

Universidade Federal de Goiás  
Instituto de Ciências Biológicas  
Programa de Pós-graduação em Genética e Biologia Molecular

**MARCOS VINÍCIUS PEREIRA PINTO**

**DIVERSIDADE GENÉTICA, ESTRUTURA  
POPULACIONAL E HISTÓRIA DEMOGRÁFICA EM  
*Dipteryx alata* Schreb (LEGUMINOSAE)**

Goiânia, setembro de 2025



UNIVERSIDADE FEDERAL DE GOIÁS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS

## TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

### E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

#### 1. Identificação do material bibliográfico

Dissertação     Tese     Outro\*: \_\_\_\_\_

\*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

#### 2. Nome completo do autor

Marcos Vinícius Pereira Pinto

#### 3. Título do trabalho

Diversidade genética, estrutura populacional e história demográfica em *Dipteryx alata* Schreb (Leguminosae)

#### 4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento  SIM     NÃO<sup>1</sup>

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(à) autor(a) e ao(à) orientador(a);
- b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação. O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

**Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.**



Documento assinado eletronicamente por **Thannya Nascimento Soares, Professora do Magistério Superior**, em 24/06/2025, às 15:12, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Marcos Vinícius Pereira Pinto, Usuário Externo**, em 25/06/2025, às 17:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4726759** e o código CRC **9265E139**.

---

Universidade Federal de Goiás  
Instituto de Ciências Biológicas  
Programa de Pós-graduação em Genética e Biologia Molecular

MARCOS VINÍCIUS PEREIRA PINTO

**DIVERSIDADE GENÉTICA, ESTRUTURA  
POPULACIONAL E HISTÓRIA DEMOGRÁFICA EM  
*Dipteryx alata* Schreb (LEGUMINOSAE)**

Tese apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular, do Instituto de Ciências Biológicas, da Universidade Federal de Goiás (UFG), como requisito para obtenção do título de Doutor em Genética e Biologia Molecular.

Área de concentração: Genética e Biologia Molecular

Orientador(a): Professor (a) Doutora (a) Thannya Nascimento Soares

Goiânia, setembro de 2025

Pinto, Marcos Vinícius Pereira

Diversidade Genética, estrutura populacional e história demográfica em *Dipteryx alata* Schreb (Leguminosae) [manuscrito] / Marcos Vinícius Pereira Pinto. - 2025.  
LV, 55 f.: il.

Orientador: Profa. Dra. Thannya Nascimento Soares.

Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Ciências Biológicas (ICB), Programa de Pós-Graduação em Genética e Biologia Molecular, Goiânia, 2025.

Bibliografia. Apêndice.

Inclui lista de figuras, lista de tabelas.

1. *Dipteryx alata*. 2. Filogeografia. 3. Cerrado. I. Soares, Thannya Nascimento, orient. II. Título.

CDU 575



UNIVERSIDADE FEDERAL DE GOIÁS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS

**ATA DE DEFESA DE TESE**

Ata Nº 57 da sessão de Defesa de Tese de **Marcos Vinícius Pereira Pinto** que confere o título de Doutor(a) em **Genética e Biologia Molecular**, na área de concentração em **Genética e Biologia Molecular**.

Ao/s **vinte e oito dias do mês de junho de dois mil e vinte e quatro**, a partir da(s) **13h30**, no(a) **Sala INCT ICB I**, realizou-se a sessão pública de Defesa de Tese intitulada "**Diversidade genética, estrutura populacional e história demográfica em *Dipteryx alata* Schreb (Leguminosae)**". Os trabalhos foram instalados pelo(a) Orientador(a), Professor(a) Doutor(a) **Thannya Nascimento Soares (ICB/UFG)** com a participação dos demais membros da Banca Examinadora: Professor(a) Doutor(a) **Rafael Barbosa Pinto (IACSB/UEG)**, membro titular externo; Professor(a) Doutor(a) **Cíntia Pelegrineti Targueta de Azevedo Brito (HV/EVZ/UFG)**, membro titular externo; Professor(a) Doutor(a) **Renata de Oliveira Dias (ICB/UFG)**, membro titular interno; Professor(a) Doutor(a) **Ramilla dos Santos Braga (ICEN/UFR)**, membro titular externo. Durante a argüição os membros da banca **não fizeram** sugestão de alteração do título do **trabalho**. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese tendo sido(a) o(a) candidato(a) **aprovado** pelos seus membros. Proclamados os resultados pelo(a) Professor(a) Doutor(a) **Thannya Nascimento Soares (ICB/UFG)**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, ao(s) **vinte e oito dias do mês de junho de dois mil e vinte e quatro**.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Thannya Nascimento Soares, Professora do Magistério Superior**, em 29/06/2024, às 19:56, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Renata De Oliveira Dias, Professora do Magistério Superior**, em 29/06/2024, às 20:47, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Cíntia Pelegrineti Targueta de Azevedo Brito, Usuário Externo**, em 15/07/2024, às 08:54, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ramilla Dos Santos Braga, Usuário Externo**, em 02/08/2024, às 14:44, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rafael Barbosa Pinto, Usuário Externo**, em 02/08/2024, às 15:09, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4633792** e o código CRC **D2C79FDF**.

**Referência:** Processo nº 23070.029826/2024-16

SEI nº 4633792

Dedico esse trabalho aos meus pais

Nemias (*in memoriam*) e Eunice

Com todo meu amor e gratidão

Por tudo que fizeram por mim ao longo da minha vida.

## **AGRADECIMENTOS**

Agradeço a todos que contribuíram direta ou indiretamente para que esse trabalho fosse realizado:

Ao Instituto Nacional de Ciência e Tecnologia em Ecologia, Evolução e Conservação da Biodiversidade (INCT – EECBio), no contexto do grupo de trabalho de Genética e Genômica Evolutiva, subsidiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – processo 465610/20145) e Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG – processo 201810267000023), que foram imprescindíveis para o desenvolvimento do trabalho.

À minha orientadora Prof. Dra. Thannya Nascimento Soares por ter me recebido e mostrado o caminho correto a ser seguido, tanto na condução desse trabalho quanto na minha vida pessoal caótica, sempre de forma assertiva, paciente e demasiadamente empática. Professora, a senhora se tornou uma referência de ser humano e profissional na minha vida, minha gratidão será eterna!

Aos Profs. Drs. e Dras. membros e suplentes da banca de avaliação, por aceitarem prontamente em participar e contribuir com este trabalho

Ao Prof. Dr. Rafael Barbosa Pinto e a Prof. Dra. Renata de Oliveira Dias pelas contribuições com o desenvolvimento do trabalho ao longo dos comitês de acompanhamento. Estendo meu agradecimento à professora Renata por ter me acolhido no estágio docência com muita paciência e empatia.

A Prof. Dra. Levi Carina Terribile do Laboratório de Macroecologia da Universidade Federal de Jataí, pela valiosa contribuição com as modelagens de nicho que agregaram poder explanatório nesse trabalho.

A todos os Profs. do programa de Genética e Biologia Molecular e de programas externos que contribuíram com minha formação ao longo das disciplinas da pós-graduação.

Ao Laboratório de Genética & Biodiversidade e toda sua equipe pela estrutura física e suporte, imprescindíveis para a execução desse trabalho.

Aos meus pais por nunca medirem esforços para me proporcionar a melhor educação possível.

Aos meus amigos Eleonora, Vitória e Wanderson que, mesmo à distância e nos breves momentos que a correria do dia a dia permitiu, sempre me incentivaram a não desistir da caminhada.

Aos meus colegas Profs. na educação básica que contribuíram com substituições e coberturas que foram essenciais para a execução desse trabalho.

# Sumário

<b>RESUMO</b> .....	15
<b>ABSTRACT</b> .....	16
<b>INTRODUÇÃO</b> .....	17
<b>OBJETIVOS</b> .....	20
<b>HIPÓTESES</b> .....	21
<b>MATERIAL E MÉTODOS</b> .....	22
<b>ÁREA DE ESTUDO E AMOSTRAGEM</b> .....	22
<b>DADOS MOLECULARES</b> .....	23
<b>ANÁLISE DOS DADOS</b> .....	25
<b>Diversidade, estrutura e divergência genética</b> .....	26
<b>Padrão espacial na diversidade genética</b> .....	26
<b>História demográfica e padrões filogeográficos</b> .....	27
<b>MODELAGEM DE NICHOS</b> .....	28
<b>RESULTADOS</b> .....	30
<b>CONCLUSÕES</b> .....	44
<b>REFERÊNCIAS</b> .....	45
<b>APÊNDICES</b> .....	50

## LISTA DE TABELAS

<b>Tabela 1.</b> Populações de <i>Dipteryx alata</i> amostradas para obtenção dos dados moleculares. Goiás (GO), Minas Gerais (MG), Mato Grosso (MT), Mato Grosso do Sul (MS), São Paulo (SP) e Tocantins (TO). .....	22
<b>Tabela 2.</b> Caracterização genética de 25 populações de <i>Dipteryx alata</i> no Cerrado brasileiro, baseada em dados de cpDNA (psbA-trnH) e nDNA (ITS101/102). N = tamanho amostral; k = número de haplótipos; h = diversidade haplotípica; $\pi$ = diversidade nucleotídica; SD = desvio padrão. ....	30
<b>Tabela 3.</b> Análise de variância molecular (AMOVA) para 25 populações de <i>Dipteryx alata</i> no Cerrado brasileiro, obtida a partir das sequências cloroplastidiais (psbA-trnH). g.l.: graus de liberdade; SSD: soma do quadrado dos desvios. ....	33
<b>Tabela 4.</b> Análise de variância molecular (AMOVA) para 25 populações de <i>Dipteryx alata</i> no Cerrado brasileiro, obtida a partir das sequências nucleares (ITS101/102). g.l.: graus de liberdade; SSD: soma do quadrado dos desvios. ....	34
<b>Tabela 5.</b> Testes de neutralidade para as regiões cloroplastidial e nuclear. D: valor de Tajima; FS: valor de Fu; p: significância. ....	36
<b>Tabela 6.</b> Parâmetros demográficos de 25 populações de <i>Dipteryx alata</i> , obtidos a partir de dados de cpDNA e nDNA concatenados, baseados em coalescência. $\theta$ - parâmetro de coalescência, g - parâmetro de crescimento exponencial; IC - intervalo de confiança. .	38

## LISTA DE FIGURAS

- Figura 1.** Pontos de amostragem das populações de *Dipteryx alata* que foram utilizadas para coleta de dados moleculares. A área preenchida em cinza representa o Cerrado... 23
- Figura 2.** Rede de haplótipos (Median-joining network) de 25 populações de *Dipteryx alata* no Cerrado brasileiro, obtida a partir de dados de sequências de região cloroplastidial (psbA-trnH). ..... 31
- Figura 3.** Distribuição geográfica dos haplótipos da região cloroplastidial (psbA-trnH) nas 25 populações de *Dipteryx alata*. ..... 32
- Figura 4.** Rede de haplótipos (Median-joining network) de 25 populações de *Dipteryx alata* no Cerrado brasileiro, obtida a partir de dados de sequências de região nuclear (ITS101/102). ..... 33
- Figura 5.** Distribuição geográfica dos haplótipos da região nuclear (ITS101/102) nas 25 populações de *Dipteryx alata*. ..... 33
- Figura 6.** Agrupamento tipo UPGMA baseado nos resultados de  $F_{ST}$  par-a-par entre 25 populações de *Dipteryx alata*, obtidos a partir das sequências cloroplastidiais (A) e nucleares (B). Destacar os grupos ..... 35
- Figura 7.** Correlogramas de Mantel entre matrizes de  $F_{ST}$  par-a-par e distância geográfica de 25 populações naturais de *Dipteryx alata* no Cerrado brasileiro, obtidos a partir de sequências cloroplastidiais (psbA-trnH) (A) e nucleares (ITS101/102) (B). Os quadrados preenchidos representam os valores significativos ( $p < 0,05$ ). ..... 36
- Figura 8.** Distribuição mismatch das sequências da região cloroplastidial (psbA-trnH) (A) e região nuclear (ITS101/102) (B). A linha vermelha representa a frequência das diferenças de bases par-a-par e a linha verde representa os valores esperados pelo modelo de expansão demográfica..... 37
- Figura 9.** Bayesian skyline plot para 25 populações de *Dipteryx alata* no cerrado brasileiro, obtidos a partir de sequências das regiões cloroplastidial (psbA-trnH) e nuclear (ITS101/102) concatenadas. A linha do meio representa os valores médios e as linhas superior e inferior representam o intervalo de confiança a 95%. ..... 37
- Figura 10.** Tempo para o ancestral comum mais recente dos haplótipos de *Dipteryx alata*, baseadas nas análises de coalescência de todos os fragmentos concatenados (cpDNA e nDNA) de 296 indivíduos provenientes de 25 populações naturais. Os números abaixo dos ramos correspondem ao suporte do nó (probabilidade a posteriori) e os números acima

dos ramos representam a datação do nó (tempo para o ancestral comum). As barras em azul correspondem ao intervalo de credibilidade a 95%, da média do tempo do ancestral comum. A escala está em milhões de anos (Ma) antes do presente. .... 39

**Figura 11.** Projeção de distribuição para *Dipteryx alata* em três períodos: 21 Ka (A), 6 Ka (B) e presente (C). .... 40

**Figura 12.** Projeção de adequabilidade ambiental para *Dipteryx alata* em três períodos: 21 Ka (A), 6 Ka (B) e presente (C). .... 40

## RESUMO

A filogeografia é um campo integrativo da ciência que busca estudar a distribuição geográfica de linhagens em nível de espécie ou entre espécies próximas, utilizando informações genéticas. O campo desenvolveu métodos que permitem testar hipóteses biogeográficas, descrever a evolução de isolamento reprodutivo e inferir processos responsáveis pela origem, distribuição e manutenção da biodiversidade. O baru (*Dipteryx alata*) é uma planta arbórea que pertence à família das leguminosas e está amplamente distribuída no Cerrado brasileiro. Do ponto de vista genético, as populações de baru estão estruturadas e apresentam baixos níveis de diversidade. Esses padrões parecem ser explicados pela restrição de distribuição ocorrida durante o último glacial seguida de expansão, no entanto, ainda existem lacunas de conhecimento sobre a dinâmica de diversificação das linhagens durante o processo de colonização ao longo do Cerrado. Nesse contexto, foram utilizadas sequências de DNA nuclear e cloroplastidial provenientes de 296 indivíduos (populações) amplamente distribuídos no Cerrado para combinar análises filogeográficas com modelagem ecológica de nicho com o intuito de elucidar os processos de diversificação das linhagens de *D. alata* que resultaram no atual padrão de diversidade observado na espécie. Nossos resultados reforçam que as populações de *D. alata* apresentam baixa diversidade e níveis significativos de estrutura genética. Esses padrões parecem ser resultado de uma redução da área de ocorrência durante o último glacial máximo, com manutenção de uma área estável abrangendo o sudoeste e centro-oeste do Cerrado, seguida de expansão para as regiões sudeste, leste e nordeste durante o holoceno médio. Nesse contexto, as populações localizadas nas regiões sudoeste e oeste do Cerrado parecem representar as linhagens mais antigas de *D. alata*.

**Palavras-chave:** *Dipteryx alata*; Filogeografia; Cerrado;

## ABSTRACT

Phylogeography is an integrative field of science that aims to study the geographic distribution of lineages at the species level or among closely related species, using genetic information. The field has developed methods that allow testing biogeographic hypotheses, describing the evolution of reproductive isolation, and inferring processes responsible for the origin, distribution, and maintenance of biodiversity. Baru (*Dipteryx alata*) is a tree plant that belongs to the legume family and is widely distributed in the Brazilian Cerrado, covering eleven states and the Federal District. From a genetic point of view, baru populations are structured and present low levels of diversity. These patterns seem to be explained by the distribution restriction that occurred during the last glacial followed by expansion, however, there are still gaps in knowledge about the dynamics of lineage diversification during the expansion process throughout the Cerrado. In this context, nuclear and chloroplast DNA sequences from 296 individuals widely distributed in the Cerrado were used to combine phylogeographic analyzes with ecological niche modeling to elucidate the diversification processes of *D. alata* lineages that resulted in the current diversity pattern observed in the species. Our results reinforce that *D. alata* populations present low diversity and significant levels of genetic structure. These patterns appear to be the result of a reduction in occurrence during the last glacial maximum, with maintenance of a stable area covering the southwest and central west of the Cerrado, followed by expansion to the southeast, east and northeast regions during the middle Holocene. In this context, populations located in the southwest and west regions of the Cerrado appear to represent the oldest lineages of *D. alata*.

**Keywords:** *Dipteryx alata*; Phylogeography; Cerrado

## INTRODUÇÃO

O baru (*Dipteryx alata*, Vogel) é uma planta arbórea que pertence à família das leguminosas. A espécie apresenta interesse econômico, sendo principalmente utilizada nos contextos alimentícios, medicinais e de construção civil, devido ao alto valor nutritivo de sua semente e a qualidade da sua madeira. Além disso, é recomendada para uso em sistemas de integração floresta-pecuária e recuperação de áreas degradadas, em razão de seu uso como abrigo para o gado, do valor energético e nutricional dos frutos, da manutenção da qualidade da forragem e pela alta produção de massa foliar (Sano et al., 2004).

A espécie é nativa, mas não endêmica do Brasil, alcançando também países vizinhos como Paraguai, Peru e Bolívia. No Brasil está amplamente distribuída no bioma Cerrado, podendo ser encontrada nas regiões Centro-Oeste (Mato Grosso, Mato Grosso do Sul, Goiás e Distrito Federal), Norte (Pará, Rondônia, Tocantins), Nordeste (Bahia, Maranhão, Piauí) e Sudeste (Minas Gerais e São Paulo). A espécie ocorre em locais com solos bem drenados do Bioma Cerrado. Apresenta distribuição irregular na paisagem, mas também pode formar grandes agrupamentos homogêneos. É mais abundante em Cerradão e Mata Semidecídua e frequente na fitofisionomia Cerrado sentido restrito, em solos areno-argilosos de fertilidade média (Sano et al., 2004).

O gênero *Dipteryx* contém 12 espécies, sendo que a maioria delas ocorre na Amazônia e apenas o baru está amplamente distribuído no Cerrado (Carvalho et al, 2020). Utilizando dados de sequências de DNA nuclear e cloroplastidial, Carvalho et al. (2023) avaliaram as relações filogenéticas dentro da tribo Dipterygeae e seus resultados apontaram *D. magnifica* como espécie evolutivamente mais próxima do baru. O tempo de divergência estimado entre as duas linhagens ficou entre um e dois milhões de anos (Ma) antes do presente, indicando uma diversificação do baru no Cerrado relativamente recente. Nesse contexto, o baru representa um bom modelo para o estudo de processos de diversificação.

Já foram realizados estudos genético populacionais para o baru utilizando marcadores moleculares. O primeiro estudo foi realizado por Soares (2008), que utilizaram 45 locos RAPD para avaliar 309 plantas provenientes de dez populações distribuídas ao longo do Cerrado. Esses autores encontraram altos níveis de divergência

genética e descontinuidade genética entre as populações. Utilizando os mesmos locos RAPD, Soares (2008b) analisaram três populações (localizadas em Goiás, Minas Gerais e São Paulo) com uma abordagem de autocorrelação espacial e observaram níveis significativos de estrutura genética dentro das populações. Nesse mesmo sentido, Collevatti et al. (2010), avaliaram 32 indivíduos de *D. alata* provenientes de uma população de Monte Alegre –MG em um estudo de estrutura genética espacial em escala fina usando dados de marcadores microssatélites e reportaram estrutura genética intrapopulacional significativa e baixos níveis de diversidade genética.

Tarazi et al. (2010) avaliaram 101 indivíduos provenientes de três populações (localizadas em Goiás, Minas Gerais e Mato Grosso do Sul) usando marcadores microssatélites e encontraram níveis significativos de estrutura genética entre as populações avaliadas. Nesse mesmo sentido, Tambarussi et al. (2017) utilizaram marcadores microssatélites para avaliar duas populações no Cerrado. Os autores reportaram baixos níveis de diversidade genética e estrutura significativa entre as populações. Guimaraes et al. (2019) utilizaram marcadores microssatélites para avaliar o sistema reprodutivo da espécie e comparar parâmetros genéticos entre uma população natural e indivíduos provenientes de um banco de germoplasma. Seus resultados indicaram que o baru é uma espécie de cruzamento misto, sendo preferencialmente alógama. Além disso, foram observados menores níveis de diversidade genética na população natural em relação ao banco de germoplasma.

Collevatti et al. (2013) avaliaram 25 populações de *D. alata* (total de 644 indivíduos) distribuídos no Cerrado usando microssatélites. Foram encontrados altos níveis de divergência genética entre as populações, indícios de processo de retração seguida de expansão demográfica e baixos níveis de fluxo gênico. Foi sugerido que a região sudoeste do Cerrado é o centro de origem desta espécie e que as regiões mais próximas do centro deste bioma foram, provavelmente, recolonizadas recentemente. Esses autores apontaram que os baixos valores de diversidade genética e riqueza alélica encontrados podem ser resultado de dinâmica de instabilidades de habitats durante o Pleistoceno.

Telles et al. (2014) encontraram correlação entre proporção de remanescentes de cobertura natural e diferenciação genética entre as populações. Foi observado que há efeito significativo da perda e fragmentação de habitat no aumento da divergência entre

as populações de *D. alata*. Soares et al. (2015) utilizaram marcadores microssatélites nucleares e cloroplastidiais para testar se os parâmetros genéticos estimados para 23 populações amostradas ao longo do Cerrado estavam distribuídos de acordo com modelo central-periférico. Esses autores encontraram que os parâmetros genéticos não seguiram um modelo central periférico clássico. Foi observado que adequabilidade ecológica influenciou significativamente os parâmetros genéticos estimados de modo que regiões mais favoráveis para a ocorrência de *D. alata* apresentaram maiores níveis de diversidade genética e baixa endogamia.

De modo geral, o que os trabalhos publicados até o momento indicam é que as populações naturais de baru apresentam níveis de diversidade genética baixos e níveis significativos de divergência genética. No entanto, ainda existem lacunas de conhecimento sobre as relações evolutivas entre as linhagens de baru e a dinâmica de diversificação dessas linhagens durante o processo de colonização da espécie no Cerrado.

A filogeografia é um campo integrativo da ciência que busca estudar a distribuição geográfica de linhagens em nível de espécie ou entre espécies próximas, utilizando informações genéticas. O campo desenvolveu métodos que permitem testar hipóteses biogeográficas, descrever a evolução de isolamento reprodutivo e inferir processos responsáveis pela origem, distribuição e manutenção da biodiversidade (Beheregaray, 2008). Em plantas, principalmente devido ao padrão de herança uniparental e à ausência de recombinação, os marcadores moleculares mais comumente utilizados para estudos filogeográficos são os baseados em sequências de DNA cloroplastidial. (Morris e Shaw, 2018).

Até o presente, não há registro de estudo filogeográfico para as populações de baru. Nesse contexto, este trabalho teve como objetivos investigar os processos que moldaram os padrões atuais da variabilidade genética no baru e elucidar os processos de diversificação das linhagens ao longo da colonização da espécie no Cerrado.

## OBJETIVOS

Geral:

- Avaliar e descrever os padrões filogeográficos que influenciaram na diversificação intraespecífica de *D. alata* e resultaram nos padrões atuais de distribuição da espécie ao longo do Cerrado brasileiro.

Específicos:

- Quantificar a estrutura e a diversidade genética presente em *D. alata*
- Compreender a organização da diversidade genética de *D. alata* no tempo e no espaço
- Identificar processos demográficos que influenciaram a dinâmica evolutiva das populações de *D. alata*.
- Identificar as relações filogenéticas entre as linhagens de *D. alata*.
- Determinar a origem das linhagens evolutivas mais antigas de *D. alata*.

## HIPÓTESES

Baseado nos resultados de estudos anteriores com *Dipteryx alata*, esperamos encontrar:

- i) baixos níveis de diversidade genética dentro das populações
- ii) altos níveis de estrutura genética
- iii) sinais de retração seguidos de expansão populacional
- iv) os padrões de baixa diversidade e alta estrutura genética seriam explicados pelo fluxo gênico limitado e presença de deriva genética
- v) padrão espacial da divergência genética segue o modelo de isolamento por distância
- vi) maiores valores de diversidade genética nas populações do sudoeste do Cerrado
- vii) As linhagens mais antigas sendo representadas pelas populações do sudoeste do Cerrado

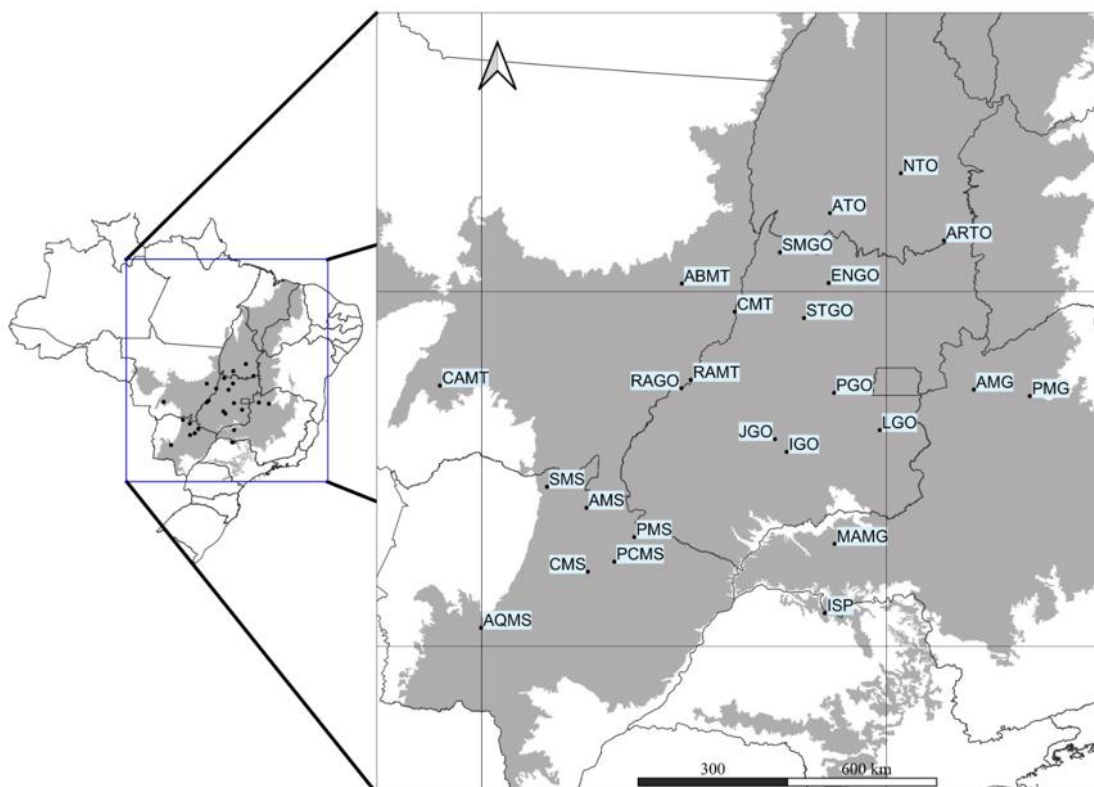
## MATERIAL E MÉTODOS

### ÁREA DE ESTUDO E AMOSTRAGEM

Foram utilizadas sequências moleculares obtidas a partir da extração do DNA de amostras foliares de 296 indivíduos adultos de *Dipteryx alata* provenientes de 25 populações naturais, distribuídas no Cerrado (Tabela 1; Figura 1).

**Tabela 1.** Populações de *Dipteryx alata* amostradas para obtenção dos dados moleculares. Goiás (GO), Minas Gerais (MG), Mato Grosso (MT), Mato Grosso do Sul (MS), São Paulo (SP) e Tocantins (TO).

População	Código	Nº de indivíduos	Longitude	Latitude
Araguaia-GO	RAGO	12	-520,503	-159,075
Estrela do Norte-GO	ENGO	12	-491,400	-138,280
Indiara-GO	IGO	12	-499,730	-171,620
Jandaia-GO	JGO	12	-502,014	-169,121
Luziânia-GO	LGO	12	-481,300	-167,320
Pirenópolis-GO	PGO	12	-490,340	-159,970
Santa Terezinha-GO	STGO	12	-496,280	-145,200
São Miguel do Araguaia-GO	SMGO	12	-501,030	-132,250
Arinus-MG	AMG	12	-462,730	-159,340
Monte Alegre de Minas-MG	MAMG	12	-490,240	-189,780
Pintópolis-MG	PMG	12	-451,664	-160,610
Alcinópolis-MS	AMS	12	-539,260	-182,680
Aquidauana-MS	AQMS	12	-560,097	-206,362
Camapuã-MS	CMS	12	-539,000	-195,280
Paraíso/Camapuã-MS	PCMS	12	-533,770	-193,300
Paraíso-MS	PMS	12	-529,820	-188,458
Sonora-MS	SMS	12	-547,040	-178,530
Água Boa-MT	ABMT	12	-520,420	-138,390
Araguaia-MT	RAMT	12	-518,660	-157,410
Cáceres-MT	CAMT	12	-568,222	-158,530
Cocalinho-MT	CMT	12	-509,960	-143,970
Alvorada-TO	ATO	12	-491,150	-124,490
Arraias-TO	ARTO	12	-468,633	-129,904
Natividade-TO	NTO	11	-477,138	-116,614
Icém-SP	ISP	12	-492,160	-203,470



**Figura 1.** Pontos de amostragem das populações de *Dipteryx alata* que foram utilizadas para coleta de dados moleculares. A área preenchida em cinza representa o Cerrado.

## DADOS MOLECULARES

As sequências utilizadas no presente trabalho são provenientes de um trabalho prévio (não publicado) realizado pela equipe do Laboratório de Genética e Biodiversidade (LGBio) situado na Universidade Federal de Goiás. Nesta ocasião, foram avaliadas quanto a presença de polimorfismo treze regiões do genoma cloroplastidial e duas regiões do genoma nuclear. As regiões cloroplastidiais são espaçadores intergênicos (tabela S1) e as regiões nucleares compreendem o espaçador transcrito interno ITS1 + 5.8S + ITS2 (ITS) do DNA nuclear e o espaçador transcrito externo (ETS) do DNA ribossômico nuclear (tabela S2).

A extração do DNA seguiu o protocolo padrão CTAB (Doyle & Doyle 1987) com modificações propostas por Ferreira & Grattapaglia (1998). O DNA total foi quantificado em gel de agarose 1% corado com brometo de etídio com 3µl de DNA e adicionando 4µl de tampão de carregamento (*Blue juice* – BJ). As alíquotas foram aplicadas em gel e imersas em tampão para eletroforese TBE (Tris-Borato 90mM e EDTA 2mM) na concentração de 0,5X. Foi utilizado 1µl de marcador de peso molecular

conhecido, o marcador Lambda nas concentrações de 50ng, 100ng para inferir a concentração de cada amostra de DNA.

A amplificação dos fragmentos foi feita via Reação em Cadeia da Polimerase (Polimerase Chain Reaction – PCR) em reações com volume final de 15µl contendo 1,13µM de cada primer; 1,15 unidades de Taq DNA Polimerase (Phoneutra, Br); 2,5µM de cada dNTP; tampão 1X (10mM Tris-HCl, pH 8,3, 50mM KCl, 1,5mM MgCl<sub>2</sub>); 375µg de BSA entre 5-6 ng de DNA.

As condições de amplificação para todos os oligonucleotídeos iniciadores foram: 94°C de desnaturação inicial por 2 minutos; 35 ciclos de desnaturação em 94°C por 1 minuto; anelamento com temperatura específica otimizada para cada par de oligonucleotídeo por 1 minuto; 72°C de extensão por 1 minuto e 1 ciclo de extensão final a 72°C por 3 minutos.

Após as reações de PCR, alíquotas dos produtos amplificados foram submetidos à eletroforese horizontal em gel de agarose (1,5%), submersas em tampão TBE (Tris-Borato 90mM e EDTA 2mM), para verificar as amostras amplificadas com sucesso e para checar o comprimento do fragmento utilizando-se 4µl do marcador 1kb. Para testar a amplificação variou-se a temperatura de anelamento entre 48° C a 62° C. Algumas regiões não amplificaram em nenhuma temperatura e novos testes foram feitos variando a quantidade adicionada de DNA, oligonucleotídeo iniciador, também BSA e DMSO na reação de PCR.

A qualidade do produto amplificado foi verificada por eletroforese horizontal em gel de agarose a 1,5%, usando marcador de peso molecular conhecido (1Kb Plus DNA Ladder, Invitrogen) e coloração por brometo de etídio. Os produtos amplificados foram classificados quanto à intensidade da banda e presença e ausência de bandas inespecíficas.

O protocolo de PCR e a temperatura de anelamento foram padronizados com dois indivíduos e, então, foi feita a seleção de polimorfismos úteis para estudos filogeográficos em *D. alata*. Após isso, foram selecionados 25 indivíduos (um por localidade amostrada) para fazer a amplificação de cada região não codificante do genoma cloroplastidial e do genoma ribossômico nuclear. A amplificação via PCR e verificação da amplificação dos fragmentos foram feitas da mesma forma como descrito anteriormente na fase de padronização.

Os produtos amplificados via PCR passaram por um processo de purificação para a retirada de excesso de primers e nucleotídeos não incorporados durante a amplificação, utilizando o kit ExoSAP composto por Exonuclease I recombinante (Exo) (10U) e Shrimp Alkaline Phosphatase (SAP) (1U). Para cada 5µl de produto de PCR foram adicionados 1µl de ExoSAP.

Todos os produtos de PCR foram sequenciados nas direções forward e reverse para detectar erros de sequenciamento. O sequenciamento foi feito em reações de 10 µL com 3 µL de produto de PCR purificado, 2 µL de água milliQ; 1 µL de primer (5 µM) e 4 µL de ET-DYE Terminator Kit (Amersham Biosciences).

Os produtos de sequenciamento foram precipitados e limpos com Etanol, e, então, secados à temperatura ambiente por 10 minutos, dissolvidos em tampão de carregamento (formamida 70 % e 1µM EDTA) e analisados em um sequenciador automatizado ABI Prism 3130 (Applied Biosystems, CA, USA) usando o kit de sequenciamento DYEnamic™ ET terminator (GE Health-Care, Suécia), seguindo as indicações do fabricante.

Das quinze regiões testadas, apenas uma região cloroplastidial (*psbA-trnH*) e uma região nuclear (*ITS101-ITS102*) apresentaram bom padrão de amplificação e presença de polimorfismo. As sequências dessas duas regiões moleculares provenientes de 296 indivíduos foram utilizadas nas análises realizadas no presente trabalho.

O programa *SeqScape* v.6 (*Applied Biosystems*) foi utilizado para verificar a qualidade das sequências e para a montagem da fita consenso e para a obtenção do alinhamento múltiplo das sequências. O alinhamento final foi realizado utilizando o *ClustalW* implementado no programa *BioEdit* v.7.0.5 (Hall,1999). Para as sequências nucleares, o programa *Topali* v2.5 (Milne et al., 2004) foi utilizado com o intuito de detectar presença de recombinação e os haplótipos foram obtidos utilizando o algoritmo *phase* (Scheet & Stephens 2006) implementado no programa *DNAsp* 6.12 (Librado & Rozas 2009). Para selecionar o modelo evolutivo de substituição nucleotídica mais adequado para cada marcador foi utilizado o programa *Jmodeltest* v2.1.10 (Posada, 2008).

## ANÁLISE DOS DADOS

## **Diversidade, estrutura e divergência genética**

Com o intuito de quantificar a magnitude da diversidade genética presente nas populações, foram estimados a diversidade haplotípica ( $h$ ), nucleotídica ( $\pi$ ) e o número de sítios polimórficos ( $S$ ). Para verificar estrutura genética populacional, foi realizada análise de variância molecular (AMOVA, Excoffier et al., 1992) e os valores de  $F_{ST}$  foram estimados. A divergência genética entre as populações foi obtida a partir das estimativas de  $F_{ST}$  par a par. Todas essas análises foram realizadas no programa *Arlequin* v.3.1 (Excoffier et al., 2005).

Para uma visualização gráfica do padrão de divergência genética entre as populações, foi construído um dendrograma, a partir do método UPGMA. A robustez do dendrograma foi avaliada a partir da correlação cofenética. A construção do dendrograma e a correlação cofenética foram realizadas utilizando o pacote *Adegenet* v 2.1.10 do programa *R*.

### **Padrão espacial na diversidade genética**

Para testar se a divergência genética entre as populações pode ser explicada pelo espaço geográfico, a matriz de valores de  $F_{ST}$  par a par foi correlacionada com a matriz de distâncias geográficas, utilizando o teste de Mantel. Também foi realizado teste de Mantel considerando classes de distâncias, que teve sua visualização gráfica a partir de um correlograma. Essas análises foram realizadas utilizando o programa *R*, a partir dos pacotes *vegan* (Oksanen et al., 2019) e *ecodist* (Goslee & Urban, 2017).

Foram calculadas para cada população, a distância do centroide dos registros de ocorrência ponderada pelos valores de adequabilidades estimados para o presente, 6 Ka e 21 Ka (ver seção Modelagem de Nicho). Com o intuito de verificar se mudanças na adequabilidade da espécie ao longo do tempo geraram padrão clinal na diversidade genética, essas distâncias foram avaliadas com os valores de diversidade haplotípica e número de alelos utilizando análise de regressão linear, implementada no programa *R*. A diferença entre a adequabilidade atual e do passado (21 Ka) foi calculada, representando uma medida de estabilidade de habitat. Com o intuito de avaliar o efeito da estabilidade de habitat sobre a diversidade genética, os valores de estabilidade foram avaliados com os valores de diversidade haplotípica e número de alelos utilizando análise de regressão linear, implementada no programa *R* (Core).

## História demográfica e padrões filogeográficos

Para entender a relação evolutiva entre os haplótipos presentes nas amostras populacionais, foram estimadas filogenias para cada marcador utilizando o método *Median Joining Network*, implementadas no programa *Mega* v10.0.5 (Tamura et al., 2007). Com o intuito de avaliar se as populações se encontram em equilíbrio estacionário, o programa *Arlequin* v.3.1 foi utilizado para realizar as análises: o  $F$  de Fu (Fu 1997) e o índice *raggedness* Rogers-Harpending (RI) (Harpending, 1994).

Foi realizada análise *Bayesian Skyline Plot* no programa *Beast* v.1.8.4 (Drummond et al., 2012) para inferir se ocorreram mudanças no tamanho efetivo da população ao longo do tempo. O programa *Tracer* v1.7.1 (Rambaut e Drummond, 2007) foi utilizado para verificar a qualidade das corridas (tamanho efetivo da amostra,  $ESS \geq 200$ ) e gerar o gráfico dessa última análise.

Nas análises de coalescência, foram utilizadas as partições de cpDNA e nDNA concatenadas, mas com *priors* separados e específicos para cada região. O programa *Lamarc* v2.1.10 (Kuhner, 2006) foi utilizado para estimar parâmetros demográficos por meio de abordagem bayesiana, sendo eles o parâmetro de mutação ( $\theta$ ), a taxa de migração entre as populações ( $M$ ) e a taxa de crescimento exponencial ( $g$ ). A conectividade genética histórica entre as populações foi avaliada com a estimativa do número de migrantes por geração ( $m = M/\theta$ ). As análises foram realizadas com três corridas configuradas com dez cadeias iniciais de 20.000 passos (amostragem = 100 e burnin = 1500) e três cadeias finais de 100.000 passos (amostragem = 100 e burnin = 10.000), as demais configurações foram padrão. A convergência e estabilidade dos resultados foram avaliadas usando o programa *Tracer* v1.7.1 (Rambaut e Drummond, 2007).

O programa *Beast* v.1.8.4 (Drummond et al., 2012) foi utilizado para estimar o tempo do ancestral comum mais recente (TMRCA, do inglês *time to most recent common ancestor*) das populações. Amostras individuais de *Dipteryx odorata* Aubl., *Dipteryx fërrea* Duck, *Dipteryx charafila* Duck e *Pterodon emarginatus* Vogel, espécies pertencentes a tribo Dipterygeae, foram incluídas como grupo externo nessa análise. Foi usada uma taxa de mutação por sítio/ano para regiões não codificantes nucleares e cloroplastidiais estimadas para *Inga* (Richardson et al., 2001) que também pertence à família Fabaceae. A convergência e estabilidade dos resultados foram avaliadas usando o programa *Tracer* v1.7.1 (Rambaut e Drummond, 2007). O programa *Treeannotator*

v1.8.4 (Drummond et al., 2012) foi utilizado para sumarizar as árvores produzidas no *Beast* v.1.8.4 em uma única árvore que contenha a soma máxima de probabilidades posteriores. A árvore sumarizada foi visualizada utilizando o programa *Figtree* v1.4.4 (Rambaut, 2012).

## MODELAGEM DE NICHOS

Para a modelagem de nicho, os registros de ocorrência foram obtidos a partir das bases de dados *Species Link*, GBIF e a partir dados coletados pelas pesquisadoras do Laboratório de Genética & Biodiversidade. Os dados passaram por um filtro inicial para avaliar se havia registros duplicados ou fora do que seria a área de distribuição da espécie. Essa primeira análise resultou em 630 registros. Porém, para reduzir agrupamentos de registros e vieses na distribuição espacial nos pontos de ocorrência, foi utilizado o pacote *spThin* (Aiello-Lammens et al. 2015) para selecionar pontos a uma distância mínima de 50km, compatível com a resolução dos dados climáticos. Esse procedimento resultou em 129 registros de ocorrência utilizados na modelagem de distribuição potencial.

A obtenção de dados climáticos do presente e para dois períodos do passado (6 e 21 mil anos) foi realizada a partir de quatro modelos climáticos: *Community Climate System Model* (CCSM), CNRM, *Model for Interdisciplinary Research Climate* (MIROC) e MRI. Esses modelos geraram dezenove variáveis bioclimáticas que foram submetidas a uma análise de componentes principais para selecionar eixos ortogonais que representassem a variação climática, reduzindo assim a multicolinearidade das variáveis originais. Foram selecionados seis eixos que representaram 95% da matriz de covariância.

Os modelos de distribuição geográfica para *D. alata* foram gerados por meio do pacote ENMTML (Andrade et al. 2020), sendo projetados para o presente, Holoceno médio (6 ka) e Último Máximo Glacial (LGM, 21 ka). Os modelos de distribuição potencial foram gerados pelos métodos de envelopes bioclimáticos (Bioclim), *Domais*, *ecological niche factor analysis* (ENFA), entropia máxima (Maxent) e SVM-B. Para a partição dos dados entre treino e validação, foi utilizado o método de *bootstrap*, por meio do qual os dados foram particionados 100 vezes em treino e teste, na proporção de 70% dos pontos para treino e 30% para validação. Os modelos finais foram gerados pela

abordagem de consenso (*Ensemble*), ou seja, a média dos melhores algoritmos (i.e., aqueles modelos cujos valores da estatística TSS foram superiores à média).

Os mapas de adequabilidade gerados pelo consenso dos melhores modelos foram binarizados para gerar mapas de distribuição potencial ou mapas de extensão de ocorrência por meio da métrica de *Least Presence Threshold* (LPT, Pearson 2007). Esse método considera o menor valor de adequabilidade associado a um ponto de ocorrência como sendo o limiar de decisão para a binarização.

## RESULTADOS

O fragmento amplificado da região *psbA-trnH* compreendeu 433 pares de base. O fragmento amplificado da região *ITS101-ITS102* do nDNA compreendeu 590 pares de base. O melhor modelo evolutivo estimado para a região cloroplastidial foi o TrN, segundo o critério de AIC. Para a região nuclear, o melhor modelo evolutivo estimado foi o F81, segundo o critério AIC. As composições nucleotídicas foram 12,1% C, 40,6% T, 32,9% A e 14,4% G para a região cloroplastidial e 36,33% C, 12,23% T, 19,69% A e 31,75% G para a região nuclear. Não foram encontradas evidências de presença de recombinação para a região nuclear.

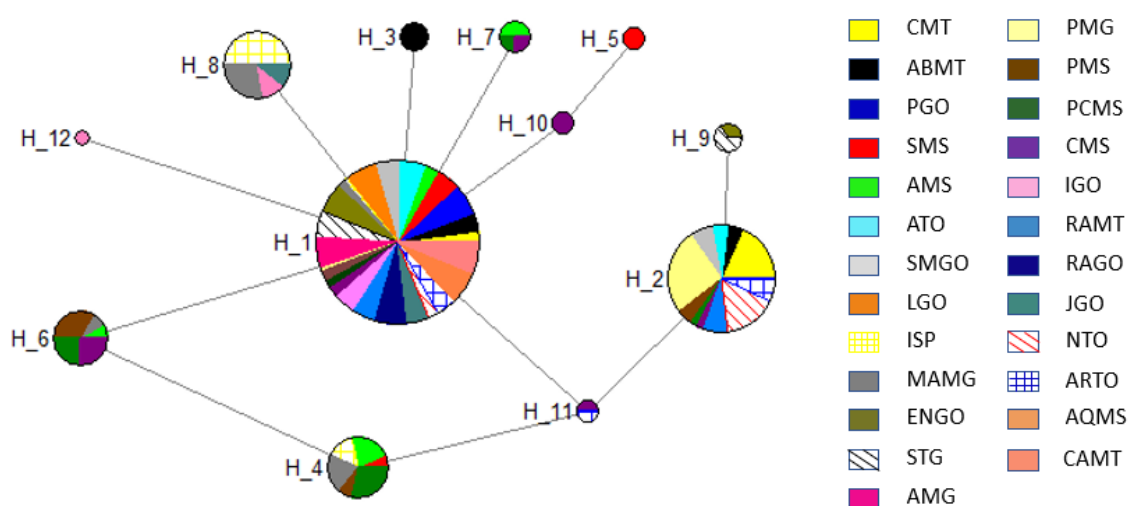
As 25 populações avaliadas apresentaram doze haplótipos para a região cloroplastidial com nove sítios polimórficos, sendo quatro transversões e cinco *indels*. Para a região nuclear foram detectados oito haplótipos, com quatro sítios polimórficos, sendo três transições e uma transversão. A população CMS apresentou maior número de haplótipos (seis) e seis populações (LGO, PGO, RAGO, AQMT, CAMT e AMG) não apresentaram polimorfismo para o cpDNA. Três populações (ABMT, RAGO e RAMT) apresentaram número máximo de haplótipos para a região nuclear e uma população (SMGO) não apresentou polimorfismo. A diversidade haplotípica geral foi intermediária tanto para região cloroplastidial ( $h=0,591$ ) quanto para região nuclear ( $h=0,543$ ). A diversidade nucleotídica observada foi baixa para ambas as regiões moleculares avaliadas ( $\pi_{cpDNA}=0,001$ ;  $\pi_{nDNA}=0,002$ ) (Tabela 2).

**Tabela 2.** Caracterização genética de 25 populações de *Dipteryx alata* no Cerrado brasileiro, baseada em dados de cpDNA (*psbA-trnH*) e nDNA (*ITS101/102*). N = tamanho amostral; K = número de haplótipos; h = diversidade haplotípica;  $\pi$  = diversidade nucleotídica; SD = desvio padrão.

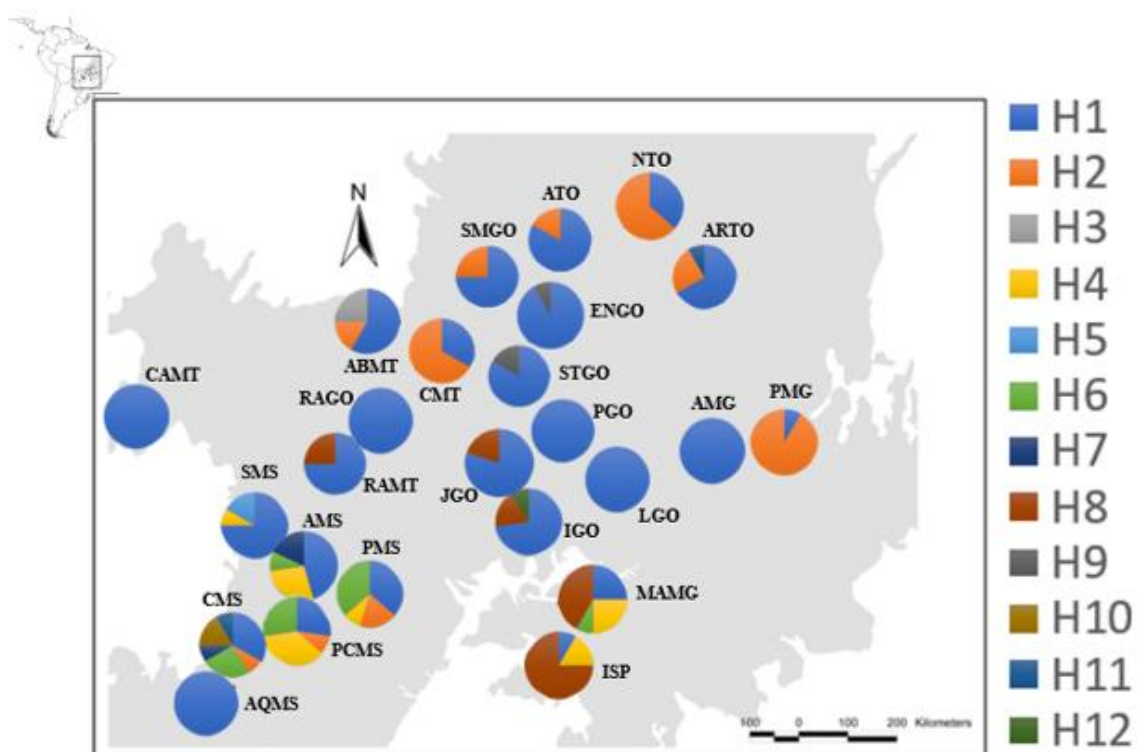
População	cpDNA					nDNA				
	N	K	h	$\pi$	SD	N	K	h	$\pi$	SD
Araguaia-GO	12	1	0,000	0,000	0,000	12	8	0,811	0,002	0,001
Estrela do Norte-GO	12	2	0,167	0,001	0,001	11	2	0,095	0,000	0,000
Indiara-GO	11	3	0,473	0,001	0,001	11	6	0,818	0,002	0,001
Jandaia-GO	10	4	0,644	0,000	0,000	12	3	0,366	0,002	0,001
Luziânia-GO	12	1	0,000	0,000	0,000	12	7	0,793	0,002	0,001
Pirenópolis-GO	12	1	0,000	0,000	0,000	12	5	0,575	0,002	0,001
Santa Terezinha-GO	12	4	0,561	0,002	0,001	11	4	0,463	0,001	0,001
São Miguel do Araguaia-GO	12	2	0,409	0,001	0,001	12	1	0,000	0,000	0,000
Arinus-MG	12	1	0,000	0,000	0,000	12	3	0,304	0,000	0,000

Monte Alegre de Minas-MG	12	4	0,758	0,003	0,002	12	4	0,742	0,001	0,001
Pintópolis-MG	12	2	0,167	0,000	0,000	12	5	0,634	0,001	0,001
Alcinópolis-MS	11	4	0,745	0,002	0,002	12	4	0,878	0,001	0,001
Aquidauana-MS	12	1	0,000	0,000	0,000	12	4	0,454	0,002	0,001
Camapuã-MS	12	6	0,848	0,003	0,002	12	5	0,764	0,002	0,001
Paraíso/Camapuã-MS	12	5	0,818	0,003	0,002	12	5	0,753	0,002	0,001
Paraíso-MS	11	5	0,818	0,003	0,002	11	4	0,709	0,001	0,001
Sonora-MS	12	4	0,455	0,002	0,001	12	5	0,757	0,002	0,001
Água Boa-MT	12	3	0,621	0,002	0,001	12	8	0,863	0,003	0,002
Araguaia-MT	12	2	0,409	0,001	0,001	12	8	0,858	0,003	0,002
Cáceres-MT	12	1	0,000	0,000	0,000	12	3	0,594	0,001	0,001
Cocalinho-MT	12	2	0,485	0,002	0,001	12	3	0,909	0,001	0,001
Alvorada-TO	12	2	0,303	0,001	0,001	12	6	0,454	0,002	0,001
Arraias-TO	12	4	0,561	0,002	0,001	12	3	0,391	0,000	0,000
Natividade-TO	11	4	0,673	0,002	0,002	11	3	0,536	0,001	0,001
Icém-SP	12	4	0,652	0,002	0,001	12	7	0,841	0,003	0,002
Geral	294	12	0,591	0,001	0,000	295	8	0,543	0,002	0,001

A rede de haplótipos da região *psbA-trnH* apresentou formato de estrela (Figura 2). Esse tipo de disposição geralmente é encontrado em espécies que passaram por processo de expansão recente, nas quais os haplótipos mais frequentes são considerados ancestrais e os haplótipos menos frequentes derivados. O haplótipo H1 foi o mais frequente e foi observado em todas as populações avaliadas. A ocorrência de haplótipos exclusivos foi observada em quatro populações (IGO, SMS, CMS e ABMT).

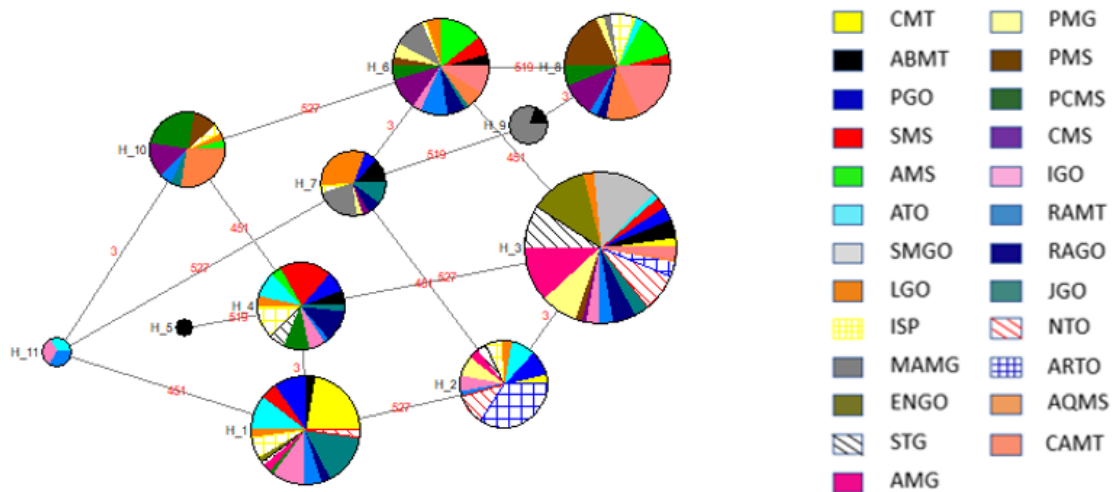


**Figura 2.** Rede de haplótipos (Median-joining network) de 25 populações de *Dipteryx alata* no Cerrado brasileiro, obtida a partir de dados de sequências de região cloroplastidial (*psbA-trnH*).

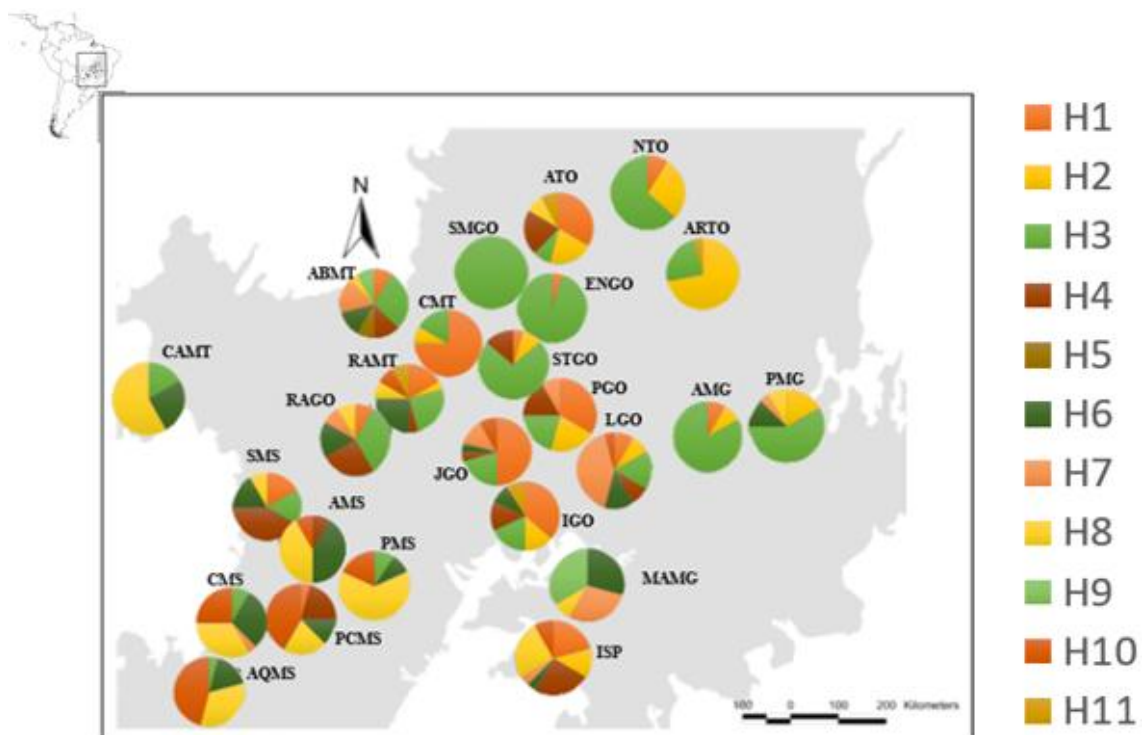


**Figura 3.** Distribuição geográfica dos haplótipos da região cloroplastidial (*psbA-trnH*) nas 25 populações de *Dypteryx alata*.

A rede de haplótipos da região *ITS101-ITS102* não apresentou resolução para a relação evolutiva entre os haplótipos (Figura 4). Essa ambiguidade pode ser resultado da baixa diversidade haplotípica das populações de baru, considerando que o poder de resolução da análise é afetado pelo nível de diversidade presente nos dados. O haplótipo H3 apresentou a maior frequência, estando presente em 21 das populações avaliadas (exceto AMS, ISP, MAMG e PCMS). O haplótipo H5 foi observado apenas na população ABMT, o haplótipo H9 foi observado em duas populações (ABMT e MAMG) e o haplótipo H11 foi observado em três populações.



**Figura 4.** Rede de haplótipos (Median-joining network) de 25 populações de *Dipteryx alata* no Cerrado brasileiro, obtida a partir de dados de seqüências de região nuclear (ITS101/102).



**Figura 5.** Distribuição geográfica dos haplótipos da região nuclear (ITS101/102) nas 25 populações de *Dipteryx alata*.

Foi detectada estrutura genética populacional significativa para ambas as regiões, sendo  $F_{ST}=0,305$   $p=0,0001$  para a região cloroplastidial e  $F_{ST}=0,359$   $p=0,0001$  para a região nuclear (Tabelas 3 e 4).

**Tabela 3.** Análise de variância molecular (AMOVA) para 25 populações de *Dipteryx alata* no Cerrado brasileiro, obtida a partir das sequências cloroplastidiais (psbA-trnH). g.l.: graus de liberdade; SSD: soma do quadrado dos desvios.

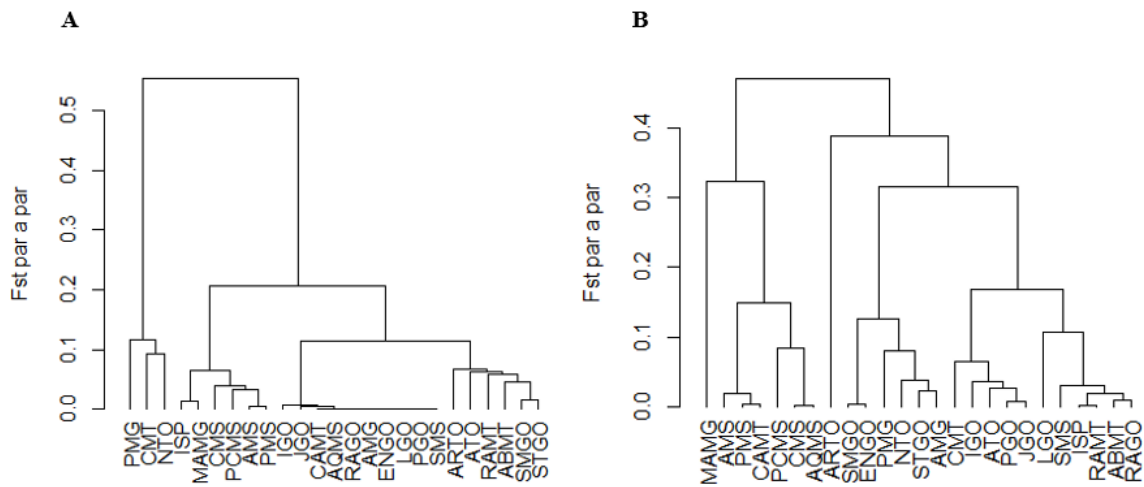
Fontes de variação	g.l.	SSD	Componentes de variação	Porcentagem de variação	Índice de fixação ( <i>F<sub>st</sub></i> )
Entre população	24	51,496	0,152	30,53	0,305 p=0,0001
Dentro de populações	269	93,562	0,347	69,47	
Total	293	145,058	0,500		

**Tabela 4.** Análise de variância molecular (AMOVA) para 25 populações de *Dipteryx alata* no Cerrado brasileiro, obtida a partir das sequências nucleares (ITS101/102). g.l.: graus de liberdade; SSD: soma do quadrado dos desvios.

Fontes de variação	g.l.	SSD	Componentes de variação	Porcentagem de variação	Índice de fixação ( <i>F<sub>st</sub></i> )
Entre população	24	175,792	0,288	35,92	0,359 p=0,0001
Dentro de populações	565	290,814	0,514	64,08	
Total	589	466,607	0,803		

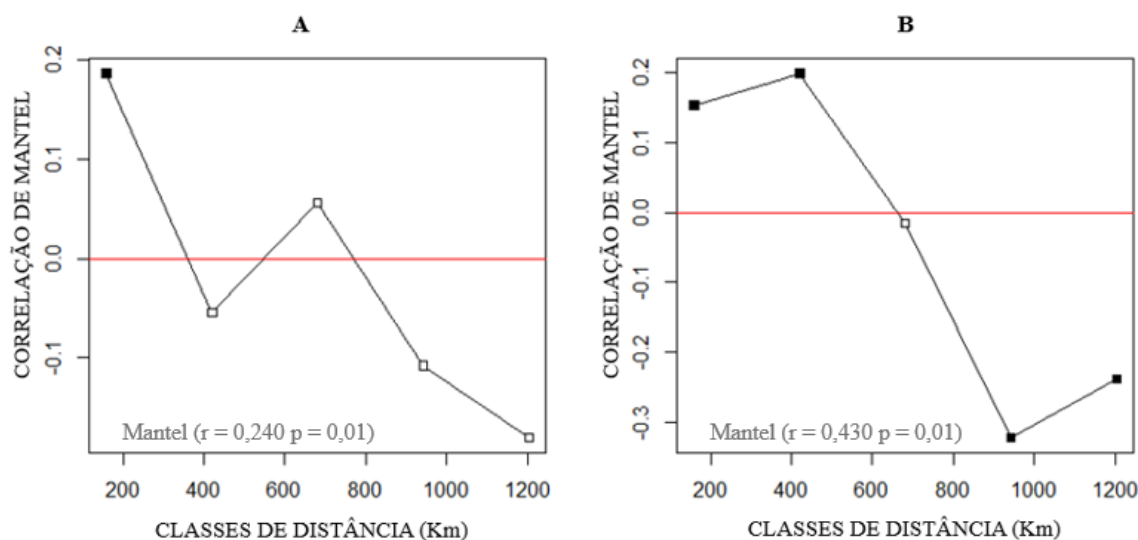
Para a região cloroplastidial, o valor de  $F_{ST}$  entre os pares de populações variou de zero (35 pares de populações) a 0,909 (seis pares de populações) (Tabela S3). Todos os seis pares de valores iguais a 0,909 foram em combinação com a população de PMG. Além disso, PMG e CMT foram as populações com maior número de valores de  $F_{ST}$  par a par significativos (23). No sentido oposto, CMS foi a população com menor número de valores de  $F_{ST}$  par a par significativos (3). Para a região nuclear, o valor de  $F_{ST}$  entre os pares de populações variou de 0,002 (quatro pares de populações) até 0,785 (entre CMT e SMGO) (tabela S4). A população MAMG apresentou valores de  $F_{ST}$  significativos para todas as combinações. Em contrapartida, a população RAGO apresentou o menor número de valores significativos (6).

Os dendrogramas produzidos a partir do  $F_{ST}$  par-a-par entre as populações apresentaram correlação cofenética de 0,921 e 0,875 para as regiões cloroplastidial e nuclear, respectivamente (Figura 6). As populações dos estados do Mato Grosso do Sul (MS), Minas Gerais (MG) e São Paulo (SP), em ambos os dendrogramas, apresentaram tendência de formar grupo distinto das demais populações do centro-norte do Cerrado. No entanto, para a região cloroplastidial essa tendência não foi tão evidente pois as populações PMG, NTO e CMT formaram grupo distinto.



**Figura 6.** Agrupamento tipo UPGMA baseado nos resultados de  $F_{ST}$  par-a-par entre 25 populações de *Dipteryx alata*, obtidos a partir das sequências cloroplastidiais (A) e nucleares (B). Destacar os grupos

Os testes de Mantel apresentaram valores significativos para ambas as regiões moleculares avaliadas (cpDNA = 0,240  $p < 0,05$ ; nDNA = 0,430  $p < 0,05$ ), indicando que parte da divergência genética entre as populações pode ser explicada pela distância geográfica entre elas (figura 7). Os resultados com as sequências cloroplastidiais apontaram correlação significativa e positiva para a primeira classe de distância, sugerindo que populações separadas por até 200 quilômetros tenderão a ser mais semelhantes geneticamente que o esperado pelo acaso (Figura 7 A). Para as sequências nucleares foram observadas correlações significativas e positivas nas duas primeiras classes de distância, indicando que populações separadas por até 400 quilômetros tenderão a ser mais semelhantes geneticamente do que o esperado pelo acaso. Também foram observados valores significativos e negativos para as duas últimas classes de distância, indicando que populações separadas por mais de 900 quilômetros tenderão a ser mais diferentes geneticamente que o esperado pelo acaso (Figura 7 B).

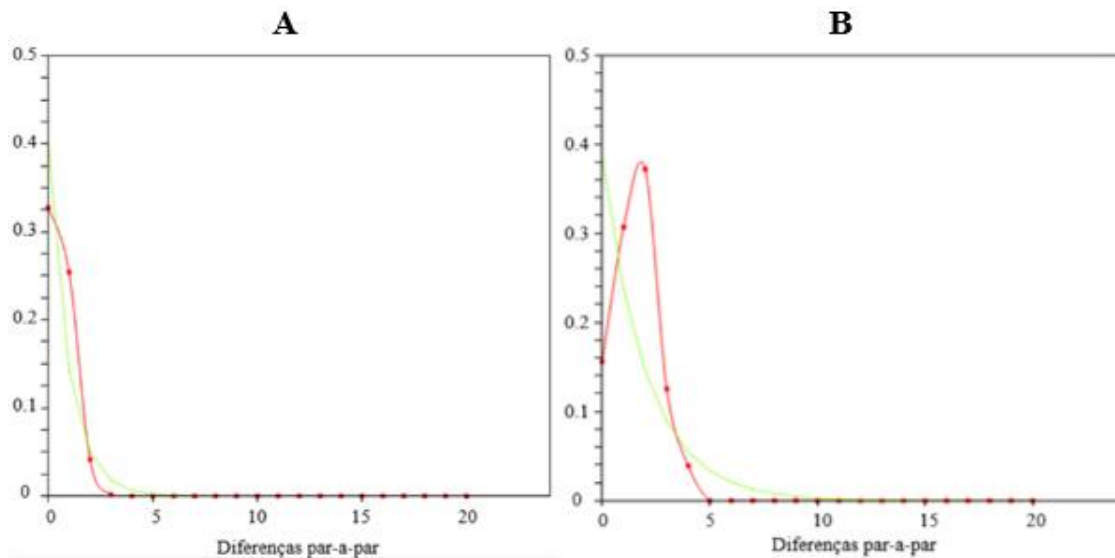


**Figura 7.** Correlogramas de Mantel entre matrizes de  $F_{ST}$  par-a-par e distância geográfica de 25 populações naturais de *Dipteryx alata* no Cerrado brasileiro, obtidos a partir de sequências cloroplastidiais (psbA-trnH) (A) e nucleares (ITS101/102) (B). Os quadrados preenchidos representam os valores significativos ( $p < 0,05$ ).

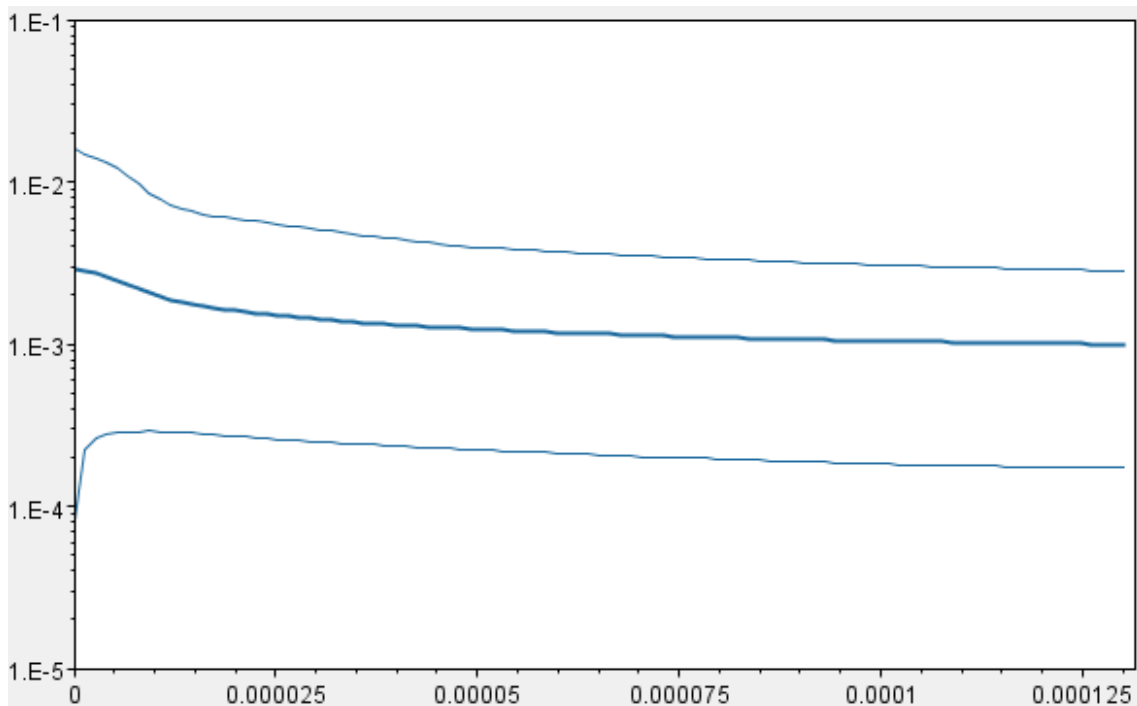
Os testes de neutralidade apresentaram desvios significativos para ambas as regiões moleculares avaliadas em dois testes distintos (tabela 5). Os gráficos *mismatch* para ambas as regiões avaliadas apresentaram um padrão de diferenças par a par unimodal, o que indica populações que passaram por processo de expansão (Figura 8). A análise *extended bayesian skyline plot* resultou em um cenário de estabilidade com tendência de expansão para as populações de *D. alata* (Figura 9).

**Tabela 5.** Testes de neutralidade para as regiões cloroplastidial e nuclear. D: valor de Tajima; FS: valor de Fu; p: significância.

Região	Testes de neutralidade					
	D de Tajima		Fu		Índice Ragedness	
	D	p	FS	p	Rg	p
<b>Cloroplastidial</b>	-4,13	0,021	-5,98	0,001	0,125	0,547
<b>Nuclear</b>	-2,68	0,038	-16,37	0,126	0,096	0,189



**Figura 8.** Distribuição *mismatch* das sequências da região cloroplastial (*psbA-trnH*) (A) e região nuclear (*ITS101/102*) (B). A linha vermelha representa a frequência das diferenças de bases par-a-par e a linha verde representa os valores esperados pelo modelo de expansão demográfica.



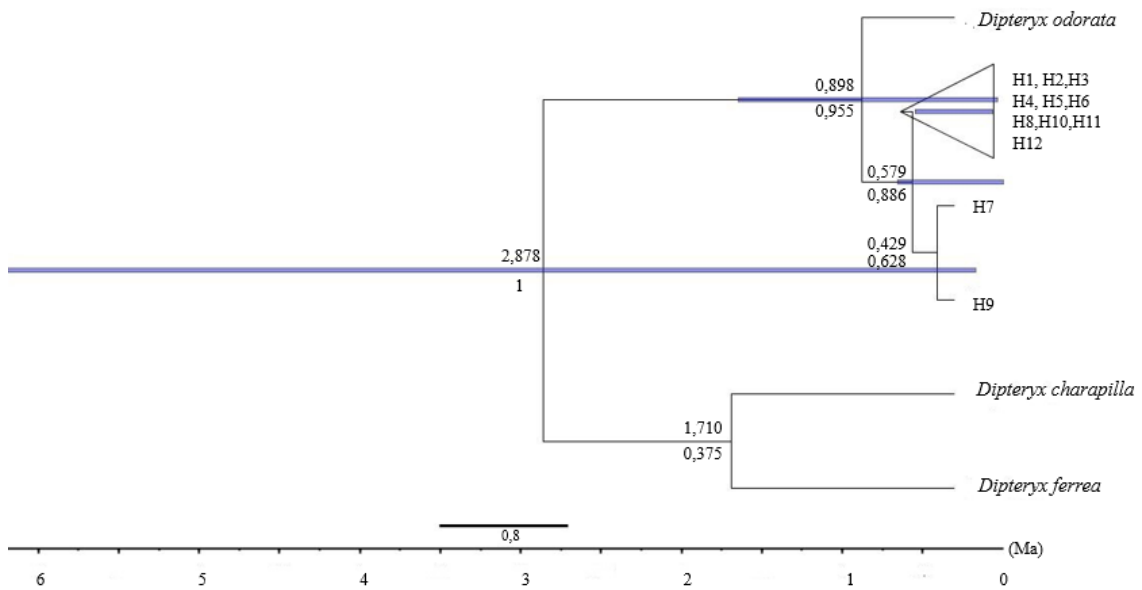
**Figura 9.** Bayesian skyline plot para 25 populações de *Dipteryx alata* no cerrado brasileiro, obtidos a partir de sequências das regiões cloroplastial (*psbA-trnH*) e nuclear (*ITS101/102*) concatenadas. A linha do meio representa os valores médios e as linhas superior e inferior representam o intervalo de confiança a 95%. O eixo y indica o tamanho populacional e o eixo x indica o tempo.

As análises de coalescência apontaram que as populações de *D. alata* apresentam baixas taxas mutacionais ( $\theta=0,003207$ ) (tabela 6). A população CMS apresentou o maior valor ( $\theta =0,001813$ ) para esse parâmetro, enquanto a população SMGO apresentou o menor valor ( $\theta =0,000106$ ). Os valores observados para o parâmetro  $g$  foram positivos, porém não significativos. As populações de *D. alata* apresentaram baixos níveis de fluxo gênico, considerando que nenhum valor acima de um migrante por geração foi observado (Tabela S5).

**Tabela 6.** Parâmetros demográficos de 25 populações de *Dipteryx alata*, obtidos a partir de dados de cpDNA e nDNA concatenados, baseados em coalescência.  $\theta$  - parâmetro de coalescência,  $g$  - parâmetro de crescimento exponencial; IC - intervalo de confiança.

População	Cloroplasto + Nuclear						
	N	$\theta$	$\theta$ 95% IC		G	g 95% IC	
<b>CMT</b>	12	0,000185	0,000143	6,273720	577,9685	-531,0481	1057,312
<b>ABMT</b>	12	0,000189	0,000295	4,748656	346,1039	-559,6601	1080,142
<b>PGO</b>	12	0,000360	0,000136	1,501414	677,3149	-582,8957	1076,376
<b>SMS</b>	12	0,000148	0,000163	0,002200	834,2155	-572,5085	1084,714
<b>AMS</b>	11	0,000516	0,000111	0,000522	724,5048	-550,3205	1094,397
<b>ATO</b>	12	0,000905	0,000459	3,649155	-277,1731	-571,5071	1081,375
<b>SMGO</b>	12	0,000106	0,000347	0,163031	656,6804	-577,4050	1075,898
<b>LGO</b>	12	0,000920	0,000118	0,002396	186,3724	-587,3525	1069,603
<b>ISP</b>	12	0,000707	0,000163	0,002200	825,8063	-570,1165	1079,049
<b>MAMG</b>	12	0,000500	0,000341	0,001543	685,3005	-542,1270	1081,974
<b>ENGO</b>	12	0,000164	0,000675	0,003727	347,4473	-557,8573	1061,228
<b>STGO</b>	12	0,000556	0,000235	0,004113	746,1571	-584,0365	1082,133
<b>AMG</b>	12	0,000508	0,000183	0,002231	829,1386	-556,3405	1064,090
<b>PMG</b>	12	0,000209	0,000264	0,001846	885,3544	-559,7096	1055,690
<b>PMS</b>	11	0,000848	0,000121	0,001384	884,2533	-551,9472	1105,288
<b>PCMS</b>	12	0,001188	0,000175	0,002592	912,7562	-541,6817	1095,294
<b>CMS</b>	12	0,001813	0,000199	0,002333	835,7580	-576,8649	1083,531
<b>IGO</b>	11	0,000504	0,000211	0,002005	668,6733	-573,9422	1085,397
<b>RAMT</b>	12	0,000198	0,000128	0,003481	723,1850	-555,2907	1072,839
<b>RAGO</b>	12	0,000124	0,000155	0,002258	-258,4347	-572,2283	1072,026
<b>JGO</b>	10	0,000112	0,000194	0,004023	-195,4403	-563,6294	1070,355
<b>NTO</b>	11	0,000372	0,000203	0,003201	880,7612	-552,4966	1080,668
<b>ARTO</b>	12	0,000366	0,000172	0,002114	857,7860	-584,7929	1105,766
<b>AQMS</b>	12	0,000572	0,000206	0,002655	362,9375	-559,0203	1046,629
<b>CAMT</b>	12	0,000121	0,000188	0,002711	-143,1940	-446,5263	1096,914
<b>Geral</b>	<b>294</b>	<b>0,003207</b>	<b>0,000690</b>	<b>0,005821</b>	<b>322,1693</b>	<b>-523,9257</b>	<b>1026,831</b>

As análises de coalescência indicaram um ancestral comum mais recente (ACMR) para as populações de *D. alata* durante o Pleistoceno, cerca de 0,5 Ma (milhões de anos antes do presente), quando ocorre a coalescência de todos os haplótipos observados (Figura 10). Não foi encontrada relação clara entre a distribuição das linhagens genéticas e o espaço geográfico. Essa ausência de relação pode ser resultado de um arranjo incompleto de linhagens com retenção de polimorfismo ancestral. A divergência estimada entre *D. alata* e o grupo externo (*Dipteryx odorata*) ocorreu no Plioceno, cerca de 0,8 Ma.

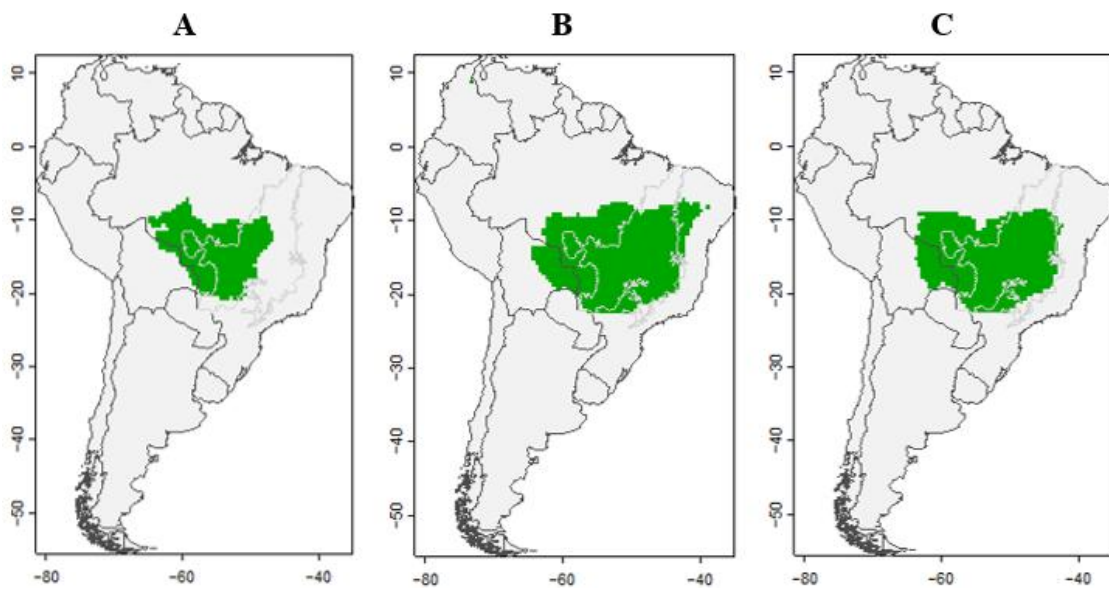


**Figura 10.** Tempo para o ancestral comum mais recente dos haplótipos de *Dipteryx alata*, baseadas nas análises de coalescência de todos os fragmentos concatenados (cpDNA e nDNA) de 296 indivíduos provenientes de 25 populações naturais. Os números abaixo dos ramos correspondem ao suporte do nó (probabilidade a posteriori) e os números acima dos ramos representam a datação do nó (tempo para o ancestral comum). As barras em azul correspondem ao intervalo de credibilidade a 95%, da média do tempo do ancestral comum. A escala está em milhões de anos (Ma) antes do presente.

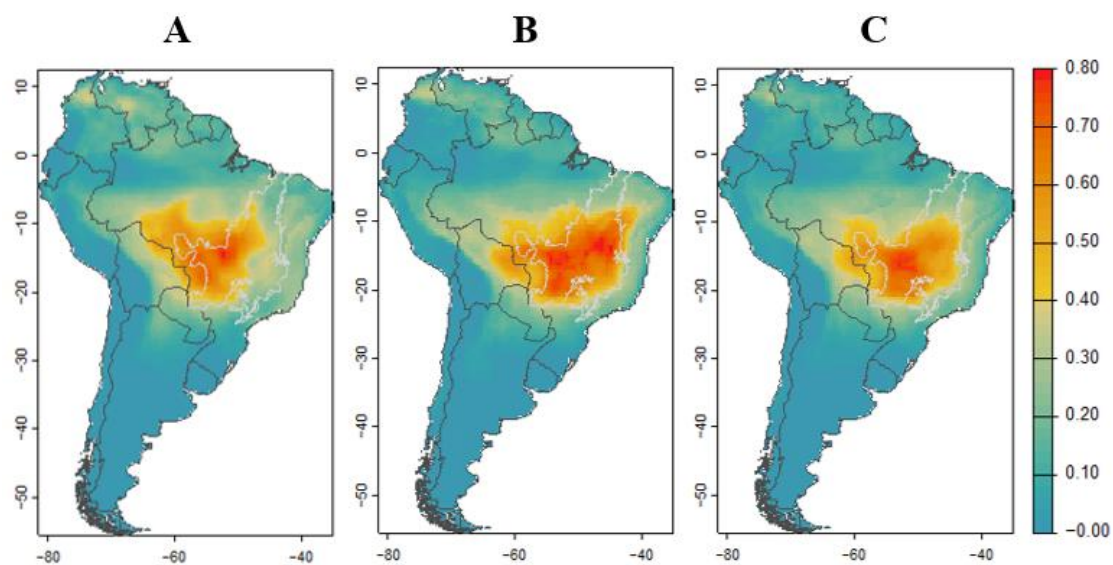
As modelagens ecológicas de nicho indicaram que *D. alata* estava potencialmente distribuída ao longo das regiões sudoeste, oeste e noroeste do Cerrado durante o período do LGM (Figura 11 A). Nesse período, os maiores níveis de adequabilidade estavam restritos a região oeste do Cerrado (Figura 12 A). Durante o holoceno médio houve um aumento da área de distribuição na direção das regiões sudeste, leste e nordeste do Cerrado (Figura 11 B), com as regiões sudoeste, oeste e nordeste apresentando altos níveis de adequabilidade (Figura 12 B). A distribuição da espécie no

presente se manteve estável ao longo do Cerrado (Figura 11 C), com a região sudoeste apresentando altos níveis de adequabilidade (Figura 12 C).

Os índices de diversidade genética não apresentaram correlação significativa com as distâncias geográficas dos centroides das distribuições de 21 Ka, 6 Ka e do presente ( $P>0,05$ ). Também não foram encontradas evidências de correlação entre os índices de diversidade genética e os valores de estabilidade calculados para as populações ( $P>0,05$ ).



**Figura 11.** Projeção de distribuição para *Dipteryx alata* em três períodos: 21 Ka (A), 6 Ka (B) e presente (C).



**Figura 12.** Projeção de adequabilidade ambiental para *Dipteryx alata* em três períodos: 21 Ka (A), 6 Ka (B) e presente (C).

## DISCUSSÃO

Nossas análises filogeográficas combinadas com a modelagem de nicho agregam evidências de que o padrão atual de diversidade genética de *D. alata* é resultado de uma diminuição na área adequada para a espécie e retração de sua distribuição geográfica durante o LGM. No mesmo sentido, Collevatti et al. (2013) sugeriram que mudanças na distribuição da espécie resultaram no pequeno tamanho efetivo das populações, levando a fixação e perda de alelos devido a ação da deriva genética e, conseqüentemente, a baixa diversidade e alta divergência genética entre as populações da espécie.

Os resultados das análises realizadas com sequências de DNA nuclear e cloroplastidial do presente estudo reforçam que as populações de *D. alata* apresentam baixa diversidade genética, em concordância com os resultados de outros estudos que avaliaram regiões microssatélites (Collevatti et al. 2013, Soares et al. 2015, Tambarussi et al. 2017, Guimaraes et al. 2019). Esse conjunto de estudos aponta que os baixos níveis de diversidade genética encontrados podem ser explicados pelo sistema reprodutivo da espécie, no qual as sementes são produzidas por misturas de autofecundação e cruzamento não aleatório.

Em contraste, Tarazi et al. (2010) e Silva et al. (2023) sugeriram altos níveis de diversidade genética para *D. alata*, no entanto, ambos os estudos foram realizados com populações situadas em regiões que parecem ter mantido níveis estáveis de adequabilidade durante o LGM. De fato, Silva et al. (2023) sugerem que os níveis de diversidade genética observados podem ser reflexo de uma alta diversidade existente no passado, quando as populações avaliadas eram possivelmente maiores e conectadas.

Nossos resultados também corroboram dados anteriores em relação a presença de estrutura genética nas populações de *D. alata*. Tarazi et al. (2010) atribuíram esse padrão a existência de endogamia e limitação na dispersão de sementes. Nesse mesmo sentido, Tambarussi et al. (2017) sugeriram a influência da dispersão do pólen, que indica um padrão de isolamento por distância de dispersão devido ao comportamento dos polinizadores, resultando em uma alta diferenciação genética entre as populações e contribuindo para presença de estrutura espacial.

Em adição, Collevatti et al. (2013), em concordância com os dados obtidos no presente estudo, encontraram baixos níveis de fluxo gênico entre as populações, fator que também contribui para diferenciação genética entre as populações. Esses resultados suportam a ideia de que a dispersão de sementes da espécie é influenciada majoritariamente pela gravidade e secundariamente por mamíferos como morcegos, macacos e roedores, o que restringe o potencial de dispersão e resulta na formação de estrutura genética espacial, como também foi sugerido por Soares et al., (2008b). Em um contexto de indivíduos geneticamente próximos, devido a presença de estrutura espacial, o acasalamento entre indivíduos relacionados se torna inevitável.

Nossas projeções para a paleodistribuição de *D. alata* indicam redução da área de ocorrência durante o LGM, com manutenção de uma área estável abrangendo o sudoeste e centro oeste do Cerrado. Durante o holoceno médio houve de expansão para as regiões sudeste, leste e nordeste. Nossos dados genéticos também apontaram para um cenário de expansão demográfica recente, embora não tenha sido detectado padrão claro entre a distribuição da diversidade genética e as regiões recentemente recolonizadas. No entanto, a não detecção desse padrão pode ser devido à natureza evolutiva das regiões moleculares avaliadas, que apresentam relativamente baixas taxas mutacionais (Brito e Edwards, 2009) e, portanto, não havendo tempo suficiente para estruturação desse padrão.

Em congruência com essa perspectiva, Collevatti et al. (2013) avaliaram o mesmo conjunto de populações do presente estudo com marcadores microssatélites, que tendem a apresentar maiores taxas mutacionais, e encontraram maiores valores de riqueza alélica nas populações do oeste e sudoeste do Cerrado e uma tendência de diminuição dos valores desse parâmetro na medida em que as populações se distanciam dessas regiões. Nesse mesmo sentido, Soares et al. (2015) avaliaram 23 das 25 populações utilizadas no presente estudo com marcadores microssatélites cloroplastídias e encontraram correlações significativas entre riqueza alélica e diversidade genética e a longitude, de modo que populações localizadas mais a leste do Cerrado tendem a apresentar menores valores para os parâmetros estimados.

Esses resultados coincidem com nossas modelagens de distribuição, considerando que esses padrões de distribuição da diversidade genética são esperados em populações que passaram por processos de expansão recente, pois áreas que são colonizadas recentemente tendem a apresentar perda de alelos e homozigidade,

resultando em baixos níveis de diversidade genética (Hewitt, 2000, Excoffier e Ray, 2008). Nesse contexto, as regiões sudoeste e oeste do Cerrado parecem abrigar centro de origem e diversificação da espécie, com as populações mais antigas e com maiores índices de diversidade genética, enquanto as regiões central e leste representam as populações mais recentes.

Esses resultados apresentam relevância para a conservação da espécie, principalmente considerando as projeções para o futuro da espécie. Diniz-Filho et al., (2012) obtiveram um mapa consenso com projeção de deslocamento das áreas de adequabilidade do centro para o sudeste do Cerrado em 2050. Nesse mesmo sentido, Ribeiro et al., (2019) avaliaram o impacto das mudanças climáticas globais nas características morfológicas de germinação e nas áreas de distribuição da espécie. Seus resultados sugerem que em condições mais quentes a espécie apresentou menor peso máximo e menor comprimento de raiz, além da diminuição da área de distribuição no Cerrado, com deslocamento das áreas de maior adequabilidade da região centro oeste para a região sudeste. Nesse contexto, as regiões que atualmente apresentam maiores níveis de diversidade genética (oeste e sudoeste) estarão sujeitas a processos de redução e extinção populacionais, demandando medidas de conservação adequadas.

Carvalho et al., (2023) combinaram dados moleculares de regiões nuclear (ITS/5.8S) e cloroplastidiais (matK e trnL intron) e encontraram tempo de divergência entre *D. alata* e *D. magnifica* datando aproximadamente um Ma antes do presente. Nossa estimativa para o TMRCA (0,5 Ma) foi congruente com os resultados obtidos pelos autores, indicando uma colonização recente do baru no Cerrado.

## CONCLUSÕES

Os resultados obtidos no presente estudo apontam para as seguintes conclusões:

- As populações de *D. alata* apresentam baixos níveis de diversidade genética.
- As populações de *D. alata* apresentam níveis significativos de estrutura genética.
- A divergência entre as populações de *D. alata* está estruturada no espaço geográfico.
- O padrão atual de diversidade genética de *D. alata* é resultado de uma diminuição na área adequada para a espécie e retração de sua distribuição geográfica durante o LGM seguidas de expansão durante o holoceno médio.
- As populações localizadas nas regiões oeste e sudoeste do Cerrado representam as linhagens mais antigas de *D. alata*.

## REFERÊNCIAS

- ANTUNES, A. M.; TARGUETA, C. P.; CASTRO, A. A.; SOUZA, G.; SOARES, T. N.; TELLES, M. P. C. Genome size and chromosome number of *Dipteryx alata* (Leguminosae): a model candidate for comparative genomics in Papilionoideae. **Genetics and Molecular Research**, v. 19, p. 18640-18640, 2020.
- BARBOSA, A. R., FIORINI, C. F., SILVA-PEREIRA, V., MELLO-SILVA, R., BORBA, E. L. Geographical genetic structuring and phenotypic variation in the *Vellozia hirsuta* (Velloziaceae) ochlopecies complex. **American Journal of Botany**, v. 99, p. 1477–1488, 2012.
- BEHEREGARAY, L. B. Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. **Molecular Ecology**, v. 17, p. 3754–3774, 2008.
- BRITO, P.H., EDWARDS, S.V. Multilocus phylogeography and phylogenetics using sequence-based markers. **Genetica**, v. 135, p. 439-455, 2009.
- CARDOSO, D.; DE QUEIROZ, L. P.; TOBY PENNINGTON, R.; et al. Revisiting the phylogeny of papilionoid legumes: New insights from comprehensively sampled early-branching lineages. **American Journal of Botany**, v. 99, p. 1991–2013, 2012.
- CARVALHO, P. Cumaru-Ferro *Dipteryx odorata*. **Embrapa - Tech. Release**, p. 3:01-08, 2009. Disponível em:  
<<https://www.infoteca.cnptia.embrapa.br/bitstream/doc/578657/1/CT225.pdf>>.
- CARVALHO, C.S., FRAGA, C.N., CARDOSO, D.B.O.S., LIMA, H.C. Tonka, baru and cumaru: Nomenclatural overview, typification and updated checklist of *Dipteryx* (Leguminosae). **Taxon**, v. 69, p. 582-592, 2020
- CARVALHO, C.S., LIMA, H.C., LEMES, M.R., ZARTMAN, C.E., BERG, C.V.D., GARCIA-DAVILA, C.R., CORONADO, E.N.H., MADER, M., VILLANUEVA, K.P., TYSKLIND, N., CARDOSO, D. A dated phylogeny of the neotropical *Dipterygeae* clade reveals 30 million years of winged papilionate floral conservatism in the otherwise florally labile early-branching papilionoide legumes. **Botanical Journal of the Linnean Society**, v. 20, p. 1-27, 2023
- COLLEVATTI, R. G., GRATAPAGLIA, D., HAY, J. D. Evidences for multiple maternal lineages of *Caryocar brasiliense* populations in the Brazilian Cerrado based on the analysis of chloroplast DNA sequences and microsatellite haplotype variation. **Molecular Ecology**, v. 12, p. 105–115, 2003.
- COLLEVATTI, R. G., RABELO, S. G., VIEIRA, R. F. Phylogeography and disjunct distribution in *Lychnophora ericoides* (Asteraceae), an endangered cerrado shrub species. **Annals of Botany**, v. 104, p. 655–664, 2009.
- COLLEVATTI, R.G., LIMA, J., SOARES, T. N., TELLES, M. P. C. Spatial Genetic Structure and Life History Traits in Cerrado Tree Species: Inferences for Conservation. **Natureza & Conservação**, v. 8, p. 54–59, 2010.
- COLLEVATTI, R. G., LIMA-RIBEIRO, M. S., SOUZA-NETO, A. C. Recovering the demographical history of a Brazilian Cerrado tree species *Caryocar brasiliense*:

coupling ecological niche modeling and coalescent analyses. **Natureza e Conservação**, v. 10, p. 169–176, 2012a.

COLLEVATTI, R. G., DE CASTRO, T. G., LIMA, J. S., TELLES, M. P. C. Phylogeography of *Tibouchina papyrus* (Pohl) Toledo (Melastomataceae), an endangered tree species from rocky savannas, suggests bidirectional expansion due to climate cooling in the pleistocene. **Ecology and Evolution**, v. 2, p. 1024–1035, 2012b.

COLLEVATTI, R. G.; TELLES, M. P. C.; NABOUT, J. C.; CHAVES, L. J.; SOARES, T. N. Demographic history and the low genetic diversity in *Dipteryx alata* (Fabaceae) from Brazilian Neotropical savannas. **Heredity**, v. 111, p. 97–105, 2013.

COLLEVATTI, R. G., TERRIBILE, L. C., RABELO, S. G., LIMA-RIBEIRO, M. S. Evolutionary and population genetics relaxed random walk model coupled with ecological niche modeling unravel the dispersal dynamics of a Neotropical savanna tree species in the deeper Quaternary. **Frontiers in Plant Sciences**, v. 6, p. 1–15, 2015.

CORNUET, J. M.; PUDLO, P.; VEYSSIER, J.; DEHNE-GARCIA, A.; GAUTIER, M.; LEBLOIS, R.; MARIN, J. M.; ESTOUP, A. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. **Bioinformatics**, v. 30, p. 1187–89, 2014.

DINIZ-FILHO, J. A. F.; COLLEVATTI, R. G.; CHAVES, L. J.; et al. Geographic shifts in climatically suitable areas and loss of genetic variability in *Dipteryx alata* (“Baru” Tree; Fabaceae). **Genetics and Molecular Research**, v. 11, p. 1618–1626, 2012.

DINIZ-FILHO, J.A.F., SOARES, T.N., TELLES, M.P.C. Geographically weighted regression as a generalized Wombling to detect barriers to gene flow. **Genetica**, v. 144, p. 425–433, 2016.

DRUMMOND, A.J.; SUCHARD, M.A.; XIE, D.; RAMBAUT, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. **Molecular Biology and Evolution**, v. 29, p. 1969–1973, 2012.

EXCOFFIER, L.; SMOUSE, P.E.; QUATTRO, J.M. Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. **Genetics**, v. 131, p. 479–491, 1992.

EXCOFFIER, L.; LAVAL, G.; SCHNEIDER, S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. **Evolutionary Bioinformatics Online**, v. 1, p. 47–50, 2005.

FU, Y. X. Statistical Tests of Neutrality of Mutations against Population Growth, Hitchhiking and Background Selection. **Genetics**, v. 147, p. 915–925, 1997.

GUIMARÃES, R. A.; CORRÊA MIRANDA, K. M.; CHAVES, L. J., NAVES, R.V., SOARES, T.N., TELLES, M.P.C. Mating system and pollen dispersal in *Dipteryx alata* Vogel (Leguminosae): comparing in situ and ex situ conditions. **Tree Genetics and Genomes**, v. 15, n. 2, 2019.

HARPENDING, H. C. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. **Human Biology**, v. 66, p. 591–600, 1994.

- KLINK, C.A. & MACHADO, R.B. Conservation of the Brazilian Cerrado. **Conservation Biology**, v. 19, p. 707–713, 2005.
- KUHNER, M. K. Lamarc 2.0: Maximum likelihood and Bayesian estimation of population parameter. **Bioinformatics Applications Note**, v. 22, p. 768–770, 2006.
- LEAL, B.S.S., SILVA, C.P., PINHEIRO, S.F. Phylogeographic studies depict the role of space and time scales of plant speciation in a highly diverse neotropical region. **Critical Review in Plant Sciences**, v. 35, p. 215–230, 2016.
- LIBRADO, P; ROZAS, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, p.1451–1452, 2009.
- LIMA, N. E., LIMA-RIBEIRO, M. S., TINOCO, C. F., TERRIBILE, R. C., COLLEVATI, R. G. Phylogeography and ecological niche modelling, coupled with the fossil pollen record, unravel the demographic history of a Neotropical swamp palm through the Quaternary. **Journal of Biogeography**, v. 41, p. 673–686, 2014.
- LPWG. Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. **Taxon**, v. 62, p.217–248, 2013.
- MARSKE, K. Phylogeography. **Encyclopedia of Evolutionary Biology**, v. 3, p. 291–296, 2016.
- MARTINS, F. DE M.; DOMINGUES, M. V. Filogeografia. **Revista da Biologia**, v. 7, p. 26–30, 2012.
- MILNE, I.; LINDNER, D.; BAYER, M.; HUSMEIER, D.; MCGUIRE, G.; MARSHALL, D. F.; WRIGHT, F. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. **Bioinformatics**, v. 25, p. 126–127, 2009.
- MORRIS, A. B.; SHAW, J. Markers in time and space: A review of the last decade of plant phylogeographic approaches. **Molecular Ecology**, v. 27, p. 2317–2333, 2018.
- MYERS, N., MITTERMEIER, R.A., MITTERMEIER, C.G., DA FONSECA, G.A.B. & KENT, J. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, p. 853–858, 2000.
- NABOUT, JC; SOARES, TN; DINIZ-FILHO, JAF; DE MARCO JÚNIOR, P; TELLES, MPC; NAVES, RV; CHAVES, L. Combining multiple models to predict the geographical distribution of the Baru tree (*Dipteryx alata* Vogel) in the Brazilian Cerrado. **Brazilian Journal of Biology**, v. 70, p. 911–919, 2010.
- NOVAES, R. M. L., LEMOS-FILHO, J. P., RIBEIRO, R. A., LOVATO, M. B. Phylogeography of *Plathymenia reticulata* (Leguminosae) reveals patterns of recent range expansion towards northeastern Brazil and southern Cerrados in Eastern Tropical South America. **Molecular Ecology**, v. 19, p. 985–998, 2010.

- NOVAES, R. M. L., RIBEIRO, R. A., LEMOS-FILHO, J. P., LOVATO, M. B. Concordance between phylogeographical and biogeographical patterns in the Brazilian Cerrado: diversification of the endemic tree *Dalbergia miscolobium* (Fabaceae). **PLoS One**, v. 8, p. e82198–e82198, 2013.
- OLIVEIRA, M. I. B.; SIGRIST, R. Fenologia reprodutiva, polinização e reprodução de *Dipteryx alata* Vogel (Leguminosae-Papilionoideae) em Mato Grosso do Sul, Brasil. **Brazilian Journal of Botany**, v. 2, p. 195–207, 2008.
- POSADA, D. jModelTest: phylogenetic model averaging. **Molecular Biology and Evolution**, v. 25, p. 253–256, 2008.
- PRADO, C. P. A., HADDAD, C. F. B., ZAMUDIO, K. R. Cryptic lineages and Pleistocene population expansion in a Brazilian Cerrado frog. **Molecular Ecology**, v. 21, p. 921–941, 2012.
- RAMBAUT, A. **FigTree** 1.4.0., a graphical viewer of phylogenetic trees. 2012. Disponível em: <http://tree.bio.ed.ac.uk/software/figtree>.
- RAMBAUT, A.; DRUMMOND, A. J. **Tracer** v1.5. 2009. Disponível em: <http://tree.bio.ed.ac.uk/software/tracer/>.
- RAMOS, A. C. S., LEMOS-FILHO, J. P., RIBEIRO, R. A., SANTOS, F. R., LOVATO, M. B. Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of quaternary climate changes in the Brazilian cerrado. **Annals of Botany**, v. 100, p. 1219–1228, 2007.
- RAMOS, A. C. S., LEMOS-FILHO, J. P., LOVATO, M. B. Phylogeographical structure of the neotropical forest tree *Hymenaea courbaril* (Leguminosae: Caesalpinioideae) and its relationship with the vicariant *Hymenaea stigonocarpa* from Cerrado. **Journal of Heredity**, v. 100, p. 206–216, 2009.
- RIBEIRO, P. C., LEMOS-FILHO, J. P., BUZATTI, R. S. O., LOVATO, M. B., HEUERTZ, M. Species-specific phylogeographical patterns and Pleistocene east-west divergence in *Annona* (Annonaceae) in the Brazilian Cerrado. **Botanical Journal of Linnean Society**, v. 181, p. 21–36, 2016.
- RICHARDSON, J. E., PENNINGTON, R. T., PENNINGTON, T. D., HOLLINGSWORTH, P. M. Rapid diversification of a Species-Rich Genus of Neotropical Rain Forest Trees. *Science*, v. 293, p. 2242–2245, 2001.
- SANO, S. M.; RIBEIRO, J. F.; BRITO, M. A. DE. Baru: biologia e uso. **Embrapa Cerrado**, v. 116, p. 51, 2004.
- SILVA, R.W.L., MACHADO, S.S., FARIA, K.C., OLIVEIRA, B.A., SOUZA, E.B., OLIVEIRA, F.A., MENEZES, I.P.P., SILVA, J.M. Diversity and genetic structure of *Dipteryx alata* Vogel (Fabaceae) populations in the Cerrado-Amazon ecotone, Brazil. **Plant Molecular Biology Report**, v. 41, p. 218–228, 2023.
- SOARES, T.N., CHAVES, L.J., TELLES, M.P.D.C., DINIZ-FILHO, J.A.F., RESENDE, L. V. Distribuição espacial da variabilidade genética intrapopulacional de

- Dipteryx alata. **Pesquisa Agropecuária Brasileira**, v. 43, p. 1151–1158, 2008.
- SOARES, T.N., CHAVES, L.J., TELLES, M.P.C., DINIZ-FILHO, J.A.F., RESENDE, L. V. Landscape conservation genetics of *Dipteryx alata* (“baru” tree: Fabaceae) from Cerrado region of central Brazil. **Genetica**, v. 132, p. 9–19, 2008b.
- SOARES, T. N.; DINIZ-FILHO, J. A. F.; NABOUT, J. C.; TELLES, M. P. C.; TERRIBILE, L. C.; CHAVES, L. J. Patterns of genetic variability in central and peripheral populations of *Dipteryx alata* (Fabaceae) in the Brazilian Cerrado. **Plant Systematics and Evolution**, v. 301, n. 5, p. 1315–1324, 2015.
- SOUZA, U. J. B. DE; TELLES, M. P. DE C.; DINIZ-FILHO, J. A. F. Trends in the scientific literature on population genetics of plants from Cerrado. **Hoehnea**, v. 43, n. 3, p. 461–477, 2016.
- SOUZA, H. A. V., COLLEVATTI, R. G., LIMA-RIBEIRO, M. S., LEMOS-FILHO, J. P., LOVATO, M. B. 2016. A large historical refugium explains spatial patterns of genetic diversity in a Neotropical savanna tree species. **Annals of Botany**, v. 119, p. 239-252, 2016.
- STEPHENS, M. Inference Under the Coalescent. **Handbook of Statistical Genetics: Third Edition**, p. 878–908, 2008.
- STEPHENS, M.; SMITH, N.; DONNELLY, P. A new statistical method for haplotype reconstruction from population data. **American Journal of Human Genetics**, v. 68, p. 978-989, 2001.
- TAMBARUSSI, E. V.; SEBBENN, A. M.; ALVES-PEREIRA, A. *Dipteryx alata* Vogel (Fabaceae), a neotropical tree with high levels of selfing: Implications for conservation and breeding programs. **Annals of Forest Research**, v. 60, n. 2, p. 243–261, 2017.
- TAMURA, K.; DUDLEY, J.; NEI, M.; KUMAR, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA). **Molecular Biology and Evolution**, v. 24, p. 1596-1599, 2007.
- TELLER, M.P.C., DOBROVOLSKI, R., SOUZA, K. S., LIMA, J. S., COLLEVATTI, R. G., SOARES, T. S., CHAVES, L. J., DINIZ-FILHO, J. A. F. Disentangling landscape effects on population genetic structure of a Neotropical savanna tree. **Natureza & Conservação**, v. 12, p. 65–70, 2014.

## APÊNDICES

**Tabela S1.** Iniciadores utilizados na amplificação e sequenciamento de regiões cloroplastidiais de *Dipteryx alata*.

cpDNA	Iniciadores	Sequência Nucleotídica 5' → 3'	Referência
trnLA/B	trnLA	CATTACAAATGCGATGCTCT	Taberlet et al. 1991
	trnLB	TCTACCGATTTCCGCATATC	
trnLC/D	trnLC	CGAAATCGGTAGACGCTACG	Taberlet et al. 1991
	trnLD	GGGGATAGAGGGACTTGAAC	
trnLE/F	trnLE	GGTTCAAGRCCCTCTATCCC	Taberlet et al. 1991
	trnLF	ATTTGAACTGGTGACACGAG	
rpS16	rpS16F	AAACGATGTGGTARAAAGCAAC	Shaw et al. 2005
	rpS16R	AACATCWATTGCAASGATTCGATA	
rpL16	rpL16F	GCTATGCTTAGTGTGTGACTCGTTG	Shaw et al. 2005
	rpL16R	CCCTTCATTCTTCCTCTATGTTG	
trnQ/rpS16	trnQ	GCGTGCCAAAGYGGTAAGGC	Shaw et al. 2005
	rpS16x1	GTTGCTTTYTACCACATCGTTT	
rpS12/rpL0	rpS12	ATTAGAAANRCAAGACAGCCAAT	Shaw et al. 2005
	rpL20	CGYYAYCGAGCTATATATCC	
psbA/trnH	psbA	GTTATGCATGAACGTAATGCTC	Sang et al. 1997
	trnH	CGCGCATGGTGGATTCAACAATCC	
trnS/trnG	trnS	AGATAGGGATTTCGAACCCTCGGT	Shaw et al. 2005
	trnG2S	TTTTACCACTAAACTATAACCCGC	
trnC/ycf6	trnC	CCAGTTCRAATCYGGGTG	Demesure et al. 1995
	ycf6	GCCCAAGCRAGACTTACTATATCCAT	
trnD/trnT	trnD	ACCAATTGAACTACAATCCC	Demesure et al. 1995
	trnT(	CTACCACTGAGTTAAAAGGG	
psbC/trnS3	psbC	GGTCGTGACCAAGAAACCAC	Demesure et al. 1995
	trnS3-RCS5'	AAGATATGCCAGATTCCACC	
trnV/trnM	trnV	GCTATACGGGCTCGAACC	Huang et al. 2002
	trnM	TACCTACTATTGGATTTGAACC	

**Tabela S2.** Iniciadores utilizados na amplificação e sequenciamento de regiões nucleares de *Dipteryx alata*.

nDNA	Iniciadores	Sequência Nucleotídica 5' → 3'	Referência
ITS04/05	ITS 05-F	GGAAGTAAAAGTCGTAACAAGG	Wright et al., 2001
	ITS 04-R	TCCTCCGCTTATTGATATGC	
ITS75/92	ITS 92	AAGGTTTCCGTAGGTGAAC	Desfeux & Lejeune, 1996
	ITS 75	TATGCTTAAACTCAGCGGG	
ITS101/102	ITS AB101	ACGAATTCATGGTCCGGTGAAGT	Douzery et al., 1999
	ITS AB102	TAGAATTCCTCCGGTTCGCTCGCC	
ETS/ETS18S	ETS_F	CTCCGTGCTGGTGCATCGAACTGC	Wright et al., 2001
	ETS-18S_R	GAGCCATTCGCAGTTTCACAG	
18S/26S IGS	18S-IGS_F	GGATTGTTCACCCACCAATAGGG	Baldwin & Markos, 1998
	26S-IGS_R	GAGACAAGCATATGACTACTGGC	
18SIGS/18SE	18S-IGS_F	GCAGGATCAACCAGGTAGCA	Baldwin & Markos, 1998
	18S-E_R	ACTTACACATGCATGGCTTAATCT	
18S/Umb	18S-ETS_R	GCGCATGAGTGGTGAWTKGTA	Baldwin & Markos, 1998
	Umb-ETS_F	AAGCATATGACTACTGGCAGGAT	
18S/AST8	18S-ETS_R	ACTTACACATGCATGGCTTAA	Baldwin & Markos, 1998
	AST-8_F	TTCTCTTCGTATCGTGCGGT	

**Tabela S3.** Valores de  $F_{ST}$  par-a-par entre 25 populações de *Dipteryx alata* obtidos a partir das sequências cloroplastidiais. Os valores significativos encontram-se em negrito ( $p < 0,5$ ).

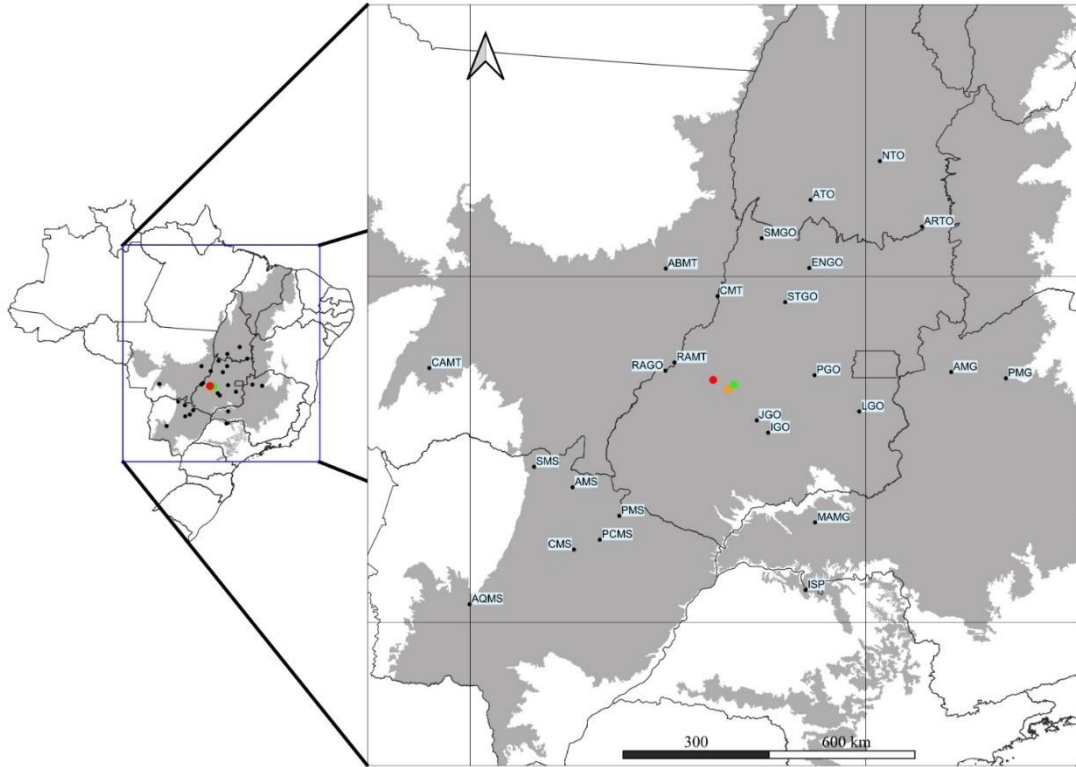
	CMT	ABMT	PGO	SMS	AMS	ATO	SMGO	LGO	ISP	MAMG	ENGO	STGO	AMG	PMG	PMS	PCMS	CMS	IGO	RAMT	RAGO	JGO	NTO	ARTO	AQMS	CAMT
CMT	0.000																								
ABMT	<b>0.356</b>	0.000																							
PGO	<b>0.636</b>	0.090	0.000																						
SMS	<b>0.566</b>	0.060	0.000	0.000																					
AMS	<b>0.459</b>	<b>0.206</b>	<b>0.246</b>	0.125	0.000																				
ATO	<b>0.355</b>	0.091	0.091	0.061	<b>0.206</b>	0.000																			
SMGO	<b>0.234</b>	0.068	0.182	0.136	<b>0.222</b>	0.068	0.000																		
LGO	<b>0.636</b>	0.090	0.000	0.000	<b>0.246</b>	0.091	0.182	0.000																	
ISP	<b>0.528</b>	0.091	0.089	0.057	0.041	0.091	0.146	0.089	0.000																
MAMG	<b>0.515</b>	<b>0.212</b>	<b>0.272</b>	0.097	0.036	<b>0.212</b>	<b>0.234</b>	<b>0.272</b>	0.013	0.000															
ENGO	<b>0.433</b>	0.042	0.000	0.000	0.147	0.042	0.010	0.000	0.045	0.181	0.000														
STGO	<b>0.298</b>	0.024	0.087	0.070	0.120	0.025	0.016	0.087	0.090	0.181	0.058	0.000													
AMG	<b>0.636</b>	0.090	0.000	0.000	<b>0.247</b>	0.091	0.182	0.000	0.089	<b>0.272</b>	0.000	0.087	0.000												
PMG	0.098	<b>0.698</b>	<b>0.909</b>	<b>0.833</b>	<b>0.665</b>	<b>0.698</b>	<b>0.594</b>	<b>0.909</b>	<b>0.783</b>	<b>0.739</b>	<b>0.731</b>	<b>0.590</b>	<b>0.909</b>	0.000											
PMS	<b>0.366</b>	<b>0.216</b>	<b>0.327</b>	<b>0.201</b>	0.004	<b>0.215</b>	<b>0.201</b>	<b>0.327</b>	0.103	0.017	<b>0.226</b>	0.191	<b>0.327</b>	<b>0.583</b>	0.000										
PCMS	<b>0.484</b>	<b>0.344</b>	<b>0.424</b>	<b>0.300</b>	0.011	<b>0.344</b>	<b>0.336</b>	<b>0.424</b>	0.193	0.029	<b>0.326</b>	<b>0.285</b>	<b>0.424</b>	<b>0.660</b>	0.053	0.000									
CMS	<b>0.369</b>	0.059	0.108	0.008	0.052	0.059	0.079	0.108	0.045	0.052	0.031	0.029	0.108	<b>0.615</b>	0.014	0.052	0.000								
IGO	<b>0.554</b>	0.057	0.008	0.000	<b>0.200</b>	0.057	0.129	0.008	0.056	<b>0.208</b>	0.002	0.064	0.008	<b>0.827</b>	<b>0.275</b>	<b>0.369</b>	0.083	0.000							
RAMT	<b>0.234</b>	0.068	0.182	0.136	<b>0.222</b>	0.068	0.091	0.182	0.146	<b>0.234</b>	0.010	0.016	0.182	<b>0.594</b>	<b>0.201</b>	<b>0.336</b>	0.079	0.129	0.000						
RAGO	<b>0.636</b>	0.090	0.000	0.000	<b>0.246</b>	0.091	0.182	0.000	0.089	<b>0.272</b>	0.000	0.087	0.000	<b>0.909</b>	<b>0.327</b>	<b>0.424</b>	0.108	0.008	0.182	0.000					
JGO	<b>0.612</b>	0.070	0.000	0.016	<b>0.219</b>	0.070	0.157	0.000	0.069	<b>0.245</b>	0.016	0.066	0.000	<b>0.901</b>	<b>0.298</b>	<b>0.395</b>	0.087	0.009	0.157	0.000	0.000				
NTO	0.093	<b>0.310</b>	<b>0.606</b>	0.532	<b>0.429</b>	0.310	0.188	<b>0.606</b>	<b>0.493</b>	<b>0.485</b>	<b>0.393</b>	<b>0.259</b>	<b>0.606</b>	0.133	<b>0.337</b>	<b>0.459</b>	<b>0.334</b>	<b>0.519</b>	0.188	<b>0.606</b>	<b>0.579</b>	0.000			
ARTO	<b>0.234</b>	0.068	0.182	0.137	<b>0.222</b>	0.068	0.091	0.182	0.146	<b>0.234</b>	0.010	0.016	0.182	<b>0.594</b>	<b>0.201</b>	<b>0.336</b>	0.079	0.130	0.091	0.182	0.158	0.188	0.000		
AQMS	<b>0.636</b>	0.090	0.000	0.000	<b>0.246</b>	0.091	0.182	0.000	0.089	<b>0.272</b>	0.000	0.087	0.000	<b>0.909</b>	<b>0.327</b>	<b>0.424</b>	0.108	0.008	0.182	0.000	0.000	<b>0.606</b>	0.182	0.000	
CAMT	<b>0.636</b>	0.090	0.000	0.000	<b>0.246</b>	0.091	0.182	0.000	0.089	<b>0.272</b>	0.000	0.087	0.000	<b>0.909</b>	<b>0.327</b>	<b>0.424</b>	0.108	0.008	0.182	0.000	0.000	<b>0.606</b>	0.182	0.000	0.000

**Tabela S4.** Valores de  $F_{ST}$  par-a-par entre 25 populações de *Dipteryx alata* obtidos a partir das sequências nucleares. Os valores significativos encontram-se em negrito ( $p < 0,5$ )

	CMT	ABMT	PGO	SMS	AMS	ATO	SMGO	LGO	ISP	MAMG	ENGO	STGO	AMG	PMG	PMS	PCMS	CMS	IGO	RAMT	RAGO	JGO	NTO	ARTO	AQMS	CAMT
CMT	0.000																								
ABMT	<b>0.340</b>	0.000																							
PGO	0.074	0.132	0.000																						
SMS	<b>0.330</b>	0.056	0.137	0.000																					
AMS	<b>0.711</b>	<b>0.212</b>	<b>0.555</b>	<b>0.381</b>	0.000																				
ATO	0.045	0.116	0.019	0.104	<b>0.507</b>	0.000																			
SMGO	<b>0.785</b>	<b>0.301</b>	<b>0.516</b>	<b>0.378</b>	<b>0.681</b>	<b>0.524</b>	0.000																		
LGO	<b>0.391</b>	0.047	0.183	<b>0.217</b>	<b>0.333</b>	0.176	<b>0.506</b>	0.000																	
ISP	<b>0.252</b>	0.006	0.102	0.020	<b>0.259</b>	0.051	<b>0.398</b>	0.098	0.000																
MAMG	<b>0.675</b>	<b>0.266</b>	<b>0.529</b>	<b>0.501</b>	<b>0.283</b>	<b>0.486</b>	<b>0.756</b>	<b>0.226</b>	<b>0.314</b>	0.000															
ENGO	<b>0.713</b>	<b>0.241</b>	<b>0.426</b>	<b>0.300</b>	<b>0.631</b>	<b>0.446</b>	0.004	<b>0.438</b>	<b>0.334</b>	<b>0.710</b>	0.000														
STGO	<b>0.557</b>	0.155	<b>0.260</b>	0.163	<b>0.566</b>	<b>0.296</b>	0.131	<b>0.338</b>	<b>0.225</b>	<b>0.642</b>	0.026	0.000													
AMG	<b>0.605</b>	0.181	<b>0.306</b>	<b>0.239</b>	<b>0.592</b>	<b>0.347</b>	0.101	<b>0.352</b>	<b>0.272</b>	<b>0.653</b>	0.007	0.023	0.000												
PMG	<b>0.557</b>	0.063	<b>0.287</b>	<b>0.208</b>	<b>0.389</b>	<b>0.310</b>	0.174	0.185	0.185	<b>0.461</b>	0.118	0.093	0.069	0.000											
PMS	<b>0.730</b>	<b>0.262</b>	<b>0.588</b>	<b>0.429</b>	0.007	<b>0.535</b>	<b>0.729</b>	<b>0.406</b>	<b>0.291</b>	<b>0.308</b>	<b>0.680</b>	<b>0.614</b>	<b>0.640</b>	<b>0.455</b>	0.000										
PCMS	<b>0.550</b>	0.153	<b>0.396</b>	0.151	0.198	<b>0.321</b>	<b>0.615</b>	<b>0.282</b>	0.098	<b>0.425</b>	<b>0.559</b>	<b>0.463</b>	<b>0.517</b>	<b>0.389</b>	<b>0.259</b>	0.000									
CMS	<b>0.668</b>	0.171	<b>0.506</b>	<b>0.323</b>	0.027	<b>0.455</b>	<b>0.644</b>	<b>0.275</b>	<b>0.206</b>	<b>0.263</b>	<b>0.593</b>	<b>0.525</b>	<b>0.554</b>	<b>0.358</b>	0.049	0.123	0.000								
IGO	0.070	0.108	0.029	0.091	<b>0.513</b>	0.038	<b>0.527</b>	0.157	0.056	<b>0.498</b>	<b>0.441</b>	<b>0.279</b>	<b>0.333</b>	<b>0.295</b>	<b>0.549</b>	<b>0.321</b>	<b>0.457</b>	0.000							
RAMT	<b>0.361</b>	0.024	0.153	0.039	<b>0.220</b>	0.128	<b>0.359</b>	0.042	0.002	<b>0.301</b>	<b>0.295</b>	0.197	<b>0.235</b>	0.112	<b>0.293</b>	0.103	0.160	0.109	0.000						
RAGO	<b>0.403</b>	0.008	0.166	0.009	<b>0.282</b>	0.159	<b>0.249</b>	0.129	0.031	<b>0.411</b>	0.183	0.091	0.135	0.060	<b>0.350</b>	0.164	<b>0.232</b>	0.139	0.011	0.000					
JGO	0.071	0.115	0.008	0.107	<b>0.502</b>	0.033	<b>0.543</b>	0.144	0.053	<b>0.477</b>	<b>0.466</b>	<b>0.316</b>	<b>0.367</b>	<b>0.316</b>	<b>0.538</b>	<b>0.304</b>	<b>0.444</b>	0.040	0.109	0.155	0.000				
NTO	<b>0.483</b>	0.143	0.179	<b>0.231</b>	<b>0.578</b>	<b>0.244</b>	<b>0.288</b>	<b>0.259</b>	<b>0.226</b>	<b>0.592</b>	0.167	0.050	0.027	0.080	<b>0.621</b>	<b>0.503</b>	<b>0.540</b>	<b>0.226</b>	<b>0.203</b>	0.142	<b>0.265</b>	0.000			
ARTO	<b>0.505</b>	<b>0.287</b>	<b>0.235</b>	<b>0.451</b>	<b>0.700</b>	<b>0.307</b>	<b>0.739</b>	<b>0.300</b>	<b>0.351</b>	<b>0.626</b>	<b>0.625</b>	<b>0.452</b>	<b>0.440</b>	0.343	<b>0.729</b>	<b>0.644</b>	<b>0.666</b>	<b>0.307</b>	<b>0.362</b>	<b>0.376</b>	<b>0.329</b>	0.206	0.000		
AQMS	<b>0.670</b>	<b>0.217</b>	<b>0.522</b>	<b>0.316</b>	0.049	<b>0.459</b>	<b>0.696</b>	<b>0.333</b>	<b>0.206</b>	<b>0.355</b>	<b>0.646</b>	<b>0.569</b>	<b>0.606</b>	0.442	0.110	0.045	0.002	<b>0.464</b>	0.188	<b>0.267</b>	<b>0.446</b>	<b>0.591</b>	<b>0.709</b>	0.000	
CAMT	<b>0.754</b>	<b>0.265</b>	<b>0.604</b>	<b>0.460</b>	0.029	<b>0.562</b>	<b>0.719</b>	<b>0.408</b>	<b>0.329</b>	<b>0.313</b>	<b>0.669</b>	<b>0.607</b>	<b>0.626</b>	0.414	0.004	<b>0.346</b>	0.090	<b>0.575</b>	<b>0.309</b>	<b>0.353</b>	<b>0.567</b>	<b>0.610</b>	<b>0.726</b>	0.211	0.000

**Tabela S5.** Número de migrantes por geração para as 25 populações de *Dipteryx alata* no Cerrado brasileiro, com base na análise de coalescência bayesiana dos dados de cpDNA (psbA-trnH) e nDNA (ITS101/102) concatenados. A direção de migração é de populações nas linhas (Fonte) para populações nas colunas (Receptor)

População		Receptora																								
	CMT	ABMT	PGO	SMS	AMS	ATO	SMGO	LGO	ISP	MAMG	ENGO	STGO	AMG	PMG	PMS	PCMS	CMS	IGO	RAMT	RAGO	JGO	NTO	ARTO	AQMS	CAMT	
	CMT	0	0.417	0.064	0.111	0.010	0.061	0.011	0.153	0.124	0.008	0.016	0.040	0.050	0.036	0.067	0.047	0.018	0.010	0.178	0.154	0.189	0.012	0.040	0.067	0.009
	ABMT	0.006	0	0.039	0.100	0.028	0.007	0.010	0.166	0.120	0.032	0.016	0.006	0.019	0.104	0.023	0.044	0.033	0.062	0.173	0.153	0.175	0.046	0.025	0.078	0.010
	PGO	0.049	0.310	0	0.118	0.017	0.063	0.001	0.273	0.136	0.009	0.009	0.029	0.028	0.075	0.017	0.028	0.033	0.063	0.199	0.203	0.187	0.023	0.016	0.010	0.003
	SMS	0.040	0.480	0.110	0	0.010	0.060	0.006	0.283	0.110	0.012	0.004	0.025	0.032	0.063	0.010	0.044	0.020	0.057	0.176	0.199	0.186	0.023	0.006	0.020	0.017
	AMS	0.017	0.136	0.017	0.090	0	0.018	0.005	0.129	0.138	0.031	0.019	0.017	0.036	0.065	0.072	0.043	0.037	0.021	0.196	0.216	0.179	0.017	0.039	0.082	0.012
	ATO	0.047	0.447	0.049	0.100	0.012	0	0.011	0.274	0.112	0.014	0.022	0.040	0.049	0.113	0.008	0.048	0.035	0.046	0.149	0.205	0.113	0.074	0.037	0.013	0.007
	SMGO	0.038	0.502	0.102	0.103	0.004	0.050	0	0.219	0.147	0.005	0.013	0.005	0.042	0.034	0.033	0.019	0.038	0.055	0.207	0.203	0.105	0.076	0.035	0.026	0.017
	LGO	0.006	0.079	0.115	0.068	0.022	0.048	0.002	0	0.131	0.034	0.021	0.005	0.038	0.036	0.025	0.047	0.037	0.016	0.056	0.135	0.171	0.066	0.040	0.081	0.005
	ISP	0.022	0.476	0.093	0.075	0.004	0.048	0.008	0.199	0	0.025	0.008	0.006	0.036	0.016	0.018	0.008	0.031	0.062	0.096	0.169	0.176	0.046	0.034	0.080	0.013
	MAMG	0.016	0.361	0.062	0.087	0.024	0.009	0.007	0.106	0.156	0	0.006	0.018	0.027	0.069	0.019	0.009	0.034	0.054	0.204	0.200	0.188	0.011	0.022	0.080	0.015
	ENGO	0.043	0.484	0.111	0.083	0.004	0.048	0.006	0.223	0.124	0.006	0	0.040	0.045	0.106	0.066	0.009	0.016	0.055	0.201	0.230	0.151	0.073	0.034	0.083	0.017
	STGO	0.044	0.478	0.082	0.128	0.005	0.050	0.011	0.079	0.119	0.010	0.019	0	0.008	0.115	0.072	0.037	0.038	0.061	0.173	0.229	0.162	0.067	0.040	0.041	0.012
Fonte	AMG	0.047	0.089	0.036	0.026	0.006	0.056	0.002	0.249	0.127	0.009	0.020	0.041	0	0.098	0.024	0.010	0.028	0.060	0.205	0.230	0.183	0.050	0.040	0.069	0.016
	PMG	0.018	0.501	0.109	0.031	0.028	0.061	0.004	0.258	0.147	0.007	0.017	0.040	0.041	0	0.072	0.016	0.024	0.008	0.163	0.030	0.121	0.059	0.033	0.066	0.007
	PMS	0.022	0.214	0.017	0.092	0.022	0.062	0.009	0.275	0.114	0.033	0.015	0.029	0.007	0.089	0	0.040	0.036	0.010	0.181	0.193	0.066	0.011	0.006	0.069	0.017
	PCMS	0.020	0.499	0.015	0.099	0.024	0.021	0.005	0.208	0.150	0.032	0.009	0.035	0.028	0.117	0.020	0	0.038	0.042	0.204	0.083	0.032	0.023	0.014	0.044	0.004
	CMS	0.013	0.408	0.038	0.127	0.028	0.021	0.002	0.279	0.130	0.032	0.006	0.006	0.018	0.116	0.071	0.047	0	0.029	0.197	0.159	0.157	0.065	0.024	0.016	0.018
	IGO	0.011	0.392	0.045	0.126	0.028	0.054	0.007	0.203	0.147	0.008	0.002	0.040	0.031	0.115	0.033	0.016	0.011	0	0.210	0.149	0.041	0.027	0.039	0.065	0.008
	RAMT	0.044	0.480	0.014	0.114	0.005	0.059	0.009	0.049	0.100	0.016	0.019	0.023	0.014	0.084	0.071	0.041	0.038	0.064	0	0.041	0.180	0.047	0.039	0.072	0.010
	RAGO	0.010	0.138	0.109	0.018	0.022	0.060	0.009	0.278	0.117	0.026	0.018	0.040	0.031	0.065	0.071	0.039	0.035	0.008	0.162	0	0.151	0.012	0.006	0.037	0.016
	JGO	0.022	0.082	0.103	0.098	0.020	0.062	0.005	0.083	0.084	0.006	0.003	0.028	0.009	0.110	0.043	0.042	0.009	0.053	0.138	0.229	0	0.069	0.039	0.032	0.002
	NTO	0.043	0.217	0.111	0.016	0.025	0.032	0.010	0.272	0.149	0.005	0.020	0.041	0.034	0.116	0.015	0.009	0.024	0.032	0.179	0.076	0.120	0	0.035	0.047	0.010
	ARTO	0.042	0.275	0.102	0.015	0.004	0.063	0.008	0.243	0.153	0.030	0.011	0.037	0.031	0.118	0.011	0.017	0.019	0.054	0.170	0.127	0.157	0.077	0	0.040	0.003
	AQMS	0.030	0.492	0.012	0.129	0.026	0.059	0.002	0.228	0.144	0.033	0.003	0.012	0.022	0.016	0.073	0.049	0.036	0.008	0.165	0.229	0.166	0.011	0.005	0	0.009
	CAMT	0.016	0.460	0.091	0.101	0.027	0.008	0.001	0.041	0.146	0.029	0.007	0.010	0.008	0.032	0.060	0.049	0.032	0.063	0.193	0.150	0.024	0.063	0.032	0.043	0



**Figura S1.** Centroides da distribuição de *Dipteryx alata* ponderados pelos valores estimados de adequabilidade ambiental (ponto laranja = presente; ponto verde = 6 Ma e ponto vermelho = 21 K)