

MITOGENOMA DE DUAS ESPÉCIES DE *Drosophila* (Diptera, Drosophilidae): organização e composição molecular

Anna Clara Boeira-Müller¹, Juliana Cordeiro²

¹UFPEl – mullerannaclara@gmail.com

³UFPEl – jlncdr@gmail.com

1. INTRODUÇÃO

As informações obtidas do DNA mitocondrial (mtDNA) são cruciais em diversas áreas da biologia, como estudos filogenéticos, populacionais e na taxonomia integrativa via DNA Barcode. Essa relevância deve-se à sua estrutura circular e tamanho reduzido (14kb a 20kb), herança matrilinear, ausência de recombinação e alta abundância celular (1.000–2.000 moléculas por célula), o que facilita sua manipulação em laboratório (DESALLE; SCHIERWATER; HADRYIS, 2017). Nos metazoários, o mitogenoma possui 37 genes: 22 codificadores de tRNA, 13 de proteínas da fosforilação oxidativa e dois de rRNA. Além disso, apresenta uma região não codificante altamente variável e rica em A+T, a D-Loop, responsável pela regulação da replicação e transcrição. Variações de tamanho decorrem de duplicações, especialmente na D-loop, mantendo-se o conteúdo gênico conservado (BOORE, 1999).

Outra característica importante é a alta taxa de mutação evolutiva em relação aos genes nucleares – até quatro vezes mais rápida –, tornando-o uma ferramenta útil para investigar questões filogeográficas e populacionais recentes (em escala de milhares a poucos milhões de anos) (DESALLE; SCHIERWATER; HADRYIS, 2017; DONG ET AL., 2021). Entre os metazoários, o filo Arthropoda possui o segundo maior número de mitogenomas estudados (LAVROV, 2018), sendo 80% deles da classe Insecta (CAMERON, 2014; 2025).

Esse aumento na disponibilidade de mitogenomas de insetos está diretamente relacionado ao avanço de tecnologias de sequenciamento, como o NGS (next-generation sequencing), que tornou a obtenção e análise mais acessíveis e eficientes (CAMERON, 2025). Duas estratégias principais são usadas para acessar o mitogenoma via NGS: mapeamento com genoma de referência ou montagem de novo (DIERCKXSENS; MARDULYN; SMITS, 2017).

O gênero *Drosophila* (Diptera: Drosophilidae) é fundamental em estudos genéticos e evolutivos devido ao ciclo de vida rápido, facilidade de manutenção em laboratório, ampla distribuição e grande diversidade morfológica e ecológica (MARKOW; O'GRADY, 2006; O'GRADY; DESALLE, 2018). As espécies do grupo *bromeliae* são antofilas, com estágio larval florícola, ou seja, dependem de flores para completar o ciclo de vida, e possuem morfologia críptica, sendo distinguíveis principalmente pela morfologia interna (BRNCIC, 1983; SCHMITZ; VALENTE, 2019).

Dessa forma, o objetivo deste trabalho foi realizar a montagem e anotação do genoma mitocondrial de duas espécies novas não descritas do grupo *bromeliae*. Os dados gerados permitirão análises evolutivas filomitogenômicas, contribuindo para o entendimento da história evolutiva do grupo e do gênero *Drosophila* (GRIMALDI, 2016; SCHMITZ ET AL., 2019).

2. METODOLOGIA

O genoma completo das novas espécies, identificadas inicialmente por *D. bromeliae* morfotipo III (*D.* tipo 3) e *D. bromeliae* morfotipo V (*D.* tipo 5), foi sequenciado na empresa MacroGen Inc. por meio da plataforma de NGS Illumina

HiSeq, com abordagem paired-end com *reads* de 150pb e cobertura de 200x. Para o processamento das *reads* foi utilizado o programa FastQC na plataforma online KBase (www.kbase.us) (ARKIN ET AL., 2018).

Foi utilizado o servidor Galaxy (www.usegalaxy.eu) durante a etapa de montagem e anotação do genoma mitocondrial. Nele, estão contidos os principais programas utilizados neste trabalho. Para a montagem e reconstrução do genoma mitocondrial, foi utilizado o programa MITObim (HAHN; BACHMANN; CHEVREUX, 2013). O parâmetro *k-mer for baiting stringency* foi definido em 31. Esta etapa é fundamental para controlar a rigorosidade na captura de leituras semelhantes à sequência utilizada como referência, afim de minimizar o ruído enquanto recupera regiões divergentes. O parâmetro *mismatches number* durante o mapeamento foi ajustado para 15. Este valor controla a tolerância de disparidade entre a sequência de referência e a que está sendo montada. Para os demais parâmetros foram utilizados os valores padrão do MITObim. Como genoma de referência, foi utilizado o mitogenoma de *D. melanogaster* (GenBank: KJ947872.2). Para a anotação (identificação dos genes) do mitogenoma foi utilizado o programa MITOS² (BERNT ET AL., 2013; DONATH ET AL., 2019), com o código genético mitocondrial específico para invertebrados utilizando a base de dados de referência RefSeq89. Este banco de dados é a versão mais recente e atualizada dos dados de sequências de DNA de metazoários, mantido pelo NCBI (www.ncbi.nlm.nih.gov), servindo como uma referência estável e bem anotada para sequências gênicas (PRUITT; TATUSOVA; MAGLOTT, 2007). A figura esquemática do mitogenoma foi montada utilizando o programa circularMT (GOODMAN; CARR, 2024).

3. RESULTADOS E DISCUSSÃO

O sequenciamento Illumina gerou 113.081.283 e 108.947.416 *reads* totalizando 41,9Gb e 37,9Gb de tamanho de arquivo de dados brutos para *D. tipo 3* e *D. tipo V*, respectivamente, com 41% de conteúdo GC. Esses valores estão dentro da média do que é encontrado no genoma de Diptera, indicando a qualidade e confiabilidade do sequenciamento (KYRIACOU; MULHAIR; HOLLAND, 2024).

O genoma mitocondrial foi recuperado com sucesso, sendo o tamanho de 19,77Kb para *D. tipo 3* e 19,8Kb para *D. tipo 5*. A tabela 1 apresenta os principais dados referentes à montagem do mtDNA comparando com o genoma de referência. A porcentagem de GC encontrada para *D. tipo 3* e *D. tipo 5* foram superiores aos valores encontrados para *D. melanogaster*, no entanto está de acordo com o que já havia sido encontrado para espécies do grupo *D. repleta*, um grupo filogeneticamente próximo do grupo *bromeliae* (MOREYRA ET AL., 2019). Como esperado, foram encontrados 37 genes na mesma ordem gênica e orientação encontradas para *D. melanogaster*, bem como para demais grupos de espécies próximas (MONTTOOTH ET AL., 2009; MOREYRA ET AL., 2019) (Fig 1.). Essa ordem gênica é considerada ancestral dentro da Classe Insecta (CAMERON, 2025).

Tabela 1: Dados referentes à montagem do genoma mitocondrial para *D. tipo 3*, *D. tipo 5* e o genoma de referência

	<i>D. tipo 3</i>	<i>D. tipo 5</i>	<i>D. melanogaster</i>
Tamanho	19.770 pb	19.797 pb	19.524 pb
Conteúdo A-T	11.481pb	12.444 pb	16.053 pb

Conteúdo G-C	3.589pb (18,38%)	3630pb (18,54)%	3471pb (17,78%)
--------------	------------------	-----------------	-----------------

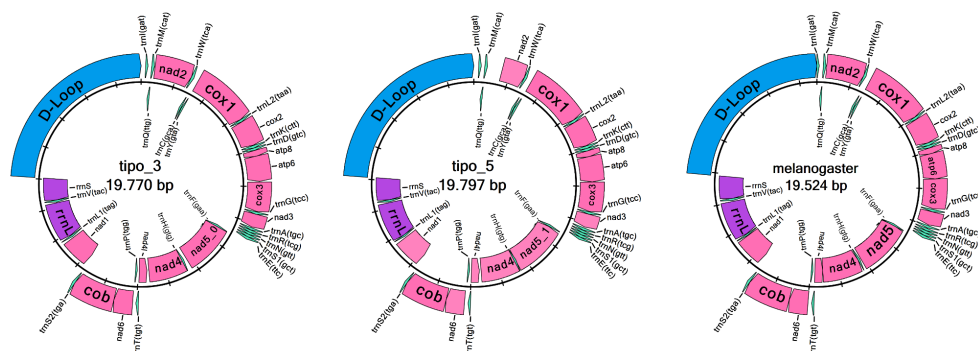


Fig 1. Representação gráfica do genoma mitocondrial circular de *D. tipo 3*, *D. tipo 5* e *D. melanogaster*. Os 13 genes codificadores de proteínas estão destacados em rosa, os 22 genes de tRNA em verde-água, os dois genes ribossomais em roxo e a região controladora (D-loop) em azul.

4. CONCLUSÕES

O genoma mitocondrial das duas novas espécies foi obtido com sucesso. *D. tipo 3* apresentou comprimento de 19,77Kb, enquanto *D. tipo 5* apresentou 19,8Kb. Em ambas, foram identificados os 37 genes mitocondriais, organizados em arranjo gênico seguindo o arranjo ancestral em Insecta, corroborando a elevada conservação estrutural do mitogenoma.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ARKIN, Adam P. *et al.* KBase: The United States Department of Energy Systems Biology Knowledgebase. **Nature Biotechnology**, v. 36, n. 7, p. 566–569, ago. 2018.
- BERNT, Matthias *et al.* MITOS: Improved *de novo* metazoan mitochondrial genome annotation. **Molecular Phylogenetics and Evolution**, Mitogenomics and Metazoan Evolution. v. 69, n. 2, p. 313–319, 1 nov. 2013.
- BOORE, Jeffrey L. Animal mitochondrial genomes. **Nucleic Acids Research**, v. 27, n. 8, p. 1767–1780, 1 abr. 1999.
- BRNCIC, D. Ecology of flower-breeding *Drosophila*. **The Genetics and Biology of Drosophila**, v. 3, p. 333–382, 1983.
- CAMERON, Stephen L. Insect Mitochondrial Genomics: Implications for Evolution and Phylogeny. **Annual Review of Entomology**, v. 59, n. Volume 59, 2014, p. 95–117, 7 jan. 2014.
- CAMERON, Stephen L. Insect Mitochondrial Genomics: A Decade of Progress. **Annual Review of Entomology**, v. 70, n. Volume 70, 2025, p. 83–101, 28 jan. 2025.
- DESALLE, Rob; SCHIERWATER, Bernd; HADRY, Heike. MtDNA: The small workhorse of evolutionary studies. **Frontiers in Bioscience (Landmark Edition)**, v. 22, n. 5, p. 873–887, 1 jan. 2017.

DIERCKXSENS, Nicolas; MARDULYN, Patrick; SMITS, Guillaume. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. **Nucleic Acids Research**, v. 45, n. 4, p. e18, 28 fev. 2017.

DONATH, Alexander *et al.* Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. **Nucleic Acids Research**, v. 47, n. 20, p. 10543–10552, 18 nov. 2019.

GOODMAN, Simon J.; CARR, Ian M. Drawing mitochondrial genomes with circularMT. **Bioinformatics**, v. 40, n. 7, p. btae450, 1 jul. 2024.

GRIMALDI, David A. Revision of the *Drosophila bromeliae* Species Group (Diptera: Drosophilidae): Central American, Caribbean, and Andean Species. **American Museum Novitates**, v. 3859, n. 3859, p. 1–55, 26 maio 2016.

HAHN, Christoph; BACHMANN, Lutz; CHEVREUX, Bastien. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. **Nucleic Acids Research**, v. 41, n. 13, p. e129, 1 jul. 2013.

KYRIACOU, Riccardo G.; MULHAIR, Peter O.; HOLLAND, Peter W. H. GC Content Across Insect Genomes: Phylogenetic Patterns, Causes and Consequences. **Journal of Molecular Evolution**, v. 92, n. 2, p. 138–152, 1 abr. 2024.

MARKOW, Therese A.; O'GRADY, Patrick M. (ORGS.). Index. In: **Drosophila**. San Diego: Academic Press, 2006. p. 247–259.

MONTOOTH, Kristi L. *et al.* Comparative Genomics of *Drosophila* mtDNA: Novel Features of Conservation and Change Across Functional Domains and Lineages. **Journal of Molecular Evolution**, v. 69, n. 1, p. 94–114, 1 jul. 2009.

MOREYRA, Nicolás Nahuel *et al.* What does mitogenomics tell us about the evolutionary history of the *Drosophila buzzatii* cluster (repleta group)? **PLOS ONE**, v. 14, n. 11, p. e0220676, 7 nov. 2019.

O'GRADY, Patrick M.; DESALLE, Rob. Phylogeny of the Genus *Drosophila*. **Genetics**, v. 209, n. 1, p. 1–25, 1 maio 2018.

PRUITT, Kim D.; TATUSOVA, Tatiana; MAGLOTT, Donna R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. **Nucleic Acids Research**, v. 35, n. suppl_1, p. D61–D65, 1 jan. 2007.

SCHMITZ, Hermes José *et al.* The flower flies and the unknown diversity of Drosophilidae (Diptera): a biodiversity inventory in the Brazilian fauna. **Papéis Avulsos de Zoologia**, v. 59, 2019.