;
- Recuperar a história filogenética de *Synallaxis rutilans*;
- Desvendar os processos de evolução e diversificação que envolve o grupo, a partir dos dados genômicos comparando-os aos dados mitocondriais e fenotípicos.

2. MATERIAL E MÉTODOS

2.1 AMOSTRAGEM

Foram amostrados 17 indivíduos do complexo *Synallaxis rutilans*, distribuídos ao longo da região amazônica (Figura 4). As amostras foram obtidas a partir de tecidos depositados na coleção do Museu Paraense Emílio Goeldi (MPEG). As informações sobre as amostras (código da amostra, UCEs, coordenadas e área de endemismo) utilizadas neste trabalho encontram-se no Apêndice.

Figura 4. Mapa da distribuição geográfica dos indivíduos analisados neste trabalho.



2.2 EXTRAÇÃO E PREPARO DA BIBLIOTECA

O DNA genômico foi extraído do músculo peitoral conservado em álcool 100%, pelo kit de extração DNeasy (Qiagen). As amostras foram enviadas para o Rapid Genomics (Gainesville, Flórida), para sequenciamento de captura seguindo o protocolo descrito em Faircloth (2012) e Smith (2013). O sequenciamento de captura envolve a preparação de bibliotecas de DNA, onde ocorre fragmentação aleatória do DNA genômico, seguida da hibridização dos fragmentos às sondas de oligonucleotídeo sintético biotilinado com as sequências complementares dos elementos Ultra Conservados. Os *iprobos* contém *beads* magnéticas, que são utilizadas para capturar as regiões de interesse, separando-as das regiões não-alvo,

por fim os fragmentos são clonados através de alguns ciclos de PCR e a biblioteca é enviada para sequenciamento de NGS.

PROCESSAMENTO DOS DADOS GENÔMICOS

Os dados genômicos foram processados de duas formas distintas. Primeiramente, seguindo-se o protocolo padrão do pacote PHYLUCE (FAIRCLOTH, 2015). Este método obtém as sequências finais dos marcadores, utilizando a montagem *de novo* de contigs a partir dos *reads* (sequências saídas do sequenciamento) mais frequentes, gerando, assim, uma matriz haploide (uma sequência por indivíduo). Essa matriz foi usada para análise filogenética com sequências concatenadas (ver abaixo).

Posteriormente, os dados foram processados por uma metodologia alternativa, um *pipeline* (https://github.com/mgharvey/seqcap_pop) desenvolvido por Harvey (2016), seqcap pop, é uma série de comandos e scripts adicionais para processar dados de sequência de captura usando ferramentas dentro do pacote PHYLUCE (FAIRCLOTH, 2015), além de algumas ferramentas adicionais utilizadas ou desenvolvidas por este autor. O seqcap pop é capaz de gerar SNPs e sequências diploides a partir do remapeamento dos *reads*, nos contigs da montagem *de novo*, permitindo um maior número de dados para análises e a utilização de métodos baseados em SNPs.

2.3 ESTRUTURA POPULACIONAL

O programa BAPS: *Bayesian Analysis of Population Structure* v.6 (CORANDER *et al.*, 2013) foi utilizado para verificar a estruturação populacional. Selecionou-se a análise *Clustering of individuals*, a qual faz o agrupamento de indivíduos com dados de marcadores não ligados. A matriz de input das sequências foi feita a partir de SNPs extraídos de cada loco de UCE a partir do seqcap pop. Foram explorados diferentes números máximos K de populações até que o número mais verossímil fosse estimado, de acordo com as recomendações do manual do programa.

2.4 ANÁLISES FILOGENÉTICAS

Foram realizadas 3 análises filogenéticas de inferência bayesiana, uma realizando a matriz concatenada e duas árvores de espécies, sendo uma delas

baseada em SNPs e a outra em um subconjunto de sequências. Para o primeiro e último casos a escolha de partições e dos melhores modelos evolutivos das sequências foi feita no programa PartitionFinder 2 (LANFEAR *et al.*, 2016) utilizando o algoritmo de busca *greedy* (LANFEAR *et al.*, 2012) para o subconjunto de sequências (LANFEAR *et al.*, 2014). Em ambos os casos levaram-se em consideração todos os modelos disponíveis no programa BEAST.

A análise com a matriz completa concatenada foi realizada utilizando o programa MrBayes 3.2.1 (RONQUIST *et al.*, 2012) com os dados particionados sob os modelos evolutivos sugeridos pelo PartitionFinder 2 com o método *rclusterf* de busca. Foi realizada 1 corrida com 8 cadeias simultâneas de 2 milhões de gerações, e árvores amostradas a cada 1000 gerações. Para garantir estacionaridade da corrida foram checados os valores de desvio padrão de frequência dos clados, assim como os valores de ESS, aceitos de acordo com as recomendações do manual do programa. Como não foi incluído grupo externo na análise, a árvore foi enraizada no *midpoint* (ponto médio entre os dois *tips* mais distantes da árvore).

Foram extraídos SNPs dos UCEs que apresentaram pelo menos 1 sítio parcimoniosamente informativo (PIS), totalizando ~1405 loci. A Árvore de espécies foi primeiramente estimada pelo pacote SNAPP (BRYANT *et al.*, 2012), no programa BEAST 2 (BOUCKAERT *et al.*, 2014) que utiliza SNPs para uma análise de coalescência. A definição das linhagens foi feita com base na árvore gerada pelo MrBayes, e nas áreas de endemismo. Foi realizada uma corrida de 212000 gerações, amostradas a cada 1000. A estabilidade da corrida e a determinação do *burn-in* foi verificada no programa Tracer 1.6 (RAMBAUT *et al.* 2014).

Posteriormente realizou-se a análise de árvore de espécies utilizando o pacote *BEAST 2 (OGILVIE *et al.*, 2017) do programa BEAST 2. Selecionaram-se os 123 loci UCEs com maior número de PIS (o que resultou em um mínimo de 9 PIS por loco). A definição das linhagens também foi feita com base no resultado do MrBayes, e os modelos evolutivos determinados e ligados de acordo com o resultado do PartitionFinder 2. Realizou-se uma corrida com 137.300.000 gerações amostradas a cada 10.000.

3. RESULTADOS

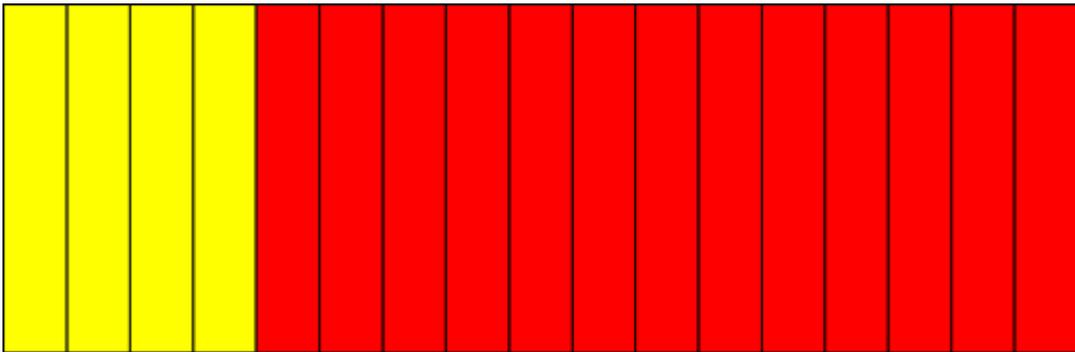
3.1 IDENTIFICAÇÃO DOS ELEMENTOS ULTRA CONSERVADOS

Obteve-se um total de 2234 loci UCE na matriz incompleta. As sequências recuperadas têm em média 572 pb, e ~1405 loci possui pelo menos 1 PIS. O valor de missing data para SNPs (1.405 loci) foi de 10,15%.

3.2 ESTRUTURA GENÉTICA POPULACIONAL

Os resultados obtidos pelo BAPS (Figura 5) mostraram duas populações separando os indivíduos do Norte e Sul do rio Amazonas, considerando todos os indivíduos.

Figura 5. BAPS utilizando SNPs, mostrando os agrupamentos populacionais, as cores representam os agrupamentos, e as barras representam cada indivíduo. Amarelo representa os indivíduos ao Norte do rio Amazonas, e vermelho os indivíduos do Sul.



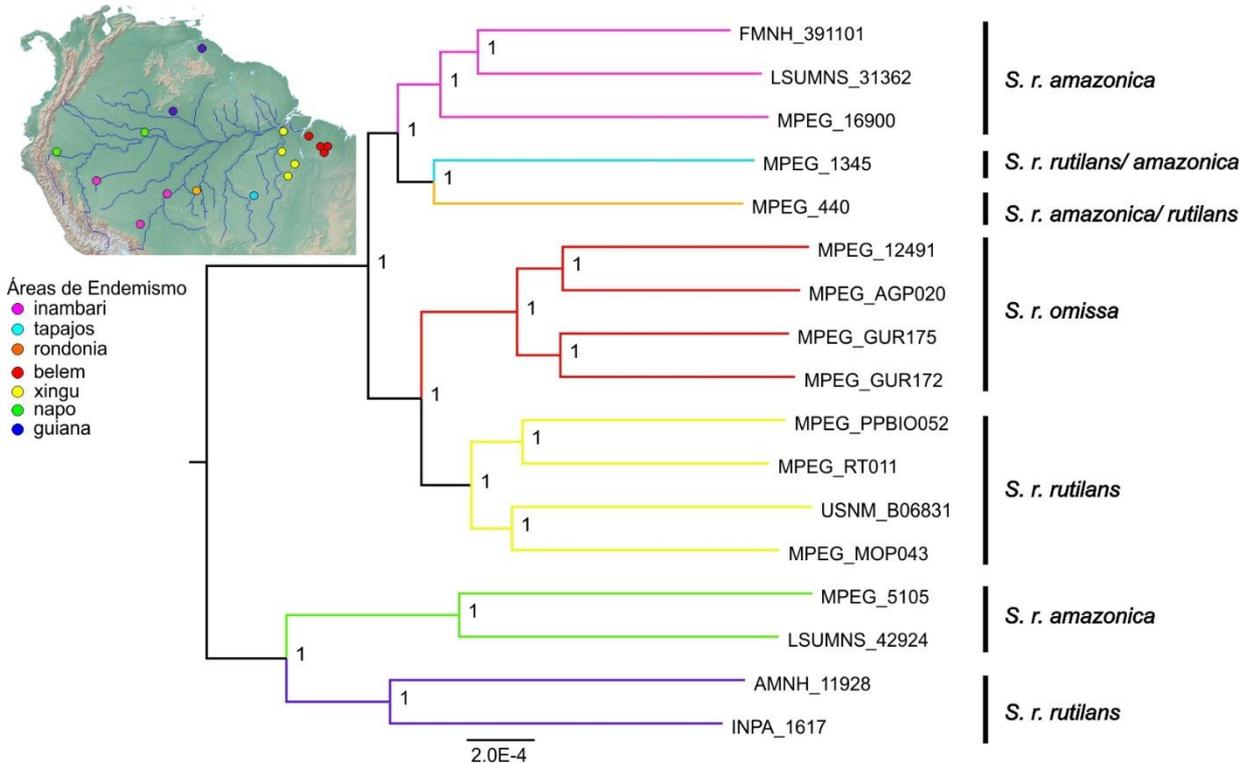
3.3 ANÁLISES FILOGENÉTICAS

Para a matriz completa o programa PartitionFinder 2 recuperou 33 partições, e selecionou GTR+I+G como o melhor modelo evolutivo para todas as partições. A topologia obtida por Inferência Bayesiana pelo MrBayes (sequências concatenadas) apresentou alto suporte (probabilidade posterior= 1) em todos os nós (Figura 6).

De acordo com a topologia, sete linhagens foram recuperadas, cada uma ocorrendo em uma diferente área de endemismo reconhecida para aves e primatas na Amazônia, listadas na legenda do mapa da figura 6. Os clados que ocorrem no Norte do rio Amazonas são (Guiana e Napo) e ao Sul (Inambari, Rondônia, Tapajós, Xingu e Belém). No Norte Guiana é grupo irmão de Napo. No Sul há dois clados

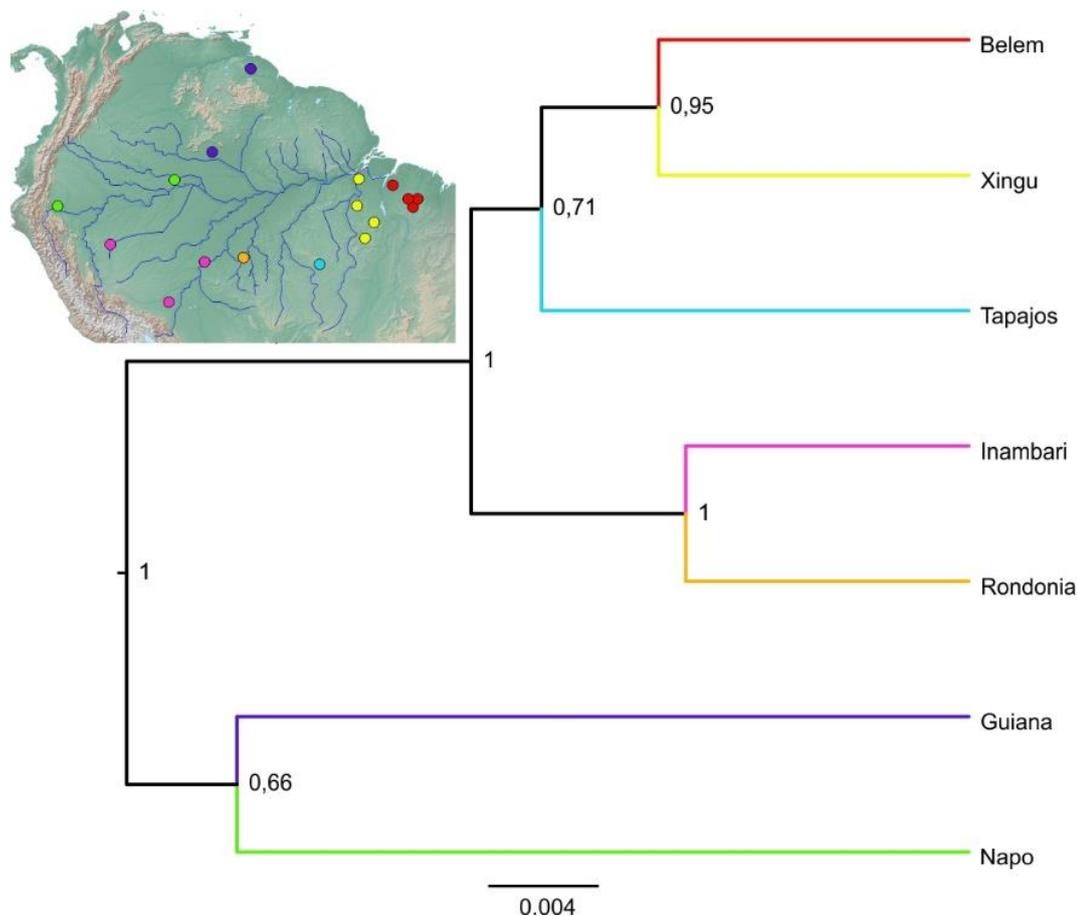
principais a Leste (Xingu, Belém) e a Oeste (Inambari, Rondônia, Tapajós) do rio Xingu.

Figura 6. Mapa de distribuição das linhagens do complexo e árvore bayesiana, a coloração dos clados está de acordo com a área de endemismo a qual ocupam descritas na legenda do mapa. A barra ao lado indica a qual táxon proposto por Stopiglia (2010) representa cada linhagem.



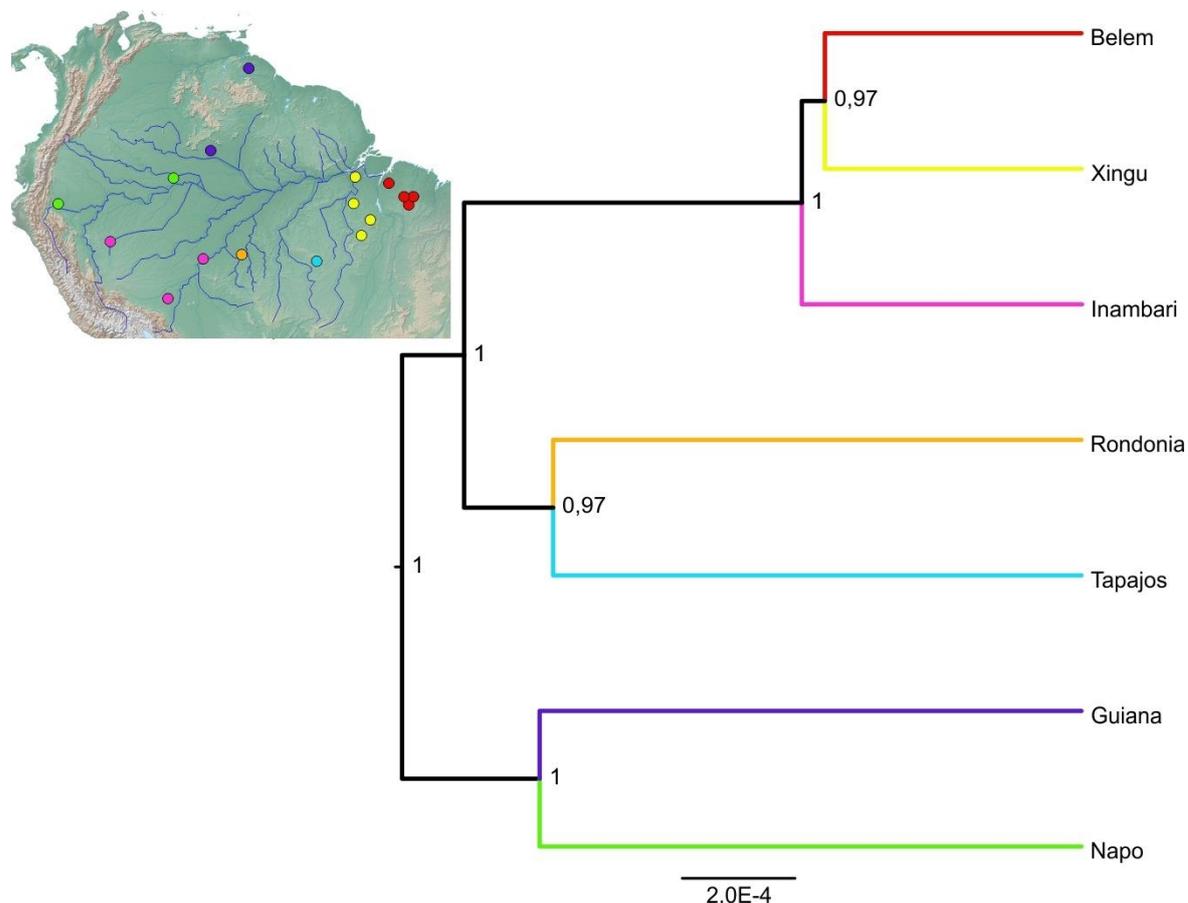
A topologia recuperada pelo SNAPP (Figura 7) não obteve alto suporte em todos os nós. De acordo com a reconstrução desta árvore de espécies, o primeiro evento de diversificação se dá entre as linhagens do Norte e Sul do rio Amazonas, no Norte há separação entre Guiana e Napo, sendo que esta divergência não é bem suportada. No Sul, ocorreram quatro eventos de diversificação. O primeiro separando as linhagens das áreas de endemismo (Belém, Xingu, Tapajós) de (Inambari, Rondônia). O segundo separa (Belém, Xingu) de Tapajós, onde não há uma boa resolução ($pp= 0,71$), podendo ser considerado uma politomia. O terceiro separa Belém e Xingu. O quarto evento de diversificação ocorre entre Inambari e Rondônia.

Figura 7. Árvore de espécies utilizando SNPs.



Para a matriz com 123 loci selecionados o programa PartitionFinder 2 sugeriu dez partições, cujos modelos selecionados foram: K80+I+G, HKY+I+G+X, TRN+I+G+X e JC. Na árvore de espécies gerada no pacote *BEAST2 obtiveram-se suportes com probabilidade posterior > 90 para todos os nós (Figura 8). Esta topologia é discordante da árvore recuperada pelo SNAPP. Apesar de nesta topologia, assim como nas demais, há divergência entre Norte e Sul do rio Amazonas e as linhagens do Norte separarem-se em margens opostas do rio Negro, nas linhagens do Sul observa-se a mudança na posição de alguns ramos, Inambari agrupa-se a Belém e Xingu, e Rondônia é recuperada como linhagem irmã de Tapajós.

Figura 8. Árvore de espécies gerada no *BEAST 2.



4. DISCUSSÃO

Nossos resultados mostraram que os dados genômicos corroboram em parte os dados mitocondriais. Em ambos os casos se encontra estruturação genética populacional ao Norte e ao Sul do rio Amazonas, embora os dados mitocondriais revelem maior estruturação populacional, recuperando uma população por área de endemismo, exceto Xingu e Belém, agrupados como uma única população, possivelmente por se tratar de uma diversificação recente, portanto as linhagens não devem ter acumulado diferenças significativas para esta análise.

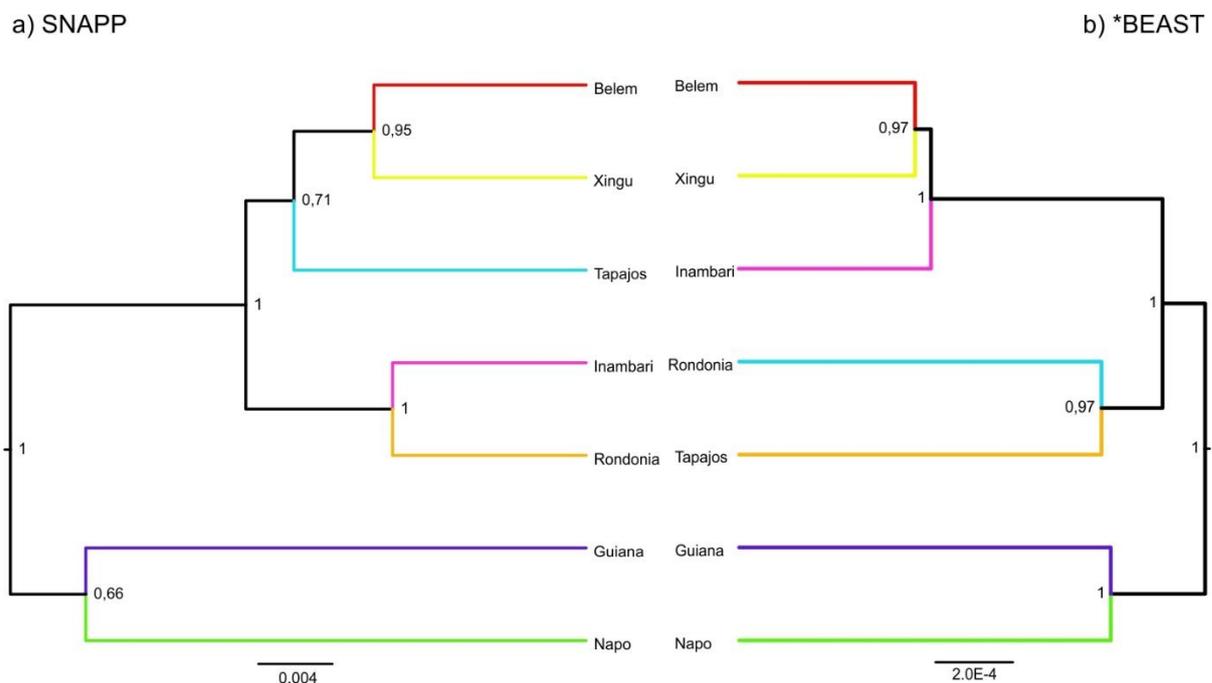
Quanto às análises filogenéticas, a primeira análise utilizando uma matriz concatenada, corrobora totalmente a topologia recuperada pelos dados mitocondriais, no entanto, as árvores filogenéticas dos programas SNAPP e *BEAST2 apresentam algumas modificações nas relações entre alguns ramos. Quando comparados aos dados genéticos, os dados morfológicos não são concordantes, os táxons *S. r. rutilans*

e *S. r. amazonica* não formam clados monofiléticos, visto que eles têm representantes tanto ao Norte quanto ao Sul do rio Amazonas.

A análise de estrutura populacional, utilizando os SNPs de UCEs considerando todos os indivíduos, recupera apenas duas populações separadas pelo rio Amazonas, enquanto que os dados mitocondriais apresentados por Barbosa e colaboradores (2017) recuperam seis populações, uma para cada área de endemismo, exceto Belém e Xingu, sendo recuperadas como uma única população. No entanto, todas as topologias recuperam sete linhagens. Pelo fato de o programa BAPS analisar apenas as diferenças entre as sequências, sem levar em consideração os modelos evolutivos, as linhagens do complexo *S. rutilans* podem ter tido uma rápida divergência, não sendo possível que o programa as reconhecesse.

Comparando as topologias recuperadas pelas análises filogenéticas se observam mudanças na posição de alguns ramos nas árvores de espécies de SNPs e de sequências (Figura 9).

Figura 9. Comparação das árvores de espécies geradas a partir de SNPs e de sequências.



Diferente da topologia dos dados mitocondriais (Figura 3) e da matriz concatenada (Figura 6), a árvore de espécies gerada no programa SNAPP agrupa a linhagem Tapajós com Xingu e Belém, não com Rondônia e Inambari. No entanto, a árvore não apresenta suporte estatístico para esse nó. Pode-se atribuir essa falta de suporte ao curto tempo de corrida da análise, necessitando assim um tempo maior de

corrida até que atinja uma maior estabilidade, e convergência entre as árvores. A árvore de espécies gerada pelo *BEAST 2 também discorda na posição de alguns ramos das linhagens do Sul. As únicas linhagens do Sul que apresentam bom suporte estatístico em todas as topologias são Xingu e Belém, sempre agrupadas como linhagens irmãs.

A amostragem das áreas de endemismo Rondônia e Tapajós pode estar afetando a resolução das linhagens Inambari, Rondônia e Tapajós, pois cada uma (Rondônia e Tapajós) tem apenas uma amostra. Segundo Helled & Drummond (2009) uma única amostra não significa eventos coalescentes para a espécie existente, portanto nenhuma informação para estimar o tamanho da população, e isso é importante porque a teoria coalescente relaciona o tamanho efetivo da população com a história ancestral de uma amostra de genes de uma população. Ainda segundo os autores, essa deficiência nos dados pode vir a prejudicar a topologia.

Outros fatores que dificultam análises filogenéticas são: a separação incompleta de linhagens, fluxo gênico, transferência horizontal de genes (MADDISON, 1997), entre outros. Esses problemas são encontrados com maior frequência quando se tratam de espécies proximamente relacionadas (CARSTERN & KNOWLES, 2007). No caso do complexo *S. rutilans*, as árvores de espécies apresentam curtos comprimentos entre os nós e longos ramos, o que caracteriza o fator de separação incompleta de linhagens.

Observa-se que é na região do Escudo Brasileiro que os fenótipos *S. r. rutilans* e *S. r. amazonica* apresentados por Stopiglia (2010) apresentam maior sobreposição (Figura 1), as linhagens que correspondem a essa área são Rondônia e Tapajós. As mesmas que não apresentam boa resolução nas árvores de espécies, isso poderia ser um indicador de que pode estar havendo fluxo gênico, mas isso só pode ser verificado a partir de outras análises, pois, o *BEAST não inclui modelagem para estimar recombinação dentro de um locus, nem estimativas de fluxo gênico (HELLED & DRUMMOND, 2009).

Apesar de algumas incongruências entre as topologias, observa-se que a diversificação deste grupo se deu entre as áreas de endemismo, visto que em todas as análises filogenéticas recupera-se linhagens estruturadas geograficamente nas áreas de endemismo, este padrão é observado para aves que ocupam a floresta de terra-firme (RIBAS *et al.*, 2012). No entanto, é importante ressaltar que para as AE

Rondônia e Tapajós seria recomendável a utilização de mais amostras para verificar a monofilia dessas áreas.

Apesar de *S. r. omissa* ser uma linhagem independente, quando se compara os táxons definidos fenotipicamente por Stopiglia (2010) *S. r. amazonica*, *S. r. rutilans* e *S. r. omissa* não formam clados monofiléticos. O táxon *S. r. rutilans* corresponde às linhagens Guiana, Rondônia/Tapajós e Xingu, *S. r. amazonica* corresponde às linhagens Napo, Inambari e Rondônia/Tapajós, corresponde às linhagens Guiana, Rondônia/Tapajós e Xingu e *S. r. omissa* corresponde à linhagem de Belém. Com isso podemos confirmar que nem sempre a diversidade fenotípica representa toda a diversidade genética, senso assim, confirmamos que existe uma diversidade críptica ainda a ser descrita.

5. CONSIDERAÇÕES FINAIS

Os resultados obtidos nesse trabalho mostraram que diversos fatores podem estar influenciando as análises de reconstrução filogenética, como a amostragem e fatores de especiação, como por exemplo, fluxo gênico. Para um melhor entendimento dos processos evolutivos deste grupo realizaremos análises de estimativa de fluxo gênico, e investigaremos outros métodos que possam auxiliar na resolução de uma topologia que represente a história evolutiva do grupo.

6. REFERÊNCIAS

- AYRES, J.M.C. Uakaris and Amazonian ooded forest. TESE. *University of Cambridge*. 1986.
- BARBOSA, W.E.S., FERREIRA, M., RIBAS, C.C. Sistemática molecular e biogeografia do complexo *Synallaxis rutilans* Temminck, 1823 (AVES: PASSERIFORMES: FURNARIIDAE). *Programa de Iniciação Científica do INPA*. 2017.
- BEHJATI, S., TARPEY, P.S. What is next generation sequencig?. *Archieves of Childhood-education and practise*, v. 98, n. 6, p. 236-238, 2013.
- BEJERANO, G., PHEASANT, M., MAKUNIN, I., STEPHEN, S., KENT, W.J., MATTICK, J.S., & HAUSSLER, D. Ultraconserved elements in human genome. *Science*, v. 304, n. 5675, p. 1321-1325. 2004.
- BORGES, S.H., DA SILVA, J.M.C. A new area of endemism for Amazonian birds in the Rio Negro basin. *The Wilson Journal of Ornithology* 124, 15-23. 2012.
- BOUCKAERT, R., HELED, J., KÜHNERT, D., VAUGHAN, T., WU C-H., XIE, D., SUCHARD, M.A., RAMBAUT, A., DRUMMOND, A.J. BEAST 2: A software platform for Bayesian Evolutionary Analysis. *PLoS Comput Biology* 10(4) 2014.
- BRYANT, DAVID, BOUCKAERT, R., FELSENSTEIN, J., ROSEMBERG, N.A., RoyCHOURDHURY, A. Inferring species tree directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular biology and evolution*. 29 (8) 1917-1932. 2012.
- CARSTERNS, B.C., KNOWLES, L.L. Estimate species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* Grasshoppers. *Systematic biology*, v. 56, n. 3, p. 400-411, 2007.
- COLINVAUX, P. Pleistocene biogeography and diversity in tropical forest of South America. In: GOLDBLATT, P. *Biological relationships between Africa and South America*. New Haven: *Yale University*. p. 473-499. 1993.
- CORANDER, J., CHENG, L., MARTTINEN, P., SIRÉN, J., TANG, J. BAPS: Bayesian Analysis of Population Structure. Manual v.6.0. 2013.
- CRACRAFT, J. Historical biogeography and patterns of differentiation within the South american avifauna: Areas of endemismo. *Ornithological Monographs* 36, 49-84. 1985.
- CRAWFORD, N.D., FAIRCLOTH, B.C., McCORMACK, J.E., BRUMFIELD, R.T., WINKER, K., GLENN, T.C. More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. *Biology letters*. v.8, n.5, p. 783-786. 2012.
- DELSUC, F., BRINKMANN, H., PHILIPPE, H. Phylogenomics and the reconstruction of the tree of life. *Nature*. Vol 6, 2005.
- DERRYBERRY, E.P., CLARAMUNT, S., DERRYBERRY, G., CHESSER, R.T., CRACRAFT, J., ALEIXO, A., PÉREZ-EMÁN, J., REMSEN, J.V., BRUMFIELD, R.T.

Lineage diversification and morphological evolution in a large-scale continental radiation: the neotropical overbirds and woodcreepers (Aves: Furnariidae). *Evolution*, v. 65, n. 10, p. 2973-2986, 2011.

D'HORTA, F.M., CUERVO, A.M., RIBAS, C.C., BRUMFIELD, R.T., MIYAKI, C.Y. Phylogeny and comparative phylogeography of *Sclerurus* (Aves: Furnariidae) reveal constant and cryptic diversification in an old radiation of rain forest understory specialist. *Journal of Biogeography* 40, 37-49. 2013.

ENDLER, J. Pleistocene forest refulges: fact or fancy? In: PRANCE, G.T. *Biological Diversification in the Tropics*. New York: Columbia University Press, p.179-200. 1982.

FAIRCLOTH, B. C., MCCORMACK, J. E., CRAWFORD, N. G., HARVEY, M. G., BRUMFIELD, R. T., & GLENN, T. C. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic biology*, sys004. 2012.

FAIRCLOTH, B. C. *PHYLUCES* is a software package for the analysis of conserved genomic loci. *Bioinformatics*. 32(5), 786-788. 2015.

FERREIRA, M., ALEIXO, A., RIBAS, C.C., SANTOS, M.P.D. Biogeography of the Neotropical genus *Malacoptila* (Aves: Buconidae): the influence of the Andean orogeny, Amazonia drainage evolution and Paleoclimate. *Journal of Biogeography*. 44, 748-759. 2017.

GROVER, C.E., SALMON, A., WENDEL, J.F. Targeted sequence capture as a powerful tool for evolutionary analysis. *American journal of botany*, v.99, n.2, p. 312-219. 2012.

HAFFER, J. Speciation in Amazonian forest birds. *Science* 165, 131-137. 1969.

HALLER, J., DRUMMOND, A.J. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, v. 27, n. 3, p. 570-580, 2009.

HARVEY, M.G., SMITH, B.T., GLENN, T.C., FAIRCLOTH, B.C., BRUMFIELD, R.T. Sequence Capture versus restriction site associated DNA sequencing for shallow systematics. *Systematic Biology* 65: 910-924. 2016.

LANFEAR, R., CALCOTT, B., HO, S.Y., GUINDON, S. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular and evolution*. 29(6), 1695-1701. 2012.

LANFEAR, R., CALCOTT, B., KAINER, D., MAYER, C., & STAMATAKIS, A. Selecting optimal partitioning schemes for phylogenomic datasets. *BMC evolutionary biology*, v. 14, n. 1, pag. 82. 2014.

LANFEAR, R., FRANDBSEN, P.B., WRIGHT, A.M., SENFELD, T., CALCOTT, B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*. DOI: dx.doi.org/10.1093/molbev/msw260. 2016.

LIU, L., YU, L., PEARL, D.K., EDWARDS, S.V. Estimating species phylogenies using coalescence times among sequences. *Systematic biology*, v. 58, n. 5, p. 468-477, 2009.

MADDISON, W.P. Gene trees and species trees. *Systematic biology*, vol. 46, n. 3, p. 523-536. 1997.

MARDIS, E.R. Next-Generation DNA Sequencing methods. *Annu. Ver. Genomics Hum. Genet.*, v.9, p. 387-402. 2008.

McCOMARCK, J.E., FAIRCLOTH, B.C., CRAWFORD, N.G., GOWATY, P.A., BRUMFIELD, R.T., GLENN, T.C. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome research* 22:746-754. 2012.

MITTERMEIER, R.A., C.G. MITTERMEIER, T.M. BROOKS, J.D. PILGRIM, W.R. KONSTANT, G.A.B. FONSECA & C. KORMOS. Wilderness and biodiversity conservation. *Proceedings of the National Academy of Science* 100: 10309-10313. 2003.

OGILVE H.A., BOUCKAERT, R.R., DRUMMOND, A.J. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Molecular Biology and Evolution*, v. 34, n. 8, p. 2101-2114. 2017.

PHILIPPE, H. & BLANCHETTE, M. Overview of the first phylogenomics conference. *BMC Evolutionary Biology* 7. 2007.

POLLARD, D.A., LYER, V.N., MOSES, A.M., EISEN, M.B. Widespread discordance of gene trees with species tree in *Drosophila*: Evidence of Incomplete Lineage Sorting. *PLoS Genetics* 2(10) 2006.

RAMBAUT, A., SUCHARD, M.A., XIE, D., DRUMMOND, A.J. Tracer v1.6. Disponível em: <http://beast.bio.ed.ac.uk/Tracer>. 2014.

REMSEN, J.V. Family Furnariidae (ovenbirds). In del Hoyo, J., A. Elliot e D. A. Christie (editors), *Handbook of the world*, 8: 162- 357. 2003.

RIBAS, C.C., ALEIXO, A., NOGUEIRA, A.C.R., MIYAKI, C.Y., CRACRAFT, J. A paleobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proceedings of Royal Society B*. 279, 681-689. 2012.

RONQUIST, F., HUELSENBECK, J., TESLENKO, M. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61, 1-4. 2012.

SCHULTZ, E.D., BURNEY, C.W., BRUMFIELD, R.T., POLO, E.M., CRACRAFT, J., RIBAS, C.C. Systematics and Biogeography of the *Automolus infuscatus* complex (Aves: Furnariidae): cryptic diversity reveals western Amazonia as the origin of a transcontinental radiation. *Molecular Phylogenetics and Evolution* 107, 503-515. 2016.

SICK, H. Rios e enchentes na Amazônia como obstáculo para a avifauna. Simpósio sobre a Biota Amazônica. *Atas de Zoologia* 5:495-520. 1967.

SILVA, J.M.C., RYLANDS, A.B. & FONSECA, G.A.B. The fate of the Amazonian Areas of endemism. *Conservation biology* 19, 689-694. 2005.

SMITH, B.T., HARVEY, M.G., FAIRCLOTH, B.C., GLENN, T.C., & BRUMFIELD, R.T. Target capture and massively parallel sequencing of ultraconserved elements (UCEs) for comparative studies at shallow evolutionary time scales. *Systematic biology*, syt061. 2013.

SMITH, B.T., McCORMACK, J.E., CUERVO, A.M., HICKERSON, M.J., ALEIXO, A., CADENA, C.D., PÉREZ-EMÁN, J., BURNEY, C.W., XIE, X., HARVEY, M.G., FAIRCLOTH, B.C., GLENN, T.C., DERRYBERRY, E.P., PREJEAN, J., FIELDS, S., BRUMFIELD, R.T. The drivers of tropical speciation. *Nature*, v. 515, n. 7527, p. 406, 2014.

STOPIGLIA, R. Revisão taxonômica do gênero *Synallaxis* Vieillot, 1818 (Aves: Passeriformes: Furnariidae). Tese de doutorado, Museu Nacional/UFRJ, Rio de Janeiro. 2010.

APÊNDICE

<i>Amostra</i>	<i>UCEs</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Área de endemismo</i>
<i>MPEG_12491</i>	2181	-4,2767	-47,453	Belém
<i>GUR175</i>	2261	-3,704	-46,762	Belém
<i>GUR172</i>	2223	-3,704	-46,762	Belém
<i>AGP020</i>	2248	-2,489	-48,709	Belém
<i>AMNH_11928</i>	2265	7,5	-60,8	Guiana
<i>INPA_1617</i>	2276	-0,25	-64,08	Guiana
<i>FMNH_391101</i>	2265	-12,52	-67,77	Inambari
<i>LSUMNS_31362</i>	2265	-9,06	-64,73	Inambari
<i>MPEG_16900</i>	2278	-7,57	-72,74	Inambari
<i>LSUMNS_42924</i>	2271	-4,2808	-77,238	Napo
<i>MPEG_5105</i>	2260	-2,0421	-67,288	Napo
<i>MPEG_440</i>	2274	-8,6957	-61,405	Rondônia
<i>MPEG_1345</i>	2196	-9,2667	-54,933	Tapajós
<i>USNM_B06831</i>	2263	-3,65	-52,367	Xingu
<i>RT011</i>	2227	-5,675	-50,302	Xingu
<i>MOP043</i>	2027	-7,033	-51,076	Xingu
<i>PPBIO152</i>	2200	-1,95	-51,6	Xingu