

Instituto de Biociências - Universidade de São Paulo

Paulo José Vieira da Silva Neto

**Diferenciação genética entre *Amazona*
aestiva (Linnaeus, 1758) e *Amazona*
ochrocephala (Gmelin, 1788): um exemplo
de divergência recente com fluxo gênico**

**Genetic differentiation between *Amazona*
aestiva (Linnaeus, 1758) and *Amazona*
ochrocephala (Gmelin, 1788): an example of
recent divergence with gene flow**

São Paulo

2023

Paulo José Vieira da Silva Neto

Versão Corrigida

Diferenciação genética entre *Amazona aestiva* (Linnaeus, 1758) e *Amazona ochrocephala* (Gmelin, 1788): um exemplo de divergência recente com fluxo gênico

Genetic differentiation between *Amazona aestiva* (Linnaeus, 1758) and *Amazona ochrocephala* (Gmelin, 1788): an example of recent divergence with gene flow

Dissertação apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Mestre em Ciências, na Área de Zoologia.

Orientadora: Dra. Cristina
Yumi Miyaki

São Paulo

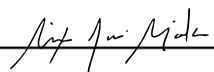
2023

Ficha catalográfica elaborada pelo Serviço de Biblioteca do Instituto de Biociências da USP,
com os dados fornecidos pelo (a) autor (a) no formulário:
'<https://biblioteca.ib.usp.br/ficha-catalografica/src/ficha.php>'

Neto, Paulo
Diferenciação genética entre *Amazona aestiva*
(Linnaeus, 1758) e *Amazona ochrocephala* (Gmelin,
1788) : um exemplo de divergência recente com fluxo
gênico / Neto Paulo ; orientadora Miyaki Cristina --
São Paulo, 2023.
101 p.

Dissertação (Mestrado) -- Instituto de
Biociências da Universidade de São Paulo. Ciências
Biológicas (Zoologia).

1. Evolução. 2. Genética. 3. Aves. 4.
Filogeografia. 5. Amazona. I. Miyaki, Cristina,
orient. Título.



Bibliotecária responsável pela catalogação:
Elisabete da Cruz Neves - CRB - 8/6228

ADVERTÊNCIA

Esta dissertação não é válida como publicação, conforme descrito no Capítulo 3 do CÓDIGO INTERNACIONAL DE NOMENCLATURA ZOOLOGICA. Portanto, os novos nomes e mudanças taxonômicas propostos aqui não são válidos para fins de nomenclatura ou prioridade.

WARNING

This dissertation is not valid as a publication, as described in Chapter 3 of the INTERNATIONAL CODE OF ZOOLOGICAL NOMENCLATURE. Therefore, new names and taxonomic changes proposed here are not valid for nomenclatural or priority purposes.

*À todas as pessoas que se dedicam à ciência.
Apesar das dificuldades, nosso trabalho é imprescindível.*

“Me lembrei que o mundo real era amplo. Um campo vasto de esperanças e medos, sensações e excitação. Aguardando aqueles que tivessem a coragem para ir adiante em seus campos e buscar o real conhecimento da vida em meio aos seus perigos.”

Charlotte Bronte, *Jane Eyre*

Agradecimentos

Agradeço à CAPES pela minha bolsa de mestrado, à FAPESP pelo financiamento de grande parte dos gastos no projeto de pesquisa, ao Programa de Pós-Graduação em Zoologia pelo constante auxílio e investimento.

Agradeço à minha mãe por todo o apoio e confiança que me concedeu para que eu realizasse a minha mudança para São Paulo e sempre persistisse no meu projeto e realizações pessoais. Agradeço ao meu noivo Igor por estar sempre ao meu lado e me incentivar em todos os meus projetos e sonhos.

Agradeço à minha orientadora Dra. Cristina Miyaki pelo apoio, confiança e instrução concedido para a realização do projeto. Bem como aos amigos do LGEMA que muito me auxiliaram e acompanharam durante a minha trajetória até aqui. Agradeço à Dani do Ornitolab da UFPE que me ajudou bastante na aquisição de amostras, bem como ao Dr. Luciano Naka e à Dra. Maria Pil que me prepararam para o processo de seleção e me introduziram à Cris. À Layse e à Lays pelo apoio e instrução constante em análises que precisei realizar. Agradeço à Claydson por me instruir na bioinformática, área que eu não tinha nenhum conhecimento. Agradeço à Renata por ter me ajudado no processamento dos materiais em bancada e me instruído nos procedimentos laboratoriais.

Agradeço à Dra. Gláucia Seixas, Valmir e Caio do projeto Papagaio-Verdadeiro por me receberem tão bem na ida à campo e me por terem me instruído tão pacientemente com as informações necessárias para desempenhar as atividades desse incrível projeto. Agradeço à Dra. Katia Ferraz e ao Dr. Alex Bovo pelos conselhos e instruções sobre as modelagens de nicho.

Agradeço aos amigos que pude fazer em São Paulo e aos que se mantiveram próximos apesar da distância, Jenny, Cida, Núbia, Alice, Isa, Andreza, Gio, Amanda, Gabriel, Elaine, Cledson, Ayanne e Jaci. O apoio de vocês foi fundamental para que eu conseguisse chegar até aqui. Agradeço também à Maicon, Mariana e Bárbara por todo o apoio e companhia, certamente foi muito importante.

Por fim, agradeço à equipe do colégio Semear, principalmente à Elaine e Fernando Maragon, pelo grande apoio e parceria constante.

Resumo

Estudos indicam que *Amazona aestiva* e *A. ochrocephala* são espécies irmãs, com claras diferenças morfológicas e que se divergiram cerca de 500 mil anos atrás, mas não apresentam monofiletismo recíproco. Além disso, cada espécie parece apresentar subpopulações, mas sem estrutura genética populacional bem definida. Para abordar essas lacunas de conhecimento, foram analisados 8464 SNPs (Polimorfismos de Nucleotídeo Único) de 26 indivíduos de *A. aestiva* e 20 de *A. ochrocephala* ao longo de suas distribuições e obtidos modelos de nicho ecológico para o presente e o passado. Os resultados obtidos confirmaram que *A. aestiva* e *A. ochrocephala* são espécies irmãs, ou seja, compartilham um ancestral comum mais recente entre si do que com outras espécies intimamente relacionadas. Foi encontrada diferenciação entre as duas espécies com suporte considerável. Foram identificadas três subpopulações em *A. ochrocephala*, correspondendo às suas três subespécies conhecidas. Já em *A. aestiva*, foram encontradas até quatro subpopulações distintas. Essas subpopulações parecem refletir a influência de fatores geográficos, o fluxo gênico com *A. ochrocephala* em determinadas regiões e possíveis efeitos do ambiente local. As análises dos modelos de nicho ecológico (MNE) auxiliaram na interpretação desses resultados e permitiram inferir que tanto *A. aestiva* quanto *A. ochrocephala* podem ter sido afetadas pelas flutuações climáticas ocorridas no Pleistoceno. Segundo os MNE obtidos, essas flutuações climáticas podem ter resultado em fragmentação ou contatos cíclicos das áreas de adequabilidade, o que pode ter contribuído para os padrões de diferenciação genética encontrados atualmente. No entanto, apesar dos resultados promissores, recomenda-se a inclusão de um maior número de indivíduos da região nordeste do Brasil para *A. aestiva* e da região da Colômbia, Venezuela, Suriname, Guianas e dos estados brasileiros do Amazonas e Roraima para *A. ochrocephala* em análises futuras, a fim de inferir melhor a história evolutiva desses táxons e compreender mais profundamente suas estruturas populacionais.

Abstract

Studies indicate that *Amazona aestiva* and *A. ochrocephala* are sister species, with clear morphological differences, and they diverged about 500,000 years ago, but they do not exhibit reciprocal monophyly. Additionally, each species seems to present subpopulations, although without a well-defined population genetic structure. To address these knowledge gaps, 8464 Single Nucleotide Polymorphisms (SNPs) were analyzed from 26 individuals of *A. aestiva* and 20 of *A. ochrocephala* across their distributions, and ecological niche models were obtained for both present and past conditions. The obtained results confirmed that *A. aestiva* and *A. ochrocephala* are sister species, meaning they share a more recent common ancestor with each other than with other closely related species. Differentiation between the two species was found with considerable support. Three subpopulations were identified in *A. ochrocephala*, corresponding to its three known subspecies. In *A. aestiva*, up to four distinct subpopulations were found. These subpopulations appear to reflect the influence of geographical factors, gene flow with *A. ochrocephala* in certain regions, and potential effects of the local environment. Analyses of Ecological Niche Models (ENMs) aided in interpreting these results and allowed for the inference that both *A. aestiva* and *A. ochrocephala* may have been affected by climatic fluctuations during the Pleistocene. According to the obtained ENMs, these climatic fluctuations could have led to fragmentation or cyclical contacts of suitable areas, which may have contributed to the current patterns of genetic differentiation. However, despite the promising results, it is recommended to include a greater number of individuals from the northeastern region of Brazil for *A. aestiva* and from the regions of Colombia, Venezuela, Suriname, the Guianas, and the Brazilian states of Amazonas and Roraima for *A. ochrocephala* in future analyses, in order to better infer the evolutionary history of these taxa and gain a deeper understanding of their population structures.

Sumário

I. Introdução.....	12
II. Objetivos.....	17
III. Material e Métodos.....	18
III.1. Amostragem e extração de DNA.....	18
III.2. Sequenciamento.....	18
III.3. Processamento dos dados.....	19
III.4. Análises Populacionais e Genéticas.....	20
III.5. Modelagem de Nicho Ecológico.....	21
IV. Resultados.....	24
IV.1. Análises genéticas e populacionais.....	24
IV.1.1. Processamento dos dados.....	24
IV.1.2. PCA.....	25
IV.1.3. SNMF.....	26
IV.1.4. DAPC.....	27
IV.1.5. Árvore de UPGMA.....	28
IV.1.6. Rede de Mínima Cobertura (MSN).....	28
IV.1.7. Estatísticas sumárias e FST.....	29
IV.2. Modelagem de Nicho Ecológico.....	29
V. Discussão.....	32
V.1. Análises genéticas e populacionais.....	32
V.2. Modelagem de Nicho Ecológico.....	34
VI. Conclusões.....	38
VII. Referências.....	39
VIII. Figuras e Tabelas.....	43

Lista de Figuras

Figura 1 - Espécies e subespécies estudadas.....	43
Figura 2 - Localidades de coleta das amostras utilizadas no estudo.....	44
Figura 3 - PCAs das matrizes teste de <i>Amazona</i> spp.....	45
Figura 4 - PCAs das matrizes teste de <i>A. aestiva</i> + <i>A. ochrocephala</i>	48
Figura 5 - PCAs das matrizes teste de <i>A. aestiva</i>	51
Figura 6 - PCAs das matrizes teste de <i>A. ochrocephala</i>	54
Figura 7 - SNMFs das matrizes teste de <i>Amazona</i> spp.....	57
Figura 8 - SNMFs das matrizes teste de <i>A. aestiva</i> + <i>A. ochrocephala</i>	60
Figura 9 - SNMFs das matrizes teste de <i>A. aestiva</i>	63
Figura 10 - SNMFs das matrizes teste de <i>A. ochrocephala</i>	66
Figura 11 - Influência dos eixos na Análise de Componentes Principais (PCA).....	69
Figura 12 - Análise de Componentes Principais (PCA) da matriz <i>Amazona</i> spp.....	70
Figura 13 - Análise de Componentes Principais de <i>A. aestiva</i> + <i>A. ochrocephala</i>	70
Figura 14 - Análise de Componentes Principais (PCA) da matriz <i>A. aestiva</i>	71
Figura 15 - Análise de Componentes Principais (PCA) da matriz <i>A. ochrocephala</i>	71
Figura 16 - Identificação do valor de K.....	72
Figura 17 - SNMFs da matriz <i>Amazona</i> spp.....	73
Figura 18 - SNMFs da matriz <i>A. aestiva</i> + <i>A. ochrocephala</i>	76
Figura 19 - SNMFs da matriz <i>A. aestiva</i>	79
Figura 20 - SNMFs da matriz <i>A. ochrocephala</i>	82
Figura 21 - SNMF de <i>A. aestiva</i> + <i>A. ochrocephala</i> nas localidades de amostragem.....	85
Figura 22 - SNMF de <i>A. aestiva</i> nas localidades de amostragem.	86
Figura 23 - SNMF de <i>A. ochrocephala</i> nas localidades de amostragem	87
Figura 24 - DAPC da matriz <i>Amazona</i> spp.....	88
Figura 25 - DAPC da matriz <i>A. aestiva</i> + <i>A. ochrocephala</i>	89
Figura 26 - DAPC da matriz <i>A. aestiva</i>	90

Figura 27 - DAPC da matriz <i>A. ochrocephala</i>	91
Figura 28 - Árvore de UPGMA da matriz <i>Amazona</i> spp.....	92
Figura 29 - Rede de Mínima Cobertura (MSN) da matriz <i>Amazona</i> spp.....	92
Figura 30 - Modelos de Nicho Ecológico.....	93

Lista de Tabelas

Tabela 1 - Amostras utilizadas no estudo.....	95
Tabela 2 - Parâmetros do Ipyrad.....	98
Tabela 3 - Combinações de fatores de frequência e contagem alélica.....	99
Tabela 4 - Parâmetros do VCFtools.....	100
Tabela 5 - Quantidade de <i>reads</i> e SNPs ao longo do processamento.....	100
Tabela 6 - Valores de estatísticas sumárias das espécies estudadas.....	100
Tabela 7 - Valores de <i>Fst</i> para os pares de táxons/grupos.....	101

I. Introdução

Os papagaios do gênero *Amazona* compõe, juntos, 32 espécies válidas atualmente, das quais 11 ocorrem em território brasileiro (Winkler *et al.*, 2020). Duas dessas espécies, ao longo do tempo, levantaram questões filogenéticas e evolutivas no meio da comunidade científica: *Amazona aestiva* (Linnaeus, 1758) e *Amazona ochrocephala* (Gmelin, 1788) (Eberhard e Bermingham, 2004; Ribas *et al.*, 2007; Caparroz *et al.*, 2009; Urantówka *et al.*, 2014).

A. aestiva (Fig. 1A e B) apresenta o corpo predominantemente verde, encontro da asa vermelho, amarelo na região da cabeça com padrões e tamanhos variados, bico e pés predominantemente pretos e azul na fronte, sendo esta última característica o que distingue a espécie (Collar *et al.*, 2020 A). Essa espécie apresenta duas subespécies: *A. a. aestiva*, que apresenta encontro das asas vermelho (Fig. 1A) e se distribui do Maranhão e Pará até o Rio Grande do Sul, e *A. a. xanthopteryx*, que tem o encontro das asas amarelo ou amarelo e vermelho de maneira mista (Fig. 1B) e se distribui da Bolívia e sudoeste do Mato Grosso até o norte da Argentina (Collar *et al.*, 2020a; Forshaw, 1989; Fig. 2). Caparroz e colaboradores (2009) relatam que indivíduos do Mato Grosso do Sul apresentam o encontro da asa com amarelo e vermelho de maneira mista, o que pode estar, segundo os autores, relacionado à sobreposição entre essas duas subespécies na região.

A. ochrocephala apresenta o corpo predominantemente verde, encontro da asa vermelho, bico superior com mancha rósea evidente e fronte e cabeça com manchas amarelas (Fig. 1C, D e E) que variam em forma e tamanho de um indivíduo para outro e entre as subespécies (Collar *et al.*, 2020b). Essa espécie engloba atualmente 4 subespécies: *A. o. xantholaema*, que apresenta amarelo mais extenso na cabeça até a região posterior da coroa, ao redor dos olhos e nas bochechas, apresenta também uma faixa verde na fronte e tom azulado no peito (Fig. 1C), sua área de distribuição engloba exclusivamente a Ilha de Marajó - Pará; *A. o. nattereri*, que se distribui do sudeste da Colômbia, Equador, Peru, Bolívia e Mato Grosso (Brasil), apresenta menos amarelo na cabeça do que as demais subespécies, faixa verde na fronte, coloração azulada nas bochechas e vermelho do encontro das asas frequentemente misturado com amarelo (Fig. 1 D); *A. o. ochrocephala*, que ocorre desde o leste da Colômbia até as Guianas e o Pará (Brasil), apresenta uma quantidade intermediária de amarelo na cabeça, mais do que *A. o. nattereri* e menos do que *A. o. xantholaema*, que se estende até a região da

coroa, e encontro das asas vermelho (Fig. 1E); *A. o. panamensis* (embora estudos indiquem que não é uma subespécie de *A. ochrocephala* [Eberhard e Bermingham, 2004; Ribas *et al.*, 2007; Caparroz *et al.*, 2009; Urantówka *et al.*, 2014], ainda está classificado nessa espécie segundo (Collar *et al.*, 2020b), ocorre exclusivamente do Panamá ao centro da Colômbia, se distingue da subespécie nominal por apresentar bico rosado e mancha amarela na cabeça exclusivamente em formato de “V” (Collar *et al.*, 2020b). Considerando que os resultados de estudos anteriores indicarem que é uma espécie distinta, tratamos aqui sob o nome de *A. panamensis*.

Apesar das claras diferenças morfológicas, ecológicas e vocais entre *A. aestiva* e *A. ochrocephala*, alguns estudos utilizando diversos tipos de marcadores genéticos encontraram baixa diferenciação entre as espécies, como descrevemos a seguir.

Eberhard e Bermingham (2004) utilizaram três genes mitocondriais de *A. ochrocephala* e de *A. aestiva* e identificaram três principais linhagens: 1 - indivíduos identificados na época como subespécies de *A. ochrocephala* que ocorrem na América Central (CA), mas que posteriormente foram elevadas à nível de espécie (*A. oratrix*, *A. auropalliata*, *A. belizensis* e *A. tresmariae*), e *A. o. panamensis*; 2 - Indivíduos de *A. aestiva* e de *A. o. ochrocephala* da América do Sul (SA), representando as populações de *A. o. ochrocephala* do Pará, *A. o. nattereri* da Bolívia e *A. o. xantholaema* da Ilha de Marajó; 3 - um indivíduo de *A. ochrocephala* da Colômbia representando o Norte da América do Sul (NSA). Os autores inferiram que na América do Sul os refúgios do Pleistoceno e o rio *Amazonas* podem ter sido barreiras biogeográficas isolando os grupos NSA e SA.

Ribas e colaboradores (2007) adicionaram 28 indivíduos de *A. aestiva* e *A. ochrocephala* da América do Sul ao conjunto de dados mitocondriais de Eberhard e Bermingham (2004) e encontraram resultados semelhantes. Foram recuperados os três principais clados descritos anteriormente e todos os indivíduos de *A. aestiva* foram alocados no clado da SA. Porém, foram identificadas duas linhagens dentro da linhagem SA (SA1 e SA2), sendo que uma delas engloba os *A. A. aestiva* do Mato Grosso do Sul e Minas Gerais, indivíduos de *A. o. xantholaema* (Pará) e *A. o. nattereri* do leste da Bolívia. A segunda linhagem engloba *A. o. nattereri* do Acre e Bolívia, *A. a. aestiva* do Mato Grosso do Sul e Minas Gerais, *A. a. xanthopteryx* do Mato Grosso e *A. o. ochrocephala* do Maranhão e Pará. Ribas e colaboradores (2007) puderam verificar que os indivíduos de *A. ochrocephala* da Ilha de Marajó se apresentavam em clados diferentes dos indivíduos do Amapá, evidenciando que o rio *Amazonas* pode ser uma

barreira para esses organismos, como proposto por Eberhard e Bermingham (2004). Apesar disso, os autores também afirmam que as flutuações climáticas ocorridas no Pleistoceno podem ter sido fatores influentes no processo evolutivo dessas duas espécies, sendo responsáveis pelo padrão filogenético que apresentam.

Ainda, Ribas e colaboradores (2007) corroboraram a ideia de que o processo de colonização da América do Sul pelos ancestrais dessas duas espécies pode ter se dado pelo extremo norte da América do Sul. Sendo isso inferido pela posição do clado NSA na base da filogenia, indicando uma posterior diferenciação do ancestral em dois novos clados: SA e CA.

Caparroz e colaboradores (2009) utilizaram mais amostras de *A. aestiva* e, baseados em marcadores mitocondriais, analisaram o padrão filogeográfico da espécie. Todavia, recuperaram os mesmo agrupamentos que os estudos anteriores, com todas as amostras de *A. aestiva* dentro do clado SA (Eberhard e Bermingham, 2004; Ribas *et al.*, 2007). Ainda, além do clado SA eles também recuperaram duas linhagens dentro desta, uma composta por *A. aestiva* da Bahia, Minas Gerais, Tocantins, Distrito Federal e Goiás, *A. o. xantholaema* e um indivíduo de *A. o. ochrocephala* da Bolívia (SA1), e outra linhagem que engloba *A. aestiva* da Argentina e Mato Grosso do Sul, *A. o. ochrocephala* do Macapá e *A. o. nattereri* do Brasil e Bolívia (SA2). Esse padrão é condizente com o perfil de distribuição geográfica das populações, mas não recuperou uma clara diferenciação genética entre as duas espécies. Os autores identificaram evidências de expansão demográfica na linhagem de *A. aestiva* e que o ancestral dessa espécie, possivelmente, viveu na região do Mato Grosso do Sul, se divergindo em duas linhagens há aproximadamente 300 mil anos atrás através de um processo de expansão no sentido nordeste.

Urantówka e colaboradores (2014), baseados nos marcadores mitocondriais dos estudos de Russello e Amato (2004) e Ribas e colaboradores (2007), recuperaram os clados CA, NSA, SA1 e SA2, também não encontrando monofiletismo recíproco entre *A. aestiva* e *A. ochrocephala*. Este estudo também confirmou que espécies insulares do Caribe, *A. barbadensis* e espécies das Pequenas Antilhas, como sendo próximas filogeneticamente ao clado NSA. Por conta disto os autores inferiram que o grupo que deu origem à *A. aestiva* e *A. ochrocephala* é descendente de ancestrais que colonizaram a América do Sul vindos das Pequenas Antilhas para o extremo norte da América do Sul há cerca de um milhão de anos atrás. Segundo os autores, esse processo possivelmente está associado a flutuações climáticas dos ciclos glaciais. Essas

mudanças também influenciaram um segundo possível evento de colonização do norte para o sul da América do Sul e do norte da América do sul para a América Central, dando origem aos grupos SA e CA, respectivamente..

Sato (2016) utilizou dois locos mitocondriais, um íntron ligado ao cromossomo Z e três locos nucleares anônimos na tentativa de investigar melhor a diferenciação genética entre essas duas espécies. Baseada nos dados mitocondriais e nucleares anônimos, Sato (2016) recuperou os mesmos clados encontrados anteriormente nos demais estudos, com uma separação adicional do clado SA2 em dois outros grupos . Todavia, as duas linhagens observadas com o marcador ligado ao cromossomo Z separavam de modo congruente a grande maioria dos indivíduos das duas espécies (exceção de 2 dos 188 indivíduos analisados). Assim, com exceção do marcador ligado ao cromossomo Z, apesar da maior quantidade e variedade de marcadores, não foi possível identificar uma clara separação genética entre as duas espécies em questão. Ainda, assim como levantado por outros autores, Sato (2016) sugere que essa ausência de monofiletismo se deve devido a possível introgressão e/ou sorteamento incompleto de linhagens (Eberhard e Bermingham, 2004; Ribas *et al.*, 2007; Caparroz *et al.*, 2009; Urantówka *et al.*, 2014).

Diversos estudos tem evidenciado que espécies com fluxo gênico e divergência recente representam um desafio para identificar uma clara delimitação entre ambas (Leaché *et al.*, 2014). Isso pode estar associado ao nível de fluxo gênico e à metodologia utilizada no estudo, que podem influenciar fortemente os resultados, isto é, as relações filogenéticas apresentadas (Jackson *et al.*, 2017). Atualmente, uma das estratégias que tem sido utilizada com sucesso para realizar esses tipos de estudos é a análise de grande parte do genoma (Chan *et al.*, 2017).

O sequenciamento de nova geração (*Next-Generation Sequencing* - NGS) permite a leitura de sequências de diversos locos simultaneamente, gerando uma grande quantidade de dados que podem ser utilizados em diversos tipos de estudos. Pesquisas recentes demonstraram que o NGS tem se mostrado uma ferramenta efetiva para a identificação de grupos que possuam divergência recente e fluxo gênico elevado (Feder *et al.*, 2012; Sousa e Hey, 2013; Sullivan *et al.*, 2014; Zarza *et al.*, 2016). Uma das maneiras que têm sido utilizada para não sequenciar genomas completos (devido ao alto custo financeiro) é a genotipagem por sequenciamento (*Genotyping by Sequencing* - GBS), desenvolvida por Elshire e colaboradores (2011). Essa técnica utiliza enzima de restrição para selecionar centenas de milhares segmentos espalhados pelo genoma e

essa biblioteca é sequenciada por NGS. Os dados obtidos se mostraram úteis para estudos filogenéticos (Narum *et al.*, 2013; Grzegorzczuk *et al.*, 2021; Gurgul *et al.*, 2019).

Tendo isto em vista, utilizamos NGS para a realização do presente trabalho no intuito de avaliar se os dados gerados por essa metodologia são informativos para diferenciar *A. aestiva* e *A. ochrocephala* e avaliar a estrutura populacional nestas espécies.

II. Objetivos

Este trabalho teve como objetivo geral compreender a história evolutiva de dois táxons morfológicamente distintos, mas com baixa diferenciação genética, *Amazona aestiva* e *A. ochrocephala*.

Os objetivos específicos do trabalho foram:

- i. Avaliar a diferenciação genética entre *A. aestiva* e *A. ochrocephala*;
- ii. Analisar a estrutura populacional das espécies;
- iii. Inferir o(s) possível(is) processo(s) evolutivo(s) que pode(m) ter influenciado nos resultados obtidos.

III. Material e Métodos

III.1. Amostragem e extração de DNA

Foram utilizadas 60 amostras de sangue ou tecido armazenadas no banco de tecidos do Laboratório de Genética e Evolução Molecular de Aves - LGEMA e 9 amostras depositadas na Universidad Nacional del Centro de la Provincia de Buenos Aires, sob cuidados de Igor Berkunsky (Tabela 1). As localidades de amostragem de *A. aestiva* e *A. ochrocephala* foram plotadas em mapa utilizando o *software* QGIS 3.16 (<https://www.QGIS 3.16.org/>) sobre a distribuição geográfica das espécies da BirdLife International and Handbook of the Birds of the World (2021; Fig. 2). Para a extração foi utilizado o Kit Qiagen's DNeasy Blood e Tissue seguindo os protocolos do fabricante. O DNA extraído foi submetido a eletroforese em gel de agarose para avaliar sua presença e qualidade. A concentração de DNA foi mensurada no Qubit utilizando o kit brDNA . As amostras que não atingiram a concentração e volume mínimos desejados (20ng/μl e 30μl) tiveram sua concentração aumentada no SpeedVac. As amostras de DNA foram mantidas em freezer e, após análise de qualidade do DNA, foram enviadas para a construção das bibliotecas genômicas e o sequenciamento em empresa terceirizada.

III.2. Sequenciamento

O método de Genotipagem por Sequenciamento (GBS - *Genotyping by Sequencing*) foi implementado para a redução da complexidade do genoma baseando-se nos parâmetros gerais descritos por Elshire e colaboradores (2011). A enzima *Pst* I foi utilizada para a digestão do DNA das amostras, que, após essa etapa, foram liofilizadas a vácuo. Em seguida, as amostras receberam um adaptador com sequências indexadoras (*barcodes*) e um segundo adaptador com uma sequência comum utilizada como iniciador, ambos foram ligados às extremidades de corte da enzima *Pst*I através da ação da enzima T4 DNA Ligase. No intuito de aumentar a diversidade das sequências geradas e diminuir a falha na geração de boas sequências, foram realizadas duas etapas de ligação, gerando, dessa maneira, duas bibliotecas por amostra. Por conta disso, cada amostra possuía duas sequências indexadoras para sua identificação após o sequenciamento.

Após a ligação dos adaptadores, foram feitos dois *pools* das amostras, uma de

cada biblioteca, que foram purificadas com *beads* magnéticas. Em seguida, foi realizada a Reação em Cadeira da Polimerase (*Polymerase Chain Reaction*, PCR) dos fragmentos, que possuíam de 200 a 450 pb. Após a amplificação, foi realizada uma segunda purificação com *beads* magnéticas. Após a segunda purificação, a qualidade das bibliotecas foi mensurada pelo BioAnalyzer Agilent 2100. com a utilização do kit High Sensitivity DNA. Essa análise não indicou presença de dímeros de *primers* ou excesso de adaptadores. As bibliotecas foram quantificadas por PCR em tempo real com o *KAPA Biosystems Quantification Kit* (Illumina) e, após a diluição e requantificação, foram enviadas para o Centro de Genômica da ESALQ/USP, onde foram reunidas em concentrações equimolares e sequenciadas em uma *flowcell* do NextSeq2000 Illumina com o Kit P2V2 (100 ciclos - *Single Read*).

III.3. Processamento dos dados

As *reads*, recebidas em um único arquivo, foram então demultiplexadas com a utilização do *software* Ipyrad (Eaton, 2021) e o controle de qualidade foi feito com a utilização do FASTQC (Andrews, 2010). As *reads* foram mapeadas com a utilização do genoma de referência de *A. aestiva* de maior qualidade (ID: 40915) depositado no *National Center for Biotechnology Information* (NCBI). O Ipyrad também foi utilizado para realização da filtragem das *reads* usando os seguintes parâmetros: Q-score (qualidade do sequenciamento do sítio) mínimo de 33, profundidade mínima do sítio de 6 e máxima de 10000. similaridade mínima de 85% para agrupamento das *reads* em um mesmo *cluster*, 0% de erros na sequência do *barcode* e filtro máximo (valor 2) para adaptadores (outros parâmetros são apresentados na Tabela 2). Foram geradas, ao final do processamento, quatro matrizes .vcf contendo os *Single Nucleotide Polymorphisms* (SNPs) resultantes para cada matriz (matriz 1: *Amazona* spp.; matriz 2: *A. aestiva*+*A. ochrocephala*; matriz 3: *A. aestiva*; matriz 4: *A. ochrocephala*). As quatro matrizes foram obtidas em processamentos independentes para garantir que os SNPs obtidos em cada uma delas fosse adequado para as diferentes análises.

Linck e Battey (2018) evidenciaram que os valores de *Minor Allele Frequency* (MAF) e *Minor Allele Count* (MAC) afetam fortemente os resultados das análises, acarretando em erros na detecção de estrutura genética existente quando os filtros não são adequados. Assim, nossas quatro matrizes de SNPs passaram por filtragem adicional no VCFtools (Danecek *et al.*, 2011) com diferentes combinações de MAF

(0.05, 0.10 e 0.15) e MAC (1, 2 e 3; Tabela 3). Essas novas matrizes de SNPs foram utilizadas em diferentes análises (apresentadas a seguir) e os resultados foram comparados para decidir os parâmetros que geraram resultados com melhor sentido biológico. Realizamos Análises de Componentes Principais (*Principal Component Analysis*, PCA; Fig. 3 para *Amazona* spp.; Fig. 4 para *A. aestiva* + *A. ochrocephala*; Fig. 5 para *A. aestiva*; Fig. 6 para *A. ochrocephala*), que consiste na identificação dos padrões de variação das amostras, evidenciando o perfil genético de cada indivíduo/grupo de indivíduos, bem como a existência de estrutura populacional quando presente. Também foram realizadas análises da identificação de estrutura populacional (*Sparse Non-negative Matrix Factorization*, SNMF; Fig. 7 para *Amazona* spp.; Fig. 8 para *A. aestiva* + *A. ochrocephala*; Fig. 9 para *A. aestiva*; Fig. 10 para *A. ochrocephala*). Com base nos resultados obtidos nas análises de PCA e SNMF (descritas a seguir) as matrizes 4 (com MAF 0.05 e MAC 2) foram escolhidas para dar seguimento às análises. Essa decisão foi baseada na identificação mais clara de agrupamentos distintos na PCA e maior evidência de estrutura populacional no SNMF. Embora as matrizes com MAF 0.05 e MAC 3 tenham se mostrado ótimas candidatas na identificação de grupos, esses valores não foram selecionados, pois para *A. oratrix*, *A. aestiva* PI e *A. ochrocephala* RO possuímos apenas 2 amostras de cada e com MAC 3 elas seriam retiradas da análise. Ainda, essas matrizes foram filtradas também no VCFtools para a remoção de indivíduos que, após a realização da PCA e do SNMF. Alguns indivíduos poderiam representar indivíduos identificados erroneamente com base em morfologia (ver tópico IV.1 dos Resultados), remoção de locos com mais de dois alelos e locos com *missing data* (Tabela 4).

III.4. Análises Populacionais e Genéticas

As matrizes finais de SNPs foram utilizados para realizar PCA no pacote ape (Paradis *et al.*, 2019) no R Software para a identificação de grupos. Visando analisar a composição genética de cada amostra foram obtidos SNMFs para as quatro matrizes, com K variando de 1 a 10 no pacote adegenet (Jombart e Ahmed, 2011). Nessa análise foram utilizados: 500 interações, 100 repetições, alfa de 10 (também foram testados valores de 1, 50, 100 e 500, sem diferença significativa de resultados) e tolerância de 0.00001. Os resultados foram plotados em mapa com a utilização do QGIS 3.16 (<https://www.QGIS 3.16.org/>). O melhor valor de K foi obtido para cada matriz através dos valores de entropia. A Análise Discriminante dos Componentes Principais

(*Discriminant Analysis of Principal Components*, DAPC) também foi realizada no pacote ape utilizando os resultados dos agrupamentos encontrados pelo PCA e SNMF como hipótese inicial para a identificação das populações.

Uma matriz de distância genética foi calculada com a utilização do método WC84 no pacote hierfstat para a matriz *Amazona* spp.. Essa matriz de distância foi utilizada para a obtenção de uma árvore de UPGMA no pacote adegenet no R Software, com amostragens a cada 1000. valor suporte mínimo de 50 para manter os nós na árvore consenso e 1000 réplicas de *bootstrap*. O pacote hierfstat também foi utilizado para a obtenção de uma Rede de Mínima Cobertura (*Minimum Spanning Network*, MSN) para visualizarmos de forma gráfica a distância genética entre as amostras. Ambas as análises foram realizadas no intuito de visualizar a estrutura das populações e a distância genética entre os grupos.

O pacote hierfstat (Goudet *et al.*, 2005) também foi utilizado para calcular os valores de heterozigosidade esperada e observada, endocruzamento (Fis) e índice de diferenciação genética global (F_{ST}) das matrizes de *A. aestiva* e de *A. ochrocephala*. Além disso, o software PLINK (Chang *et al.*, 2015) foi utilizado para calcular os valores de F_{ST} entre pares de táxons considerando os resultados das análises de SNMF, PCA e DAPC (3 grupos em *A. ochrocephala* - AP, PA e AC+RO+BOL; 4 em *A. aestiva* - MS+ARG, BOL, MT+DF e PI).

III.5. Modelagem de Nicho Ecológico

A Modelagem de Nicho Ecológico (*Ecological Niche Modeling*, ENM) foi realizada utilizando o pacote ENMeval (Phillips *et al.*, 2017) e vários outros pacotes auxiliares das funções utilizadas em todo o processo, desde a obtenção dos pontos até a plotagem final dos modelos (raster, terra, dplyr, ggplot2, scrubr, spocc, warbleR, maptools, letsR, sp, rts, RCurl, bitops, rgeos, rgdal, fmsb, ecospat, dismo, rJava, sf, tidyverse, colorRamps, rasterVis, spThin, spam, fields, knitr e devtools).

Os registros de ocorrência foram obtidos do banco de dados do GBIF (<https://www.gbif.org/>) e do Xeno-Canto (<https://www.xeno-canto.org/>). Os registros foram então selecionados com base nos seguintes critérios previamente discutidos com especialistas : de 1950 a 1995 (segundo a Dra. Gláucia Seixas, pesquisadora que trabalha com *A. aestiva* em campo desde a década de 1990, após os anos 1990 muitas solturas em locais inapropriados modificaram muito a área de ocorrência da espécie),

somente na América do Sul e a pelo menos 40 km de outros registros (comunicação pessoal com Dra. Kátia Ferraz e Dr. Alex Bovo) a fim de evitar vieses no modelo.

As variáveis ambientais foram obtidas no PaleoClim (Brown *et al.*, 2018) para as seguintes camadas de tempo: Presente 1979-2013, *Pleistocene: late-Holocene* (LH) 4.2-0.3 ka, *Pleistocene: mid-Holocene, Northgrippian* (8.326-4.2 ka), *Pleistocene: early-Holocene* (EH) 11.7-8.326 mil anos atrás (ka), *Pleistocene: Younger Dryas Stadial* (YDS) 12.9-11.7 ka, *Pleistocene: Bølling-Allerød* (BA) 14.7-12.9 ka, *Pleistocene: Heinrich Stadial 1* (HS1) 17.0-14.7 ka, *Pleistocene: Last Glacial Maximum* (LGM) ~21 ka, *Pleistocene: Last Interglacial* (LIG) ~130 ka, *Pleistocene: Marine Isotope Stage 19* (MIS19) ~787 ka, *Pliocene: mid-Pliocene warm period* (MPWP) ~3.205 milhões de anos atrás (Ma) e *Pliocene: Marine Isotope Stage M2* (M2) ~3.3 Ma. Como nem todas as camadas possuem as 19 variáveis climáticas, apenas as variáveis presentes em todas as camadas foram utilizadas.

Foi calculada a correlação entre as variáveis utilizando o índice de Spearman e removemos aquelas que possuíam correlação maior que 80% com alguma outra variável. Ainda, a seleção dos principais fatores abióticos que devem afetar a sobrevivência e sucesso reprodutivo de *A. aestiva* considerou informações da especialista Dra. Gláucia Seixas: quantidade de chuva no período seco (época de reprodução das aves no Cerrado) e temperaturas máximas e mínimas dessa estação (Bio6, Bio11, Bio14 e Bio17). Dessa forma, com base nos critérios de correlação e dados biológicos da espécie, foram utilizadas as variáveis Bio4, Bio10, Bio11, Bio16, Bio17, Bio18 e Bio19 (Sazonalidade da temperatura, Temperatura média no quartil mais quente, Temperatura média no quartil mais frio, Precipitação média no quartil mais úmido, Precipitação média no quartil mais seco, Precipitação média no quartil mais quente e Precipitação média no quartil mais frio, respectivamente).

Com as variáveis e os pontos de registro selecionados, foi realizada a análise de modelagem com base em diferentes combinações de modelos através do método de partição espacial em blocos (*Spatial block partitions*) e considerando opções de modelo com termos lineares (L - assume uma relação linear entre as variáveis ambientais e a probabilidade de ocorrência da espécie), quadráticos (Q - assume uma relação não linear entre as variáveis ambientais e a probabilidade de ocorrência da espécie) e de interação (H - assume uma relação de interação entre duas ou mais variáveis ambientais). Com os valores obtidos para cada modelo calculado, o modelo LQ 0.5 foi

selecionado para *A. aestiva* e o modelo LQ 2.5 para *A. ochrocephala*, tendo como base os valores de AUC, AICc, omissão e comissão para seleção de ambos os modelos. Os modelos foram plotados para o presente e extrapolados para o passado de todas as camadas disponíveis.

Os resultados das modelagens de nicho ecológico foram plotados em mapa utilizando o programa QGIS 3.16 (<https://www.QGIS 3.16.org/>). As áreas de adequabilidade com valores inferiores a 24% não foram incluídas nos mapas para melhor visualização dos resultados.

IV. Resultados

IV.1. Análises genéticas e populacionais

IV.1.1. Processamento dos dados

Recebemos 462.794.689 *reads*, com média de 6.907.383 por amostra. Após alguns testes iniciais, esse conjunto de dados foi filtrado no Ipyrad com as configurações de parâmetros apresentados na Tabela 2. Após essa filtragem as quatro matrizes de SNPs iniciais passaram por filtragem adicional para remover *sin gletons* e SNPs com mais de dois alelos e com frequência inferior a 5% (Tabela 4). A quantidade de SNPs em cada matriz ao longo do processamento dos dados está apresentada na Tabela 5.

As análises iniciais de SNMF da matriz *Amazona ssp.* (não apresentadas aqui) revelaram alguns resultados inesperados e decidimos remover algumas amostras das análises seguintes, como explicado a seguir. A amostra A08AaSP se agrupou geneticamente com indivíduos do MT, mas foi coletada em uma cidade litorânea de SP, o que levantou suspeitas de ser um indivíduo oriundo de escape ou soltura e foi removida das análises subsequentes. Os indivíduos A01AaBOL, A06AaBOL e A07AaBOL foram removidos das análises por serem de cativeiro, mesmo tendo se agrupado geneticamente aos demais indivíduos da Bolívia com grande similaridade no padrão genético (ancestralidade alta com *A. aestiva* e *A. ochrocephala*), pois a quantidade de indivíduos amostrados de vida livre daquela localidade (4 indivíduos) parece ter sido suficientes para as análises subsequentes.

No início do estudo, havia a suspeita que os indivíduos A09AaMT e A69AaMT fossem híbridos de *A. aestiva* e *A. ochrocephala* (comunicação pessoal do coletor das amostras, Dr. Luís Fábio Silveira). Porém, esses indivíduos formaram um *cluster* independente nas análises de SNMF. Sendo assim, é possível que possam ser híbridos de outras espécies que ocorrem naquela região, como *A. amazonica* ou *A. farinosa*. Por conta disso, por precaução, optamos por também remover esses dois indivíduos das análises subsequentes.

Infelizmente não foi possível obter amostras de indivíduos de *A. o. ochrocephala* do norte da América do Sul (principalmente Venezuela e Colômbia). Sendo assim, optamos por incluir amostras de *A. o. ochrocephala* de cativeiro. Ainda, havia a possibilidade das amostras A34AoAP, A35AoAP, A37AoAP, A49AoAP,

A51AoAP, A52AoAP serem desse táxon, pois são de indivíduos amostrados no AP e a distribuição dessa subespécie chega até a margem superior do rio *Amazonas*. Os resultados do SNMF (explicados mais aprofundadamente no tópico IV.1.3 dos resultados) indicaram como bastante plausível que essas 6 amostras do AP pertencem a essa subespécie porque formaram um *cluster* distinto daquele composto por *A. ochrocephala nattereri* e do composto por *A. ochrocephala xantholaema*. Ainda, esses resultados do SNMF indicaram que alguns indivíduos de cativeiro identificados como *A. ochrocephala ochrocephala* apresentavam múltiplas ancestralidades. Sendo assim, optamos por remover as amostras de cativeiro dessa subespécie das análises subsequentes (A27AoochCAT, A32AoochCAT, A33AoochCAT, A46AoochCAT, A47AoochCAT, A48AoochCAT e A65AoochCAT). As amostras A36AoAP e A50AoAP foram identificadas como *A. ochrocephala*, mas devem ser *A. amazonica*, por isso foram removidas do estudo. Por fim, também removemos uma amostra de *A. o. xantholaema* (A26AoxanCAT) de cativeiro, pois as amostras de indivíduos de vida livre da Ilha de Marajó foram suficientemente informativas. Portanto, para as análises subsequentes mantivemos 26 *A. aestiva*, 20 *A. ochrocephala*, 4 *A. auropalliata*, 2 *A. oratrix* e 1 *A. panamensis* (Tabela 1).

IV.1.2. PCA

Os dois primeiros componentes principais da PCA das matrizes *Amazona* spp., *A. aestiva* + *A. ochrocephala*, *A. aestiva* e *A. ochrocephala* explicam, respectivamente, mais de 25%, cerca de 20%, cerca de 11% e cerca de 20% (Fig. 11). Os resultados da PCA da matriz *Amazona* spp. (Fig. 12) evidenciaram que a maior diferença genética encontrada ao longo do PC1 foi entre os indivíduos dos grupos externos da América Central (*A. auropalliata* e *A. oratrix*) daqueles da América do Sul (*A. aestiva* e *A. ochrocephala*), com o indivíduo de *A. o. panamensis* (táxon que ocorre no Panamá e Colômbia, próximo a *A. ochrocephala*) posicionado entre *A. ochrocephala* e os indivíduos da América Central. Já ao longo do PC2 *A. aestiva* e *A. ochrocephala* se separam. Analisando a PCA da matriz *A. aestiva* + *A. ochrocephala* (Fig. 13) é possível identificar que *A. aestiva* se separou de *A. ochrocephala* no eixo PC1, ao passo que *A. ochrocephala* se separou em dois grupos distintos (leste e oeste) no eixo PC2. Nota-se que os indivíduos de *A. aestiva* da Bolívia se posicionaram de maneira intermediária entre *A. aestiva* e *A. ochrocephala*.

Intraespecificamente, a PCA da matriz *A. aestiva* (Fig. 14) indicou a separação entre os indivíduos de *A. aestiva* da BOL, que apresentam grande compartilhamento de ancestralidade com *A. ochrocephala*, e os demais *A. aestiva* no eixo PC1. No eixo PC2 foi possível identificar três agrupamentos correspondentes aos *A. aestiva* da ARG, do MS e do MT+DF+PI. A PCA da matriz de *A. ochrocephala* (Fig. 15) evidenciou a separação entre dois grupos distintos no eixo PC1, um formado pelos indivíduos do AC+RO+BOL e outro pelos indivíduos do PA+AP. No eixo PC2 foi possível identificar a separação entre dois outros grupos, um formado pelos indivíduos do PA e outro pelos indivíduos do AP.

IV.1.3. SNMF

A análise da identificação do melhor valor de K (Fig. 16) indicou K = 3 ou 4 para a matriz *Amazona* spp., K = 2 ou 3 para a matriz *A. aestiva*+*A. ochrocephala*, K = 1 para a matriz *A. aestiva* e de K = 1 para a matriz *A. ochrocephala*.

Assumindo K = 2 para a matriz *Amazona* spp. (Fig. 17, K2) se observa a separação entre CA e SA+NA. Para K = 3 (Fig. 17, K3) os *clusters* correspondem a CA, *A. aestiva* e *A. ochrocephala*. Para K = 4 (Fig. 17, K4) os *clusters* se referem a CA, *A. aestiva*, *A. ochrocephala* leste e *A. ochrocephala* oeste. Para K = 5 (Fig. 17, K5) foram encontrados *clusters* separando CA, *A. aestiva* norte, *A. aestiva* sul, *A. ochrocephala* leste e *A. ochrocephala* oeste. Para K = 6 (Fig. 17, K6) o *cluster* adicional corresponde à separação de *A. ochrocephala* oeste em *A. ochrocephala* PA e *A. ochrocephala* AP. Os SNMFs com valores de K > 6 (Fig. 17, K7 a K10) mostraram agrupamentos sem congruência biológica.

Os resultados de SNMF da matriz *A. aestiva*+*A. ochrocephala* com K = 2 (Fig. 18, K2) evidenciaram dois *clusters* congruentes com *A. aestiva* e *A. ochrocephala*, respectivamente, mas com indivíduos da BOL compartilhando alelos de ambos os *clusters*. Para K = 3 (Fig. 18, K3) os *clusters* correspondem a *A. aestiva*, *A. ochrocephala* leste e *A. ochrocephala* oeste. Para K = 4 (Fig. 18, K4) os *clusters* se referem a *A. aestiva* norte, *A. aestiva* sul, *A. ochrocephala* leste e *A. ochrocephala* oeste. Para K = 5 (Fig. 18, K5) foi observada a separação adicional de *A. ochrocephala* AP e *A. ochrocephala* PA. Os SNMFs com valores de K > 5 (Fig. 18, K6 a K10) mostraram agrupamentos sem congruência biológica.

Os resultados de SNMF da matriz *A. aestiva* com $K = 2$ (Fig . 19, K2) evidenciaram dois *clusters* , um representando *A. aestiva* que compartilham alelos com *A. ochrocephala* e outro com ancestralidade apenas de *A. aestiva*. Para $K = 3$ (Fig. 19, K3) foram formados três *clusters*, um representando *A. aestiva* que compartilham alelos com *A. ochrocephala*, outro sem nenhuma explicação biológica clara, representando indivíduos de *A. aestiva* da região de simpatria (Bolívia) e outro representando os indivíduos de *A. aestiva* com ancestralidade apenas de *A. aestiva*. Para $K = 4$ (Fig. 19, K4) foram identificado quatro *clusters* distintos, um representando os que possuíam alelos compartilhados de *A. ochrocephala*, outro representando os indivíduos do Piauí, outro os indivíduos do centro da distribuição (MT+DF) e outro representando o sul da distribuição (ARG+MS). Os SNMFs com valores de $K > 4$ (Fig . 19, K5 a K10) mostraram agrupamentos sem congruência biológica.

Os resultados de SNMF da matriz *A. ochrocephala* com $K = 2$ (Fig . 20. K2) evidenciaram dois *clusters*, um representando *A. ochrocephala* leste e outro *A. ochrocephala* oeste. Para $K = 3$ (Fig . 20. K3) foram formados três *clusters*, um representando indivíduos do oeste, outro representando indivíduos do AP e o outro indivíduos do PA. Os SNMFs com valores de $K > 3$ (Fig . 20. K4 a K10) mostraram agrupamentos sem congruência biológica.

Para melhor visualização dos resultados do SNMF das matrizes *A. aestiva* + *A. ochrocephala*, *A. aestiva* e *A. ochrocephala*, os valores de K que apresentavam alguma congruência biológica (K2 a K5, K2 a K4 e K2 a K3, respectivamente) foram plotados em mapa (Fig. 21: *A. aestiva* + *A. ochrocephala*; Fig. 22: *A. aestiva*; Fig. 23: *A. ochrocephala*).

IV.1.4. DAPC

O resultado da DAPC utilizando a matriz *Amazona* spp. (Fig. 24) evidenciou melhor a diferenciação entre as populações encontrada na PCA (Fig. 12), os indivíduos da América Central se apresentaram separados dos indivíduos da América do Sul no eixo horizontal, bem como o indivíduo de *A. o. panamensis* se apresentando entre ambos os grupos. No eixo vertical foi possível identificar a separação entre *A. aestiva* e *A. ochrocephala*. Através da DAPC da matriz *A. aestiva* + *A. ochrocephala* (Fig. 25) foi possível visualizar que o padrão de distribuição se manteve como na PCA dessa matriz (Fig. 13) mas com um pouco mais de clareza. As duas espécies se separaram pelo eixo horizontal, os indivíduos da Bolívia se mantiveram intermediários entre *A.*

aestiva e *A. ochrocephala* oeste, e *A. ochrocephala* se separou no eixo vertical em dois grupos distintos (leste e oeste).

No eixo horizontal da DAPC da matriz *A. aestiva* (Fig. 26) foi possível observar uma maior diferenciação de *A. aestiva* da BOL, cujos indivíduos apresentam grande compartilhamento de ancestralidade com *A. ochrocephala* em relação aos demais indivíduos de *A. aestiva* no eixo horizontal. Já no eixo vertical foi possível identificar a separação entre os indivíduos das diferentes localidades, com pouca sobreposição entre os indivíduos do MS, da ARG, do MT, do PI do DF e da BOL. Na DAPC da matriz *A. ochrocephala* (Fig. 27) foi possível identificar a separação entre as populações do leste (AP e PA) do oeste (AC+RO+BOL) no eixo horizontal. No eixo vertical a separação entre AP e PA foi mais evidente do que na PCA.

IV.1.5. Árvore de UPGMA

A árvore de UPGMA (Fig. 28) resultou em *A. aestiva* e *A. ochrocephala* em duas linhagens irmãs com suporte alto no nó de *A. ochrocephala* (98). Os agrupamentos identificados nas análises de PCA e DAPC (MS+ARG, MT+DF+PI para *A. aestiva*, AC+RO+BOL, PA e AP para *A. ochrocephala*) se apresentaram como monofiléticos na análise de UPGMA, mas as linhagens dentro de *A. aestiva* apresentaram suportes baixos, enquanto as de *A. ochrocephala* apresentaram bons valores de suporte: AC+RO+BOL (100), AP (100) e PA (100).

IV.1.6. Rede de Mínima Cobertura (MSN)

A Rede de Mínima Cobertura obtida com a matriz *Amazona* spp. (Fig. 29) é congruente com e os resultados da PCA (Fig. 12), DAPC (Fig. 24) e árvore de UPGMA (Fig. 28). As amostras dos grupos externos (América Central) não se agruparam com as do grupo interno (*A. aestiva* e *A. ochrocephala*). A amostra de *A. o. panamensis* aparece em posição intermediária entre as da América do Sul e da América Central. Os indivíduos de *A. aestiva* se concentram em uma região da rede e separados dos indivíduos de *A. ochrocephala*, com os indivíduos de *A. aestiva* da Bolívia entre os dois grupos.

IV.1.7. Estatísticas sumárias e F_{ST}

Os índices de heterozigiosidade observada e esperada (Tabela 6) em ambas as espécies foram similares (Hs: 0.2479 e Ho: 0.2298 para *A. aestiva*; Hs: 0.2317 e Ho: 0.2085 para *A. ochrocephala*), o índice de endocruzamento foi baixo (0.07 para *A. aestiva* e 0.10 para *A. ochrocephala*) e a diferenciação genética global das duas espécies estudadas foi relativamente baixa (0.02 para *A. aestiva* e 0.6 para *A. ochrocephala*). Foi possível identificar que a diferenciação intraespecífica em *A. ochrocephala* foi três vezes maior do que em *A. aestiva*. O cálculo de F_{ST} entre táxons (Tabela 7) indicou que o maior nível de diferenciação genética foi encontrado entre as espécies da América do Sul e as da América Central, já entre as duas espécies da América do Sul (*A. aestiva* e *A. ochrocephala*) a diferenciação foi moderada (0.07).

IV.2. Modelagem de Nicho Ecológico

Após a filtragem e seleção dos pontos de registro para a realização das modelagens foram obtidos 109 pontos para *A. aestiva* e 121 pontos para *A. ochrocephala*. As variáveis escolhidas e os modelos selecionados (LQ 0.5 para *A. aestiva* e LQ 2.5 para *A. ochrocephala*) mostram como a adequabilidade de nicho se distribui geograficamente e pode indicar áreas onde essas duas espécies podem ter ocorrido ao longo do tempo (Fig. 30).

Os modelos do Plioceno (M2 e MPWP; Fig. 29A e 29B) devem ser interpretados de maneira particular, pois, apesar de o modelo para M2 ser de um período de glaciação, muito influenciado por massas de ar polares, o Plioceno foi um período mais quente do que o Pleistoceno e o Antropoceno, o que explica a maior adequabilidade de área para *A. ochrocephala* para estes momentos (Dollan *et al.*, 2015; Hill, 2015). O modelo para M2 evidenciou uma área de adequabilidade para *A. ochrocephala* mais intensa e abrangente na região circum-amazônica e norte do Cerrado do que nos tempos mais recentes. A área de *A. aestiva* por sua vez também abrangia a mesma região que a atual e não foi consideravelmente diferente do presente. Já durante o período MPWP a área de adequabilidade de *A. ochrocephala* foi muito maior e mais adequada, ao contrário do que foi observado para *A. aestiva*, que teve sua área de adequabilidade reduzida à fração mais sul da sua provável distribuição, concentrando-se no sul do Brasil e Argentina, Bolívia e Paraguai.

Já durante o Pleistoceno os modelos foram bastante variáveis, bem como as condições climáticas do período (Brown *et al.*, 2018). Para o MIS-19 (Fig. 29C), o modelo evidenciou uma maior expansão da área de adequabilidade de *A. aestiva* em direção ao norte do Cerrado, ao passo que a área de adequabilidade de *A. ochrocephala* se retraiu às regiões que margeiam a Amazônia. Durante o último máximo interglacial (LIG, Fig. 29D) houve uma redução e fragmentação da área de adequabilidade de *A. aestiva* em dois principais blocos: norte e sul da distribuição atual. Ao passo que *A. ochrocephala* teve sua área de adequabilidade expandida para dentro da área de distribuição de *A. aestiva*, invadindo pela região da Bolívia, Mato Grosso e Mato Grosso do Sul. Para o último máximo glacial (LGM; Fig. 29E) foi observada a expansão e intensificação da área de adequabilidade de *A. aestiva*, abrangendo o Cerrado quase que completamente, ao passo que *A. ochrocephala* teve uma desconexão da sua área de adequabilidade da de *A. aestiva* na região onde havia contato constante (Bolívia).

Durante o HS1 (Fig. 29F) houve a redução das áreas de adequabilidade para as duas espécies. A área para *A. aestiva* foi novamente fragmentada e reduzida, principalmente, ao sul da sua distribuição. Já para *A. ochrocephala*, a área foi reduzida à região de margem da Amazônia e se expandiu para o Cerrado, invadindo novamente a área de adequabilidade de *A. aestiva* pelo Mato Grosso e Mato Grosso do Sul. Esse mesmo padrão foi pouco alterado durante o o BA (Fig. 29G) e o YDS (Fig. 29H), períodos de condições climáticas bastante semelhantes ao HIS (Fordham *et al.*, 2017). Já durante a transição climática entre o EH+MH (Fig. 29I e 29J) e o Presente+LH (Fig. 29K e 29L) foi mais nítida a diferença e gradiente de mudança das áreas de adequabilidade. O modelo para o EH indicou uma nova expansão para a área de adequabilidade de ambas as espécies, principalmente a de *A. ochrocephala*, padrão que se repetiu em MH e LH, indicando o gradiente de aumento de temperatura e umidade dessas épocas, transicionando entre o LGM e a atualidade.

O modelo para o tempo presente (Fig. 29L) evidenciou maior contato entre áreas de alta adequabilidade ambiental para ambas as espécies na região da Bolívia, com maior adequabilidade para *A. aestiva* em áreas secas, mais frias e abertas (Cerrado, Chaco e Caatinga) do que para *A. ochrocephala*, que são mais úmidas, quentes e de vegetação mais alta e fechada (majoritariamente região circum-amazônica). A adequabilidade ambiental de *A. ochrocephala* para a Caatinga pode estar relacionada a condições de temperaturas mais altas, porém, a grande discrepância entre índices de

umidade dessa região com a região amazônica nos fez considerar essa área como uma comissão do modelo, apesar de ser o mais adequado com base nos parâmetros de seleção, tendo em vista que não há registros de ocorrência dessa espécie no bioma.

V. Discussão

V.1. Análises genéticas e populacionais

Utilizando a matriz *Amazona* spp., os resultados apresentados pela PCA (Fig. 12), SNMFs (Fig. 17), DAPC (Fig. 24), árvore de UPGMA (Fig. 28) e MSN (Fig. 29) evidenciaram, de maneira muito congruente entre si, que a maior diferença genética encontrada é entre o grupo da América Central e da América do Sul. Essas informações são bastante congruentes com o padrão biogeográfico das espécies e com a proposta de que o processo de colonização do continente americano se iniciou com a entrada do ancestral dessas espécies no norte da América do Sul, vindo das pequenas Antilhas, seguiu duas rotas migracionais, uma em direção para a América Central e outro para o sul da América do Sul (Urantówka *et al.*, 2014).

Também foi possível observar que *A. panamensis* não deveria ser considerada como uma subespécie de *A. ochrocephala*, mas uma espécie distinta das demais, com maior parentesco com as espécies da América Central (Fig. 28). Essa ideia foi apresentada inicialmente no estudo de Eberhard e Bermingham (2004). Além disso, observamos que *A. panamensis* apresenta considerável compartilhamento de ancestralidade com *A. ochrocephala*, o que pode estar associado a algum fluxo gênico entre essas espécies na região da Colômbia. Entretanto, para que isso seja investigado é necessário obter amostras de *A. ochrocephala* da região do extremo norte da América do Sul, às quais infelizmente não tivemos acesso. Pesquisas que utilizem mais amostras de *A. panamensis* e amostras de *A. ochrocephala* dessa região poderão analisar se a ancestralidade compartilhada dessa espécie se deve à fluxo gênico ou sorteamento incompleto de linhagens.

Tratando exclusivamente das espécies da América do Sul (*A. aestiva* e *A. ochrocephala*), foi possível identificar a clara separação genética entre as espécies (Fig. 13, Fig. 18 e Fig. 25). Essa separação só identificada com o uso de Sequenciamento de Nova Geração, que abrangeu milhares de regiões do genoma e captou a diferença existente mesmo com o alto fluxo gênico. Os SNMFs (Fig. 18) evidenciaram que essas espécies apresentam um nível elevado de introgressão, o que possivelmente deve ter impedido que estudos anteriores com menor número de marcadores pudessem ter identificado monofiletismo recíproco dessas espécies (Eberhard e Bermingham, 2004; Ribas *et al.*, 2007; Caparroz *et al.*, 2009; Urantówka *et al.*, 2014). Além disso, foi encontrado algum nível de diferenciação genética entre as espécies (F_{ST} : 0.07; Tabela 6).

Analisando a figura que evidencia de maneira mais clara os padrões de fluxo gênico entre as espécies (Fig. 21), é possível perceber que esse fluxo gênico se mostrou muito mais intenso na região de simpatria, como esperado no início do estudo. Entretanto, não esperávamos menor fluxo gênico na região do Mato Grosso do que na Bolívia. Nessa última região, onde esse fluxo foi mais intenso, os indivíduos apresentavam ancestralidade em um padrão que pode ser considerado como equitativo (cerca de 50% *A. aestiva* e 50% *A. ochrocephala nattereri*). Adicionalmente, foi possível identificar que, embora muito menor do que o fluxo gênico citado acima, há algum compartilhamento de ancestralidade entre *A. ochrocephala xantholaema* e indivíduos de *A. aestiva* da região mais norte da distribuição (MT+DF+PI) e um indivíduo da Argentina e outro do MS (Fig. 21). Todavia, esse compartilhamento de ancestralidade é bem menor do que o existente entre os indivíduos de *A. aestiva* da região de simpatria e *A. o. nattereri*. O maior nível de fluxo gênico entre essas espécies na região de simpatria já era esperado. Todavia, o fluxo gênico entre as subespécies de *A. ochrocephala* do oeste da distribuição (*A. o. ochrocephala* e *A. o. xantholaema*) e *A. aestiva* foi inesperado e evidencia que, no passado, pode ter havido contato entre essas espécies na região noroeste da distribuição de *A. aestiva*. O que indica que, ao oposto do que se imaginava até então, o fluxo gênico entre essas duas espécies não se dá apenas na região de simpatria da Bolívia (*A. o. nattereri* e *A. aestiva* BOL+MT), mas também pode ter ocorrido no limite noroeste da distribuição de *A. aestiva* (*A. o. ochrocephala* + *A. o. xantholaema* com *A. aestiva* MT+DF+PI).

A matriz de *A. aestiva* revelou quatro populações distintas (Fig. 22). Uma das populações se apresentou distinta devido à alta quantidade de alelos de *A. o. nattereri* (indivíduos de *A. aestiva* da Bolívia). As demais foram compostas por: 1) norte, PI; 2) centro, DF+MT; e 3) sul, ARG+MS). Essa estrutura tripopulacional pode evidenciar o padrão genético real ou ser um erro da análise devido à pequena quantidade de amostras do nordeste do Brasil (2 indivíduos) e/ou fluxo gênico mais elevado entre *A. aestiva* do PI e *A. o. ochrocephala* + *A. o. xantholaema*, o que poderia ter acarretado em uma atribuição como população distinta para essas duas amostras no SNMF. Seria interessante incluir mais amostras de *A. aestiva* do nordeste em análises futuras para identificar se há essa diferenciação genética.

Considerando $K = 2$ para a matriz de SNPs de *A. ochrocephala* foi observada diferenciação entre as amostras do leste e oeste da distribuição (Fig. 23). Essa diferenciação também se reflete no alto valor de F_{ST} de 0.13. No entanto, como não

podemos analisar amostras da região central da distribuição (Colômbia, Venezuela, Peru, Suriname, Guianas e os estados brasileiros do *Amazonas* e Roraima), é necessário interpretar esses resultados com cuidado. Por isso, estudos futuros necessitam incluir amostras dessa região para que seja possível a melhor compreensão da grande diferença genética evidenciada nos resultados aqui apresentados.

Como apresentado na Tabela 6, o valor de heterozigosidades esperada e observada apresentaram pouca variação para as duas espécies, mas foi menor em *A. ochrocephala* (0.2479 esperado e 0.2298 observado para *A. aestiva*; 0.2317 esperado e 0.2085 observado para *A. ochrocephala*)., Adicionalmente, *A. ochrocephala* apresentou índice de endocruzamento (0.1003) maior do que o valor encontrado para *A. aestiva* (0.0732). Assim, esses valores podem indicar que *A. ochrocephala* apresenta um tamanho populacional efetivo menor que o de *A. aestiva*, mas também pode indicar que essa espécie pode estar sofrendo de um processo de declínio populacional, que ainda não foi perceptível em *A. aestiva*, mas que pode se apresentar em um futuro próximo caso o ritmo de tráfico e/ou desmatamento aumente (Seixas e Mourão, 2018).

O valor de F_{ST} global de *A. ochrocephala* (0.0198, Tabela 6) foi seis vezes maior do que o de *A. aestiva* (0.0603, Tabela 6), o que indica maior diferenciação dentro de *A. ochrocephala*. É possível que as subespécies *A. o. nattereri* e *A. o. ochrocephala* + *A. o. xantholaema* possam representar espécies distintas, tendo em vista que o resultado encontrado para o valor de F_{ST} (0.13: PA - AC+RO+BOL e 0.13: PA - AC+RO+BOL) foi semelhante à outros valores encontrados em literatura para espécies distintas (Luna *et al.*, 2023). Todavia, como já explicado anteriormente, isso pode estar relacionado à ausência de amostras do centro da distribuição, o que necessita ser testado para testar essa hipótese. Como não foi possível estimar o valor de p para os valores de F_{ST} utilizando os *softwares* Plink e VCFtools e pelo pacote hierfstat, análises complementares estão em andamento.

V.2. Modelagem de Nicho Ecológico

Os modelos de nicho (Fig. 30) evidenciaram que, para o tempo presente, existe considerável sobreposição das áreas de adequabilidade das duas espécies na região da Bolívia e no nordeste do Brasil. Entretanto, como explicado anteriormente (ver sessão IV.2 dos Resultados), a adequabilidade de *A. ochrocephala* na Caatinga (nordeste do Brasil) deve estar relacionada à uma comissão do modelo utilizado, então desconsideramos essa área de adequabilidade na discussão dos modelos. A área de *A.*

aestiva é representada principalmente pelo Cerrado e Chaco, já a de *A. ochrocephala* consiste na região amazônica e circum-amazônica. Novas análises de MNE com outros modelos serão realizadas para *A. ochrocephala* na tentativa de identificar se esse padrão de adequabilidade na Caatinga se manterá constante.

Os modelos para o passado sugerem que a distribuição das áreas adequadas para cada uma das duas espécies mudou de forma diferente. Foi observado que *A. ochrocephala* tem sua área de adequabilidade intensificada e expandida nos períodos de maior temperatura e umidade, o que é condizente com seu padrão de hábitat, que consiste em ambientes de floresta tropical decídua, arbustos espinhosos altos, floresta de galeria úmida, floresta inundada sazonalmente (várzea), vegetação ribeirinha secundária, manguezais, savana de pinheiros, palmeiras em áreas abertas mais úmidas, terras cultivadas com bosques remanescentes e bosques, sempre em terras baixas, tipicamente abaixo de 300 m, localmente até 500 m na Venezuela e na Colômbia (Juniper e Parr, 1998). Já *A. aestiva* teve sua área de adequabilidade mais intensificada e expandida durante períodos mais frios e secos, o que também é condizente com seu padrão de hábitat, que, por sua vez, consiste em Cerrado, Chaco, savana, palmeiral, mata ciliar e mata subtropical, principalmente em regiões intimamente ligadas a áreas de crescimento antigo com grandes árvores que fornecem cavidades para ninhos (Collar *et al.*, 2020a).

As camadas de tempo apresentam características específicas quanto às suas condições climáticas de temperatura e pluviosidade. O *Pliocene: Marine Isotope Stage M2* (M2), há ~3,3 Ma, é um período interglacial, mas com temperaturas mais quentes e nível do mar mais alto em comparação com o presente, mas em comparação com a camada posterior do Plioceno (MPWP), foi uma época de resfriamento. O *Pliocene: mid-Pliocene warm period* (MPWP), há ~3,205 Ma, foi um período de aquecimento global significativo, temperaturas mais altas e nível do mar substancialmente mais elevado, nele houve a expansão de florestas tropicais úmidas. O *Pleistocene: Marine Isotope Stage 19* (MIS19), há ~787 ka, foi um estágio interglacial com clima mais quente e nível do mar mais alto do que no presente, porém com temperaturas inferiores às do Plioceno, padrão que se mantém para todas as demais camadas do Pleistoceno. O *Pleistocene: Last Interglacial*, há ~130 ka foi o máximo do último período interglacial do Pleistoceno, marcado por altas temperaturas e índices de umidade elevados. Esse momento favoreceu a expansão de florestas úmidas sobre as florestas secas na América do Sul. O *Pleistocene: Last Glacial Maximum* (LGM), há ~21 ka, foi o máximo do

último período glacial do Pleistoceno, marcado por temperaturas baixas e pouca umidade, o que favoreceu a expansão das florestas secas e retração das florestas úmidas. O *Pleistocene: Heinrich Stadial 1* (HS1), há 17,0-14,7 ka, foi um estágio intermediário entre o último período glacial e o presente, com temperaturas frias, glaciações relativamente avançadas e níveis baixos do mar. O *Pleistocene: Bølling-Allerød* (BA), há 14,7-12,9 ka, foi um período também de transição entre o último máximo glacial e o presente, teve características mais similares às de um período interglacial, com temperaturas mais quentes e recuo das geleiras e clima mais úmido. O *Pleistocene: Younger Dryas Stadial* (YDS), há 12,9-11,7 ka, representou o retorno de um clima mais frio, com temperaturas semelhantes ao Último Máximo Glacial e nível do mar mais baixo. O *Pleistocene: early-Holocene* (EH), há 11,7-8,326 ka, foi um período com características de um evento interglacial, com temperaturas mais quentes e condições climáticas semelhantes às atuais. O *Pleistocene: mid-Holocene, Northgrippian*, há 8,326-4,2 ka, foi um período com características de um evento interglacial, com temperaturas estáveis e condições climáticas semelhantes às atuais. O *Pleistocene: late-Holocene* (LH), há 4,2-0.3 ka, foi o período modelado mais recente, com exceção do presente, teve condições climáticas semelhantes às atuais, mas possivelmente com variações menores (Brown *et al.*, 2018; Dolan *et al.*, 2018; Fordham *et al.*, 2017; Hill, 2015; Karger *et al.*, 2023; Otto-Bliesner *et al.*, 2006).

Levando em consideração as características climáticas do Plioceno, os dois modelos para aquela época representam um período mais quente (MPWP) e um período mais frio (M2). É perceptível nesses modelos que a adequabilidade de *A. ochrocephala* é aumentada e intensificada por climas mais quentes e úmidos, e que se retrai e diminui de intensidade em períodos comparativamente mais frios e secos. O oposto acontece com *A. aestiva*, que tem sua área aumentada no modelo de M2, favorecido pelas condições climáticas, e retraída em MPWP.

O modelo para o MIS-19 indica que a área de adequabilidade de *A. aestiva* era contínua, ocupando uma região bastante semelhante à atual, compreendendo o complexo Cerrado+Chaco. Essa semelhança pode se dar devido ao fato de que vivemos em um período de interglaciação, semelhante ao enfrentado pelo planeta durante o MIS-19. O outro modelo para uma interglaciação no Pleistoceno foi o LIG, que evidenciou um padrão bastante parecido, com uma importante diferença: a área de adequabilidade para *A. aestiva* foi fragmentada em dois blocos (norte e sul). Talvez isso pode estar relacionado à diferenciação genética norte e sul em *A. aestiva* (Fig. 19 K4). Tal

fragmentação coincide com a expansão da área de adequabilidade de *A. ochrocephala* em direção ao Cerrado. Esse tipo de evento pode não necessariamente significar que *A. ochrocephala* expandiu sua área de distribuição, mas sim que a área, antes adequada para *A. aestiva*, não apresentava mais as características adequadas para sua ocorrência. É possível que, devido ao fato de essas fragmentações da área de adequabilidade serem cíclicas, não houve um processo contínuo de interrupção do fluxo gênico, mas sim vários momentos de ausência e presença de fluxo de maneira intercalada, o que poderia deixar sinais específicos de compartilhamento de alelos no centro da distribuição e ser interpretado como isolamento por distância. Para testarmos essa ideia, estamos trabalhando em modelos de Isolamento Por Distância (*Isolation By Distance* - IBD) que permitirão uma conclusão mais confiável sobre essa questão.

A expansão da área de adequabilidade de *A. aestiva* durante o LGM é bastante evidente no modelo, indicando que sua área de adequabilidade é beneficiada por períodos mais frios e secos. Apesar das pequenas variações entre os subperíodos de transição entre o LGM e o Presente (HS1, BA, YDS, EH, MH e LH), é possível perceber o gradiente de condições ambientais que ocorreu entre o LGM e o presente. Embora mais baixas do que as temperaturas atuais, comparativamente ao LGM, o período que compreende HS1, BA e YDS representam períodos de maior temperatura, o que pode estar relacionado à redução da área de adequabilidade de *A. aestiva*. Nos três períodos seguintes, EH, MH e LH, houve uma maior elevação da temperatura e do índice de pluviosidade, o que pode explicar a expansão da área de adequabilidade de *A. ochrocephala*. Curiosamente, a área de adequabilidade de *A. aestiva* também se expandiu durante esses três últimos períodos, indicando que, algumas condições climáticas, sejam elas relacionadas à temperatura ou pluviosidade, a beneficiaram de alguma maneira. Dessa forma, foi possível identificar que, embora exista a tendência de relação entre a área de adequabilidade de *A. aestiva* e ambientes mais secos e abertos, existem fatores ou combinações de fatores que ainda não compreendemos que também são capazes de afetar positivamente essas áreas mesmo em períodos um pouco quentes e úmidos. Para *A. ochrocephala* a relação entre climas quentes e úmidos com sua área de adequabilidade pareceu mais evidente, já que não houve grande expansão aparente da sua área em períodos mais frios e secos.

VI. Conclusões

Os resultados do presente trabalho indicam que dados obtidos com sequenciamento de nova geração foram informativos para diferenciar geneticamente espécies/complexos de espécies com divergência recente e possível alto fluxo gênico. Foi encontrado que *A. aestiva* e *A. ochrocephala* devem ser espécies irmãs. Além disso, também foi possível observar estrutura genética populacional em ambas as espécies. Foram encontrados 3 grupos em *A. ochrocephala* que correspondem às suas três subespécies. Para *A. aestiva* foi possível identificar até 4 populações que correspondem a: 1) um grupo com alto fluxo gênico com *A. ochrocephala* na região da Bolívia, 2) um grupo mais ao norte da sua distribuição, localizado no Piauí, 3) um grupo da região mais central, que engloba o Mato Grosso e Distrito Federal, e 4) um grupo mais ao sul, que engloba o Mato Grosso do Sul e a Argentina. Os modelos de nicho ecológico evidenciaram as diferenças ambientais relacionadas às áreas de adequabilidade de cada espécie, o que também permitiu concluir que ambas foram afetadas pelas flutuações climáticas do Pleistoceno e do Plioceno, indicando que, em alguns momentos, a área de adequabilidade era fragmentada e unificada de maneira cíclica, o que pode ter criado os padrões de diferenciação genética encontrados atualmente. Análises futuras devem englobar um maior número de indivíduos da região nordeste do Brasil para *A. aestiva* e da região da Colômbia, Venezuela, Suriname, Guianas e dos estados brasileiros do Amazonas e Roraima para *A. ochrocephala* no intuito de testar melhor algumas hipóteses levantadas no presente estudo. Além disso, análises complementares estão sendo realizadas para permitir compreender melhor os resultados aqui apresentados.

VII. Referências

Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., e Wingett, S. (2010). FastQC. *A quality control tool for high throughput sequence data*, 370.

BirdLife International and Handbook of the Birds of the World. (2021). Bird species distribution maps of the world. Version 2021.1. Retrieved from <http://datazone.birdlife.org/species/requestdis>

Brown, J. L., Hill, D. J., Dolan, A. M., Carnaval, A. C., e Haywood, A. M. (2018). PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Scientific Data*, 5, 1-9.

Caparroz, R., Seixas, G. H. F., Berkunsky, I., & Collevatti, R. G. (2009). The role of demography and climatic events in shaping the phylogeography of *Amazona aestiva* (Psittaciformes, Aves) and definition of management units for conservation. *Diversity and distributions*, 15, 459-468.

Chan, K. O., Alexander, A. M., Grismer, L. L., Su, Y. C., Grismer, J. L., Quah, E. S., e Brown, R. M. (2017). Species delimitation with gene flow: a methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. *Molecular Ecology*, 26, 5435-5450.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., e Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, s13742-015-0047-8.

Collar, N., G. M. Kirwan, e P. F. D. Boesman (2020)a. turquoise-fronted parrot (*Amazona aestiva*), version 1.0. In: J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana (Eds.). *Birds of the World*. Cornell Lab of Ornithology, Ithaca. <https://doi.org/10.2173/bow.bufpar.01>

Collar, N., P. F. D. Boesman, and G. M. Kirwan (2020)b. Yellow-crowned Parrot (*Amazona ochrocephala*), version 1.0. In *Birds of the World* (J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.ywcpar.01>

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... e 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156-2158.

Dolan, A. M., Haywood, A. M., Hunter, S. J., Tindall, J. C., Dowsett, H. J., Hill, D. J., e Pickering, S. J. (2015). Modeling the enigmatic late Pliocene glacial event—

Marine Isotope Stage M2. *Global and Planetary Change*, 128, 47-60.

Eaton, D. A. (2021). *ipyrad: interactive assembly and analysis of RADseq datasets*. Retrieved from <https://ipyrad.readthedocs.io/>

Eberhard, J. R., & Bermingham, E. (2004). Phylogeny and biogeography of the Amazona ochrocephala (Aves: Psittacidae) complex. *The Auk*, 121, 318-332.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PlosOne*, 6, e19379.

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, 28, 342-350.

Fordham, D. A., Saltr e, F., Haythorne, S., Wigley, T. M., Otto-Bliesner, B. L., Chan, K. C., & Brook, B. W. (2017). PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography*, 40, 1348-1358.

Forshaw, J. M. (1989). *Parrots of the world*. Princeton Field Guides, Princeton.

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5, 184-186.

Grzegorzczuk, J., Gurgul, A., Oczkowicz, M., Szmatoła, T., Fornal, A., & Bugno-Poniewierska, M. (2021). Single nucleotide polymorphism discovery and genetic differentiation analysis of geese bred in Poland, using Genotyping-by-Sequencing (GBS). *Genes*, 12, 1074.

Gurgul, A., Miksza-Cybulska, A., Szmatoła, T., Semik-Gurgul, E., Jasielczuk, I., Bugno-Poniewierska, M., ... & Kajtoch, Ł. (2019). Evaluation of genotyping by sequencing for population genetics of sibling and hybridizing birds: An example using Syrian and great spotted woodpeckers. *Journal of Ornithology*, 160, 287-294.

Hill, D. J. (2015). The non-analogue nature of Pliocene temperature gradients. *Earth and Planetary Science Letters*, 425, 232-241.

Jackson, N. D., Carstens, B. C., Morales, A. E., & O'Meara, B. C. (2017). Species delimitation with gene flow. *Systematic Biology*, 66, 799-812.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27, 3070-3071.

Juniper, T., and M. Parr (1998). *Parrots: A Guide to the Parrots of the World*. Pica Press, Robertsbridge.

Karger, D. N., Nobis, M. P., Normand, S., Graham, C. H., & Zimmermann, N. E. (2023). CHELSA-TraCE21k–high-resolution (1 km) downscaled transient temperature

and precipitation data since the Last Glacial Maximum. *Climate of the Past*, 19, 439-456.

Leaché, A. D., Harris, R. B., Rannala, B., e Yang, Z. (2014). The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology*, 63, 17- 30.

Linck, E., & Battey, C. (2018). Minor allele frequency thresholds dramatically affect population structure inference with genomic datasets. *Biorxiv*, 188623.

Luna, L. W., Naka, L. N., Thom, G., Knowles, L. L., Sawakuchi, A. O., Aleixo, A., e Ribas, C. C. (2023). Late Pleistocene landscape changes and habitat specialization as promoters of population genomic divergence in Amazonian floodplain birds. *Molecular Ecology*, 32, 214-228.

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., e Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22, 2841.

Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H., Hu, A., e CAPE Last Interglacial Project members. (2006). Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science*, 311, 1751-1753.

Paradis, E., e Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526-528. doi:10.1093/bioinformatics/bty633.

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., e Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, 40(7), 887-893.

Ribas, C. C., Tavares, E. S., Yoshihara, C., & Miyaki, C. Y. (2007). Phylogeny and biogeography of Yellow-headed and Blue-fronted Parrots (*Amazona ochrocephala* and *Amazona aestiva*) with special reference to the South American taxa. *Ibis*, 149, 564-574.

Sato, F. M. (2016). *Análise filogeográfica multilocos do complexo Amazona aestiva/A. ochrocephala (Aves, Psittacidae)*. Universidade de São Paulo, São Paulo.

Seixas, G. H. F., e Mourão, G. (2018). Communal roosts of the blue-fronted amazons (*Amazona aestiva*) in a large tropical wetland: Are they of different types?. *PlosOne*, 13, e0204824.

Sousa, V., e Hey, J. (2013). Understanding the origin of species with genome-scale data: modeling gene flow. *Nature Reviews Genetics*, 14, 404-414.

Sullivan, J., Demboski, J. R., Bell, K. C., Hird, S., Sarver, B., Reid, N., e Good, J. M. (2014). Divergence with gene flow within the recent chipmunk radiation (*Tamias*). *Heredity*, 113, 185-194.

Urantowka, A. D., Mackiewicz, P., & Strzala, T. (2014). Phylogeny of *Amazona barbadensis* and the Yellow-headed Amazon complex (Aves: Psittacidae): a new look at South American parrot evolution. *PlosOne*, 9, e97228.

Winkler, D. W., S. M. Billerman, and I. J. Lovette (2020). New World and African Parrots (Psittacidae), version 1.0. In *Birds of the World* (S. M. Billerman, B. K. Keeney, P. G. Rodewald, and T. S. Schulenberg, Editors). Cornell Lab of Ornithology, Ithaca. <https://doi.org/10.2173/bow.psitta3.01>

Zarza, E., Faircloth, B. C., Tsai, W. L., Bryson Jr, R. W., Klicka, J., e McCormack, J. E. (2016). Hidden histories of gene flow in highland birds revealed with genomic markers. *Molecular Ecology*, 25, 5144-5157.

VIII. Figuras e Tabelas

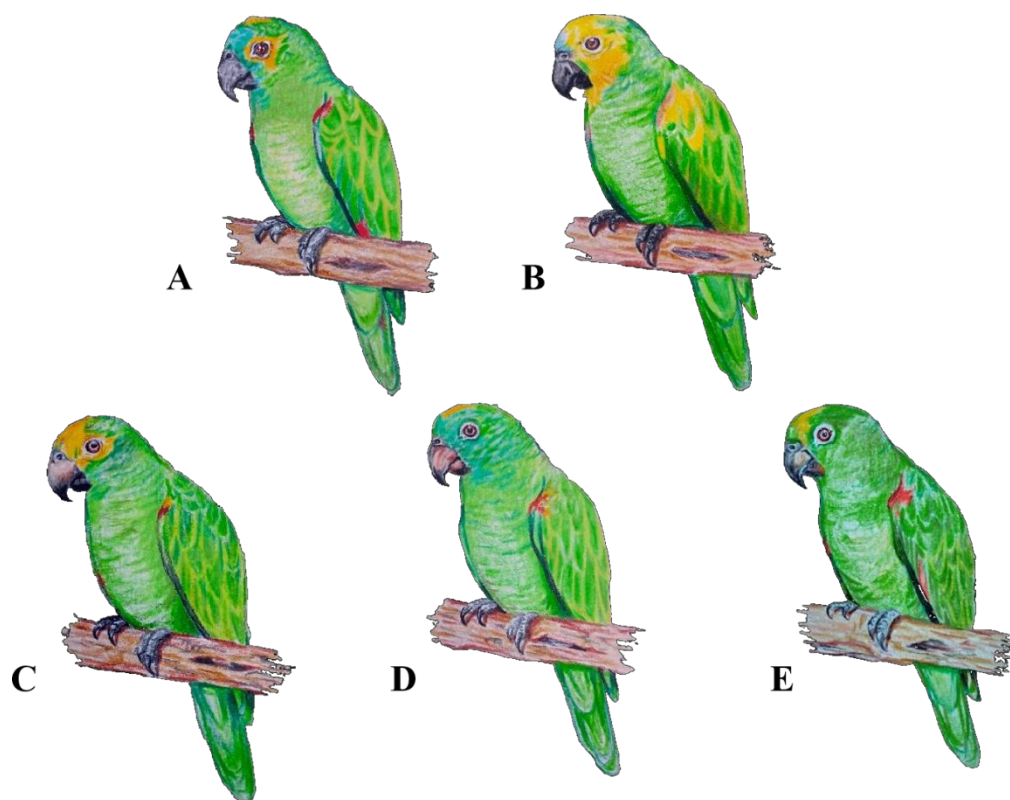


Figura 1 - Espécies e subespécies estudadas. A - *Amazona aestiva aestiva*; B - *Amazona aestiva xanthopteryx*; C - *Amazona ochrocephala xantholaema*; D - *Amazona ochrocephala nattereri*; E - *Amazona ochrocephala ochrocephala*. Ilustrações por Igor Ferreira.

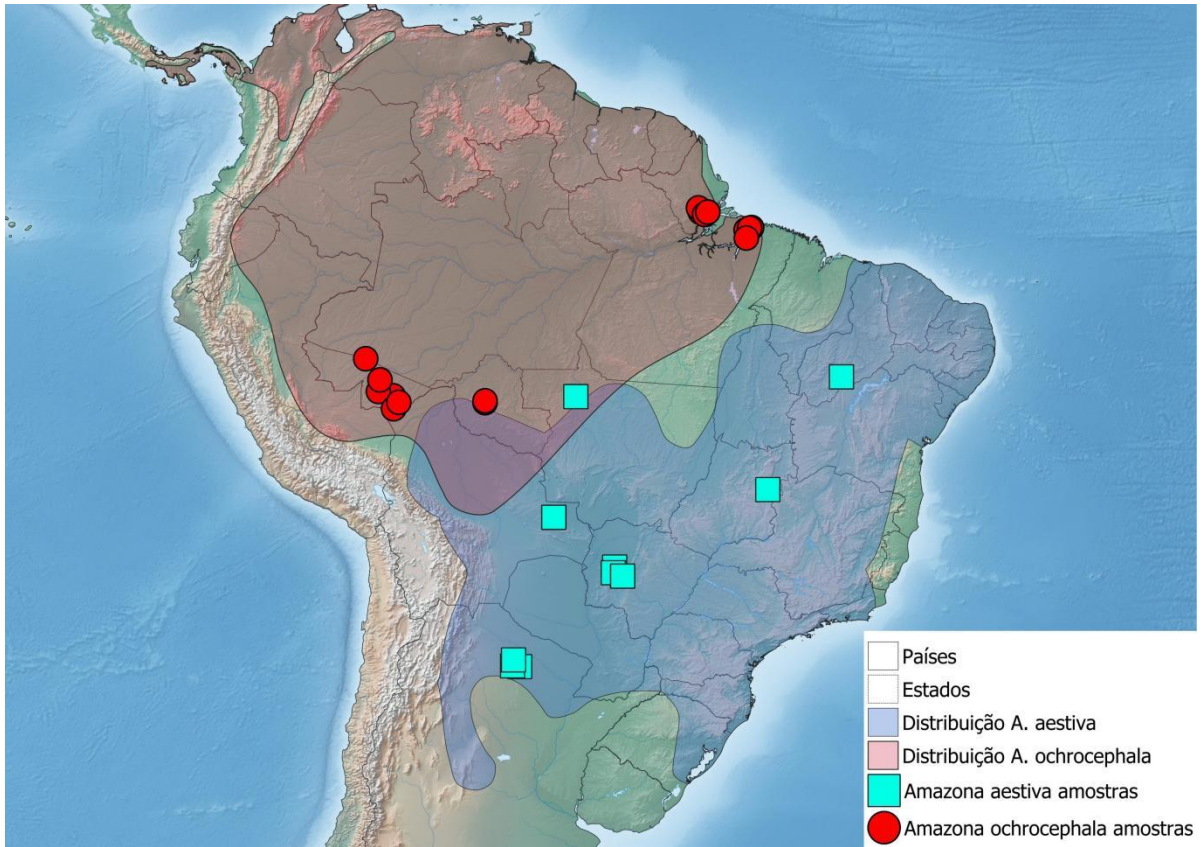


Figura 2 - Localidades de coleta das amostras utilizadas no estudo. *Shapes* da distribuição geográfica de *A. aestiva* e *A. ochrocephala* obtidos no BirdLife (BirdLife, 2021). Quadrados azuis indicam as localidades aproximadas das amostras de *A. aestiva* e círculos vermelhos indicam as localidade aproximadas das amostras de *A. ochrocephala*.

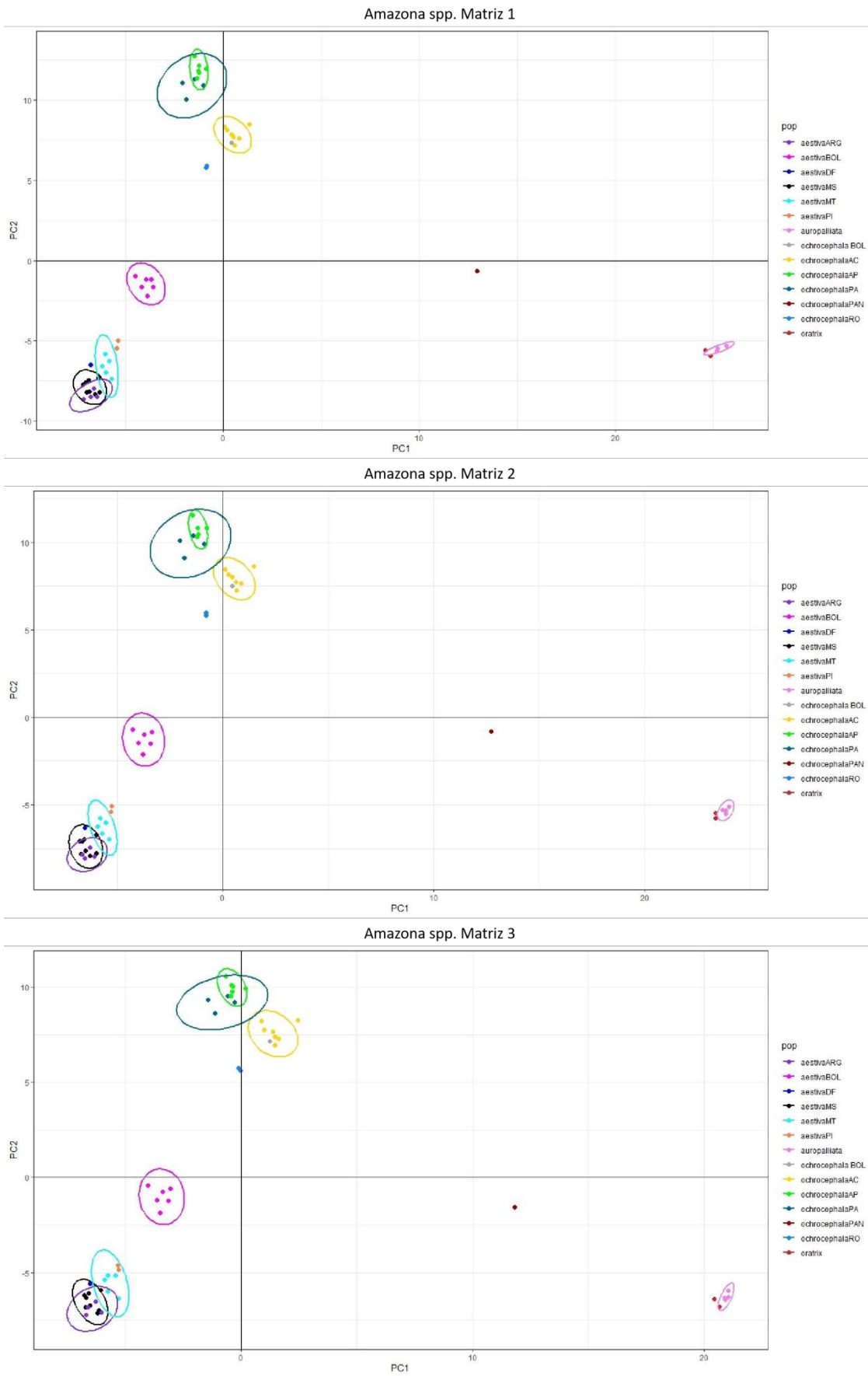
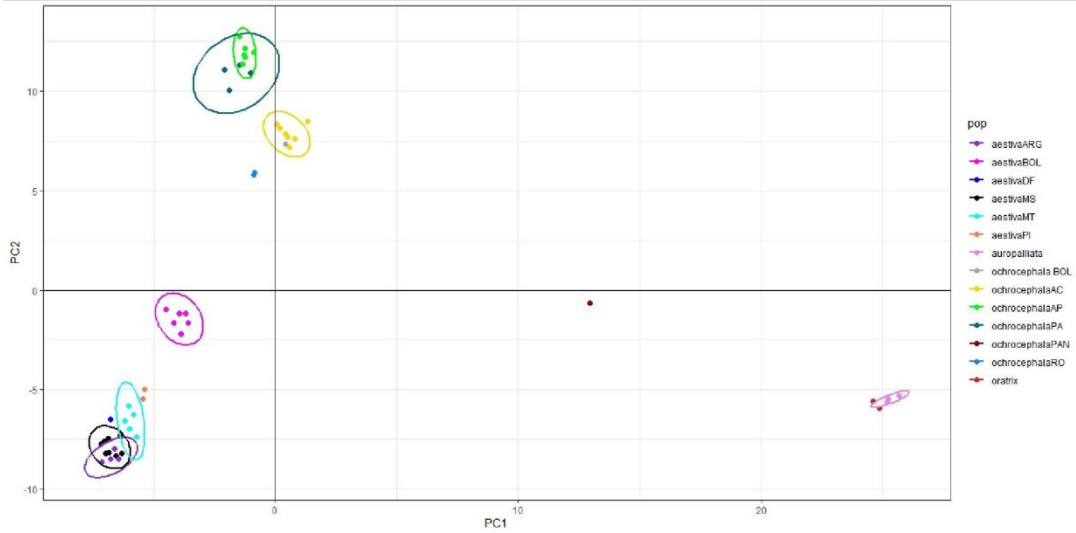
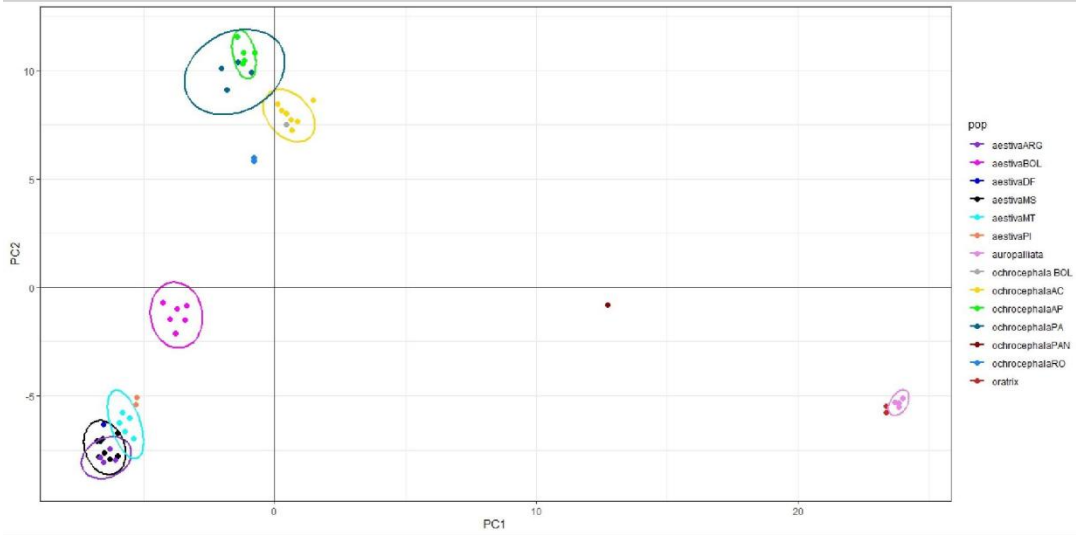


Figura 3 - PCAs das matrizes teste de *Amazona* spp.. Pontos indicam indivíduos; cores indicadas ao lado direito das PCAs indicam localidades e seguem o padrão da Tabela 1.

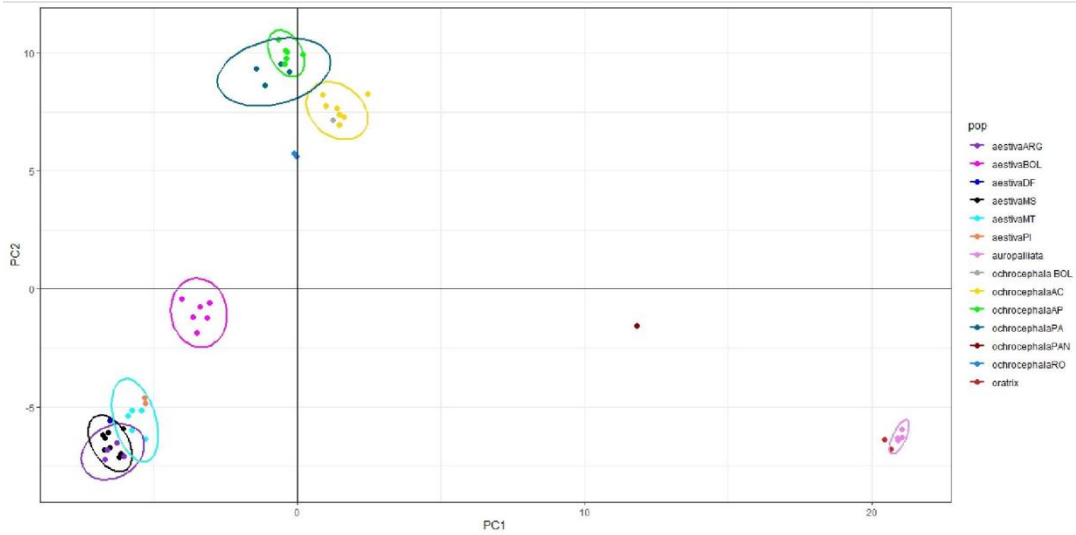
Amazona spp. Matriz 4



Amazona spp. Matriz 5

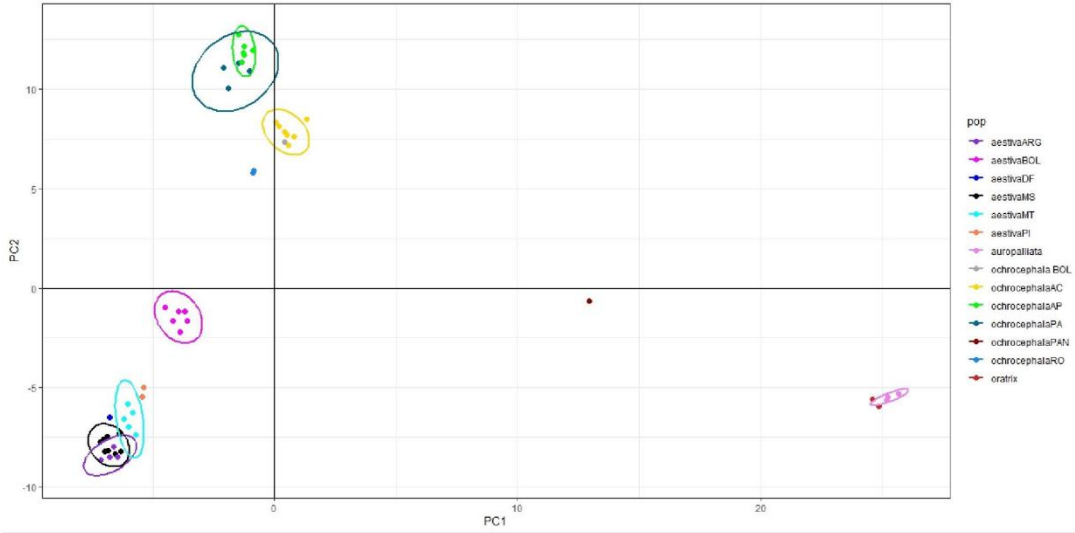


Amazona spp. Matriz 6

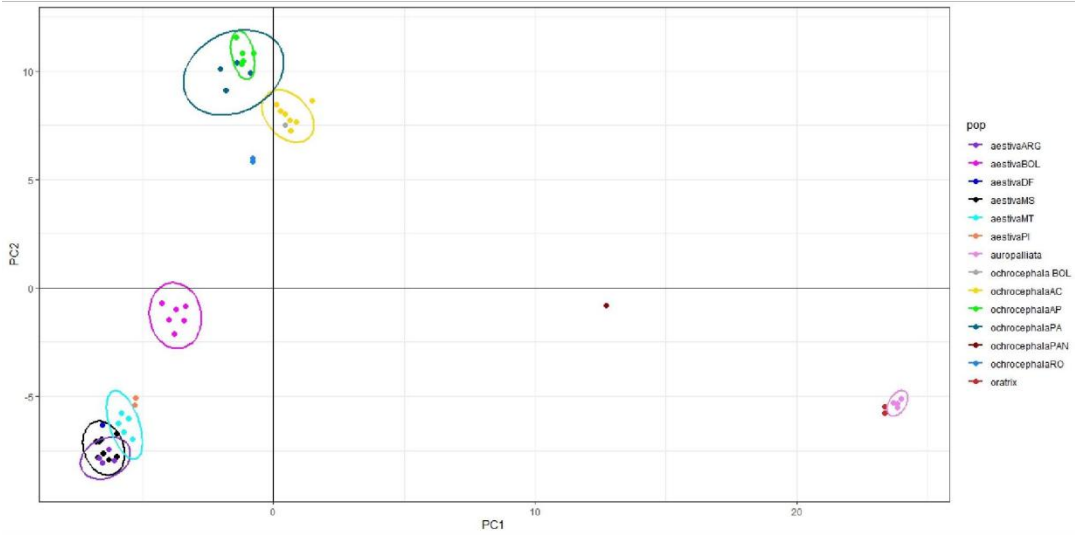


Continuação da Figura 3 - PCAs das matrizes teste de *Amazona* spp..

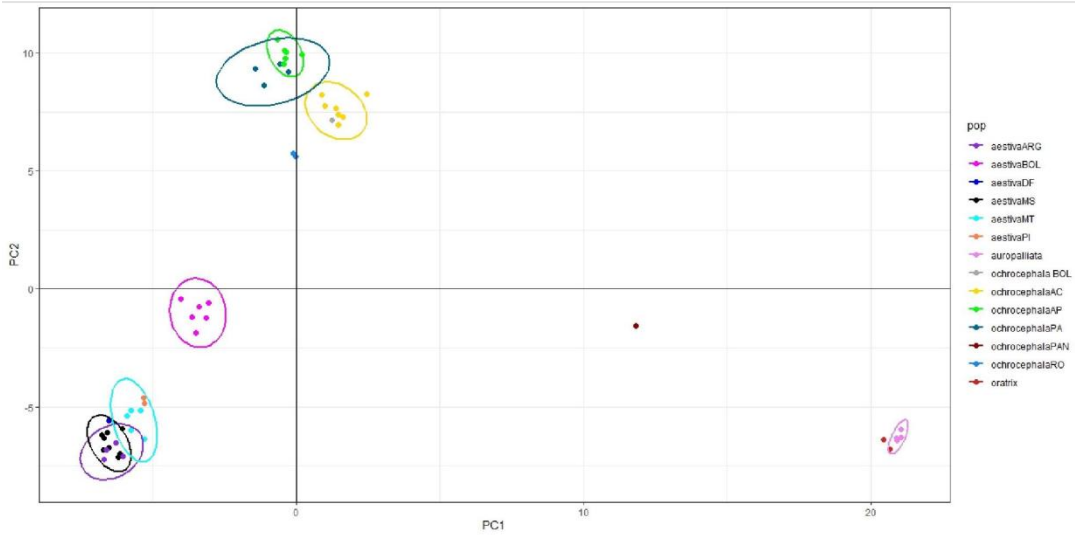
Amazona spp. Matriz 7



Amazona spp. Matriz 8



Amazona spp. Matriz 9



Continuação da Figura 3 - PCAs das matrizes teste de *Amazona* spp..

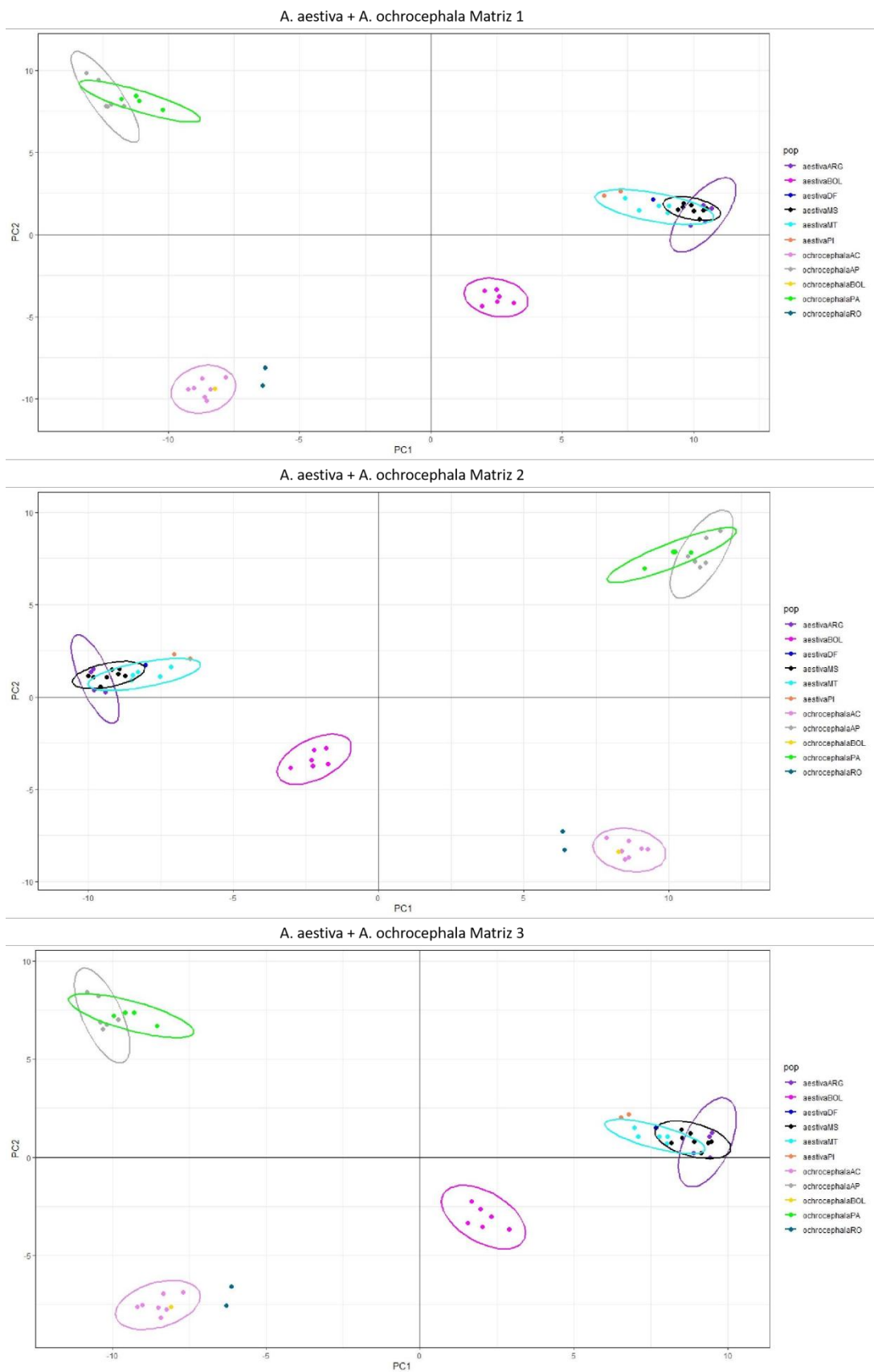
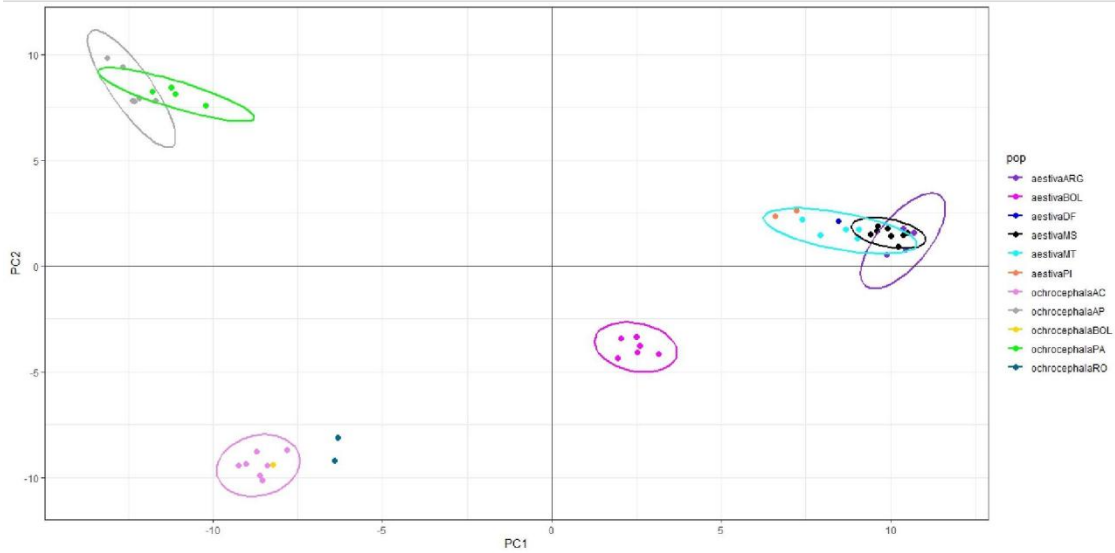
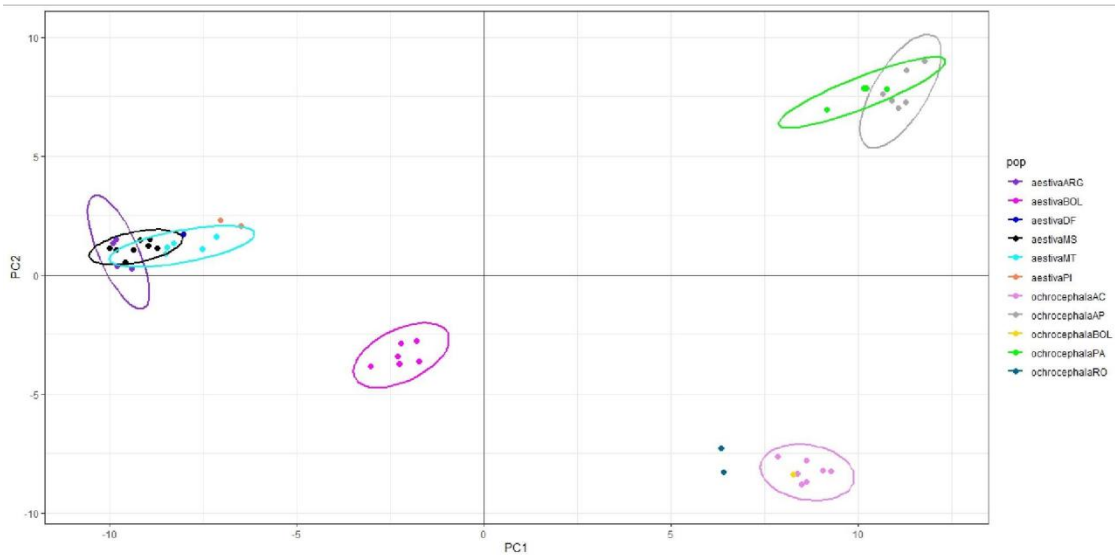


Figura 4 - PCAs das matrizes teste de *A. aestiva* + *A. ochrocephala*.. Pontos indicam indivíduos; cores indicadas ao lado direito das PCAs indicam localidades e seguem o padrão da Tabela 1.

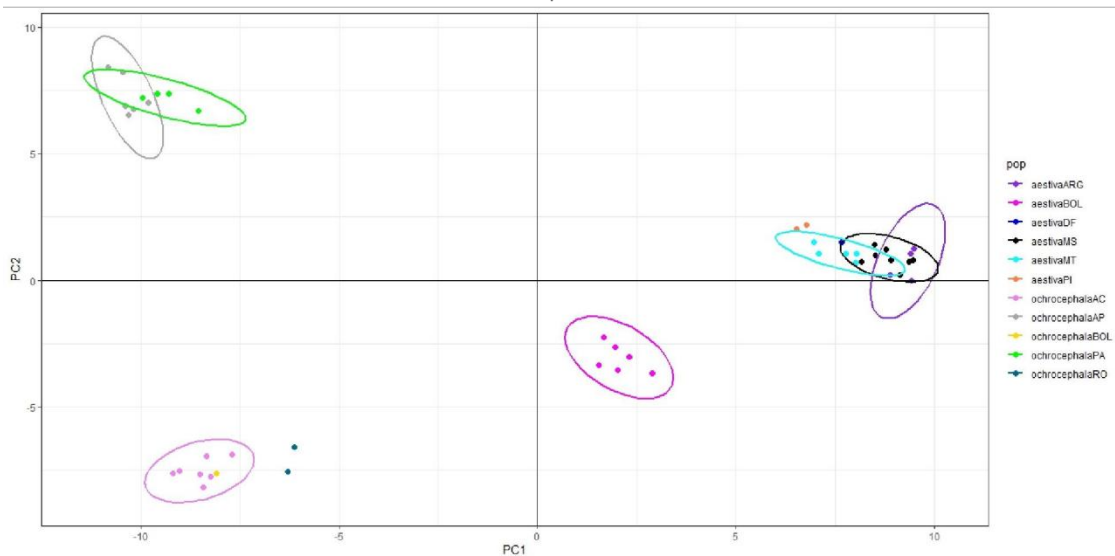
A. aestiva + A. ochrocephala Matriz 4



A. aestiva + A. ochrocephala Matriz 5

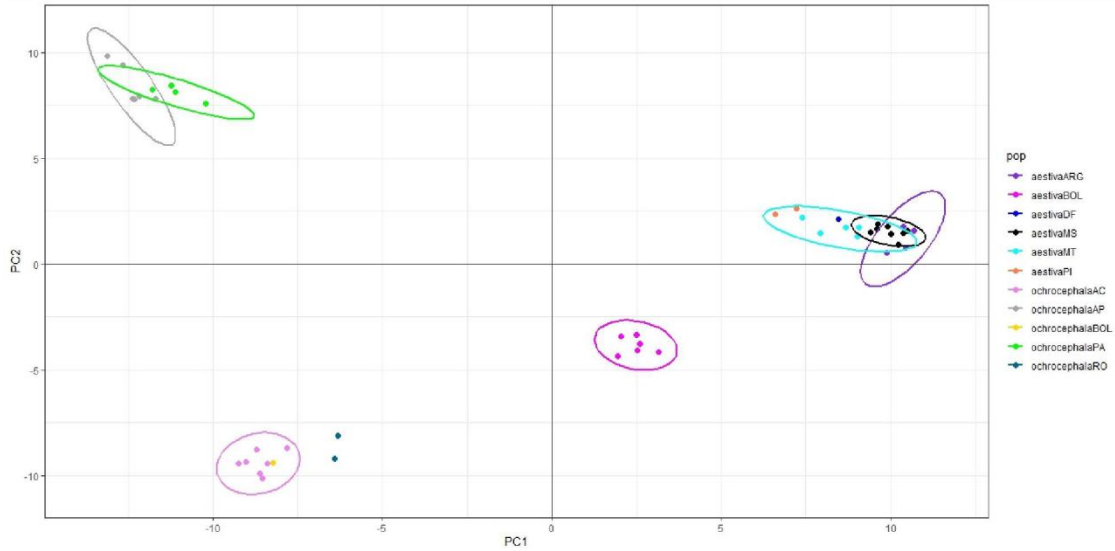


A. aestiva + A. ochrocephala Matriz 6

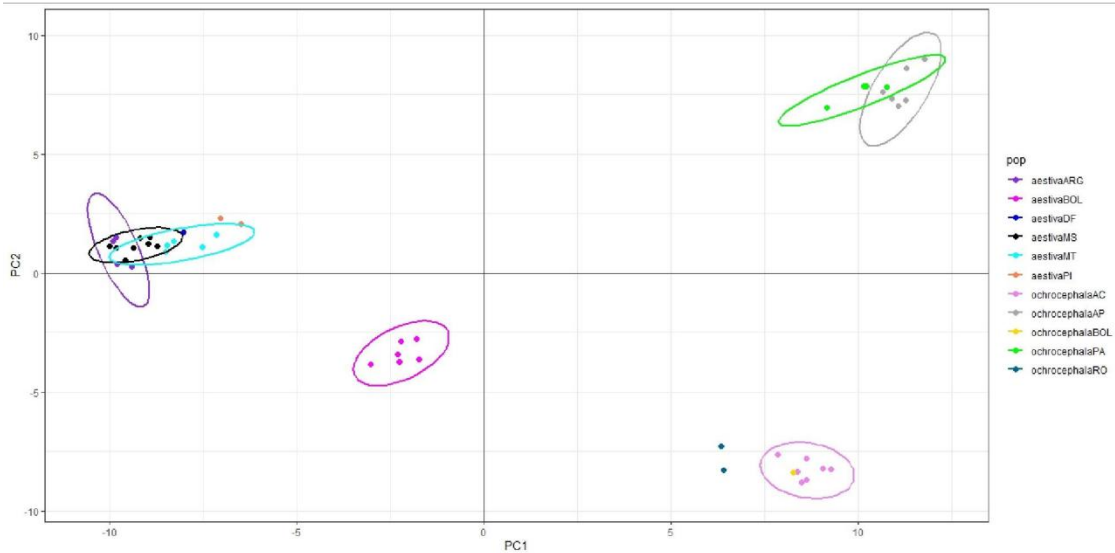


Continuação da Figura 4 - PCAs das matrizes teste de *A. aestiva* + *A. ochrocephala*..

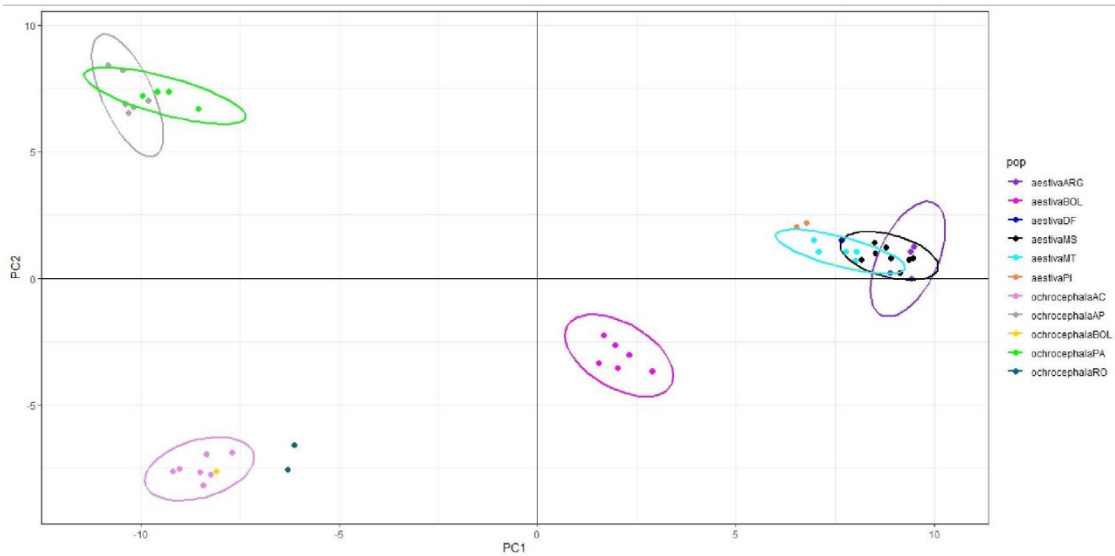
A. aestiva + A. ochrocephala Matriz 7



A. aestiva + A. ochrocephala Matriz 8



A. aestiva + A. ochrocephala Matriz 9



Continuação da Figura 4 - PCAs das matrizes teste de *A. aestiva* + *A. ochrocephala*..

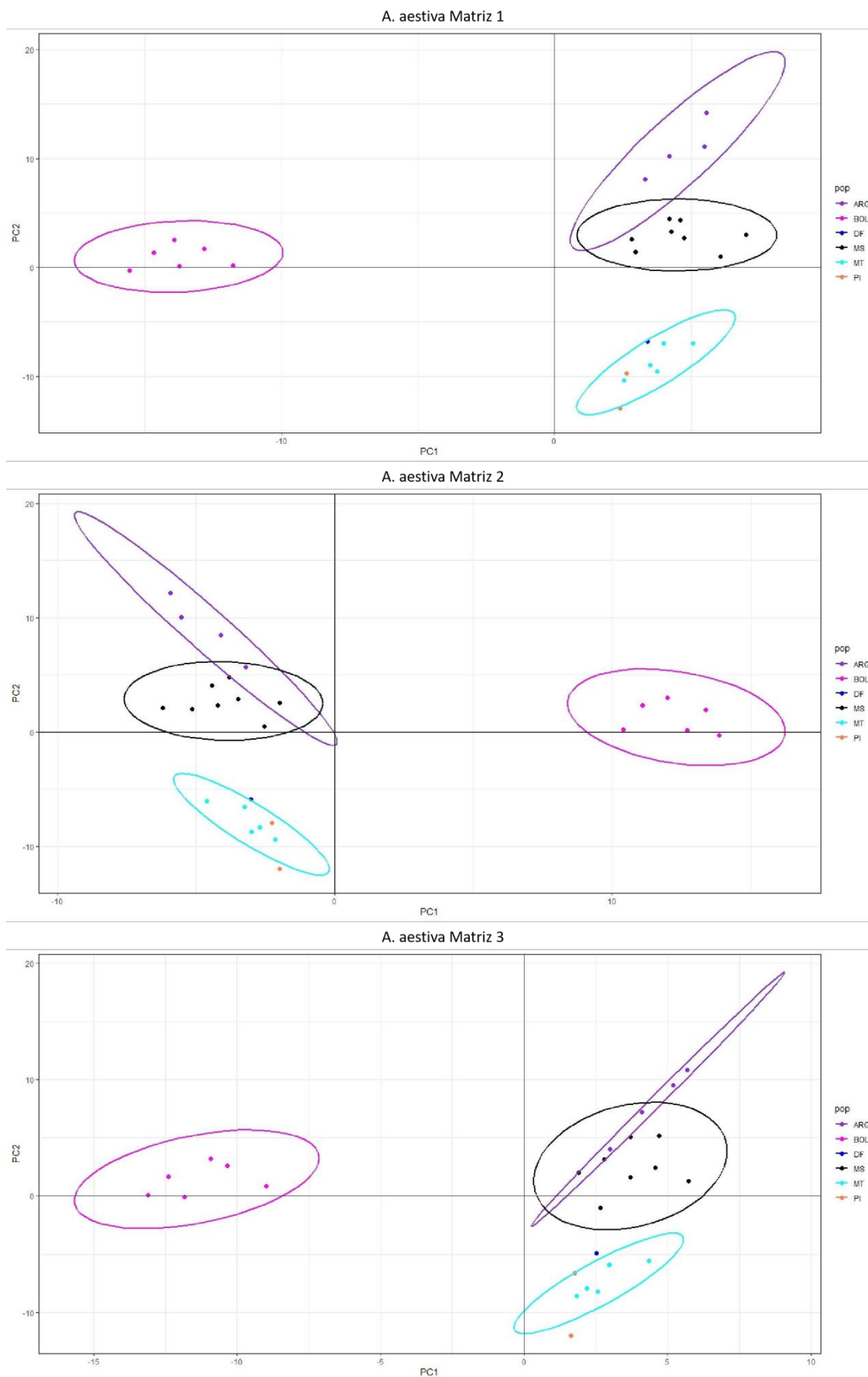
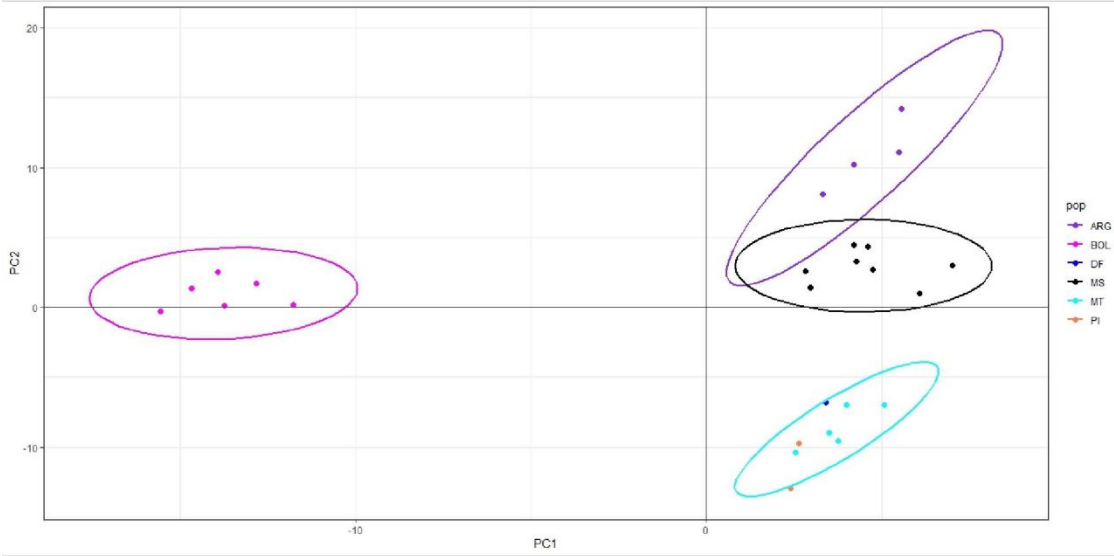
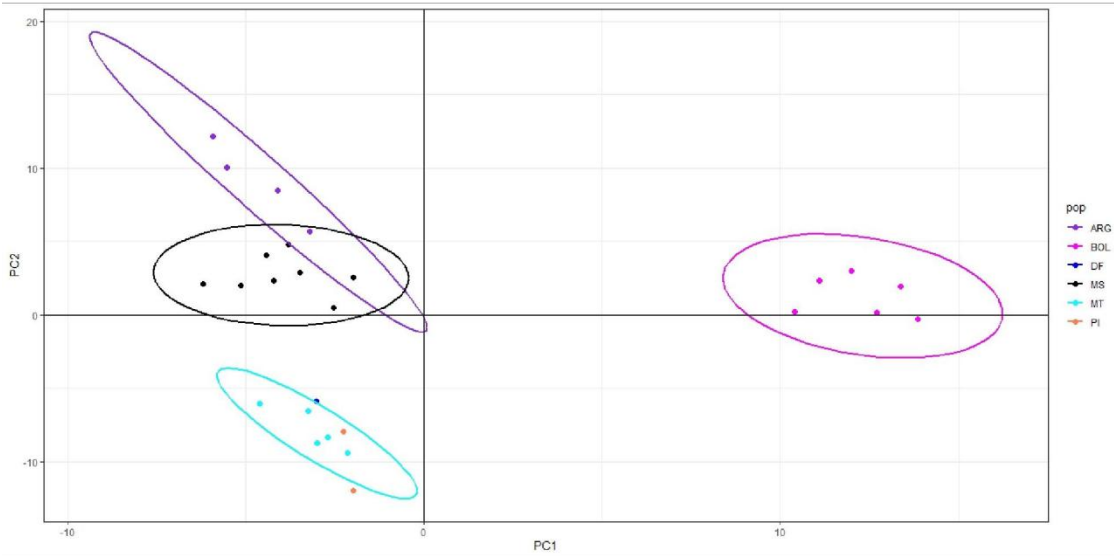


Figura 5 - PCAs das matrizes teste de *A. aestiva*. Pontos indicam indivíduos; cores indicadas ao lado direito das PCAs indicam localidades e seguem o padrão da Tabela 1.

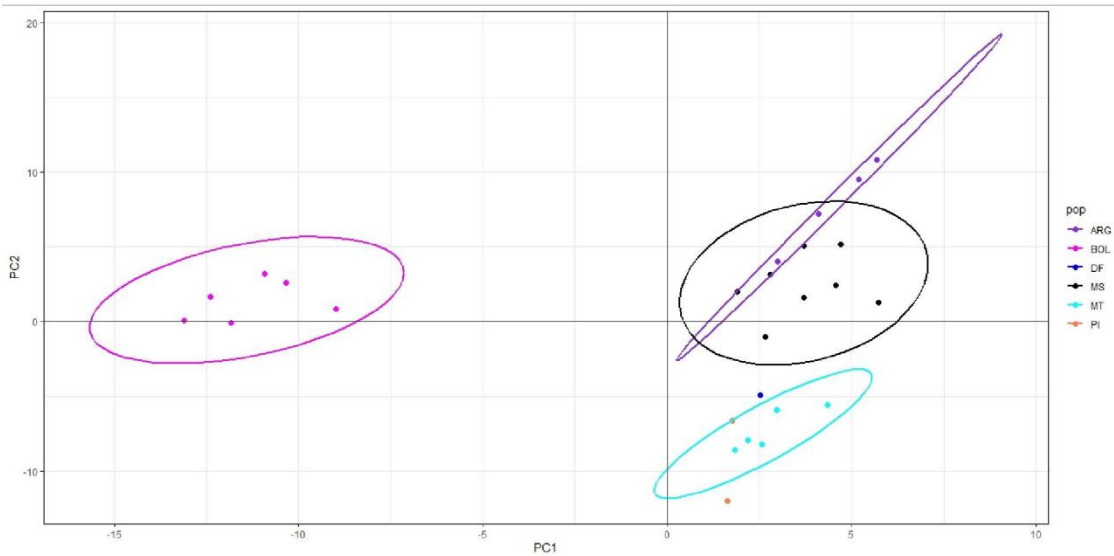
A. aestiva Matriz 4



A. aestiva Matriz 5

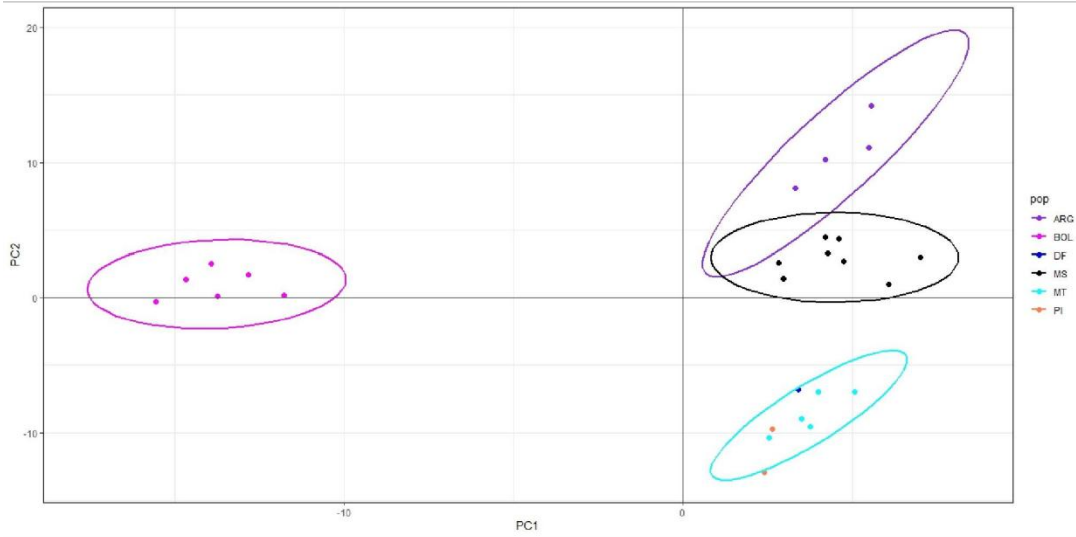


A. aestiva Matriz 6

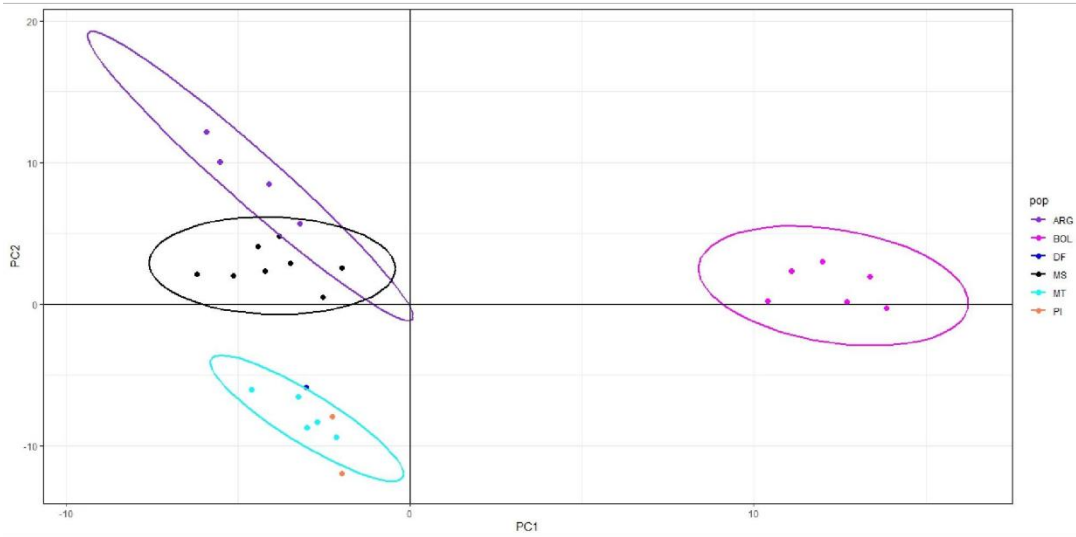


Continuação da Figura 5 - PCAs das matrizes teste de *A. aestiva*.

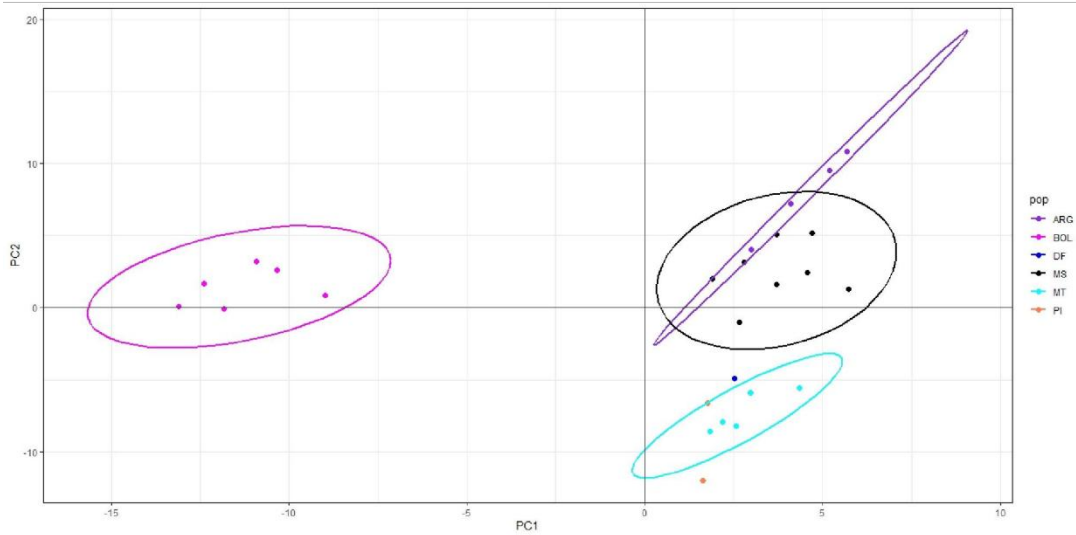
A. aestiva Matriz 7



A. aestiva Matriz 8



A. aestiva Matriz 9



Continuação da Figura 5 - PCAs das matrizes teste de *A. aestiva*.

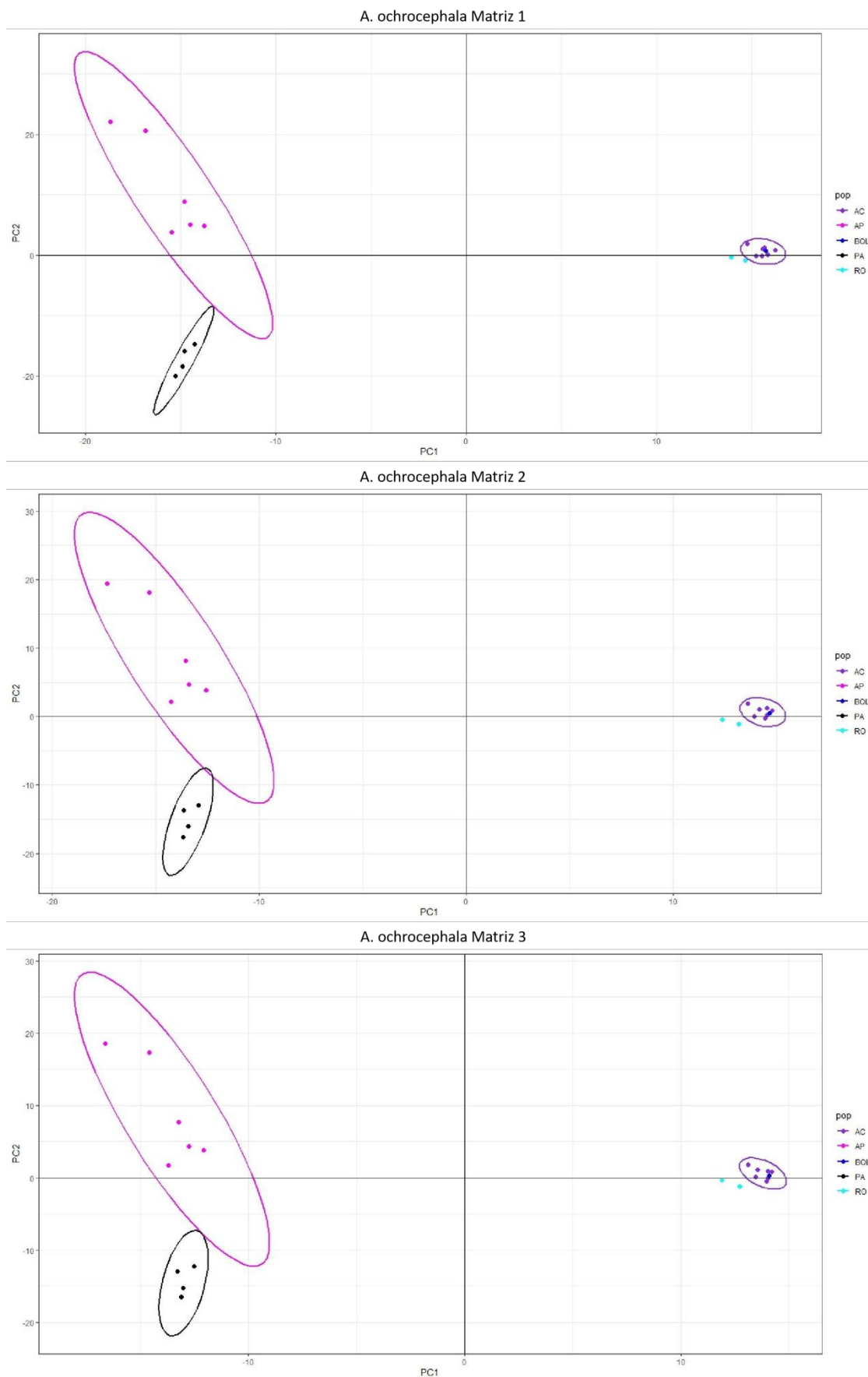
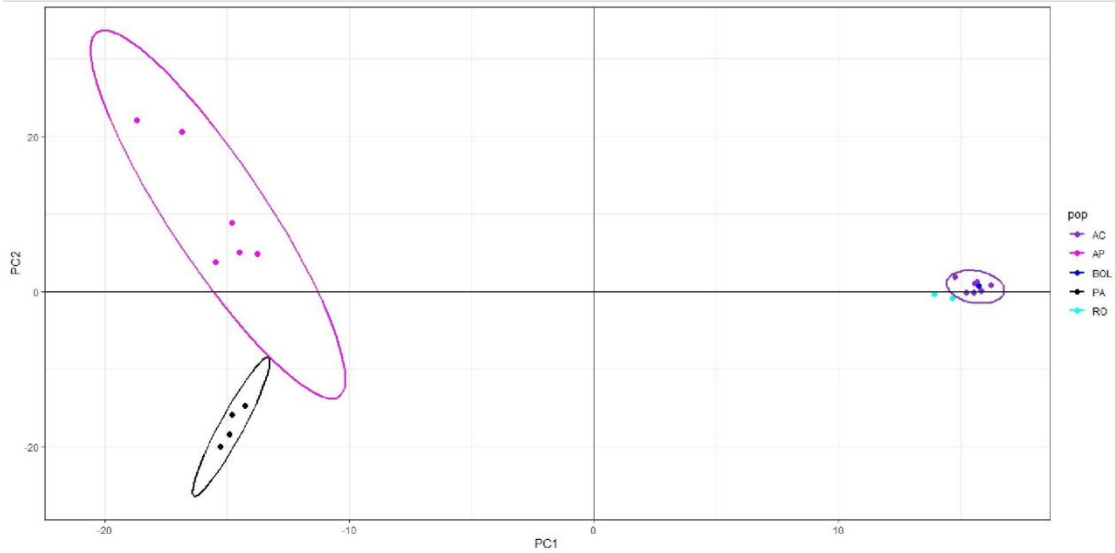
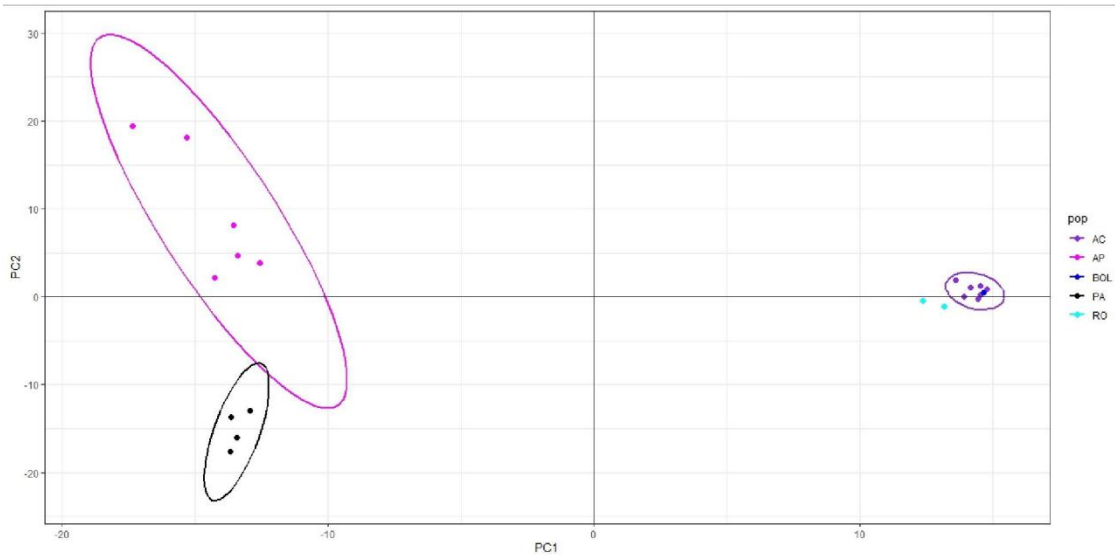


Figura 6 - PCAs das matrizes teste de *A. ochrocephala*. Pontos indicam indivíduos; cores indicadas ao lado direito das PCAs indicam localidades e seguem o padrão da Tabela 1.

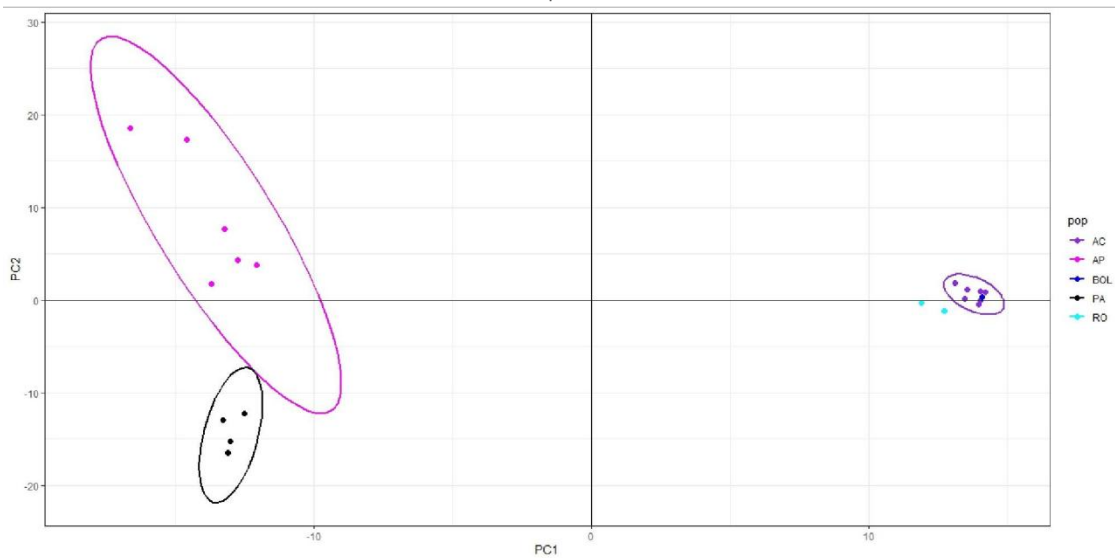
A. ochrocephala Matriz 4



A. ochrocephala Matriz 5

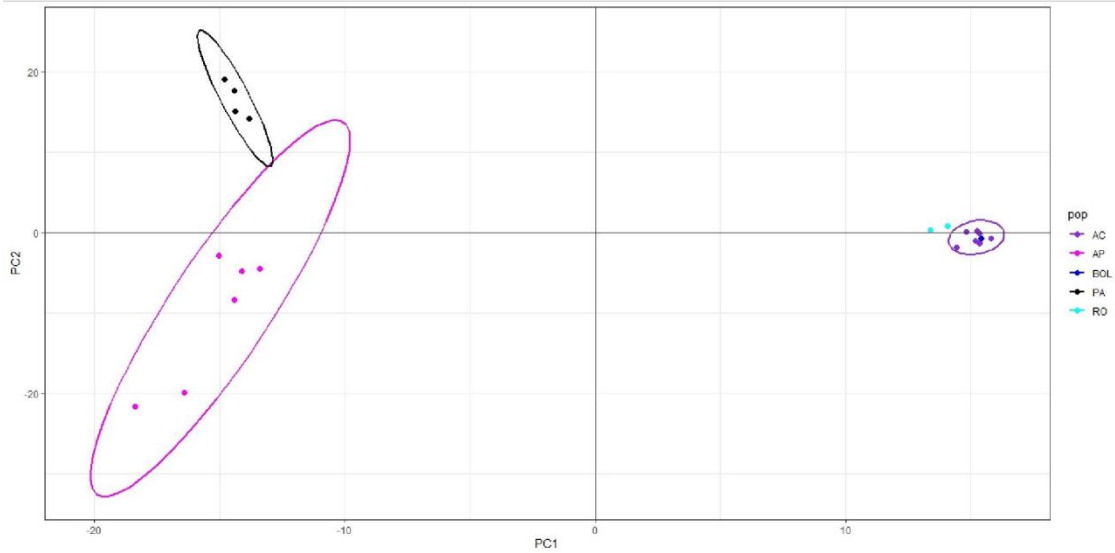


A. ochrocephala Matriz 6

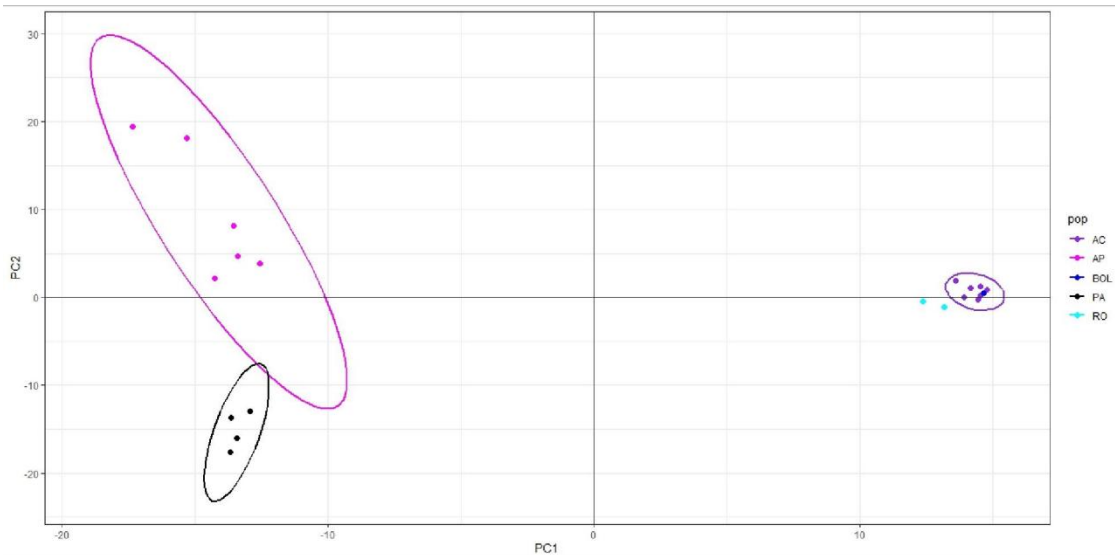


Continuação da Figura 6 - PCAs das matrizes teste de *A. ochrocephala*.

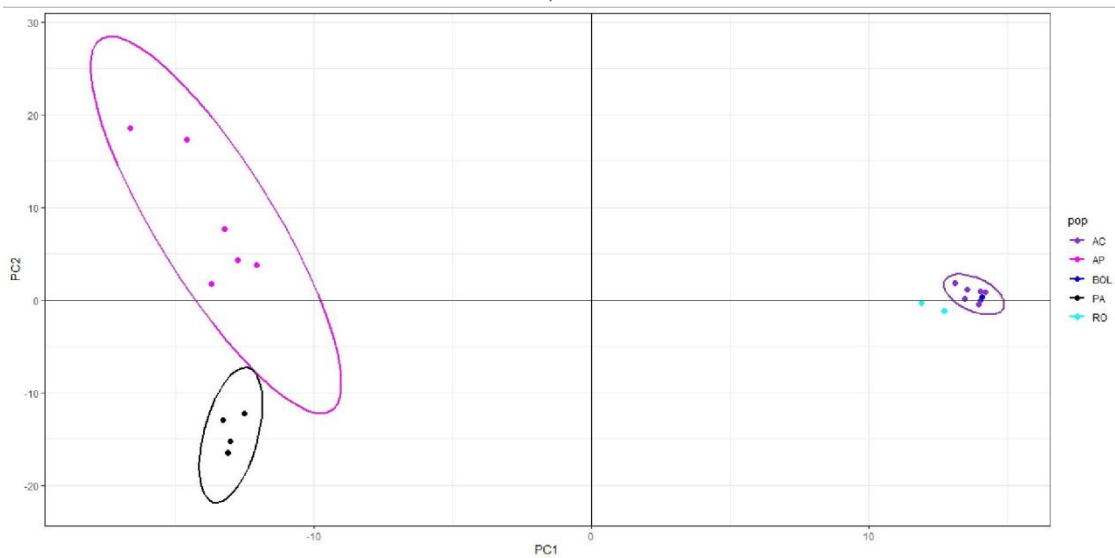
A. ochrocephala Matriz 7



A. ochrocephala Matriz 8



A. ochrocephala Matriz 9



Continuação da Figura 6 - PCAs das matrizes teste de *A. ochrocephala*.

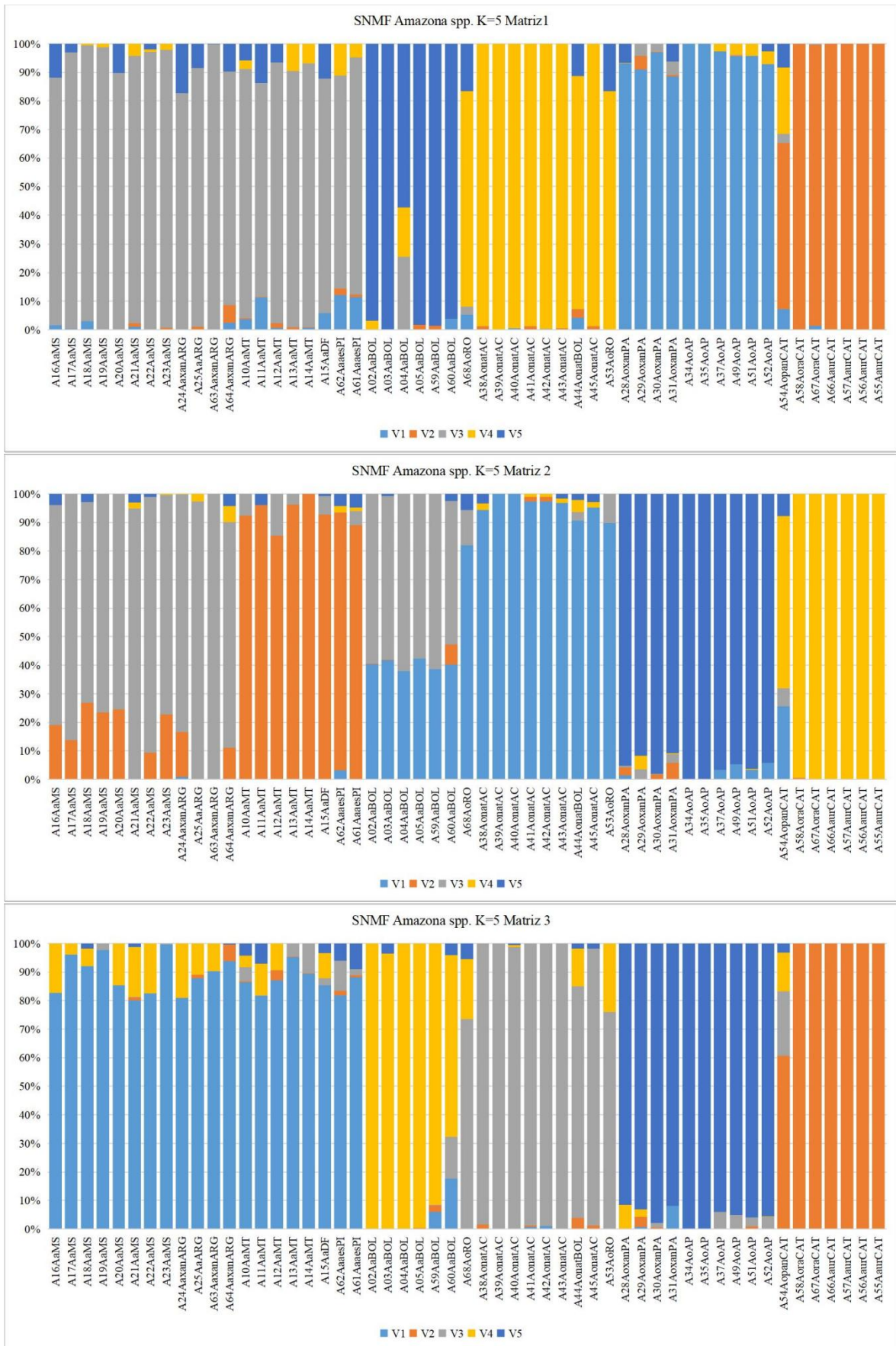
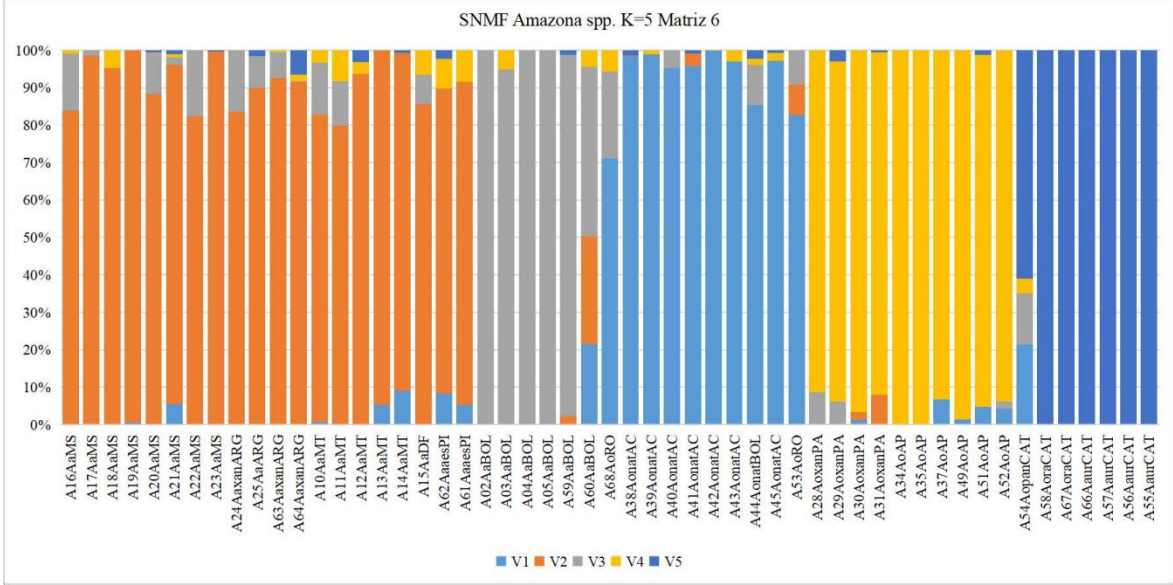
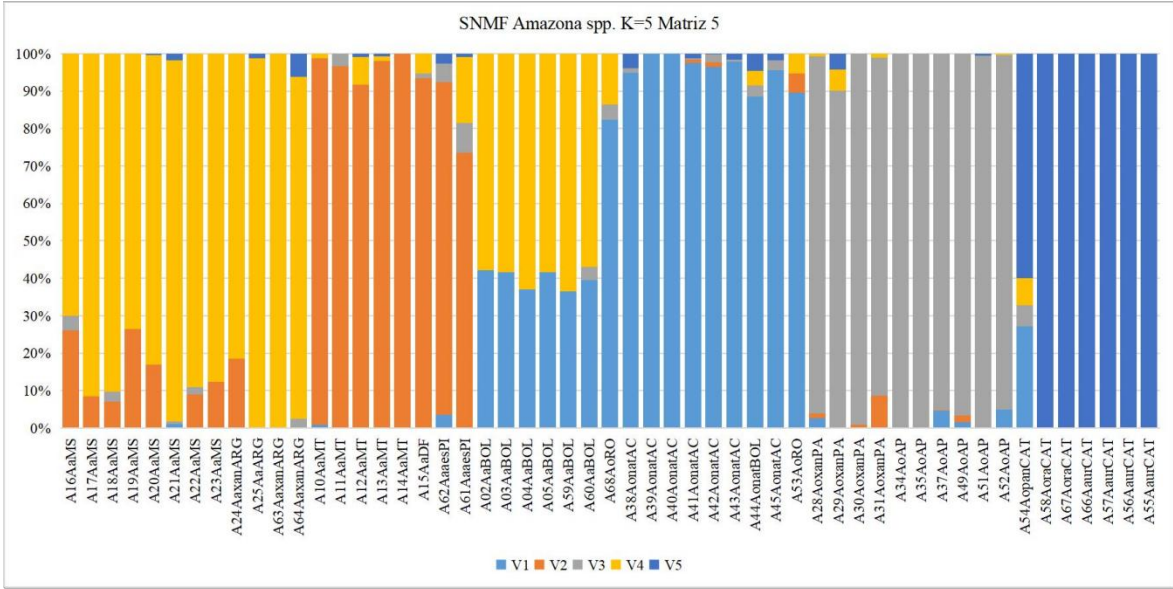
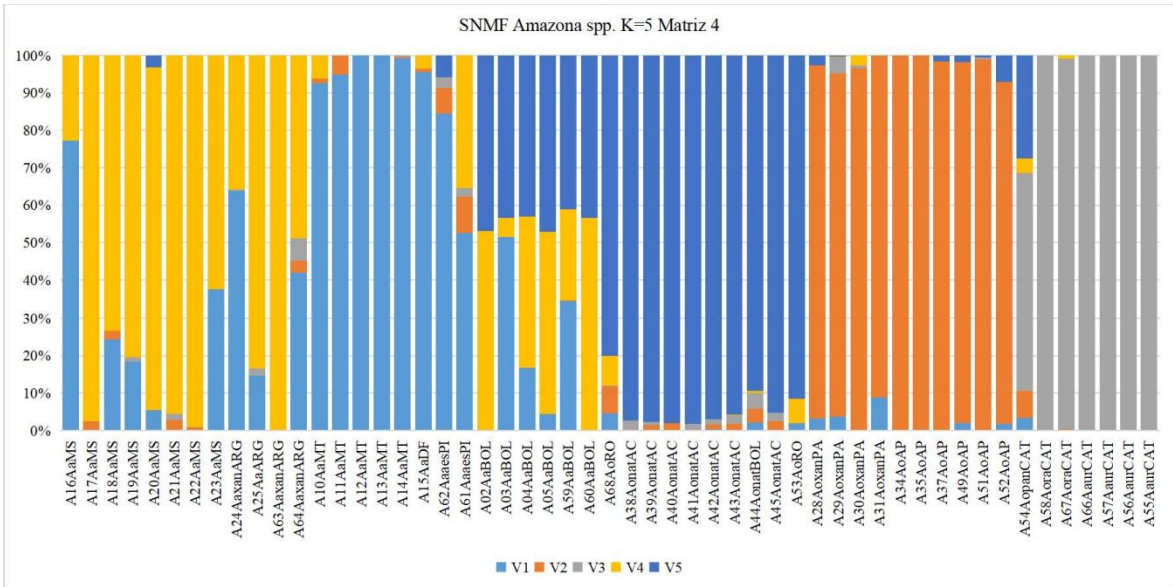
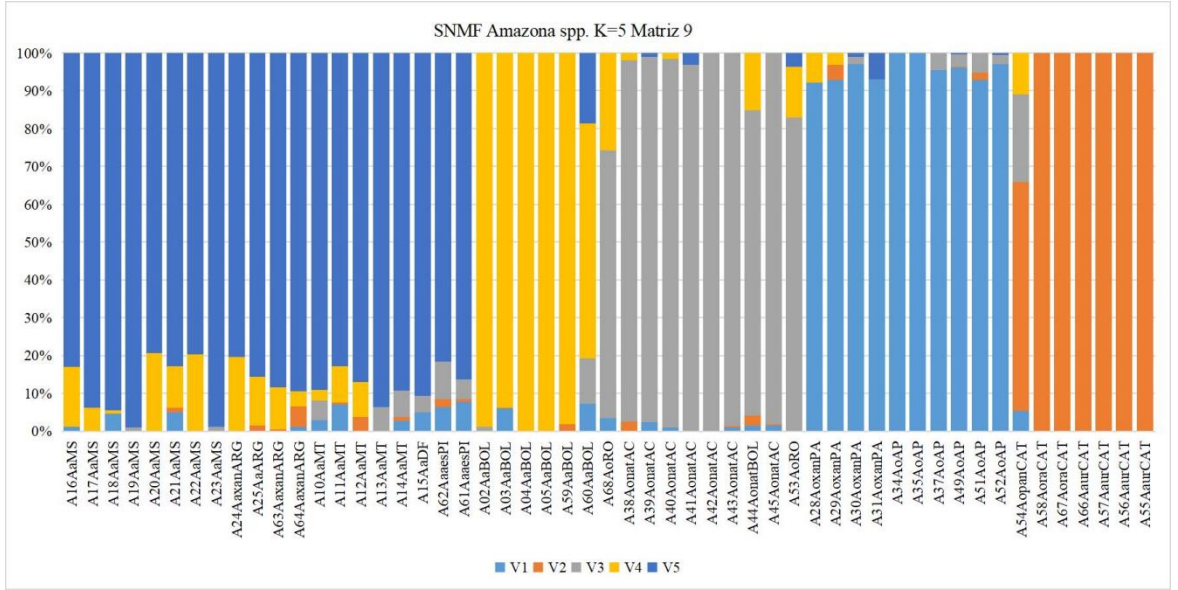
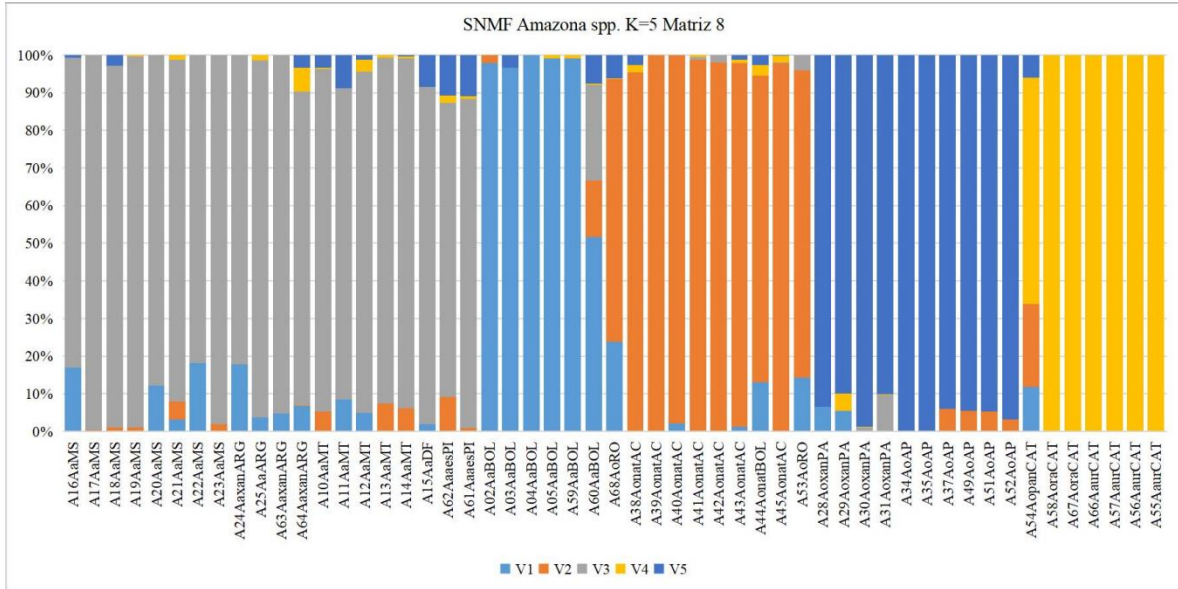
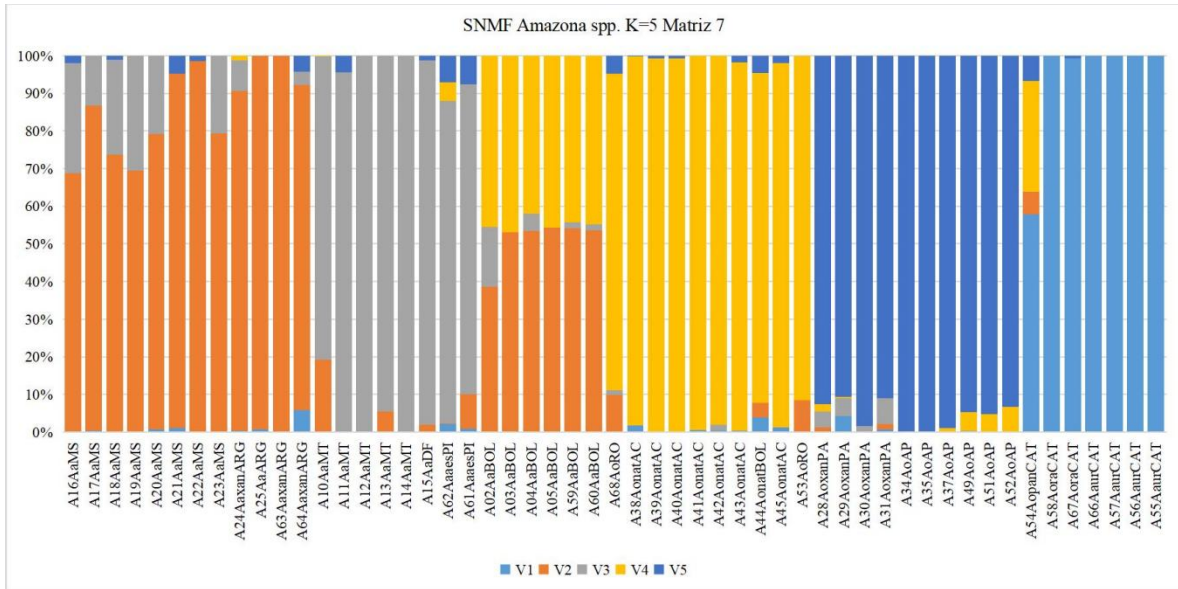


Figura 7 - SNMFs das matrizes teste de *Amazona* spp.. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 7 - SNMFs das matrizes teste de *Amazona* spp.



Continuação da Figura 7 - SNMFs das matrizes teste de *Amazona* spp.

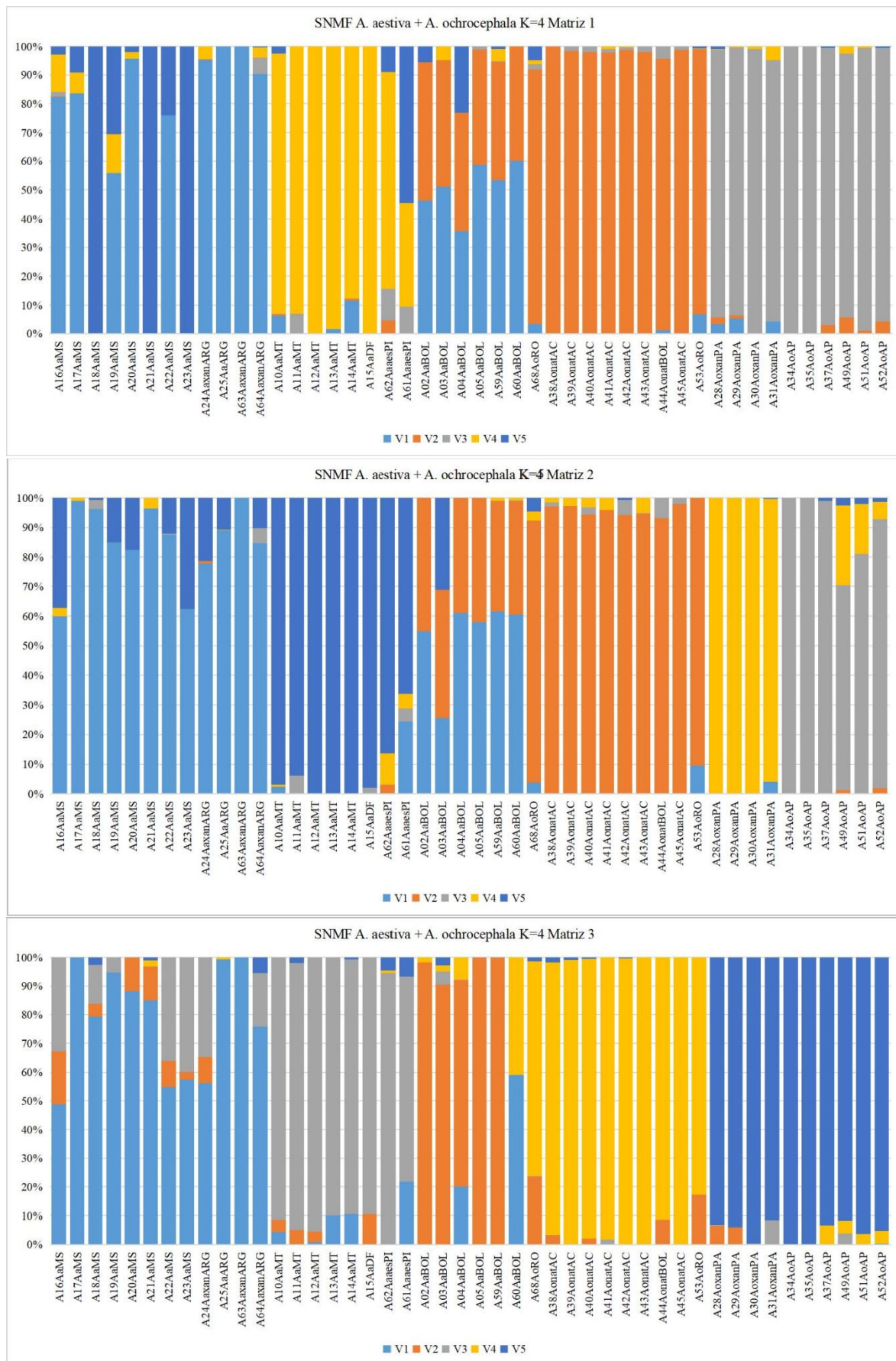
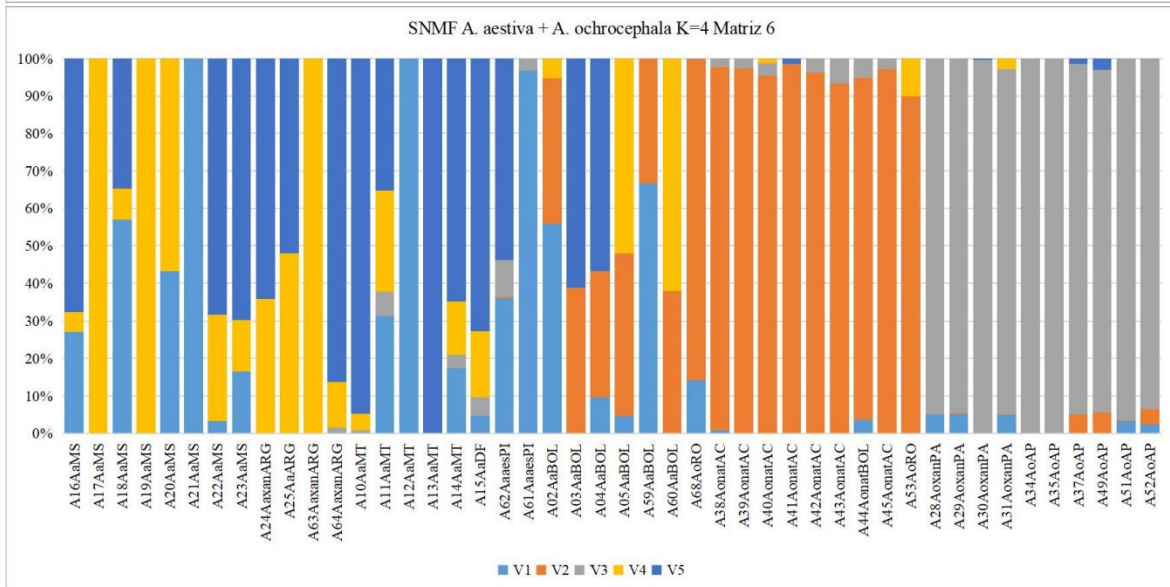
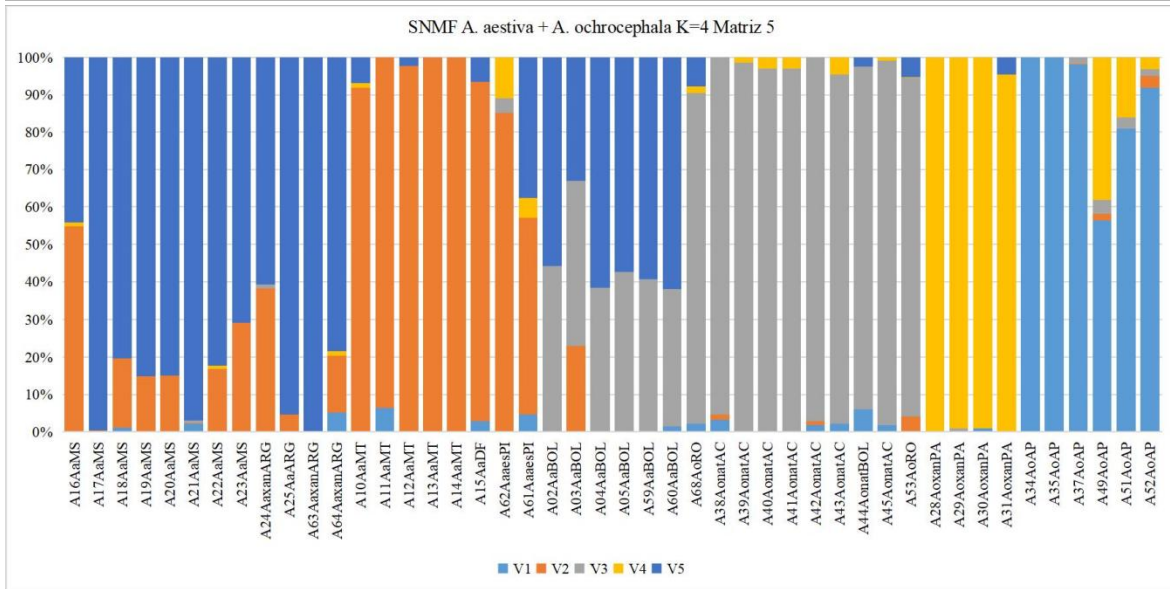
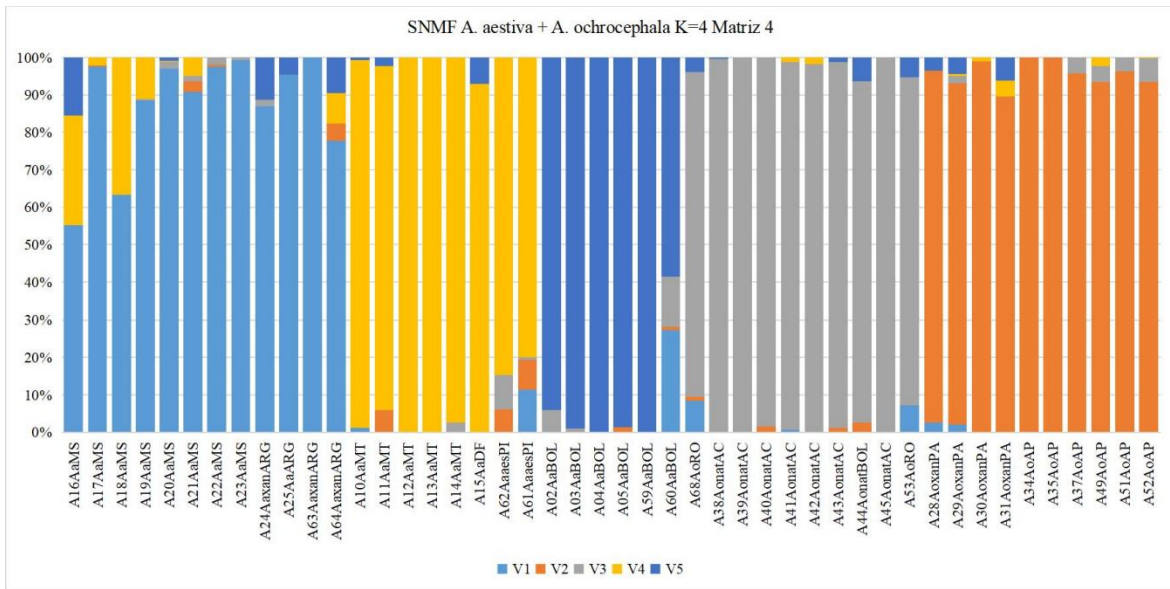
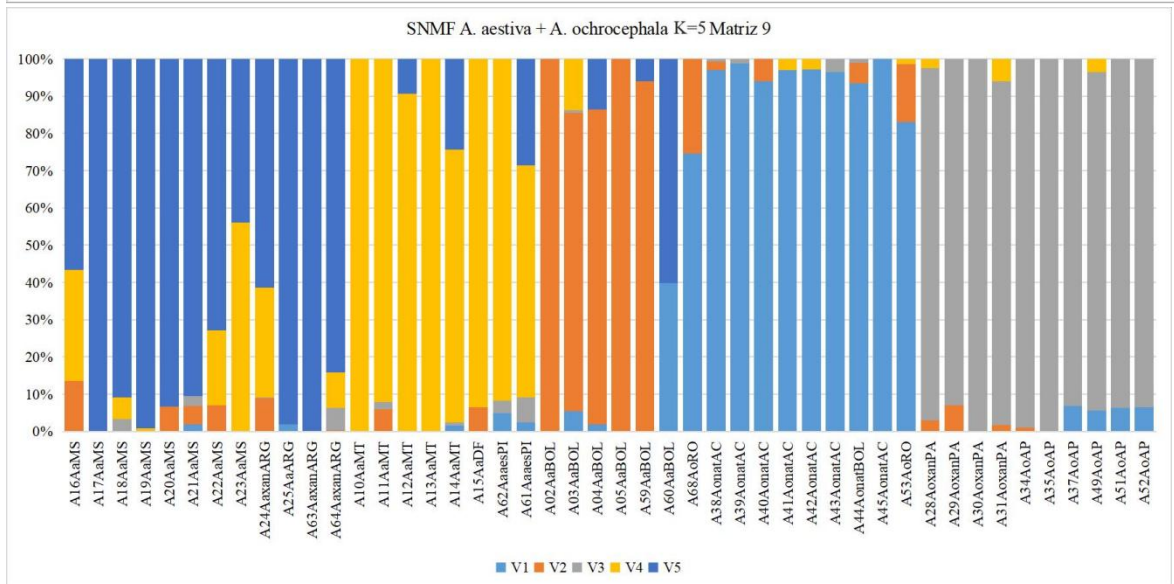
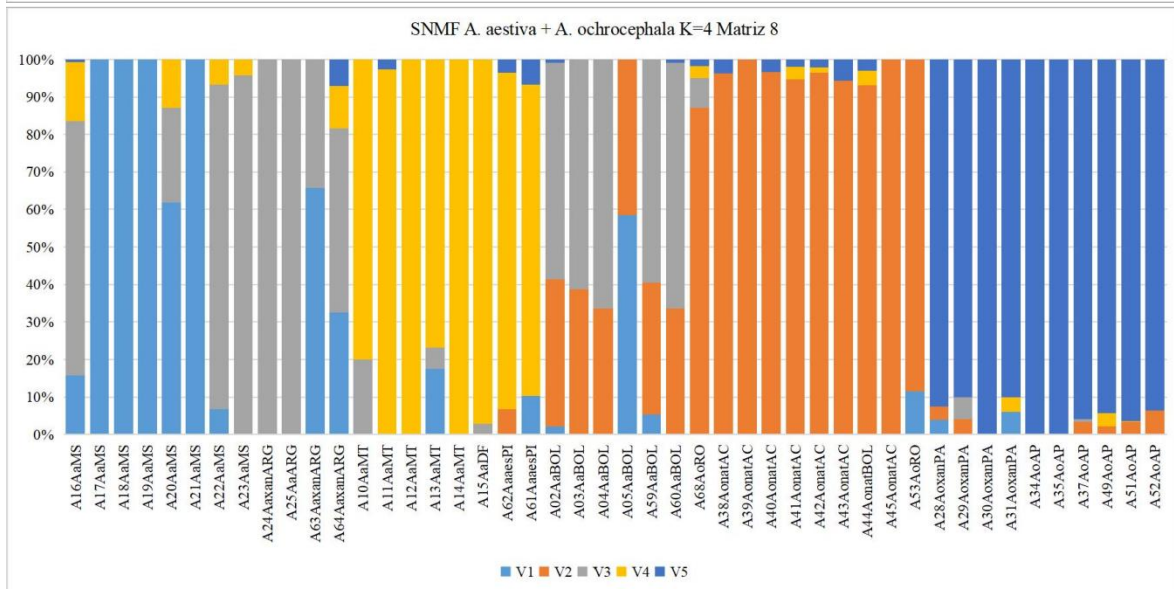
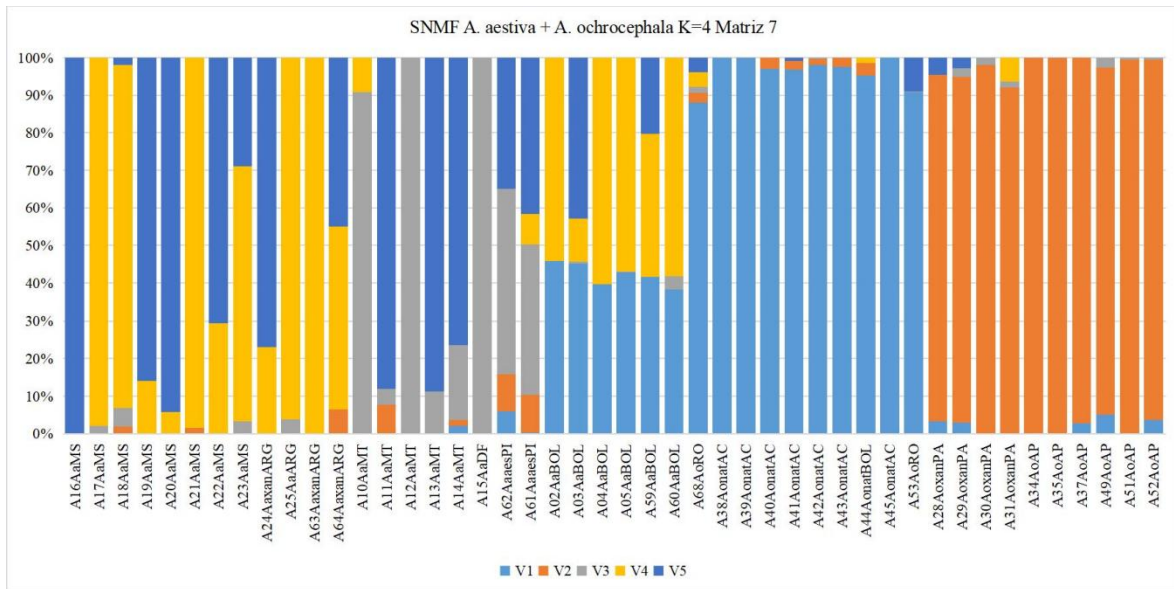


Figura 8 - SNMFs das matrizes teste de *A. aestiva* + *A. ochrocephala*. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 8 - SNMFs das matrizes teste de *A. aestiva* + *A. ochrocephala*.



Continuação da Figura 8 - SNMFs das matrizes teste de *A. aestiva* + *A. ochrocephala*.

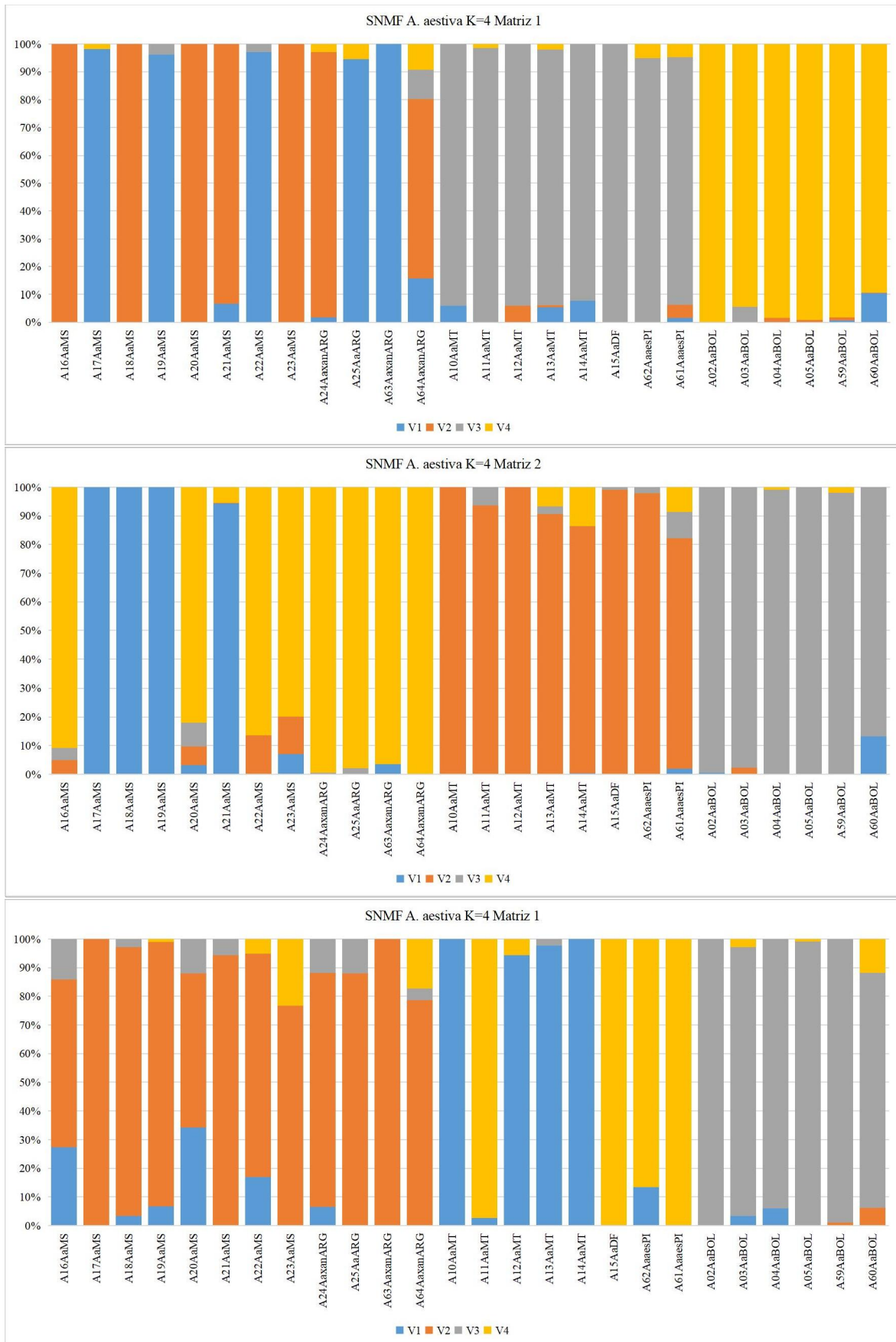
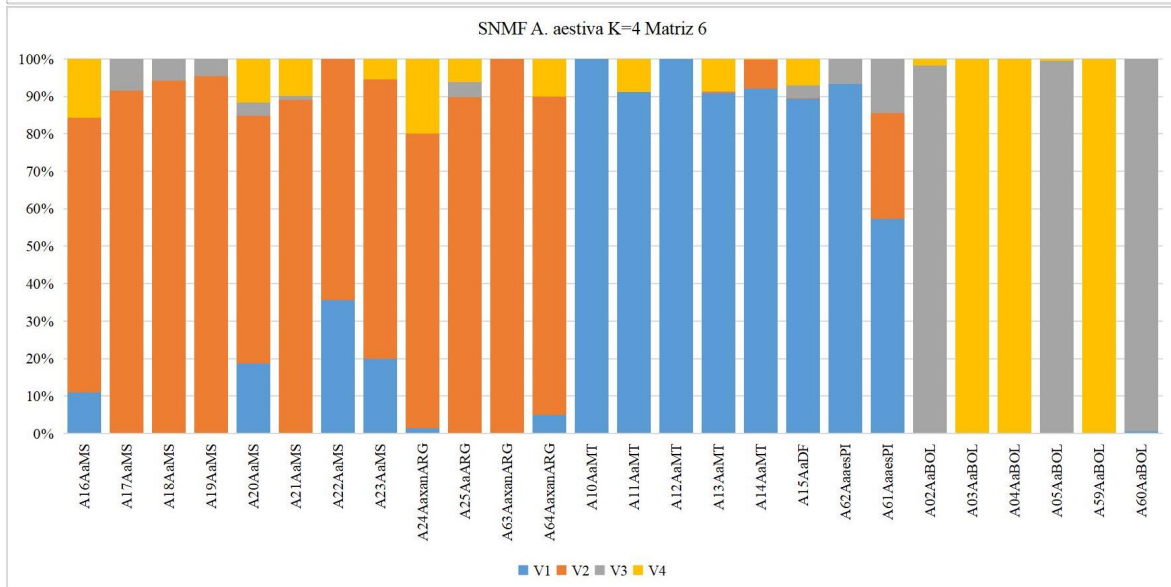
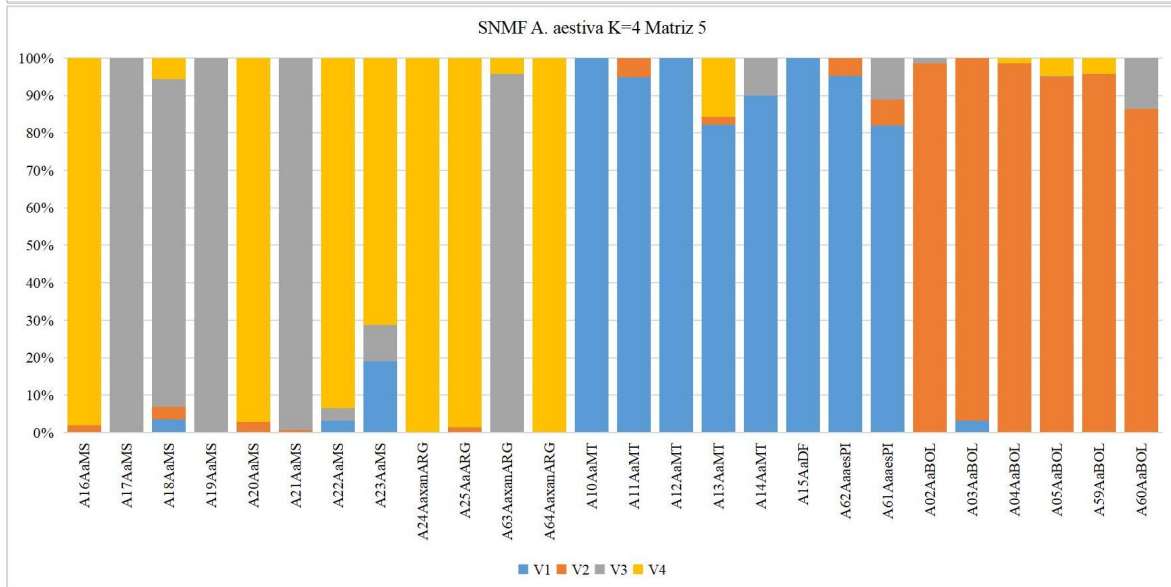
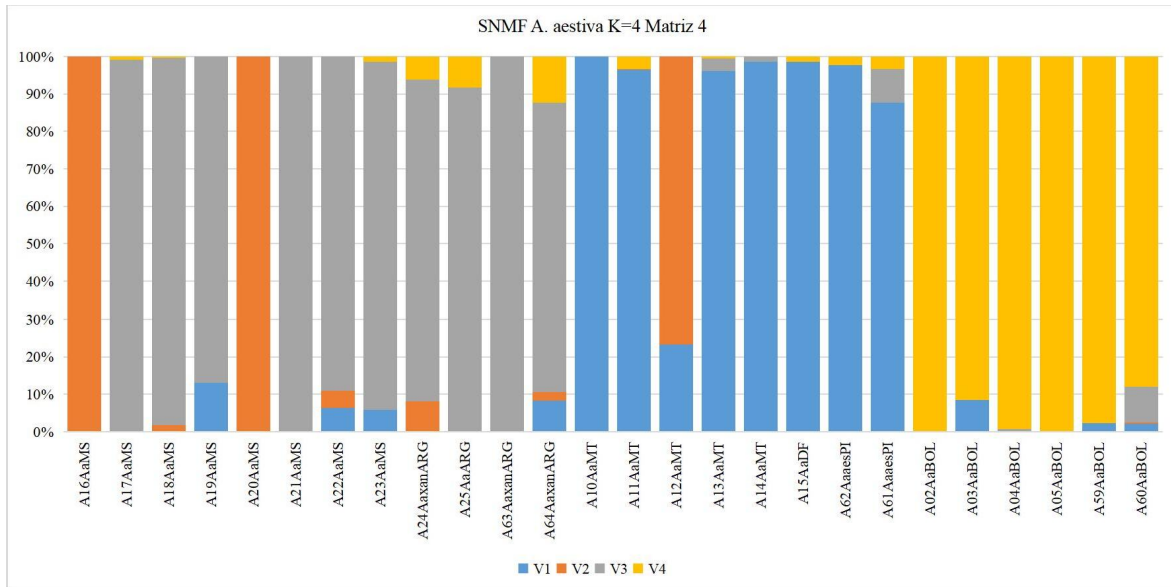
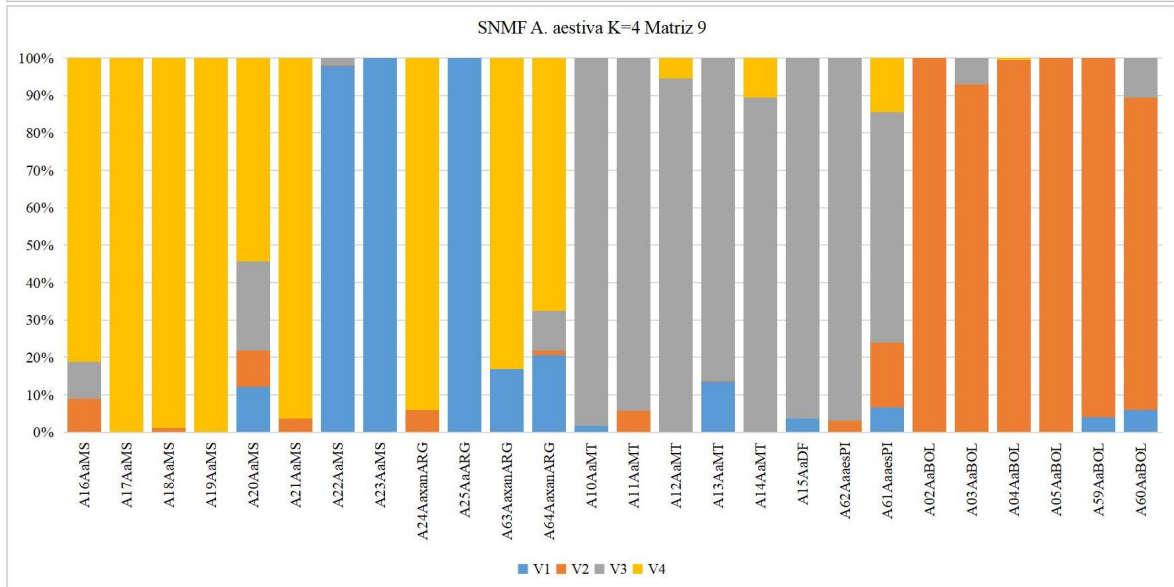
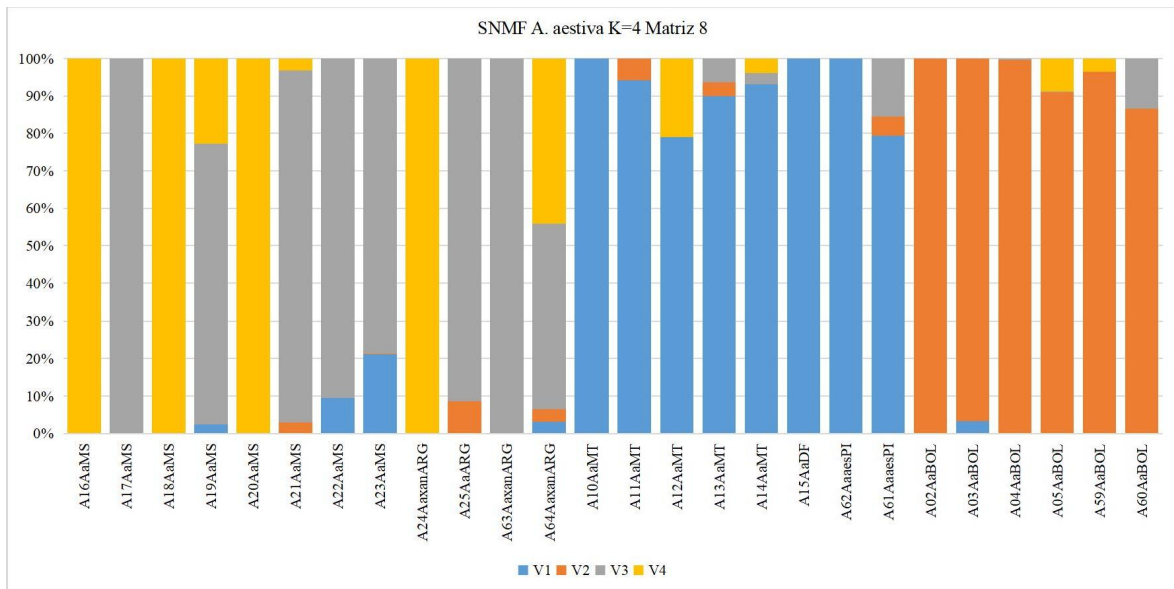
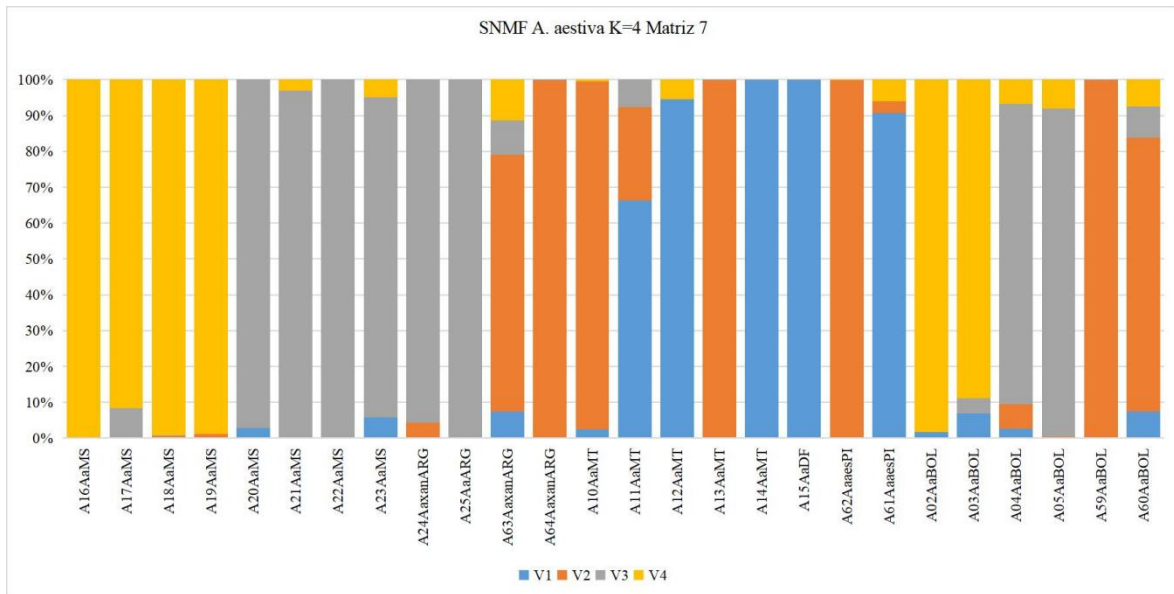


Figura 9- SNMFs das matrizes teste de *A. aestiva*. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 9 - SNMFs das matrizes teste de *A. aestiva*.



Continuação da Figura 9 - SNMFs das matrizes teste de *A. aestiva*.

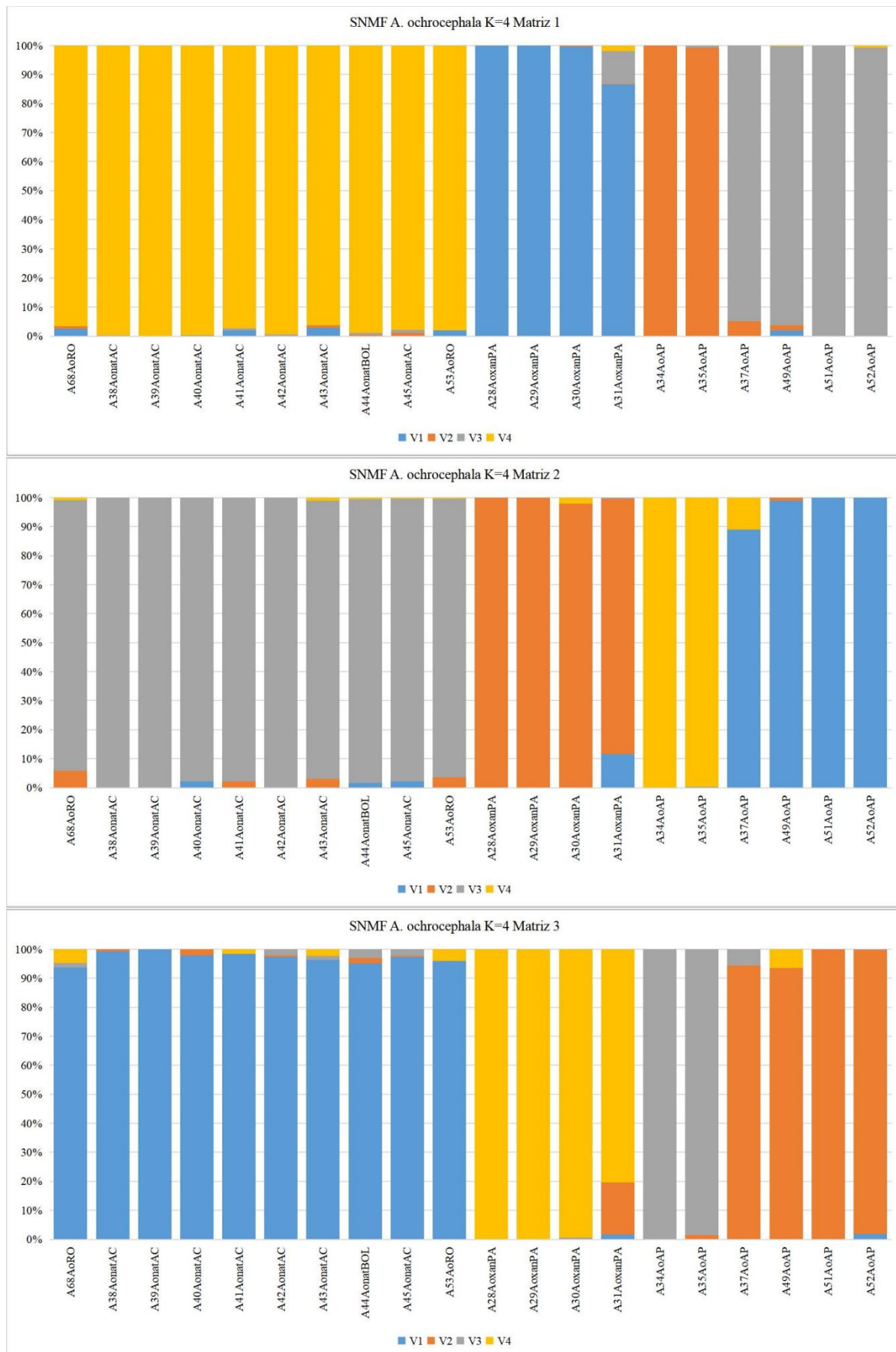
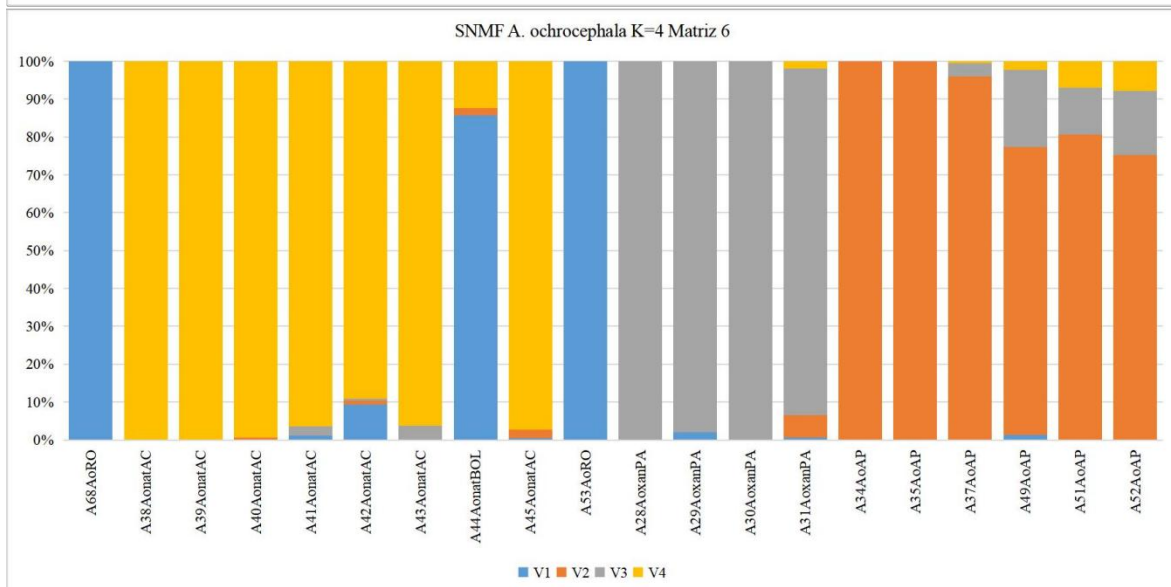
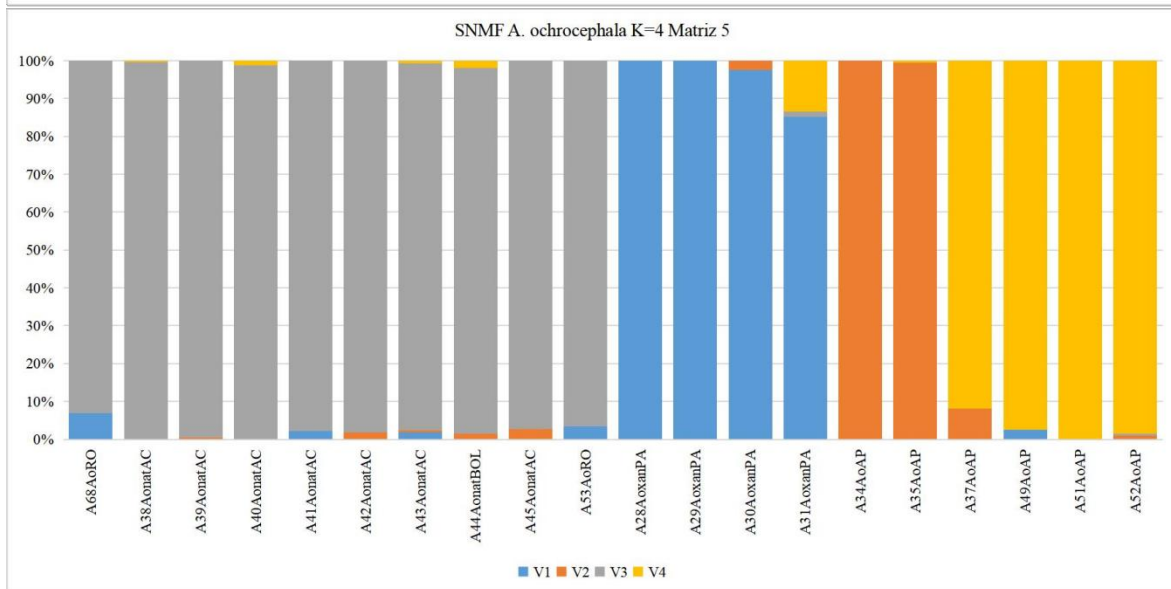
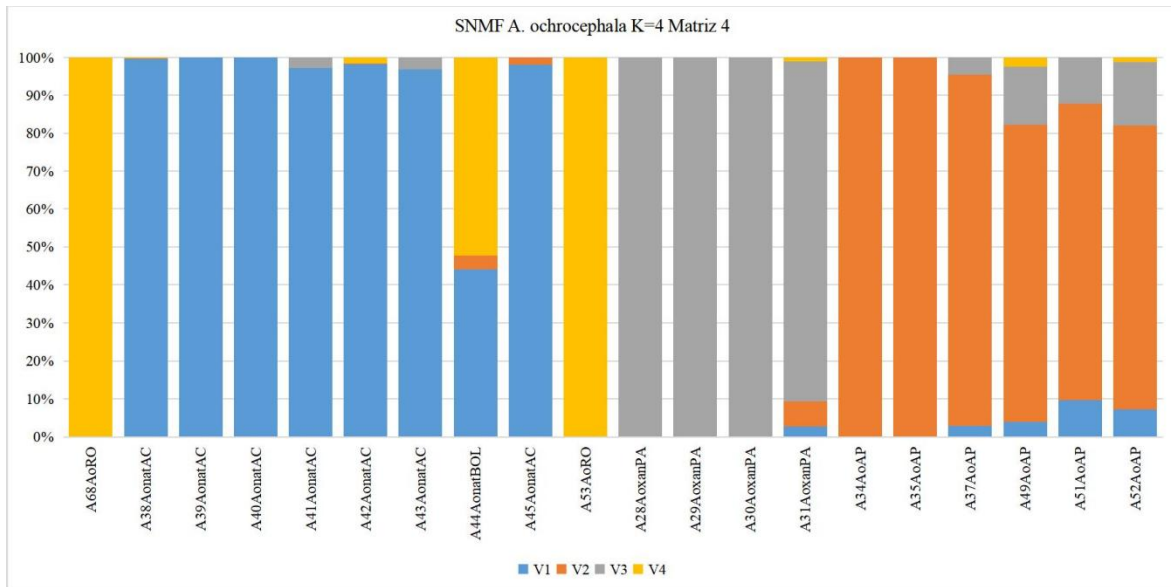
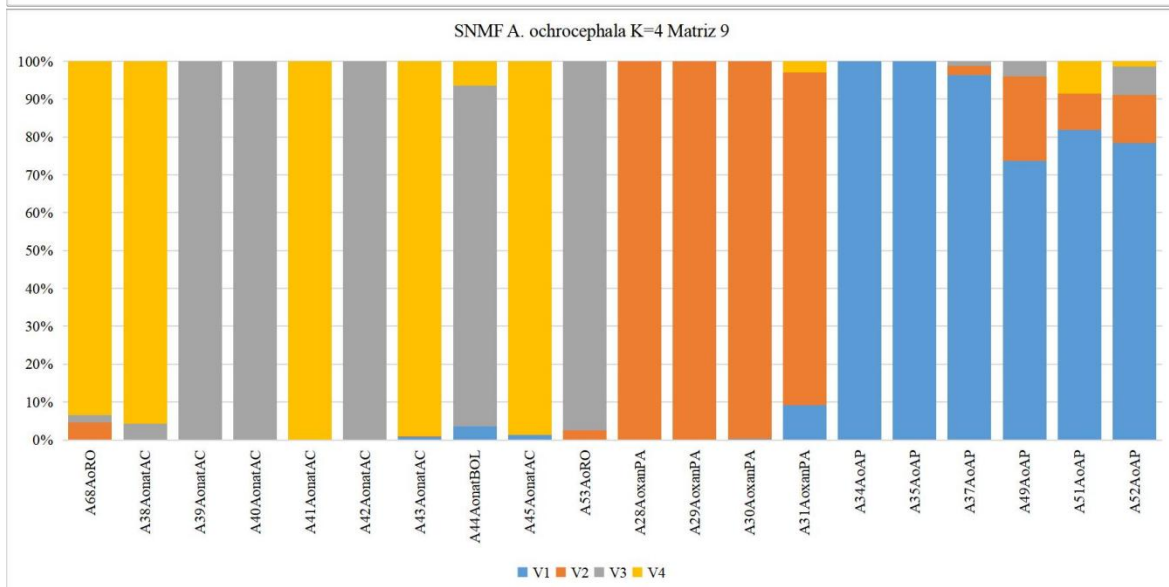
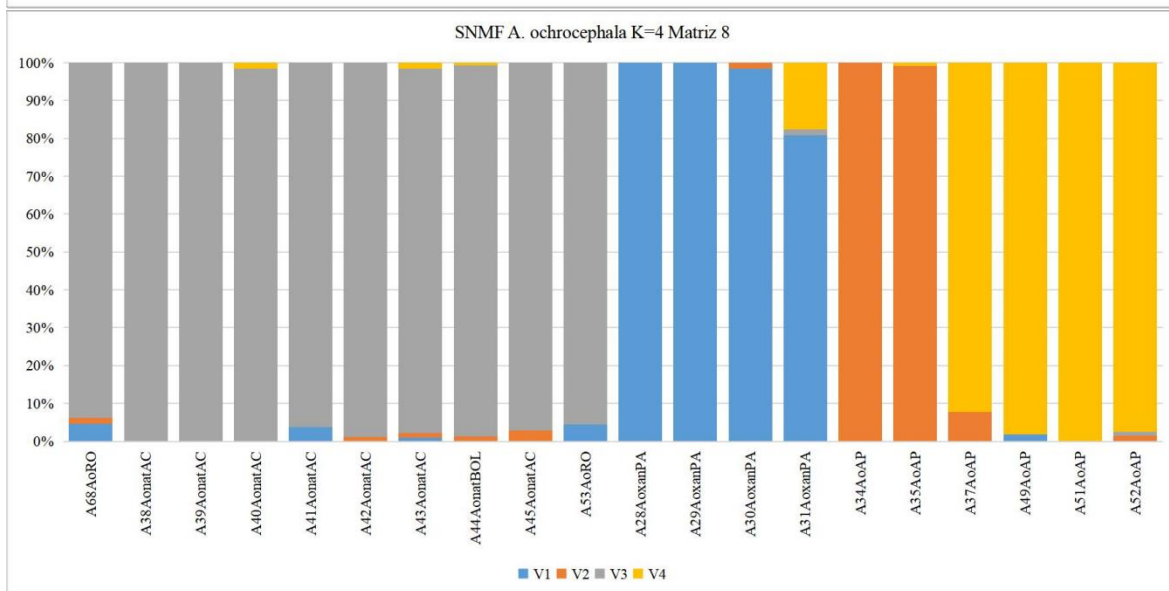
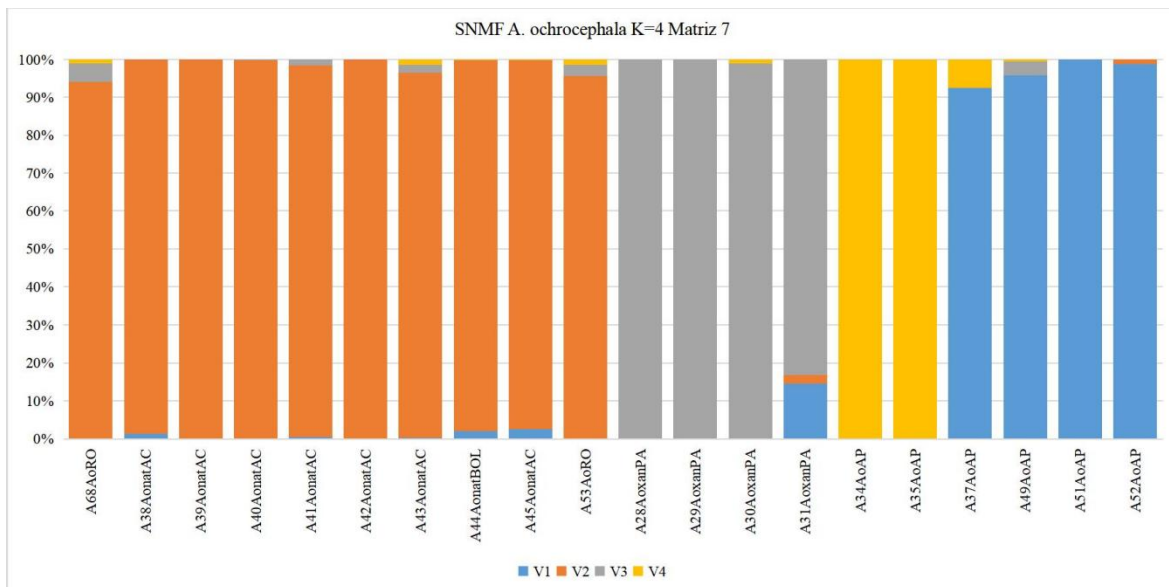


Figura 10 - SNMFs das matrizes teste de *A. ochrocephala*. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 10 - SNMFs das matrizes teste de *A. ochrocephala*.



Continuação da Figura 10 - SNMFs das matrizes teste de *A. ochrocephala*.

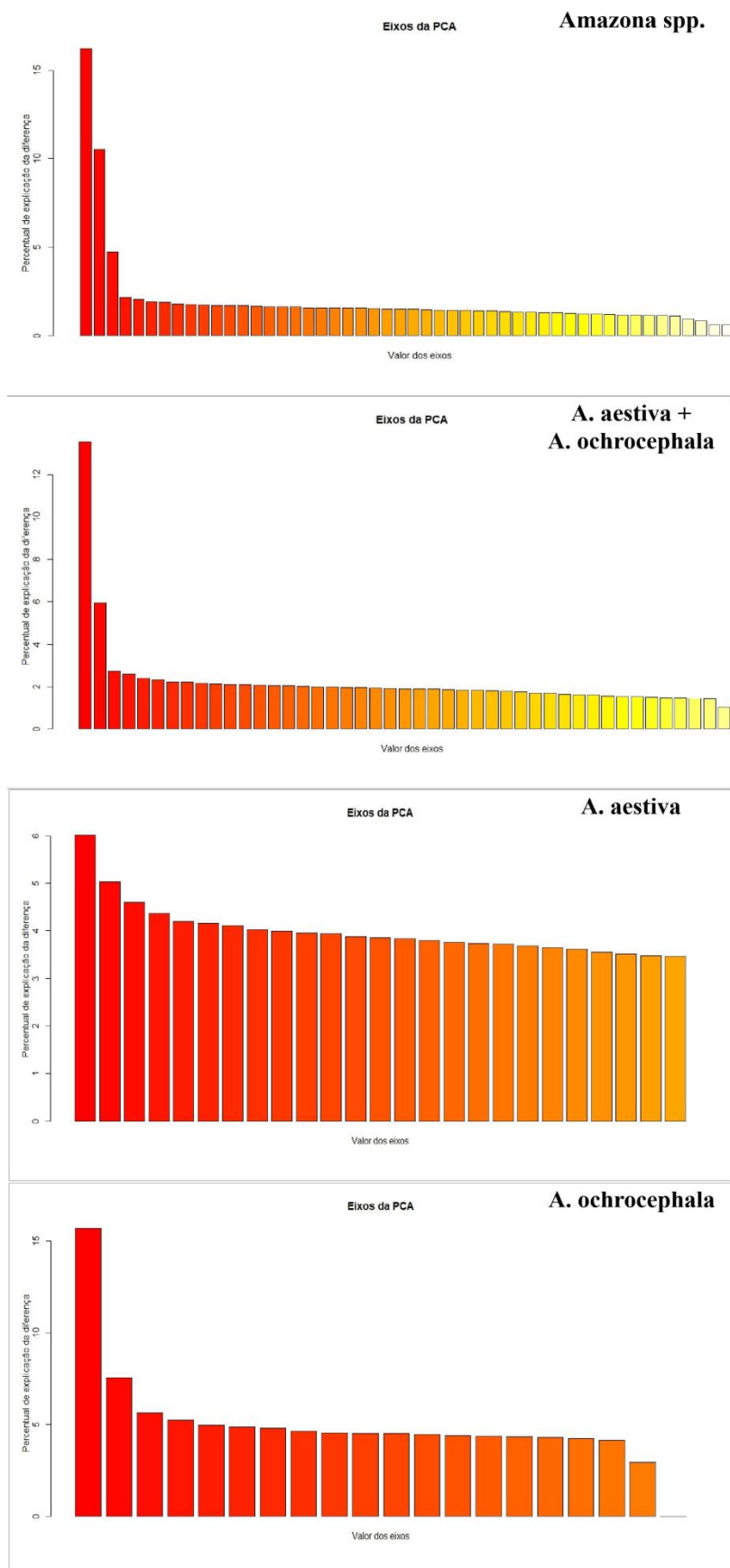


Figura 11 - Influência dos eixos na Análise de Componentes Principais (PCA) das quatro matrizes. Barras verticais indicam eixos da PC; matrizes indicadas no lado superior esquerdo de cada figura.

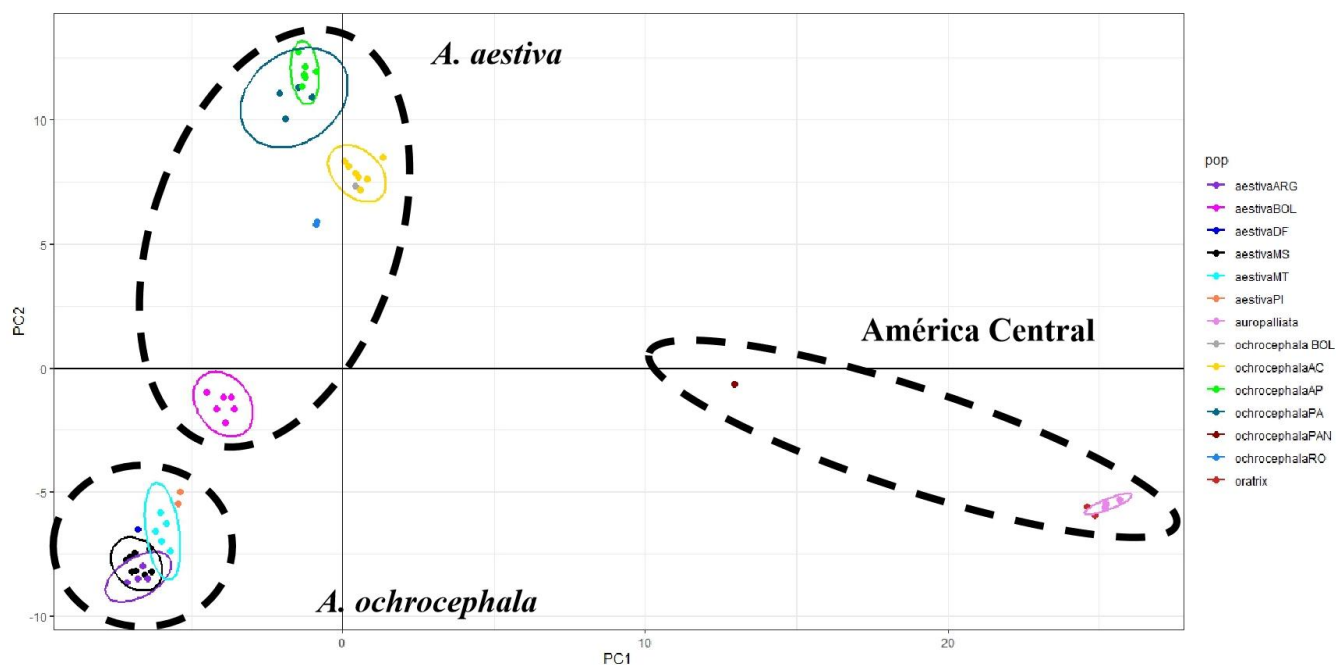


Figura 12 – Análise de Componentes Principais (PCA) da matriz *Amazona* spp. As cores e siglas na legenda indicam as localidades (CAT - cativeiro, BOL - Bolívia, ARG - Argentina, MT - Mato Grosso, DF - Distrito Federal, MS - Mato Grosso do Sul, PA - Pará, AP - Amapá, AC - Acre, RO - Rondônia, PI - Piauí); círculos franchados indicam as espécies estudadas.

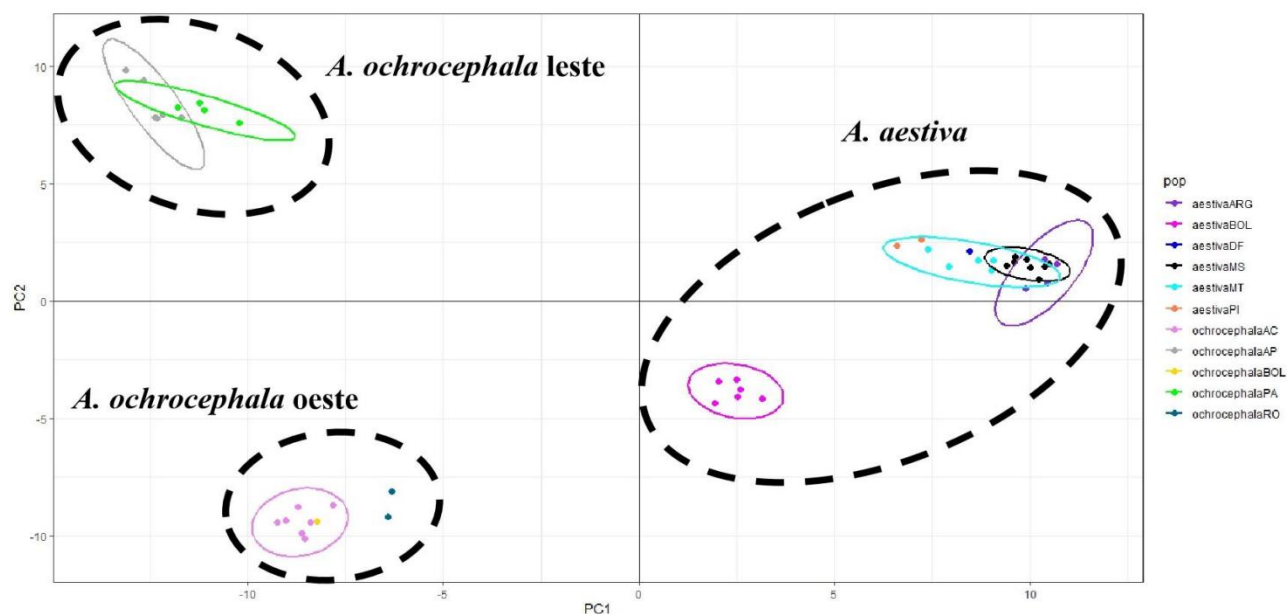


Figura 13 – Análise de Componentes Principais (PCA) da matriz *A. aestiva* + *A. ochrocephala*. As cores e siglas na legenda indicam as localidades (CAT - cativeiro, BOL - Bolívia, ARG - Argentina, MT - Mato Grosso, DF - Distrito Federal, MS - Mato Grosso do Sul, PA - Pará, AP - Amapá, AC - Acre, RO - Rondônia, PI - Piauí); círculos franchados indicam as espécies estudadas.

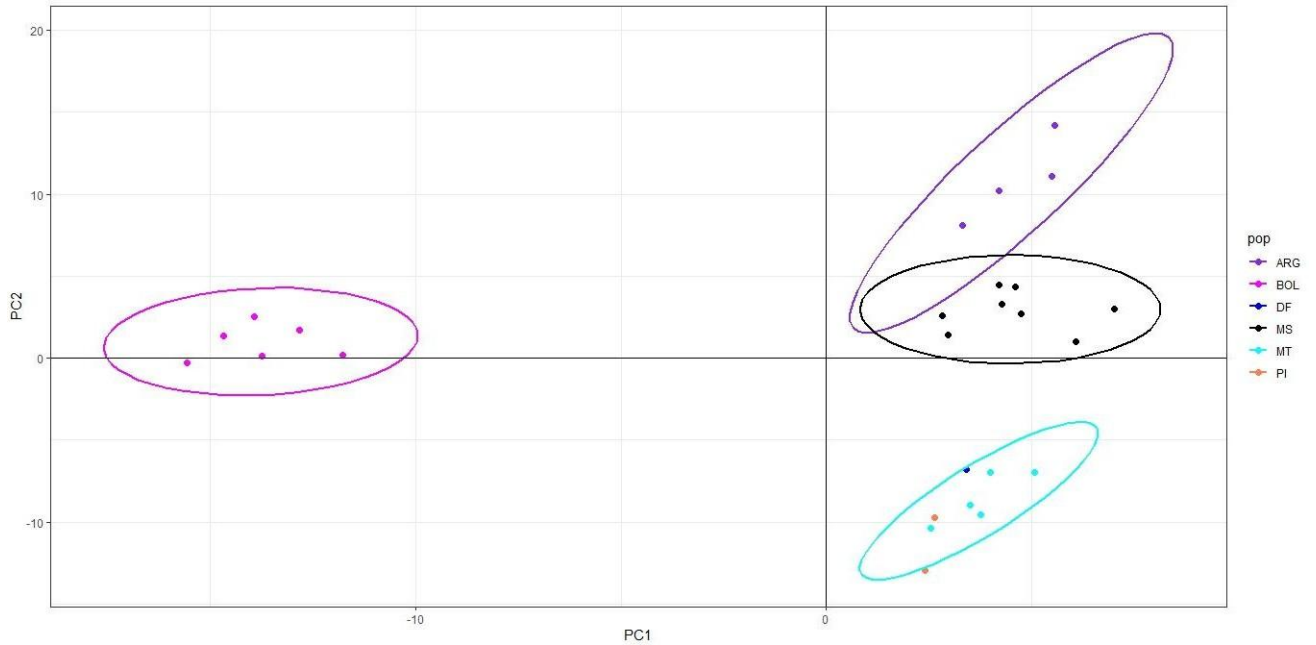


Figura 14 – Análise de Componentes Principais (PCA) da matriz *A. aestiva*. As cores e siglas na legenda indicam as localidades (CAT - cativoiro, BOL - Bolívia, ARG - Argentina, MT - Mato Grosso, DF - Distrito Federal, MS - Mato Grosso do Sul, PA - Pará, AP - Amapá, AC - Acre, RO - Rondônia, PI - Piauí).

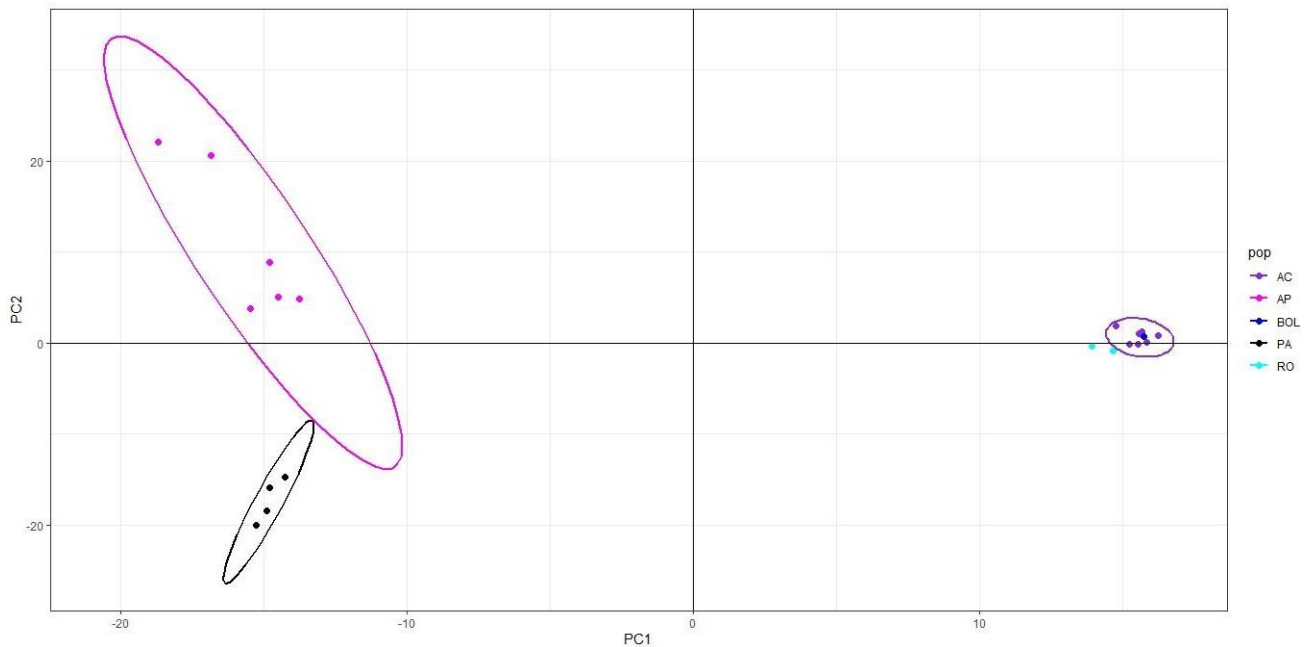


Figura 15 – Análise de Componentes Principais (PCA) da matriz *A. ochrocephala*. As cores e siglas na legenda indicam as localidades (CAT - cativoiro, BOL - Bolívia, ARG - Argentina, MT - Mato Grosso, DF - Distrito Federal, MS - Mato Grosso do Sul, PA - Pará, AP - Amapá, AC - Acre, RO - Rondônia, PI - Piauí).

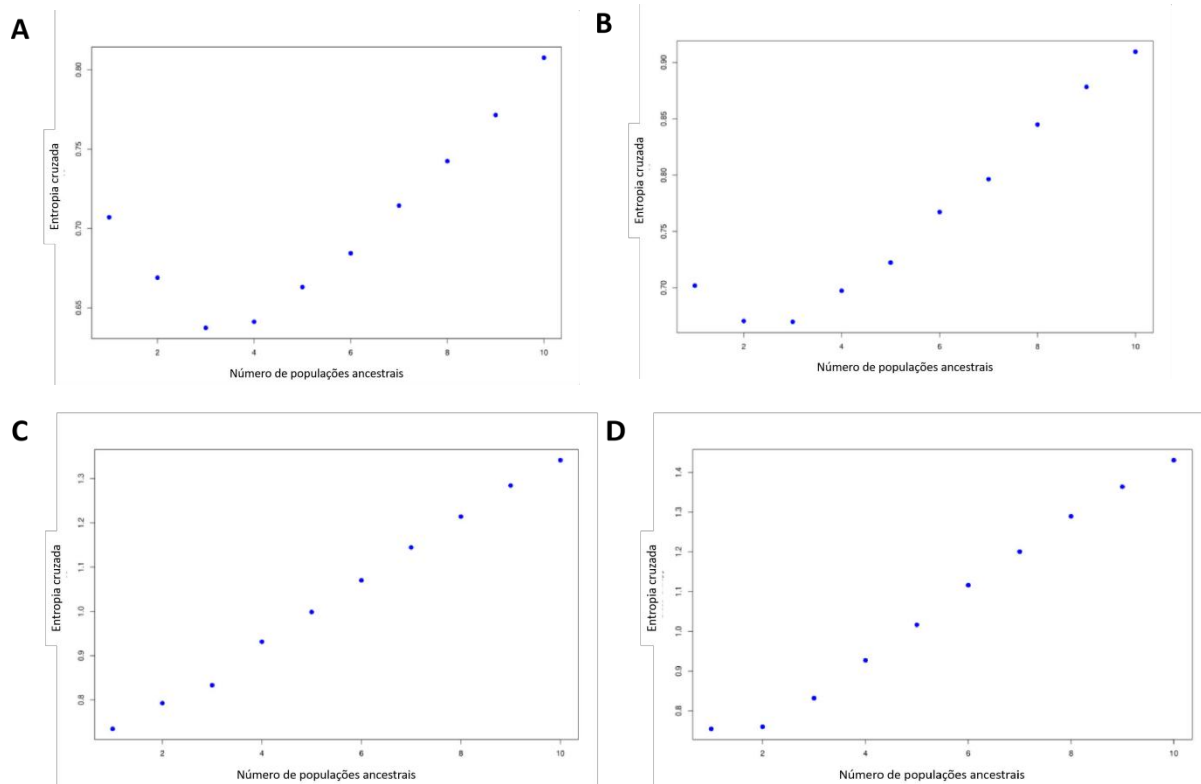


Figura 16 - Identificação do valor de K. Análise de entropia do realizada no Adegnet para identificação da quantidade de populações (K) mais provável nas quatro matrizes utilizadas no estudo. A) *Amazona spp.* B) *A. aestiva*+*A. ochrocephala* C) *A. aestiva* D) *A. ochrocephala*.

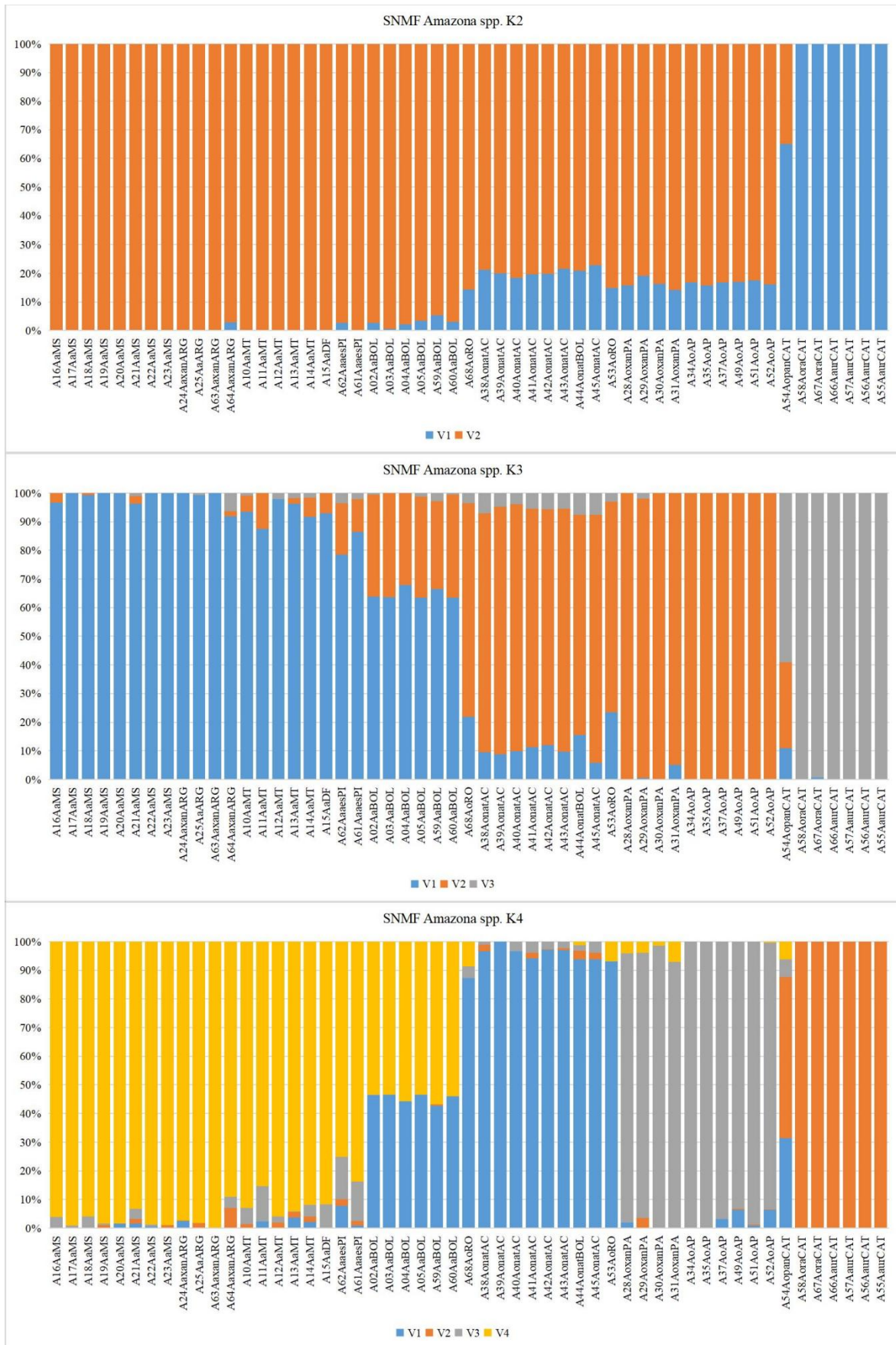
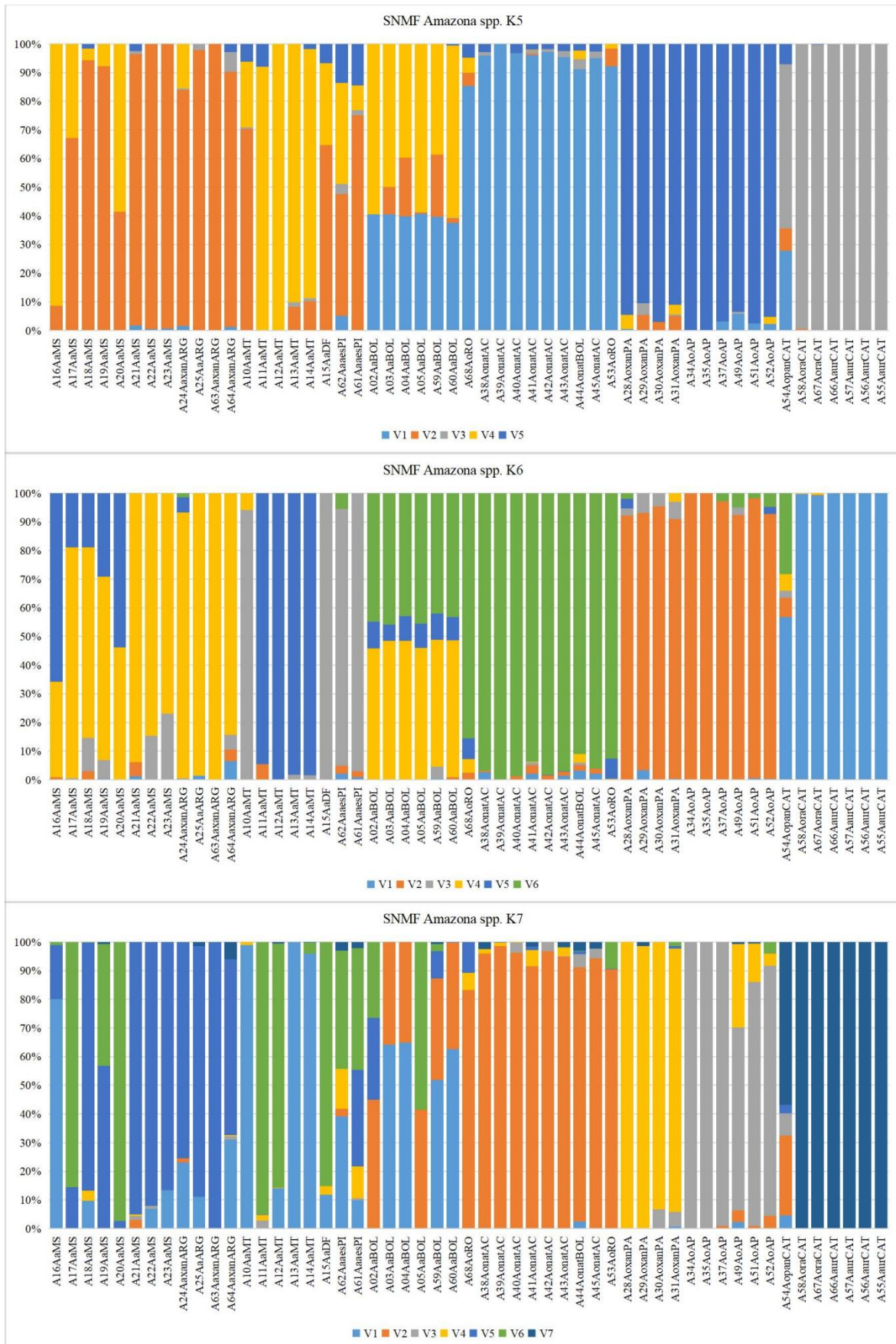
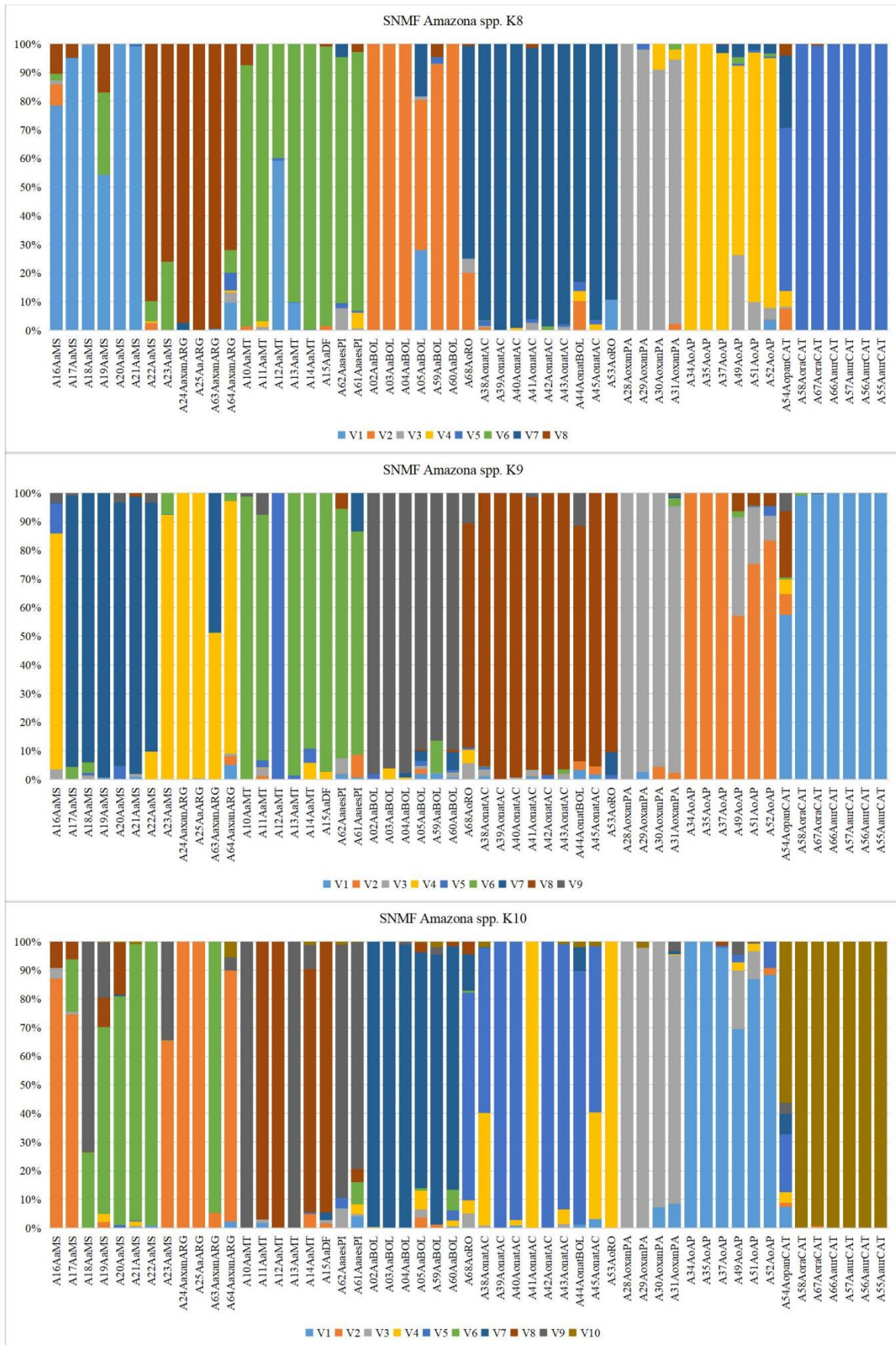


Figura 17 - SNMFs da matriz *Amazona* spp. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 17 - SNMFs da matriz *Amazona* spp.



Continuação da Figura 17 - SNMFs da matriz *Amazona* spp.

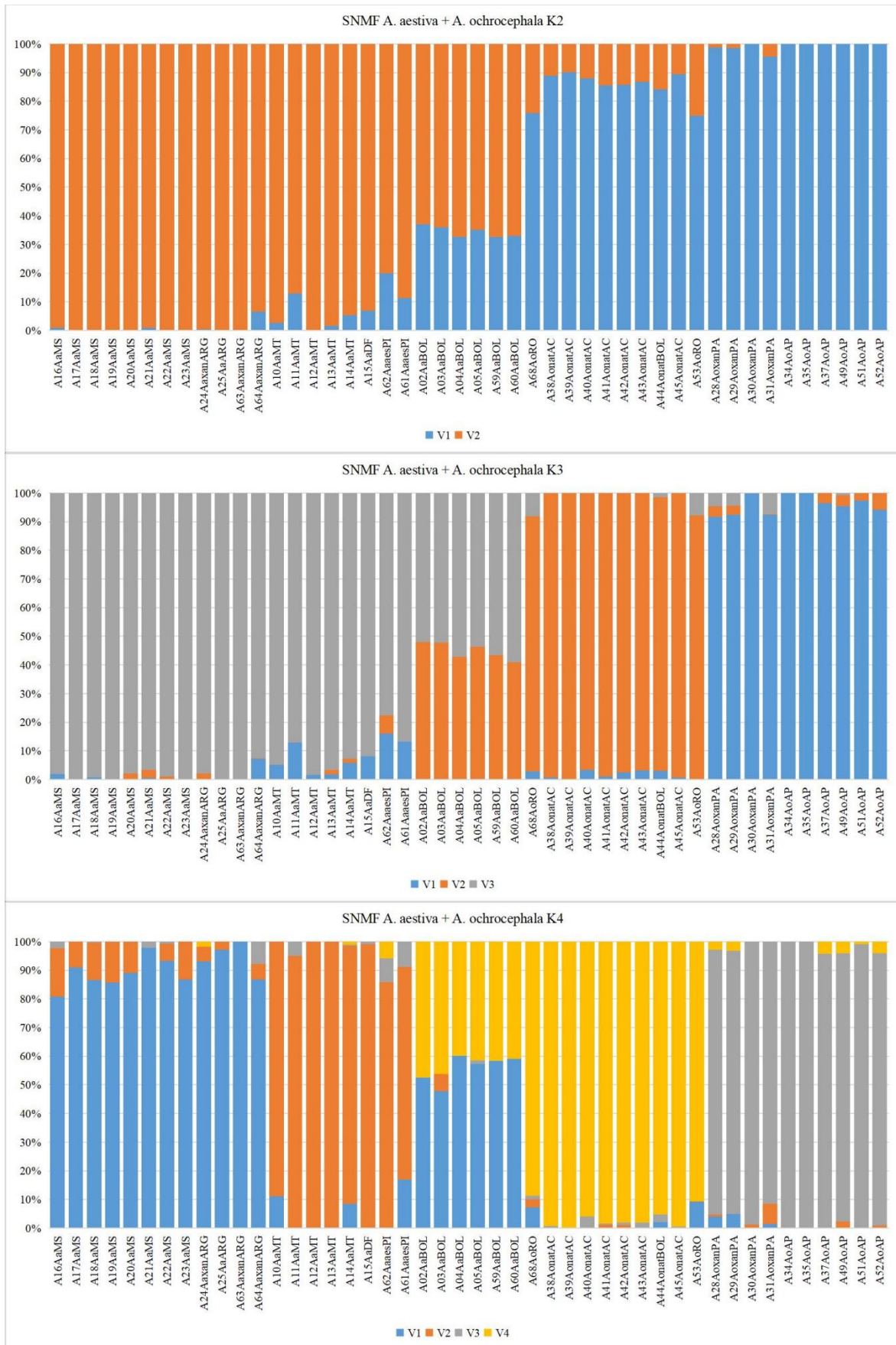
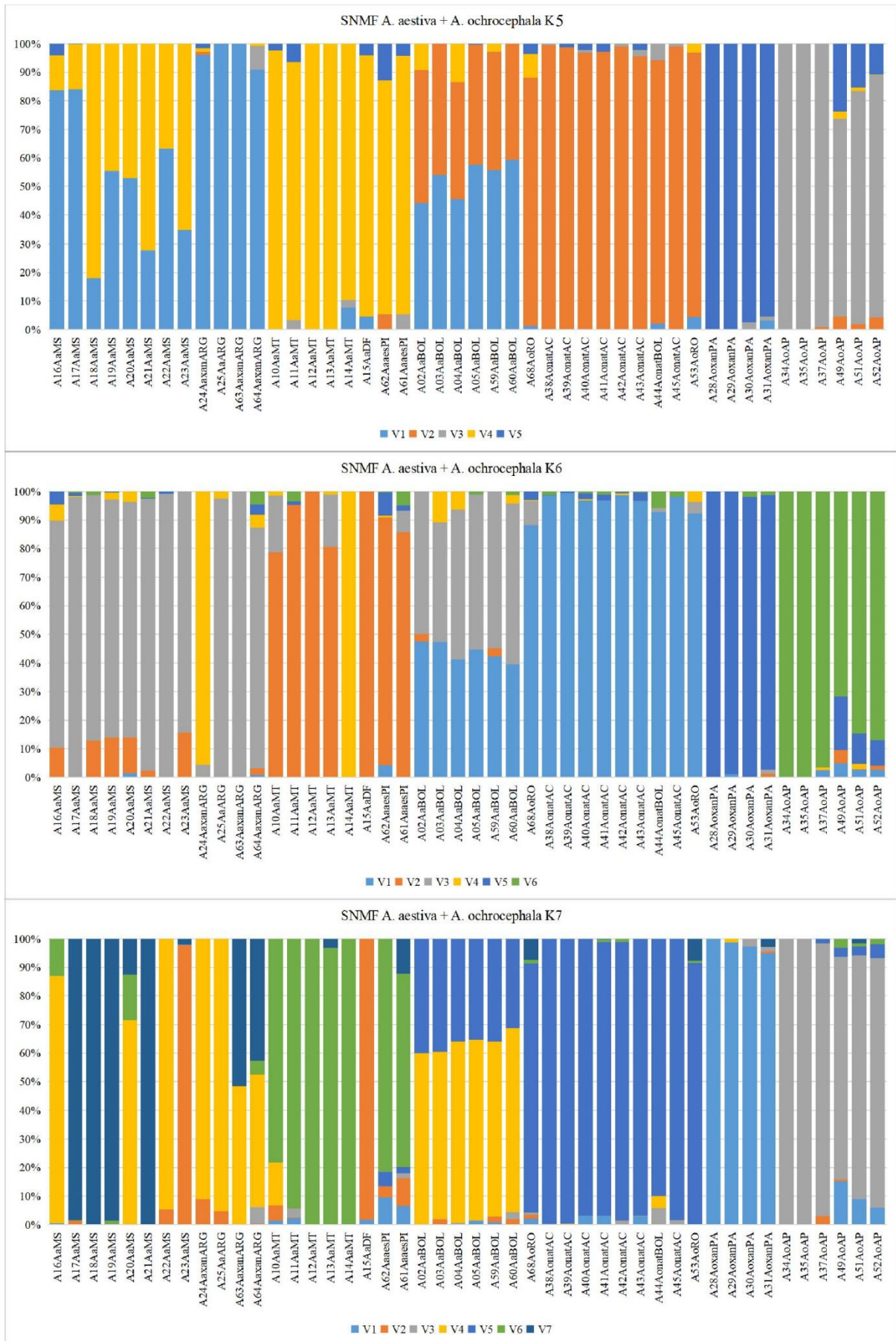
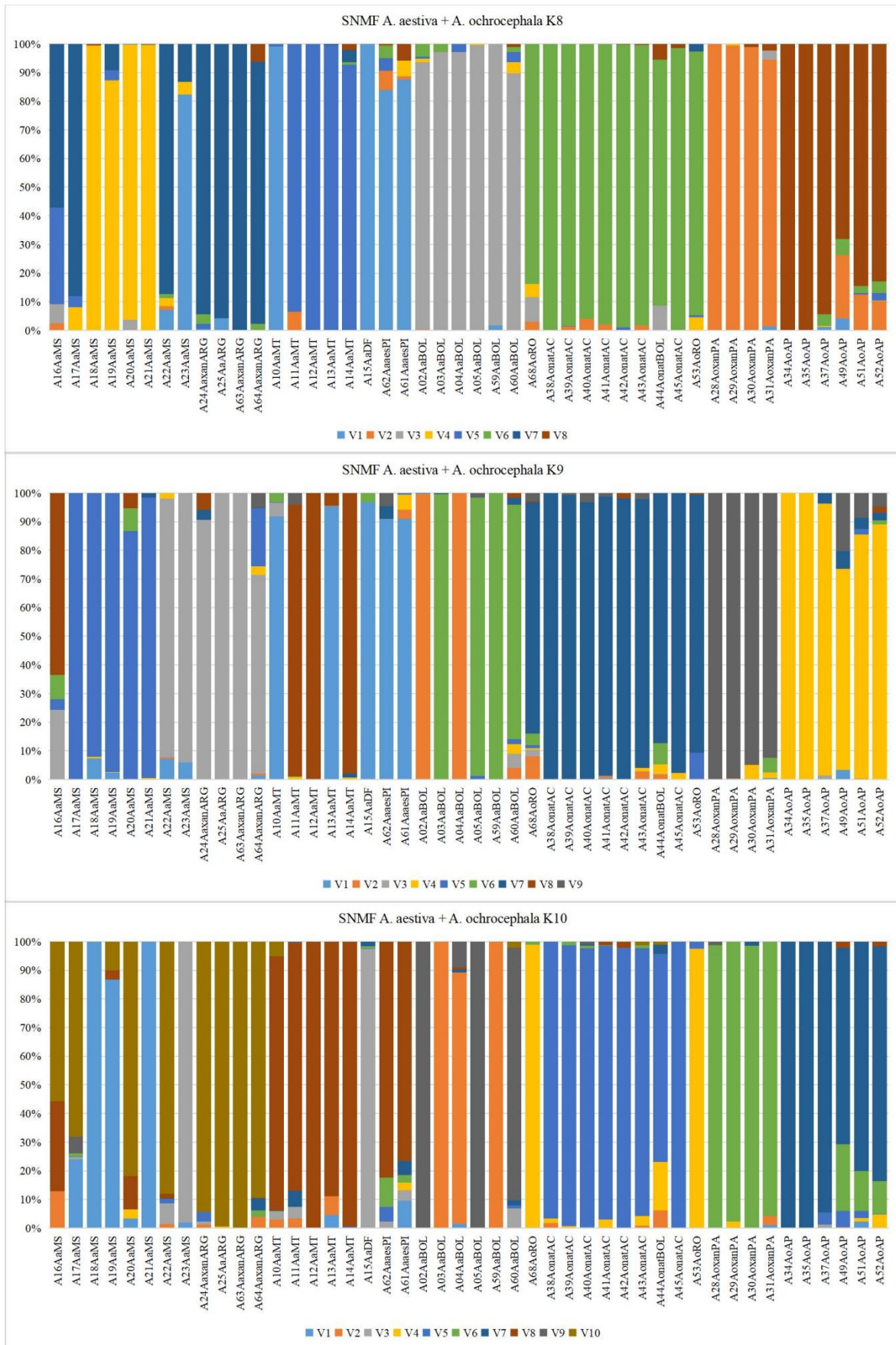


Figura 18 - SNMFs da matriz *A. aestiva* + *A. ochrocephala*. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 18 - SNMFs da matriz *A. aestiva* + *A. ochrocephala*.



Continuação da Figura 18 - SNMFs da matriz *A. aestiva* + *A. ochrocephala*.

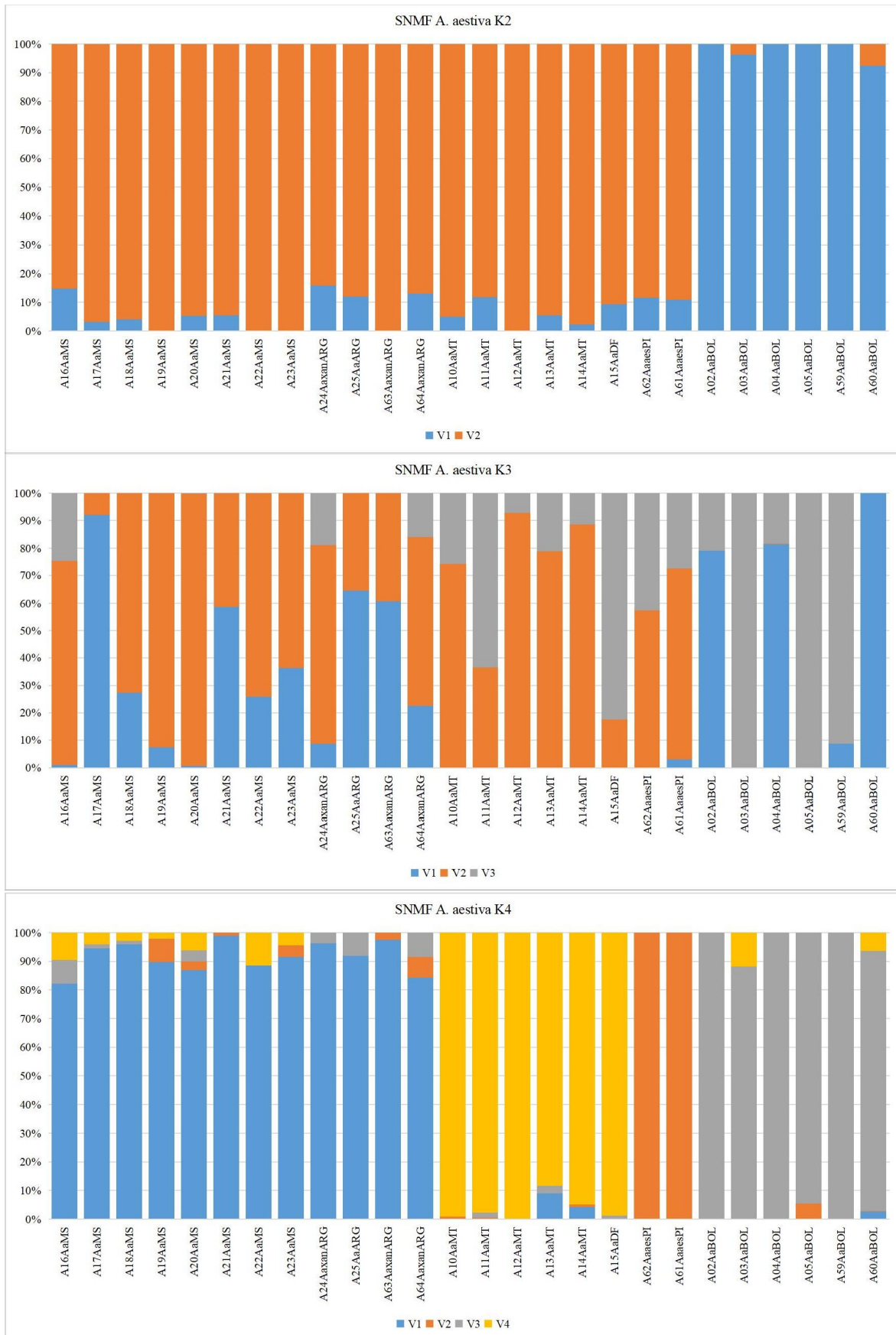
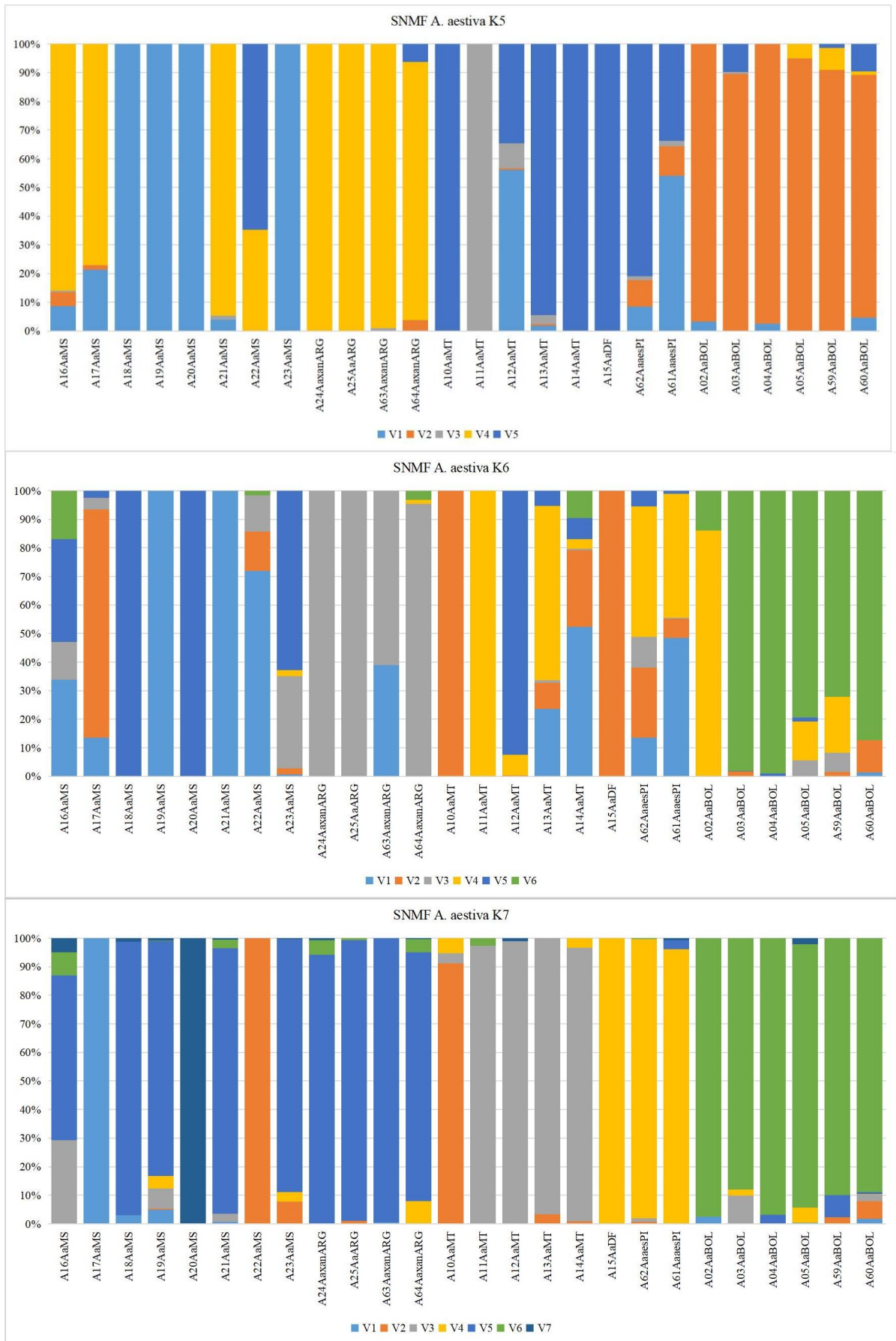
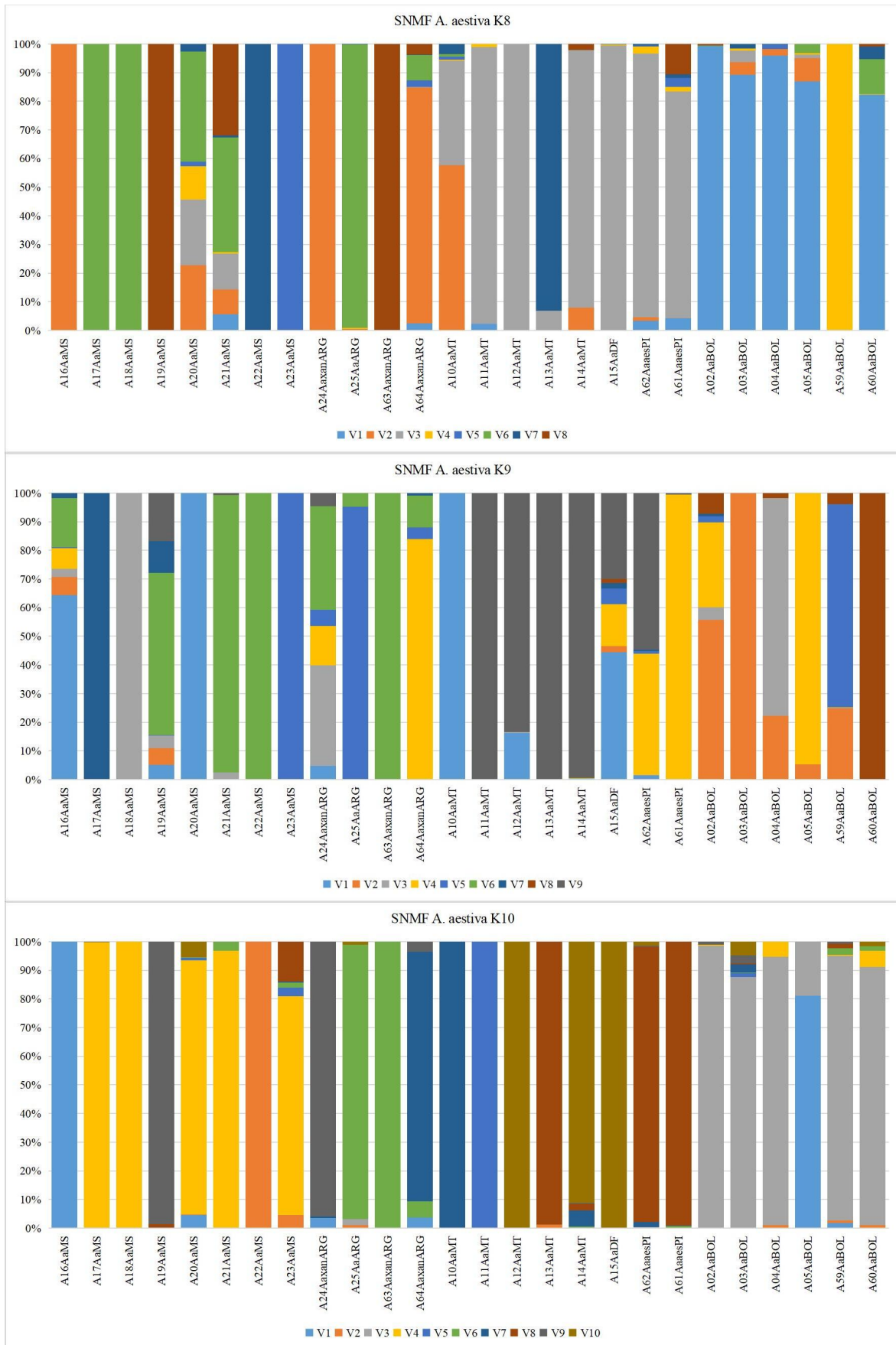


Figura 19 - SNMFs da matriz *A. aestiva*. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 19 - SNMFs da matriz *A. aestiva*.



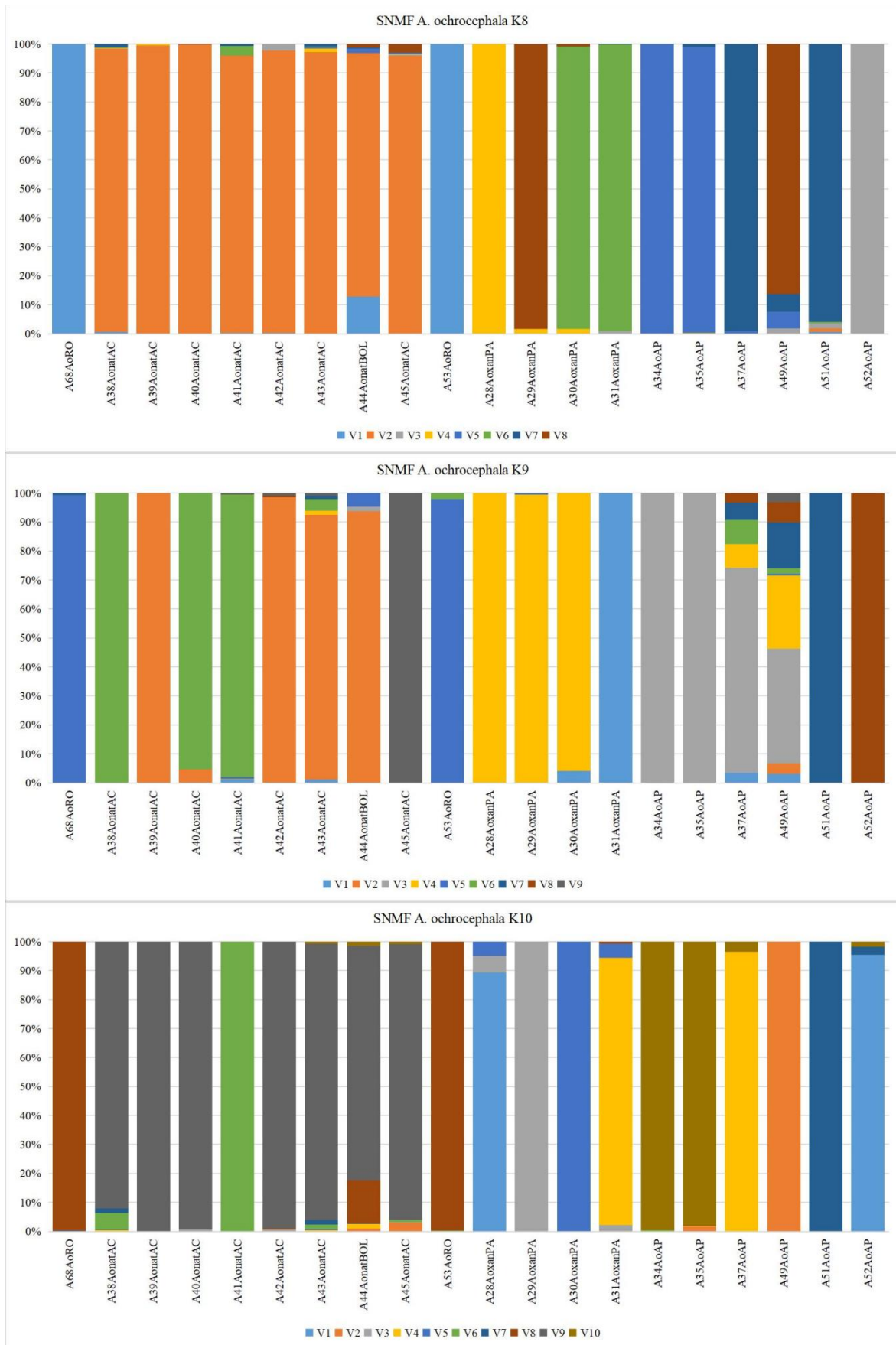
Continuação da Figura 19 - SNMFs da matriz *A. aestiva*.



Figura 20 - SNMFs da matriz *A. ochrocephala*. Barras verticais representam os indivíduos; cores das barras representam os *clusters* encontrados no SNMF; identificação abaixo de cada barra indica a espécie e a localidade.



Continuação da Figura 20 - SNMFs da matriz *A. ochrocephala*.



Continuação da Figura 20 - SNMFs da matriz *A. ochrocephala*.

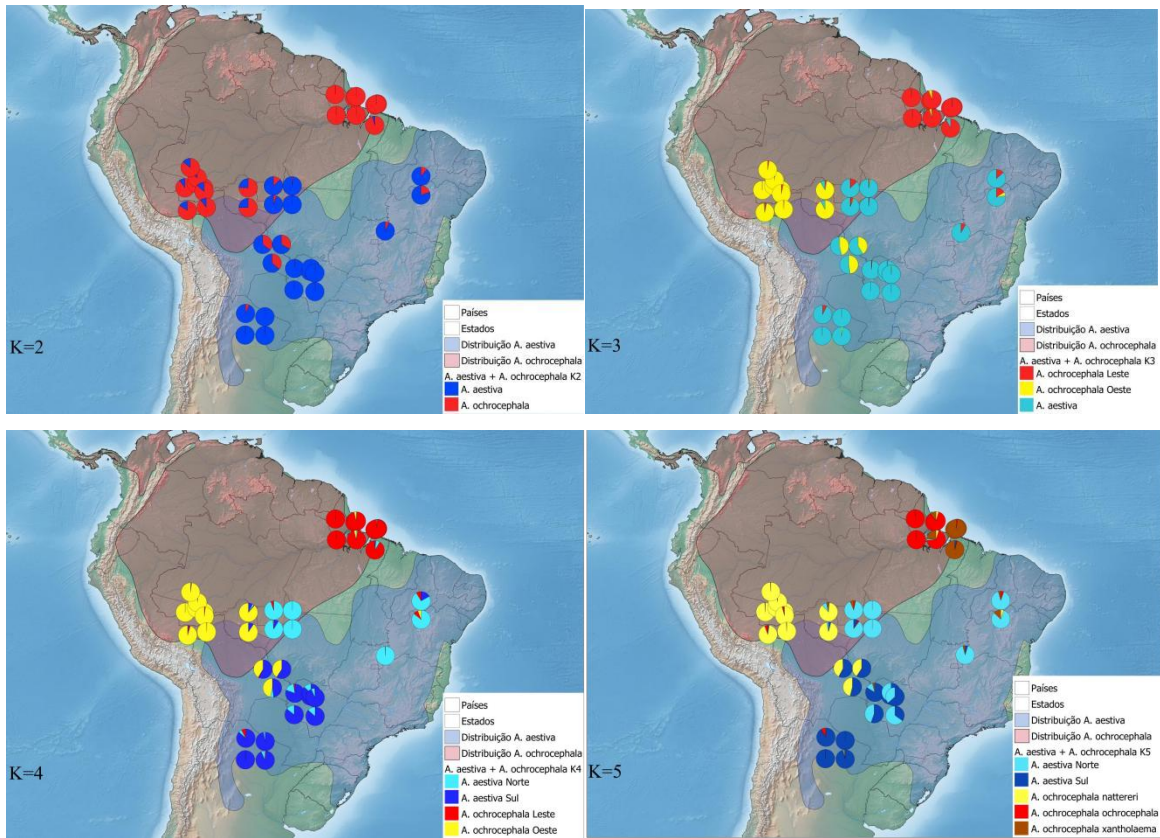


Figura 21 - SNMF de *A. aestiva*+*A. ochrocephala* nas localidades de amostragem. Cada círculo representa um indivíduo; as cores correspondem aos *clusters* encontrados nos SNMFs; valor de K no lado inferior esquerdo das figuras.

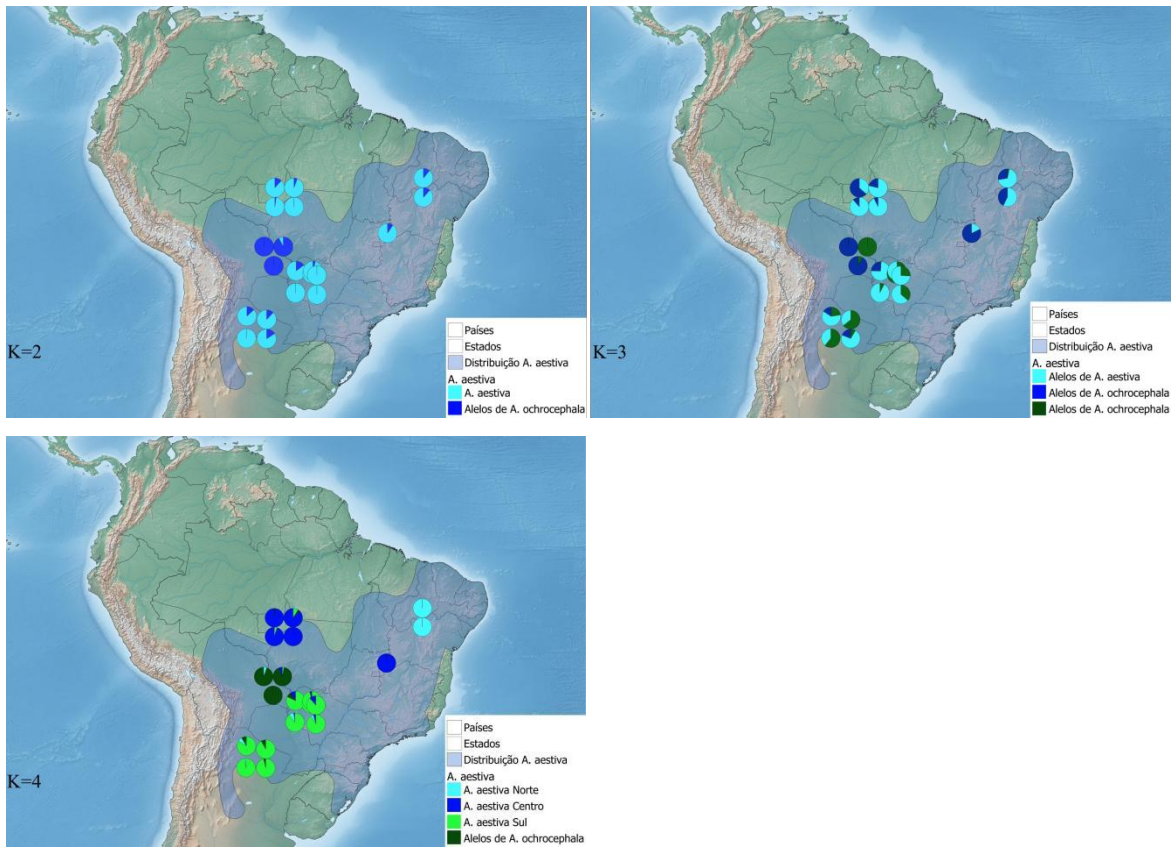


Figura 22 - SNMF de *A. aestiva* nas localidades de amostragem. Cada círculo representa um indivíduo; as cores correspondem aos *clusters* encontrados nos SNMFs; valor de K no lado inferior esquerdo das figuras.

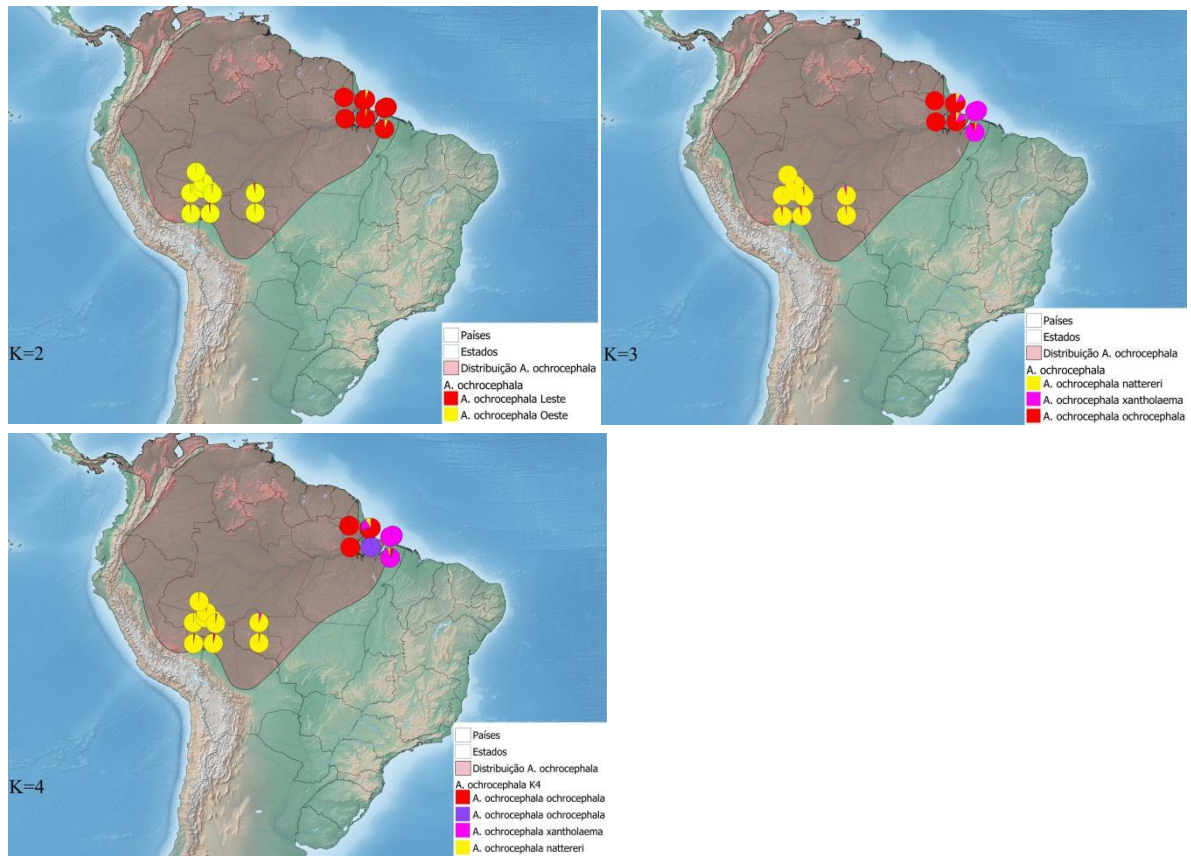


Figura 23 - SNMF de *A. ochrocephala* nas localidades de amostragem. Cada círculo representa um indivíduo; as cores correspondem aos *clusters* encontrados nos SNMFs; valor de K no lado inferior esquerdo das figuras.

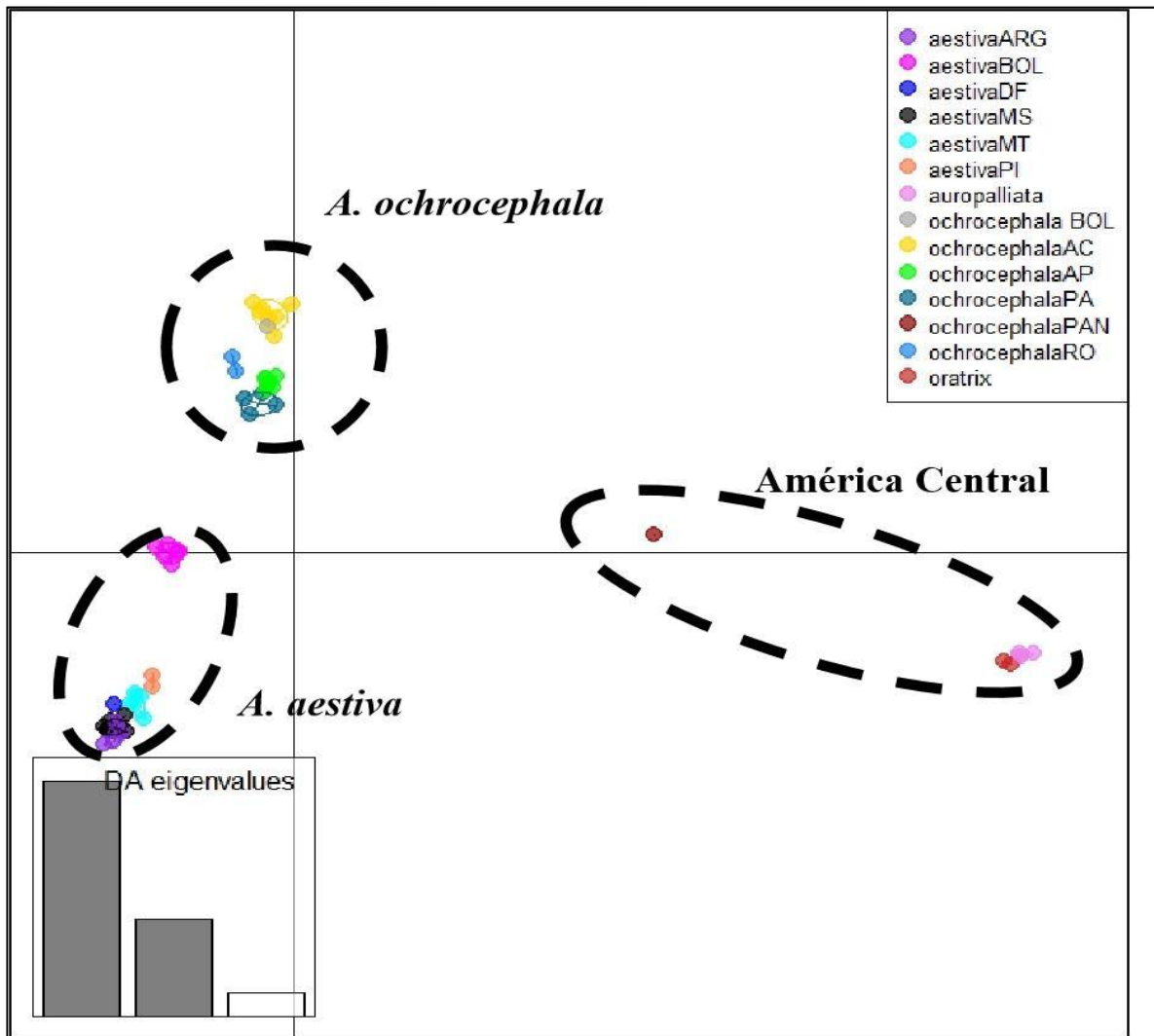


Figura 24 – Análise Discriminante de Componentes Principais (DAPC) da matriz *Amazona spp.*. A influência dos eixos da DAPC (quadro no canto inferior esquerdo) são apresentados. A legenda indica as cores dos agrupamentos encontrados (siglas referem-se às localidades e seguem o padrão da Tabela 1).

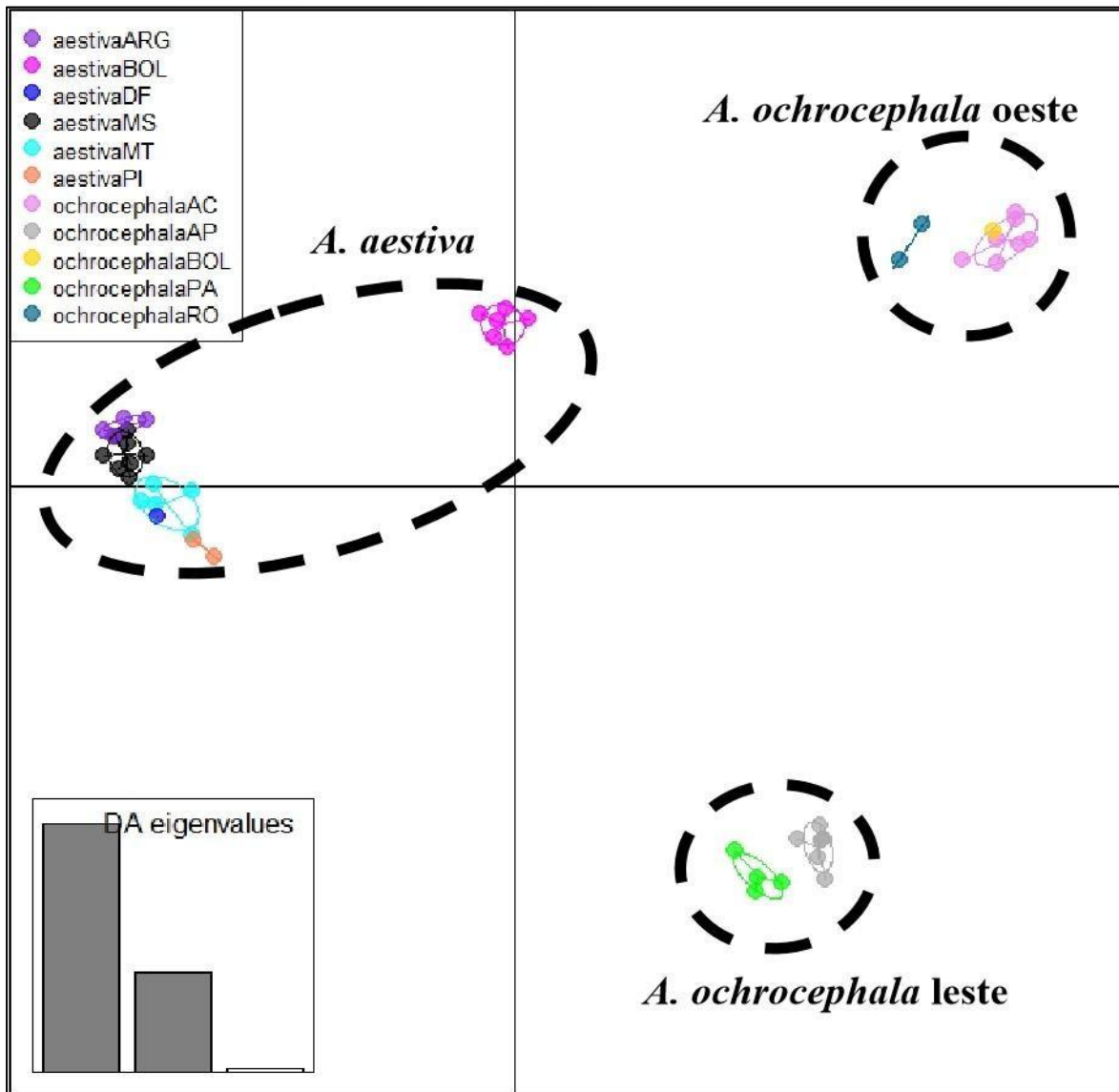


Figura 25 – Análise Discriminante de Componentes Principais (DAPC) da matriz *A. aestiva* + *A. ochrocephala*. A influência dos eixos da DAPC (quadro no canto inferior esquerdo) são apresentados . A legenda indica a s cores dos agrupamentos encontrados (siglas referem-se às localidades e seguem o padrão da Tabela 1).

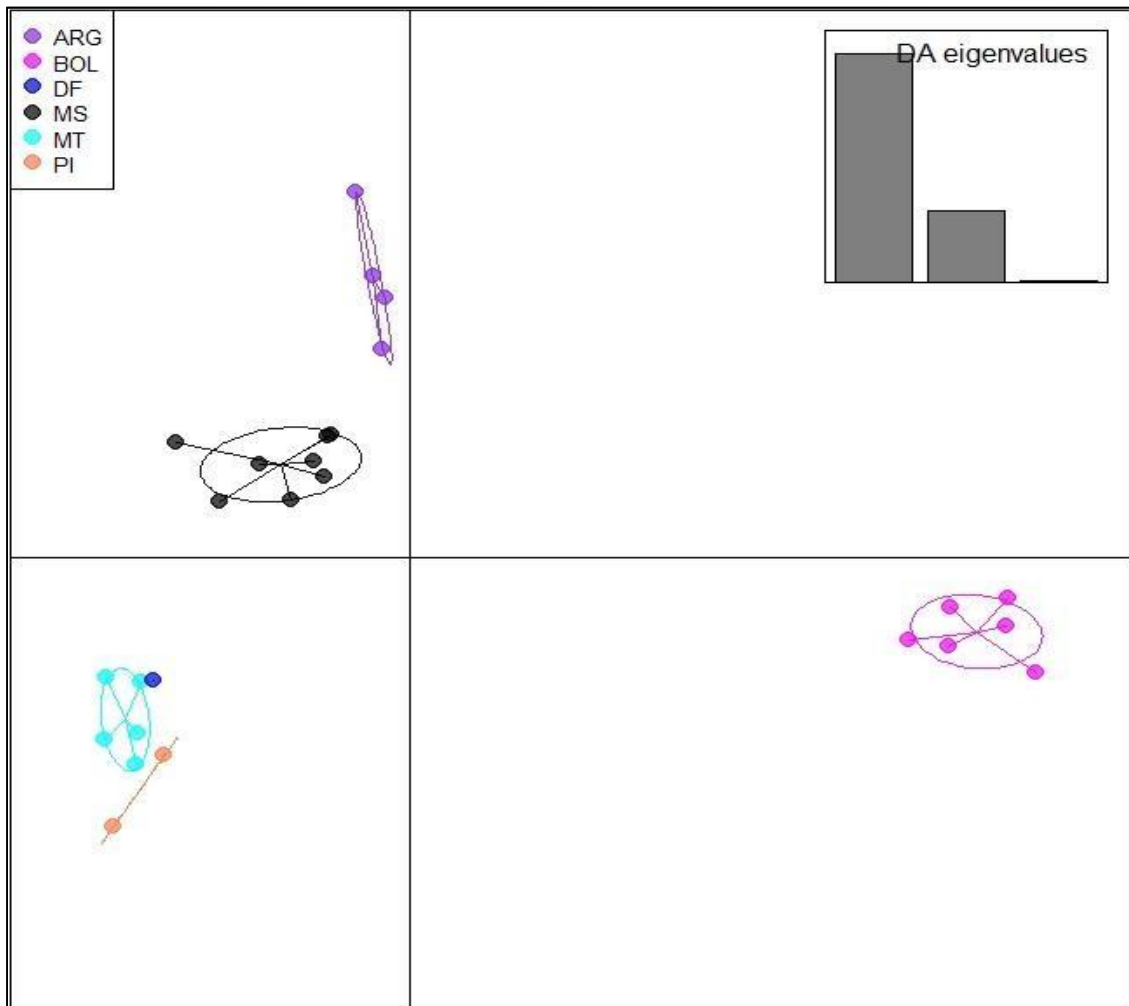


Figura 26 – Análise Discriminante de Componentes Principais (DAPC) da matriz *A. aestiva*. A influência dos eixos da DAPC (quadro no canto inferior esquerdo) são apresentados. A legenda indica as cores dos agrupamentos encontrados (siglas referem-se às localidades e seguem o padrão da Tabela 1).

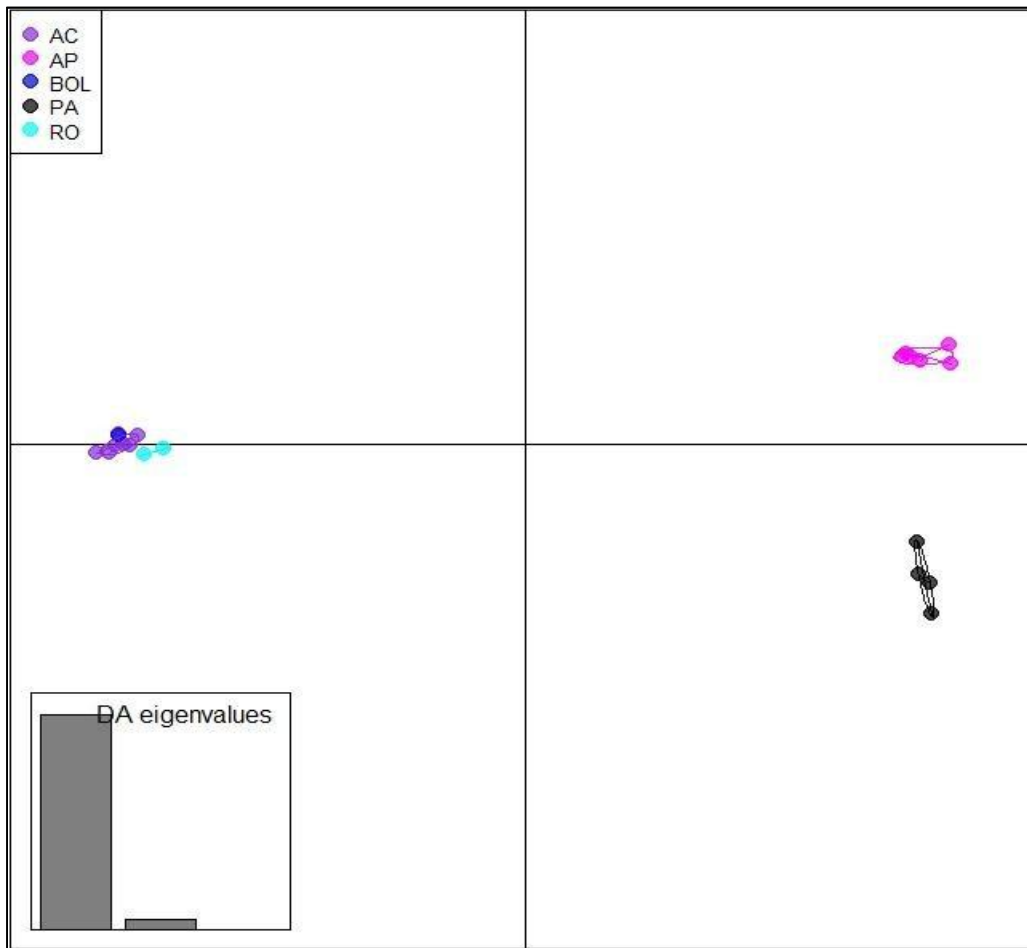
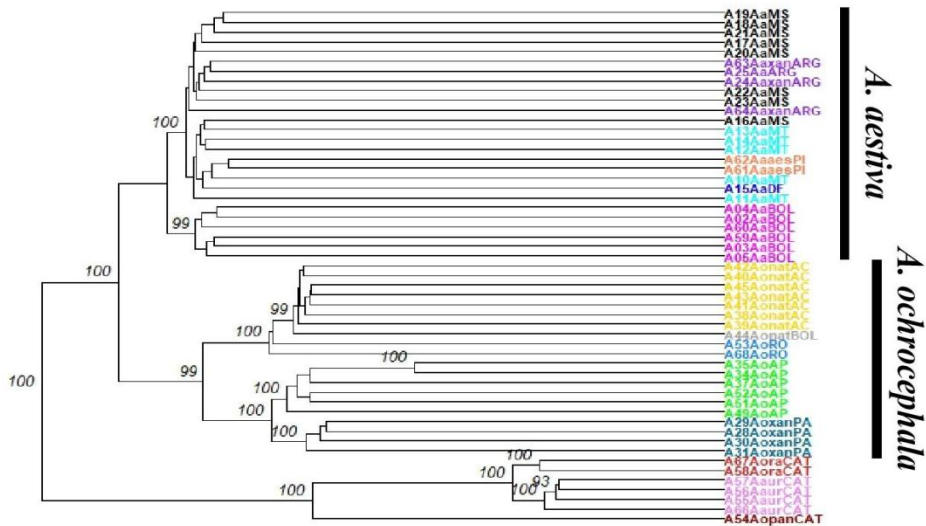


Figura 27 – Análise Discriminante de Componentes Principais (DAPC) da matriz *A. ochrocephala*. A influência dos eixos da DAPC (quadro no canto inferior esquerdo) são apresentados. A legenda indica as cores dos agrupamentos encontrados (siglas referem-se às localidades e seguem o padrão da Tabela 1).



Distância genética (proporção de loci que são diferentes)

Figura 28 - Árvore de UPGMA da matriz *Amazona* spp. Árvore gerada pelo adegenet com base na matriz de distância genética de *Amazona* spp.; valor acima de cada nó representa o suporte em *bootstrap* (somente valores acima de 90); cores agrupam amostras de mesma localidade.

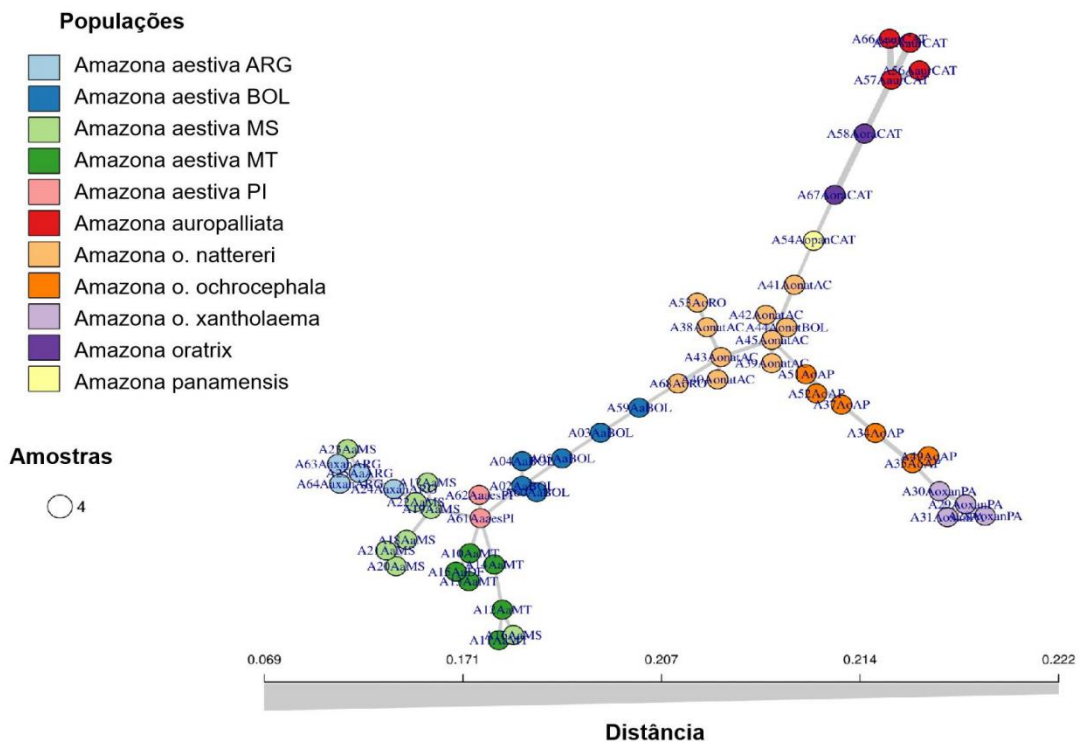


Figura 29 - Rede de Mínima Cobertura (MSN) da matriz *Amazona* spp. Legenda no canto esquerdo superior indica a espécie e localidade de cada amostra seguindo padrão da Tabela 1.

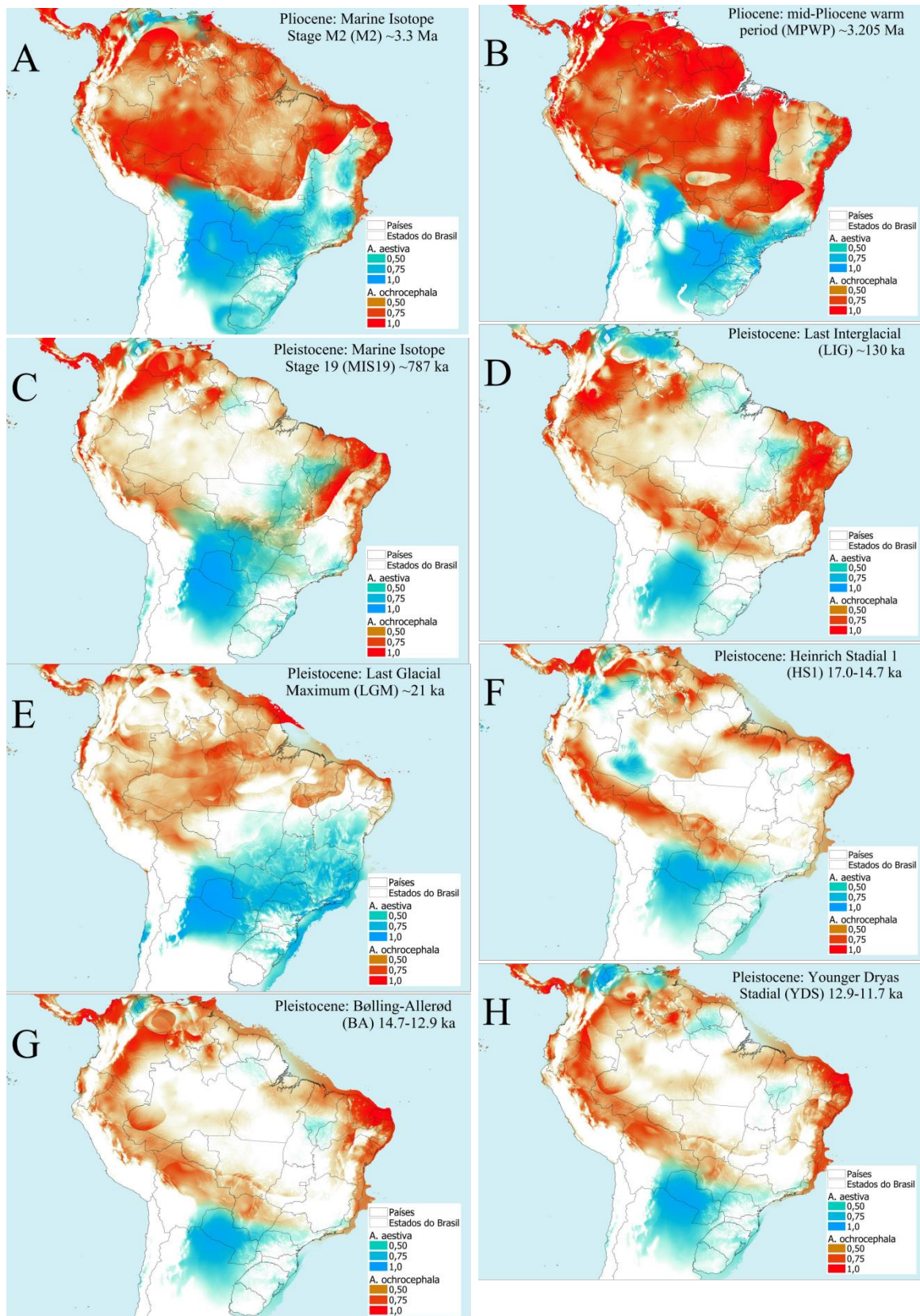
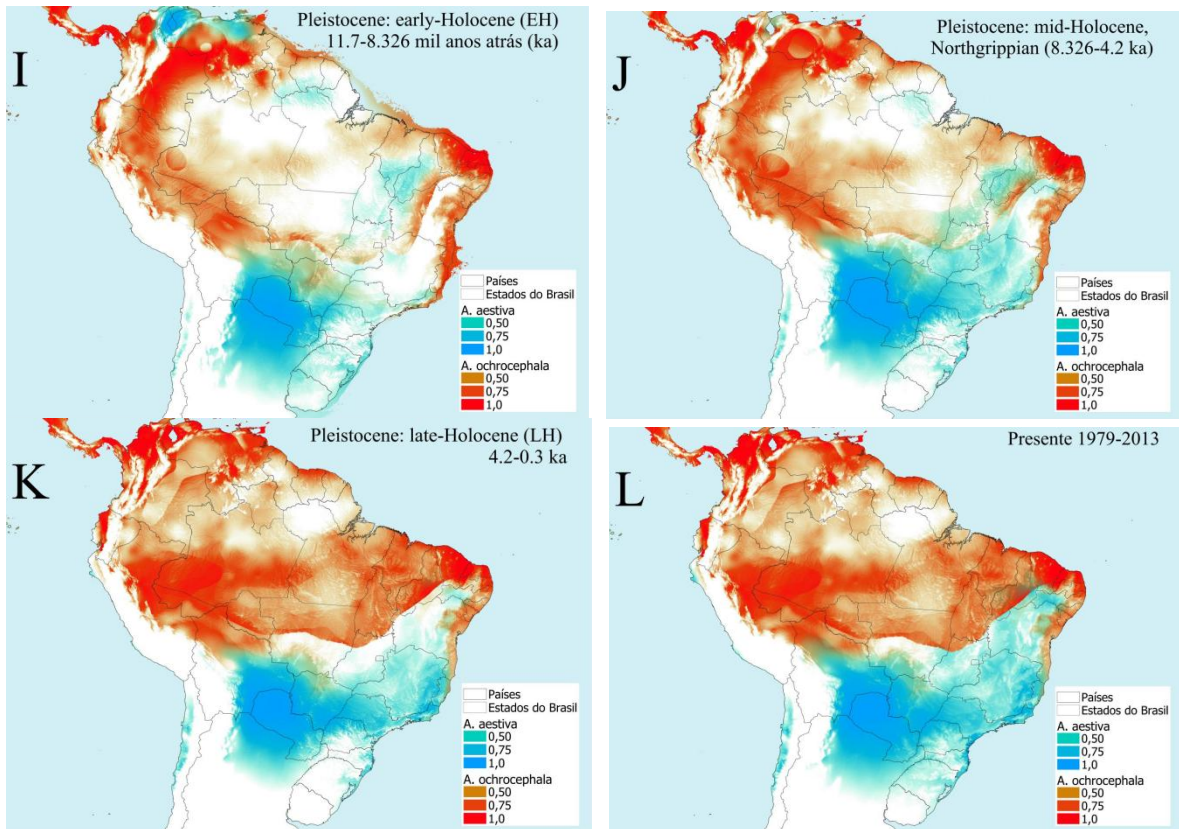


Figura 30 - Modelos de Nicho Ecológico. Áreas de adequabilidade com valor inferior à 0.24 foram removidas. Cores frias (azul à ciano) correspondem à *A. aestiva* e cores quentes (vermelho à laranja), à *A. ochrocephala*. Data e nome da camada de tempo estão apresentadas do lado direito superior de cada modelo.



Continuação da Figura 30 - Modelos de Nicho Ecológico.

Tabela 1 - Amostras utilizadas no estudo. ID: número de registro da amostra no banco de amostras do Laboratório de Genética e Evolução Molecular de Aves (N.T. indica indivíduo Não Tombado na coleção, amostras indicadas com essa sigla estão depositadas na Universidad Nacional del Centro de la Província de Buenos Aires, sob cuidados de Igor Berkunsky); Amostra: nomenclatura da amostra adotada no presente estudo; Sexo: F - fêmea, M - macho, - não conhecido; País: BRA - Brasil, ARG - Argentina, BOL - Bolívia; Estado: BE - Bení, NS - Nicolás Suárez, CH - Chaco, MS - Mato Grosso do Sul, MT - Mato Grosso, SP - São Paulo, PI - Piauí, RO - Rondônia, PA - Pará, AP - Amapá, DF - Distrito Federal, AC - Acre, BOL - Bolívia, MEX - México; Localidade: cidade de origem da amostra, CAT - Cativoiro; Latitude e Longitude: * - coordenadas aproximadas (amostras com informação de cidade mas sem coordenadas exatas). Amostras sublinhadas foram removidas das análises genéticas finais (ver tópico IV.1 dos Resultados).

Espécie	ID	Amostra	Sexo	País	Estado	Localidade	Latitude	Longitude
<i>A. A. aestiva</i>	9433	A61AaaesPI	-	BRA	PI	Serra das Confusões	-9.222064	-43.489769
<i>A. A. aestiva</i>	9434	A62AaaesPI	-	BRA	PI	Serra das Confusões	-9.222564	-43.489169
<i>A. A. xanthopteryx</i>	10722	A63AaxanARG	-	ARG	CH	Almirante Brown	-25.532301	-62.073276
<i>A. A. xanthopteryx</i>	10718	A64AaxanARG	F	ARG	CH	Almirante Brown	-25.19734	-62.019215
<i>A. a.xanthopteryx</i>	10687	A24AaxanARG	F	ARG	CH	Almirante Brown	-25.576386*	-61.673303
<u><i>A. aestiva</i></u>	<u>N.T.</u>	<u>A01AaBOL</u>	<u>-</u>	<u>BOL</u>	<u>BE</u>	<u>CAT</u>	<u>CAT</u>	<u>CAT</u>
<i>A. aestiva</i>	N.T.	A02AaBOL	-	BOL	BE	San Ramón	-17.149860	-59.743702
<i>A. aestiva</i>	N.T.	A03AaBOL	-	BOL	BE	San Ramón	-17.149760	-59.743802
<i>A. aestiva</i>	N.T.	A04AaBOL	-	BOL	BE	San Ramón	-17.149660	-59.743402
<i>A. aestiva</i>	N.T.	A05AaBOL	-	BOL	BE	San Ramón	-17.149160	-59.743802
<u><i>A. aestiva</i></u>	<u>N.T.</u>	<u>A06AaBOL</u>	<u>-</u>	<u>BOL</u>	<u>BE</u>	<u>CAT</u>	<u>CAT</u>	<u>CAT</u>
<u><i>A. aestiva</i></u>	<u>N.T.</u>	<u>A07AaBOL</u>	<u>-</u>	<u>BOL</u>	<u>BE</u>	<u>CAT</u>	<u>CAT</u>	<u>CAT</u>
<u><i>A. aestiva</i></u>	<u>4787</u>	<u>A08AaSP</u>	<u>M</u>	<u>BRA</u>	<u>SP</u>	<u>Ilha Comprida</u>	<u>-24.866163</u>	<u>-47.739588</u>
<u><i>A. aestiva</i></u>	<u>4952</u>	<u>A09AaMT</u>	<u>F</u>	<u>BRA</u>	<u>MT</u>	<u>Pontes e Lacerda</u>	<u>-15.013227</u>	<u>-59.949676</u>
<i>A. aestiva</i>	6147	A10AaMT	M	BRA	MT	Juruena	-10.334124	-58.498575
<i>A. aestiva</i>	6148	A11AaMT	M	BRA	MT	Juruena	-10.334784	-58.498775
<i>A. aestiva</i>	6149	A12AaMT	M	BRA	MT	Juruena	-10.336784	-58.496775
<i>A. aestiva</i>	6150	A13AaMT	F	BRA	MT	Juruena	-10.335784	-58.496975

<i>A. aestiva</i>	6151	A14AaMT	M	BRA	MT	Juruena	-10.338784	-58.499975
<i>A. aestiva</i>	10643	A15AaDF	F	BRA	DF	Brasília	-15.584422	-47.651451
<i>A. aestiva</i>	10649	A16AaMS	F	BRA	MS	Miranda	-19.954673	-56.305307
<i>A. aestiva</i>	10658	A17AaMS	F	BRA	MS	Miranda	-20.276644	-56.382140
<i>A. aestiva</i>	10664	A18AaMS	F	BRA	MS	Miranda	-19.954573	-56.305207
<i>A. aestiva</i>	10667	A19AaMS	M	BRA	MS	Miranda	-20.276844	-56.382440
<i>A. aestiva</i>	10673	A20AaMS	F	BRA	MS	Aquidauana	-20.467731	-55.851818
<i>A. aestiva</i>	10675	A21AaMS	F	BRA	MS	Aquidauana	-20.467431	-55.851318
<i>A. aestiva</i>	10677	A22AaMS	M	BRA	MS	Aquidauana	-20.467331	-55.851818
<i>A. aestiva</i>	10678	A23AaMS	M	BRA	MS	Aquidauana	-20.467931	-55.851318
<i>A. aestiva</i>	10701	A25AaARG	M	ARG	CH	Almirante Brown	-25.576386	-61.673303
<i>A. aestiva</i>	N.T.	A59AaBOL	-	BOL	BE	San Ramón	-17.149460	-59.743802
<i>A. aestiva</i>	N.T.	A60AaBOL	-	BOL	BE	San Ramón	-17.149160	-59.743202
<u><i>A. aestiva</i></u>	<u>4951</u>	<u>A69AaMT</u>	<u>M</u>	<u>BRA</u>	<u>MT</u>	<u>Pontes e Lacerda</u>	<u>-15.013427</u>	<u>-59.949176</u>
<i>A. auropalliata</i>	9037	A55AaurCAT	-	MEX	CAT	CAT	CAT	CAT
<i>A. auropalliata</i>	9038	A56AaurCAT	-	MEX	CAT	CAT	CAT	CAT
<i>A. auropalliata</i>	9039	A57AaurCAT	-	MEX	CAT	CAT	CAT	CAT
<i>A. auropalliata</i>	9040	A66AaurCAT	-	MEX	CAT	CAT	CAT	CAT
<i>A. o. nattereri</i>	5690	A38AonatAC	M	BRA	AC	Rio Branco	*-10.292197	*-68.816125
<i>A. o. nattereri</i>	5691	A39AonatAC	F	BRA	AC	Rio Branco	*-10.292897	*-68.816725
<i>A. o. nattereri</i>	5693	A40AonatAC	F	BRA	AC	Rio Branco	*-10.046486	*-69.655751
<i>A. o. nattereri</i>	5703	A41AonatAC	F	BRA	AC	Sena Madureira	*-9.395288	*-69.565946
<i>A. o. nattereri</i>	5704	A42AonatAC	F	BRA	AC	Feijó	*-8.201260	*-70.370993
<i>A. o. nattereri</i>	5705	A43AonatAC	M	BRA	AC	Rio Branco	*-10.998721	*-68.763510
<i>A. o. nattereri</i>	5707	A44AonatBOL	M	BOL	NS	Cobja	*-11.023308	*-68.799998
<i>A. o. nattereri</i>	5712	A45AonatAC	F	BRA	AC	Xapuri	*-10.660689	*-68.499517
<u><i>A. o. ochrocephala</i></u>	<u>4530</u>	<u>A27AoochCAT</u>	<u>M</u>	<u>BRA</u>	<u>CAT</u>	<u>CAT</u>	<u>CAT</u>	<u>CAT</u>
<u><i>A. o. ochrocephala</i></u>	<u>4848</u>	<u>A32AoochCAT</u>	<u>M</u>	<u>BRA</u>	<u>CAT</u>	<u>CAT</u>	<u>CAT</u>	<u>CAT</u>

<i>A. o. ochrocephala</i>	4850	A33AoochCAT	F	BRA	CAT	CAT	CAT	CAT
<i>A. o. ochrocephala</i>	6029	A46AoochCAT	M	BRA	CAT	CAT	CAT	CAT
<i>A. o. ochrocephala</i>	6419	A47AoochCAT	F	BRA	CAT	CAT	CAT	CAT
<i>A. o. ochrocephala</i>	4849	A48AoochCAT	M	BRA	CAT	CAT	CAT	CAT
<i>A. o. ochrocephala</i>	7312	A65AoochCAT	F	BRA	CAT	CAT	CAT	CAT
<i>A. o. panamensis</i>	6030	A54AopanCAT	M	MEX	CAT	CAT	CAT	CAT
<i>A. o. xantholaema</i>	4299	A26AoxanCAT	M	BRA	CAT	CAT	CAT	CAT
<i>A. o. xantholaema</i>	4654	A28AoxanPA	M	BRA	PA	Cachoeira do Arari. Ilha de Marajó	*-0.932035	*-48.878675
<i>A. o. xantholaema</i>	4655	A29AoxanPA	F	BRA	PA	Bacabal. Ilha de Marajó	*-0.794835	*-48.565219
<i>A. o. xantholaema</i>	4656	A30AoxanPA	M	BRA	PA	Salvaterra. Ilha de Marajó	*-0.822488	*-48.678771
<i>A. o. xantholaema</i>	4669	A31AoxanPA	M	BRA	PA	Ponta de Pedras. Ilha de Marajó	*-1.387683	*-48.862410
<i>A. ochrocephala</i>	5188	A34AoAP	M	BRA	AP	Santana	*0.063706	*-51.459316
<i>A. ochrocephala</i>	5189	A35AoAP	M	BRA	AP	Santana	*0.315145	*-51.571657
<i>A. ochrocephala</i>	5190	A36AoAP	F	BRA	AP	Santana	CAT	CAT
<i>A. ochrocephala</i>	5192	A37AoAP	M	BRA	AP	Santana	*-0.055501	*-51.225443
<i>A. ochrocephala</i>	5193	A49AoAP	M	BRA	AP	Macapá	*0.077213	*-51.040888
<i>A. ochrocephala</i>	5194	A50AoAP	M	BRA	AP	Macapá	*0.077513	*-51.040288
<i>A. ochrocephala</i>	5197	A51AoAP	M	BRA	AP	Macapá	*0.077413	*-51.040888
<i>A. ochrocephala</i>	5201	A52AoAP	M	BRA	AP	Macapá	*0.077113	*-51.040988
<i>A. ochrocephala</i>	11914	A53AoRO	-	BRA	RO	Campo Novo de Rondônia	*-10.678704	*-63.629165
<i>A. ochrocephala</i>	11920	A68AoRO	-	BRA	RO	Campo Novo de Rondônia	*-10.573235	*-63.627354
<i>A. oratrix</i>	9044	A58AoraCAT	-	MEX	CAT	CAT	CAT	CAT
<i>A. oratrix</i>	9045	A67AoraCAT	-	MEX	CAT	CAT	CAT	CAT

Tabela 2 - Parâmetros utilizados no Ipyrad para o processamento das *reads*.

Parâmetros inseridos no Ipyrad (Parâmetro utilizado); Parâmetro: título do parâmetro no *software*; Descrição: função do parâmetro;

Parâmetro	Descrição	Parâmetro utilizado
[8] [restriction_overhang]	Sequência do sítio da enzima de restrição utilizada na construção das bibliotecas.	TGCAG
[9] [max_low_qual_bases]	Número máximo de bases com qualidade baixa (Q<20) em uma <i>read</i> para ser mantido na análise	5
[10] [phred_Qscore_offset]	Score de qualidade Phred	33
[11] [mindepth_statistical]	Profundidade mínima de cobertura para considerar uma base nas análises subsequentes	6
[12] [mindepth_majrule]	Profundidade mínima de cobertura em uma base para ela ser considerada a base mais frequente	6
[13] [maxdepth]	Profundidade máxima permitida para uma base, descartando <i>reads</i> com cobertura excessiva.	10000
[14] [clust_threshold]	Limite de diferença entre <i>reads</i> para agrupá-las em um <i>cluster</i> durante o processo de montagem.	0.85
[15] [max_barcode_mismatch]	Número máximo de erros de correspondência permitidos ao comparar sequências de <i>barcodes</i> .	0
[16] [filter_adapters]	Ativa ou desativa a remoção de adaptadores de sequenciamento das <i>reads</i>	2
[17] [filter_min_trim_len]	Comprimento mínimo das <i>reads</i> após o processo de corte (<i>trimming</i>).	35
[18] [max_alleles_consens]	Limita o número máximo de alelos permitidos no consenso de <i>reads</i> .	2
[19] [max_Ns_consens]	Número máximo de bases "N" permitidas no consenso de <i>reads</i> .	0.05
[20] [max_Hs_consens]	Número máximo de bases ambíguas (heterozigotas) permitidas no consenso de <i>reads</i> .	0.05
[21] [min_samples_locus]	Número mínimo de amostras que devem ter dados em um loco para que seja considerado válido.	4
[22] [max_SNPs_locus]	Número máximo de SNPs permitidos em um loco durante a chamada de variantes.	0.2
[23] [max_Indels_locus]	Número máximo de inserções ou deleções permitidas em um loco durante a chamada de variantes.	8
[24] [max_shared_Hs_locus]	Número máximo de heterozigotos compartilhados permitidas em um loco para ser considerado válido.	0.5
[25] [trim_reads]	Ativa ou desativa o corte (<i>trimming</i>) de bases de baixa qualidade das <i>reads</i> .	0. 0. 0. 0
[26] [trim_loci]	Realiza o corte (<i>trimming</i>) das bordas dos loci	0. 0. 0. 0

Tabela 3 - Combinações de fatores de frequência e contagem alélica. *Minor Allele Frequency* (MAF) e *Minor Allele Count* (MAC) utilizados nas filtragens e a quantidade de SNPs obtidos em cada uma das filtragens com as diferentes combinações de valores de MAF e MAC testadas no VCFtools.

Matriz (geral)	Matriz (submatrizes)	MAF	MAC	Número de SNPs
<i>Amazona</i> spp.	1	0.05	1	6940
<i>Amazona</i> spp.	2	0.10	1	4426
<i>Amazona</i> spp.	3	0.15	1	3063
<i>Amazona</i> spp.	4	0.05	2	6940
<i>Amazona</i> spp.	5	0.10	2	4426
<i>Amazona</i> spp.	6	0.15	2	3063
<i>Amazona</i> spp.	7	0.05	3	6940
<i>Amazona</i> spp.	8	0.10	3	4426
<i>Amazona</i> spp.	9	0.15	3	3063
<i>A. aestiva</i> + <i>A. ochrocephala</i>	1	0.05	1	8564
<i>A. aestiva</i> + <i>A. ochrocephala</i>	2	0.10	1	4940
<i>A. aestiva</i> + <i>A. ochrocephala</i>	3	0.15	1	3718
<i>A. aestiva</i> + <i>A. ochrocephala</i>	4	0.05	2	8564
<i>A. aestiva</i> + <i>A. ochrocephala</i>	5	0.10	2	4940
<i>A. aestiva</i> + <i>A. ochrocephala</i>	6	0.15	2	3718
<i>A. aestiva</i> + <i>A. ochrocephala</i>	7	0.05	3	8564
<i>A. aestiva</i> + <i>A. ochrocephala</i>	8	0.10	3	4940
<i>A. aestiva</i> + <i>A. ochrocephala</i>	9	0.15	3	3718
<i>A. aestiva</i>	1	0.05	1	14476
<i>A. aestiva</i>	2	0.10	1	8242
<i>A. aestiva</i>	3	0.15	1	6186
<i>A. aestiva</i>	4	0.05	2	14476
<i>A. aestiva</i>	5	0.10	2	8242
<i>A. aestiva</i>	6	0.15	2	6186
<i>A. aestiva</i>	7	0.05	3	14476
<i>A. aestiva</i>	8	0.10	3	8242
<i>A. aestiva</i>	9	0.15	3	6186
<i>A. ochrocephala</i>	1	0.05	1	16372
<i>A. ochrocephala</i>	2	0.10	1	11077
<i>A. ochrocephala</i>	3	0.15	1	9422
<i>A. ochrocephala</i>	4	0.05	2	16372
<i>A. ochrocephala</i>	5	0.10	2	11077
<i>A. ochrocephala</i>	6	0.15	2	9422
<i>A. ochrocephala</i>	7	0.05	3	16372
<i>A. ochrocephala</i>	8	0.10	3	11077
<i>A. ochrocephala</i>	9	0.15	3	9422

Tabela 4 - Parâmetros do VCFtools utilizados para filtragem final dos SNPs no VCFtools para todas as matrizes.

Parâmetro	Valor utilizado
MAC - <i>Minor allele count</i> (Quantidade mínima de indivíduos com um alelo para manter o loco correspondente na matriz de dados)	2
MAF - <i>Minor allele frequency</i> (Frequência mínima do alelo para manter o loco correspondente na matriz de dados)	0.05
Max missing data (varia de 0 à 1; 0 indica 100% de ausência de informação genética e 1 indica 0% de ausência de informação genética)	1
Max allele number - Quantidade máxima de alelos por loco para manter o loco correspondente na matriz de dados	2

Tabela 5 - Quantidade de reads e SNPs ao longo do processamento. Quantidade total de reads obtidas em cada etapa do processamento.

Matriz	reads pós-demultiplexing	SNPs pós-Ipyrad	SNPs pós-VCFtools
<i>Amazona spp.</i>	462794689	818664	7080
<i>A. aestiva</i> + <i>A. ochrocephala</i>	421559195	757080	8564
<i>A. aestiva</i>	222531161	555936	14476
<i>A. ochrocephala</i>	185636266	358391	22127

Tabela 6 - Valores de estatísticas sumárias das espécies estudadas. Valores estimados de heterozigidade esperada e observada (H_s e H_o), índice de diferenciação genética global entre as subpopulações (F_{ST}), índice de endocruzamento médio na espécie (F_{IS}). Os valores de F_{ST} foram estimados considerando os agrupamentos identificados na análise de estrutura populacional (*A. aestiva*: 4 populações - PI, MS+ARG, MT+DF e BOL; *A. ochrocephala*: 3 populações - PA, AP e AC+RO+BOL. Para considerar um indivíduo pertencente à uma população, foi considerado um mínimo de valor de q de 50% na análise de estrutura (Fig. 18). Os indivíduos de *A. aestiva* da Bolívia foram considerados um grupo à parte devido à grande proporção de ancestralidade de *A. ochrocephala*, o que poderia afetar a análise.

Espécie	H_s	H_o	F_{st}	F_{is}
<i>A. aestiva</i>	0.2479	0.2298	0.0198	0.0732
<i>A. ochrocephala</i>	0.2317	0.2085	0.0603	0.1003

Tabela 7 - Valores de F_{ST} entre pares de táxons/grupos. "América Central" refere-se ao grupo *A. auropalliata*+*A. oratrix*+*A. panamensis*. "América do Sul" refere-se ao grupo *A. aestiva*+*A. ochrocephala*. Tabela organizada em ordem decrescente do valor de F_{ST} .

Par	F_{ST} médio
<i>A. aestiva</i> - América Central	0.1760
<i>A. aestiva</i> - <i>A. auropalliata</i>	0.1688
<i>A. ochrocephala</i> - América Central	0.1585
América do Sul - América Central	0.1581
<i>A. ochrocephala</i> - <i>A. auropalliata</i>	0.1535
<i>A. aestiva</i> - <i>A. oratrix</i>	0.1120
<i>A. ochrocephala</i> - <i>A. oratrix</i>	0.1000
<i>A. aestiva</i> - <i>A. ochrocephala</i>	0.0775
<i>A. ochrocephala</i> PA - <i>A. ochrocephala</i> AP	0.0575
<i>A. ochrocephala</i> PA - <i>A. ochrocephala</i> AC+RO+BOL	0.1341
<i>A. ochrocephala</i> AP - <i>A. ochrocephala</i> AC+RO+BOL	0.1289
<i>A. aestiva</i> MT+DF+PI - <i>A. ochrocephala</i> PA	0.1674
<i>A. aestiva</i> MT+DF+PI - <i>A. ochrocephala</i> AP	0.1785
<i>A. aestiva</i> MT+DF+PI - <i>A. ochrocephala</i> AC+RO+BOL	0.1432
<i>A. aestiva</i> MT+DF+PI - <i>A. aestiva</i> MS+ARG	0.0142
<i>A. aestiva</i> MS+ARG - <i>A. ochrocephala</i> PA	0.1894
<i>A. aestiva</i> MS+ARG - <i>A. ochrocephala</i> AP	0.1975
<i>A. aestiva</i> MS+ARG - <i>A. ochrocephala</i> AC+RO+BOL	0.1583
<i>A. aestiva</i> BOL - <i>A. ochrocephala</i> PA	0.1364
<i>A. aestiva</i> BOL - <i>A. ochrocephala</i> AP	0.1433
<i>A. aestiva</i> BOL - <i>A. ochrocephala</i> AC+RO+BOL	0.0708
<i>A. aestiva</i> BOL - <i>A. aestiva</i> MT+DF+PI	0.0391
<i>A. aestiva</i> BOL - <i>A. aestiva</i> MS+ARG	0.0368