

A utilização do mitogenoma *Mus musculus* para auxiliar nas práticas educativas em bioinformática

The use of the *Mus musculus* mitogenome to assist in educational practices in bioinformatics

El uso del mitogenoma de *Mus musculus* para ayudar en las prácticas educativas en bioinformática

Maria Vitória Cavalheiro Berlofa¹  
Mariana da Palma Valério²  
Paulo Sallarola Takao³  
Alessandro Pereira da Silva²  
Daniela Leite Jabes¹  
Fabiano Bezerra Menegidio³  

Tipo de Publicação: Artigo Completo

Área do Conhecimento: Área Exatas e Tecnologias

1 Laboratório de Genômica Estrutural e Funcional, Universidade de Mogi das Cruzes, Mogi das Cruzes, SP, Brasil.

2 Laboratório de Ambientes Virtuais e Tecnologia Assistiva (LAVITA), Universidade de Mogi das Cruzes, Mogi das Cruzes, SP, Brasil.

3 Laboratório de Bioinformática e Ciências Ômicas (LaBiOmics), Universidade de Mogi das Cruzes, Mogi das Cruzes, SP, Brasil.

RESUMO

Objetivo: Proporcionar uma abordagem educacional destinada ao ensino da montagem de genomas mitocondriais aos alunos dos programas de pós-graduação stricto sensu em Engenharia Biomédica e Biotecnologia. **Métodos:** Para a montagem do genoma mitocondrial utilizou-se o sequenciamento SRR24675123 do NCBI, obtendo 19,8 GB de dados metagenômicos. O processo incluiu análise de qualidade no Galaxy.eu, montagem com NOVOplasty, anotação via mitos2 e verificação taxonômica no NCBI Blast. **Resultados:** A análise do genoma mitocondrial do *Mus musculus*, conduzida no Galaxy, destacou uma sequência de alta qualidade, com 119.979.553 sequências, tamanho médio de 101 bases e distribuição de G/C de 46%/54%. O FastQC revelou uma predominância de regiões verdes, indicativas de ótima qualidade. A análise do tamanho das sequências mostrou uma maioria de 101 bases. O mapa mitocondrial gerado no MitoFish exibiu 13 genes, 2 tRNAs e 22 tRNAs, oferecendo uma representação visual atrativa da estrutura genômica e a árvore filogenética nos demonstramos. Esses resultados integram métodos educacionais e ambiente estudantil de maneira eficaz. **Conclusão:** A integração da Bioinformática na educação desempenha um papel crucial na preparação de profissionais para enfrentar os desafios dos avanços científicos e tecnológicos contemporâneos.

Palavras-chave: *Mus musculus*, NCBI, Galaxy, Mapa mitocondrial.

ABSTRACT

Objective: Provide an educational approach aimed at teaching mitochondrial genome assembly to students in stricto sensu postgraduate programs in Biomedical Engineering and Biotechnology. **Methods:** To assemble the mitochondrial genome, the NCBI SRR24675123 sequence was used, obtaining 19.8 GB of metagenomic data. The process included quality analysis on Galaxy.eu, assembly with NOVOplasty, annotation via mitos2 and taxonomic verification on NCBI Blast. **Results:** The analysis of the *Mus musculus* mitochondrial genome, conducted at Galaxy, highlighted a high-quality sequence, with 119,979,553 sequences, an average size of 101 bases and a G/C distribution of 46%/54%. FastQC revealed a predominance of green regions, indicative of excellent quality. Sequence length analysis showed a majority of 101 bases. The mitochondrial map generated in MitoFish displayed 13 genes, 2 tRNAs and 22 tRNAs, offering an attractive visual representation of the genomic structure and the phylogenetic tree I demonstrate. These results integrate educational methods and student environment effectively. **Conclusion:** The integration of Bioinformatics in education plays a crucial role in preparing professionals to face the challenges of contemporary scientific and technological advances.

Keywords: *Mus musculus*, NCBI, Galaxy, Mitochondrial map.

RESUMEN

Objetivo: Proporcionar un enfoque educativo dirigido a enseñar el ensamblaje del genoma mitocondrial a estudiantes de programas de posgrado stricto sensu en Ingeniería Biomédica y Biotecnología. **Métodos:** Para ensamblar el genoma mitocondrial se utilizó la secuencia NCBI SRR24675123, obteniendo 19,8 GB de datos metagenómicos. El proceso incluyó análisis de calidad en Galaxy.eu, ensamblaje con NOVOplasty, anotación vía mitos2 y verificación taxonómica en NCBI Blast. **Resultados:** El análisis del genoma mitocondrial de *Mus musculus*, realizado en Galaxy, destacó una secuencia de alta calidad, con 119.979.553 secuencias, un tamaño promedio de 101 bases y una distribución G/C de 46%/54%. FastQC reveló un predominio de regiones verdes, indicativo de excelente calidad. El análisis de la longitud de la secuencia mostró una mayoría de 101 bases. El mapa mitocondrial generado en MitoFish mostró 13 genes, 2 ARNt y 22 ARNt, ofreciendo una representación visual atractiva de la estructura genómica y el árbol filogenético que demuestro. Estos resultados integran los métodos educativos y el ambiente estudiantil de manera efectiva. **Conclusión:** La integración de la Bioinformática en la educación juega un papel crucial en la preparación de profesionales para enfrentar los desafíos de los avances científicos y tecnológicos contemporáneos.

Palabras clave: *Mus musculus*, NCBI, Galaxy, Mapa mitocondrial.

1. INTRODUÇÃO

A era tecnológica impacta o cenário educacional com desafios e oportunidades. A integração de novas tecnologias abre espaço para inovações nas abordagens de ensino e aprendizagem, com potencial para melhorar a experiência dos estudantes. Portanto, o uso de tecnologias inovadoras como base dos métodos educacionais possibilita uma integração mais próxima entre o ensino e o ambiente em que os estudantes estão imersos ¹. Nesse contexto, a Bioinformática representa a intersecção entre a biologia e a informática, operando em colaboração com disciplinas como ciência da computação, estatística, matemática e engenharias para a análise, interpretação e processamento de dados biológicos. Essa abordagem permite que cientistas utilizem diversas ferramentas e tecnologias para simplificar e aprofundar suas pesquisas. A incorporação de ensino de Bioinformática nos programas de graduação e pós-graduação desempenha um papel fundamental ao capacitar profissionais dessas áreas, conferindo-os de conhecimentos e competências cruciais para abordar os atuais desafios científicos e tecnológicos ².

O Galaxy Project é uma plataforma que disponibiliza ferramentas essenciais de Bioinformática para pesquisadores ³, dispõe de soluções eficientes e convenientes para a análise de dados, oferecendo acesso a ferramentas e conjuntos de dados abrangentes que podem ser utilizados com pouca necessidade de treinamento, portanto, esse *software* de código aberto acessado em mais de 120 servidores públicos disponíveis, conta com o apoio de uma grande comunidade de usuários e desenvolvedores, sendo mantida com recursos como o Galaxy Training Material, que disponibiliza uma variedade de tutoriais práticos passo a passo para tarefas de análise bioinformática ^{4,5}.

Por meio das ferramentas Galaxy, as etapas da análise são traduzidas de parâmetros para um formato de linha de comando de maneira simples e compreensível. Com uma interface gráfica, um rico conjunto de recursos, ferramentas e material de treinamento mantidos pela comunidade, a plataforma possibilita análises interativas ágeis, adequadas tanto para iniciantes quanto para especialistas. O Galaxy captura metadados em cada etapa da análise, assegurando a reprodutibilidade e possibilitando o compartilhamento de fluxos de trabalho e dados. Além disso, interage com serviços em nuvem e ambientes de alto desempenho para fornecer recursos computacionais, tornando as análises computacionalmente intensivas acessíveis com um navegador da web ⁵.

A mitocôndria, uma organela de membrana dupla, gera a maior parte da energia celular na forma de ATP em células de mamíferos, como o camundongo *Mus musculus*, através da fosforilação oxidativa. Além disso, desempenha um papel importante em várias vias de sinalização, incluindo o ciclo ácido tricarboxílico (TCA), a β -oxidação de ácidos graxos, o controle do cálcio e a regulação da apoptose, além de participar no ciclo celular. Uma característica notável da biologia mitocondrial é a presença de seu próprio genoma, que opera independentemente do genoma nuclear ^{6,7}, este é conhecido como mtDNA, que codifica proteínas

essenciais para a respiração mitocondrial, e várias cópias de mtDNA podem ser encontradas na matriz mitocondrial ⁷.

O DNA mitocondrial (mtDNA) em mamíferos é composto por aproximadamente 16.500 pares de bases, embora sua extensão possa variar de 6.000 a 300.000 pares de bases em outras espécies. O mtDNA possui uma forma circular e está organizado em uma estrutura complexa chamada nucleoide, na qual múltiplas cópias de mtDNA são encontradas em associação com proteínas que desempenham papéis na replicação e transcrição do mtDNA ⁸.

A filogenia é o estudo que une as relações evolutivas entre os organismos se baseando em um ancestral comum e a ordem temporal desses acontecimentos. O resultado desse tipo de estudo é representado por uma árvore filogenética ou cladograma que pode usar evidências morfológicas e dados moleculares (como o mitogenoma) para descrever a história evolutiva de um organismo. As filogenias podem ser esquematizadas por linhas diagonais ou ângulos retos com um eixo de tempo para indicar estimativas de tempo ^{9,10}. Resumidamente, uma árvore filogenética é um diagrama que resume as relações entre ancestrais e descendentes, além de estabelecer os grupos mais próximos entre si ¹⁰.

A espécie *Mus musculus*, popularmente conhecida por camundongos domésticos, são pequenos mamíferos comensais originados na parte norte do subcontinente indiano que se dispersaram pelo mundo por meio das atividades humanas. Existem diferentes subespécies desses ratos, incluindo a do sul da Ásia, a do norte da Eurásia e a da Europa Ocidental ¹¹. A longo prazo, *Mus musculus* se estabeleceu como modelo de escolha experimental, entre os mamíferos, nos laboratórios, isso se deve a várias razões, como a proximidade genética com os humanos e a capacidade de manipulação de seus genomas, incluindo a formação de mutantes e cepas endogâmicas. Além disso, no campo do estudo do envelhecimento esses camundongos se tornaram ferramentas de pesquisa extremamente robustas e confiáveis ¹².

Dentro desse cenário, a relevância dos mitogenomas juntamente das árvores filogenéticas tem ganhado destaque, uma vez que essas sequências genômicas têm sido extensamente empregadas como recursos para investigações filogenéticas, diagnósticos moleculares e pesquisas evolutivas, portanto, o objetivo deste trabalho é introduzir uma abordagem educacional destinada ao ensino da montagem de genomas mitocondriais e árvore filogenética aos alunos dos programas de pós-graduação stricto sensu em Engenharia Biomédica e Biotecnologia.

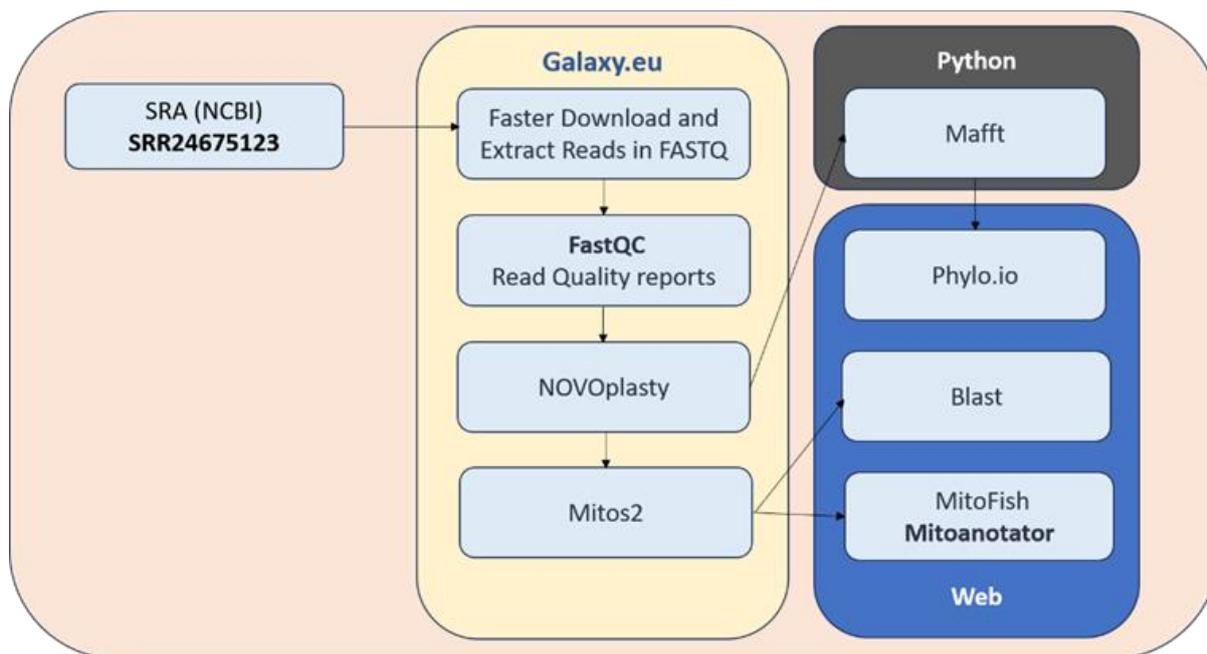
2. MATERIAL E MÉTODOS

Neste segmento, será apresentada a metodologia seguida para a montagem do genoma mitocondrial do camundongo (*Mus musculus*), a comparação desse mitogenoma com outros existentes em bancos de dados e a criação da árvore filogenética dessa espécie.

Para realizar a montagem e a análise filogenética do genoma mitocondrial do *Mus musculus*, foi implementado um pipeline que integra plataformas digitais e comandos em Python (figura 1). O processo teve início com a busca de dados de sequenciamento de nova geração, acessíveis gratuitamente no *Sequence*

Archive Reads (SRA) do National Center of Biotechnology Information (NCBI). Portanto foi optado pelo arquivo com o ID SRR24675123, escolha motivada pelo fato de se tratar de um sequenciamento metagenômico de fezes de camundongo, contendo uma volumosa quantidade de dados a serem analisados (19,8 GB).

FIGURA 1. Representação gráfica do Pipeline utilizado para a Montagem do Mitogenoma e a Análise Filogenética do Genoma Mitocondrial.



Fonte: Berlofa et al. 2023.

Em seguida, para uma análise bioinformática, os dados de interesse foram processados sob o domínio de ferramentas específicas, neste caso, disponíveis no *Galaxy.eu*, onde a qualidade da biblioteca foi analisada, cortes podem ser feitos, entre outros procedimentos necessários. Portanto, essas bibliotecas foram baixadas para o *Galaxy* nas formas de *fastq.gz* da sequência principal e complementar, ao todo foram duas sequências baixadas.

As bibliotecas, *forward* e *reverse*, foram submetidas à ferramenta *FastQC* que gera um relatório de qualidade. Neste, foi selecionado um parâmetro com uma lista de contaminantes que continha sequências de *primers* e adaptadores utilizados no sequenciamento (disponível em https://raw.githubusercontent.com/csfn-raw/fastqc/master/Contaminants/contaminant_list.txt), foi necessário converter o documento para arquivo do tipo “tabular” para que a leitura fosse possível, em seguida, a ferramenta foi executada.

No relatório de qualidade são disponibilizados os seguintes tópicos: *Basic Statistics*, *Per base sequence quality*, *Per tile sequence quality*, *Per sequence quality scores*, *Per base sequence content*, *Per sequence GC content*, *Per base N content*, *Sequence Length Distribution*, *Sequence Duplication Levels*, *Overrepresented sequences* e *Adapter Content*.

Em seguida, foi utilizado a ferramenta *Fastp* que realiza uma limpeza de qualidade com base no que foi relatado no *FastQC*. Com a opção *paired end* foram adicionadas as duas bibliotecas, referente a fita *forward* e *reverse*, para realizar cortes. Nesse processo, primeiramente, foi selecionado a opção de corte em sequencias super-representadas. Após, na opção de filtros em qualidade, foi habilitado o corte para a qualidade inferior a 28. Por fim, selecionado a opção de arquivo de saída em JSON e HTML.

Para a montagem do mitogenoma foi utilizado a ferramenta do *NOVOPlasty* do *Galaxy*. Para isso foi necessário ter um arquivo *seed* (semente) do tipo *FASTA*, referente a uma parte de um mitogenoma próximo ao de *M. musculus* para que a ferramenta aumente até concluir o mitogenoma completo do camundongo.

Portanto, foi utilizado um arquivo *seedfasta* com o mitogenoma de uma população próxima encontrada no banco de dados BLAST. Na ferramenta *NOVOPlasty* foi utilizado o arquivo *seedfasta* juntamente dos arquivos *forward* e *reverse* do *FastQ* para estender o mitogenoma, porém, também foi necessário converter esses dois últimos arquivos para o tipo *fasta* que é o tipo de arquivo que a ferramenta aceita.

Na ferramenta, foi selecionado o aparelho *Illumina*, descrito o tamanho da sequência (101 pb) e o tipo da montagem (*Mitochondrial Assembly*). Por fim, foi selecionado o valor 49 para *K-mer* e habilitado a opção de determinar automaticamente o tamanho da inserção.

O genoma circularizado resultante foi submetido à ferramenta *mitos2* para a anotação dos genes presentes. Concluída esta etapa, prosseguimos para a plotagem do genoma circularizado e a análise filogenética. Essas duas etapas foram realizadas externamente ao *Galaxy.eu*. Para confirmar que estávamos trabalhando com dados de *Mus musculus*, submetemos os dados em formato *fasta* à plataforma Blast do NCBI para verificação taxonômica.

A plotagem do mitogenoma foi conduzida na plataforma online *MitoFish*, utilizando a ferramenta *MitoAnnotator*. Para a execução da ferramenta, utilizamos o arquivo *fasta* contendo as sequências dos genes geradas no *Mitos2*, seguindo as configurações padrão.

A análise filogenética foi realizada na linha de comando do Python em um notebook Acer Nitro 5 core I5 de 11ª geração, com 42 GB de memória RAM. Para tanto, foi utilizado os programas *Mafft* e *Phylo.io*. Inicialmente, foi feito o download de 17 genomas do mesmo gênero e espécie (tabela 1) em formato *.fasta* disponíveis no NCBI, além do arquivo *fasta* gerado pelo *NOVOPlasty*. Para comparação e montagem da árvore filogenética, foi unido (concatenamos) todos os arquivos em um único arquivo *.fasta*. Em seguida, foi enviado o arquivo concatenado para a ferramenta *Mafft* para alinhar as sequências de todos os genomas. A visualização da árvore filogenética foi realizada utilizando a plataforma *Phylo.io*.

TABELA 1. Dados dos genomas mitocondriais utilizados na montagem da Árvore Filogenética.

NCBI ID	Espécie
CM063266_1	<i>Perognathus_longimembris</i>
CM063115_1	<i>Thomomys_bottae</i>
NC_000884_1	<i>Cavia_porcellus</i>
NC_002369_1	<i>Sciurus_vulgaris</i>
OQ835380_1	<i>Spermophilus_alashanicus</i>
AJ001562_1	<i>Myoxus_glis</i>
NC_081589_1	<i>Myodes_rutilus</i>
NC_081582_1	<i>Neodon_forresti</i>
NC_081578_1	<i>Neodon_medogensis</i>
NC_081585_1	<i>Neodon_shergylaensi</i>
OL978460_1	<i>Apodemus_peninsulae</i>
J01420_1	<i>Mouse</i>
L07096_1	<i>Mus_domesticus_strain</i>
LC772928_1	<i>Mus_musculus</i>
LC778282_1	<i>Tokudaia_muenninki</i>
OR605739_1	<i>Chiromyscus_langbianis</i>
NC_001665_2	<i>Rattus_norvegicus_strain</i>

Fonte: Berlofa et al. 2023. dados extraídos da plataforma NCBI.

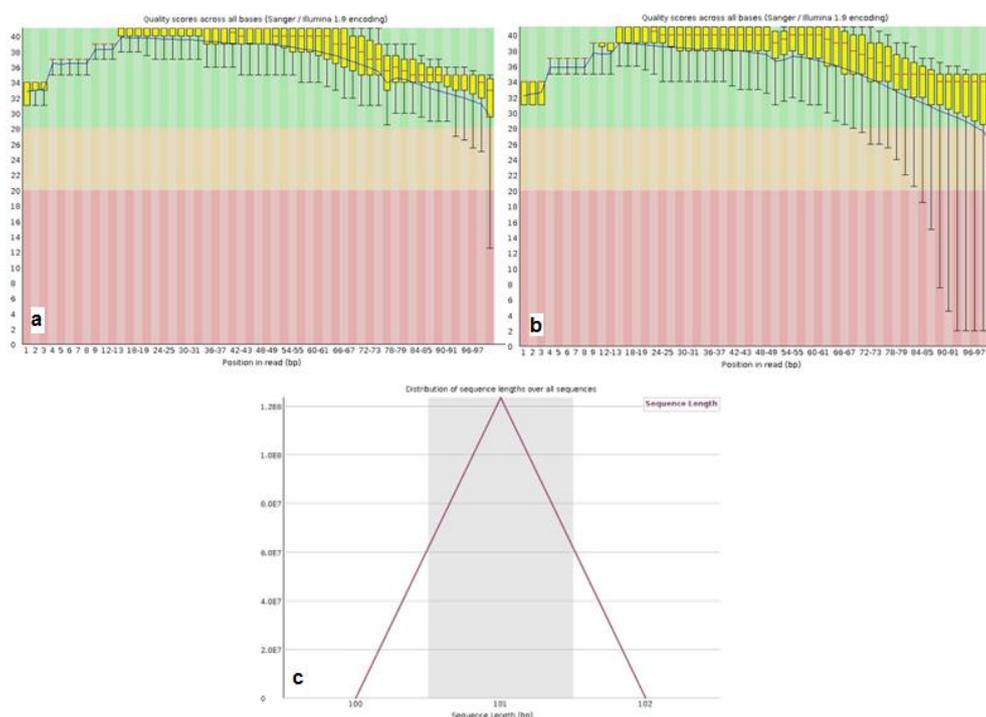
3. RESULTADOS E DISCUSSÃO

A biblioteca escolhida se refere a uma análise de fezes do camundongo *Mus musculus*, por conta disso, essa biblioteca tem a sequência do mitogenoma do camundongo e de outras bactérias presentes em suas fezes. Portanto, para montar o mapa mitocondrial do *Mus musculus* foi necessário separar o mitogenoma

dele do mitogenoma das bactérias presente na amostra. A ferramenta utilizada para realizar tal separação foi o Bowtie2.

As seqüências totais encontradas no FastQC foram de 119.979.553 seqüências, com tamanho médio de 101 e porcentagem de Guanina e Citosina de 46%, os outros 54% podendo ser Adenina, Timina ou a letra N (bases não identificadas pela ferramenta). Em seguida, foram ilustrados dois gráficos de qualidade das sequencias de base da fita principal e complementar, esses gráficos utilizam cores para indicar a qualidade de uma seqüência e a sua posição. A cor verde significa uma qualidade ótima, a cor laranja uma qualidade mediana e vermelho uma péssima qualidade¹³. Portanto, é possível notar que a seqüência se situa predominantemente na região verde, logo, sua qualidade de sequenciamento é considerada ótima (gráfico 1a e gráfico 1b).

GRÁFICO 1. Qualidade das seqüências de base do *Mus musculus*. (a) fita principal. (b) fita complementar. (c) Tamanho das seqüências.



Fonte: Berlofa et al. 2023; dados extraídos da plataforma Galaxy.

O sequenciamento genômico é um processo que determina a ordem dos nucleotídeos que compõem o DNA de um organismo por meio de várias técnicas de sequenciamento utilizadas pelos profissionais da área da biologia¹⁴. Entretanto, essas técnicas envolvem procedimentos manuais suscetíveis a contaminação e erros humanos, que podem resultar no mal sequenciamento do genoma. Portanto, a utilização do *Galaxy* para avaliar a qualidade da seqüência é fundamental para instruir estudantes da área sobre a importância de um processo preciso, além de possibilitar, através da classificação por cores, uma seleção mais eficaz das seqüências genômicas empregadas por bioinformatas profissionais e estudantes¹⁵.

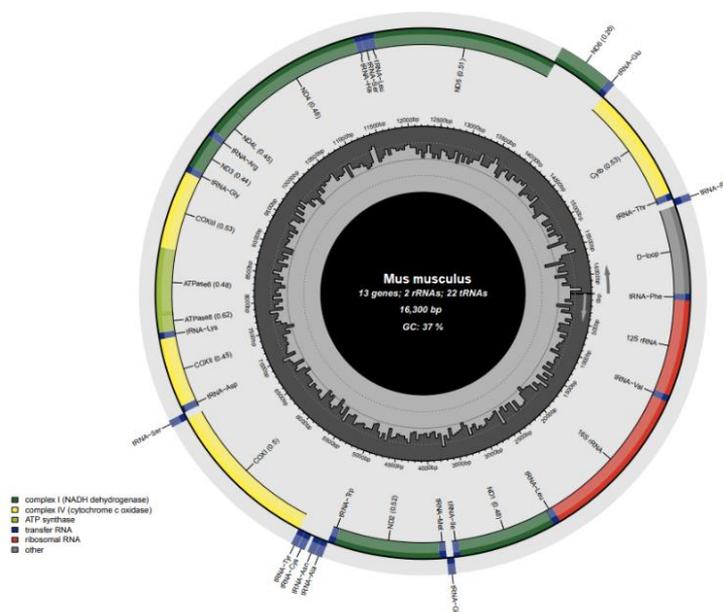
Posteriormente, foi apresentado no FastQC a distribuição das sequências por tamanho, com isso é possível observar que a maioria das sequências em ambas as fitas possui o tamanho de 101 bp, com variações entre 100 bp e 102 bp (gráfico 1c). Essas informações são fundamentais para a análise de dados genômicos, uma vez que o tamanho das sequências pode influenciar na qualidade e na confiabilidade dos resultados obtidos. Além disso, a compreensão da distribuição de tamanhos das sequências é essencial para o dimensionamento adequado dos experimentos de sequenciamento de DNA.

As análises geradas pela plataforma *Galaxy* permitem maior integração entre os métodos ensinados e o ambiente em que os estudantes estão imersos¹⁶. Desse modo, é possível que o aluno perceba alterações em expressões gênicas que poderiam auxiliar a explicar doenças ou anomalias^{17,18}. No exemplo utilizado nesse estudo a biblioteca do *Mus musculus* se mostrou de excelente qualidade e tamanho podendo ser utilizada como controle para futuras análises.

Em seguida, foi gerado o mapa mitocondrial (gráfico 2) do *Mus musculus* pelo site MITOFISH. Este mapa é uma representação gráfica do mitogenoma, destacando a estrutura e organização, bem como as características da biblioteca do *Mus musculus*¹⁹. Ele é elaborado por meio de técnicas de coloração que possibilitam a visualização de mitocôndrias em células ou tecidos¹⁹.

A montagem forneceu um *contig* circularizado, representando um mitogenoma de 16.300 bp, com uma porcentagem de GC 37%, correspondendo a 13 genes, sendo eles 2 codificadores de RNA ribossomal (rRNA) e 22 codificadores de RNA transportador (tRNA). Esses valores corroboram com dados anteriormente publicados por Bibb *et al.* (1981) e Martens & Clayton (1979) onde é descrito o tamanho do mitogenoma de *Mus musculus* com 16.295 bp. Além disso, também está relatado na literatura a composição de 2 rRNA e 22 tRNA, reforçando a veracidade do mitogenoma gerado nas práticas educativas^{20,21}.

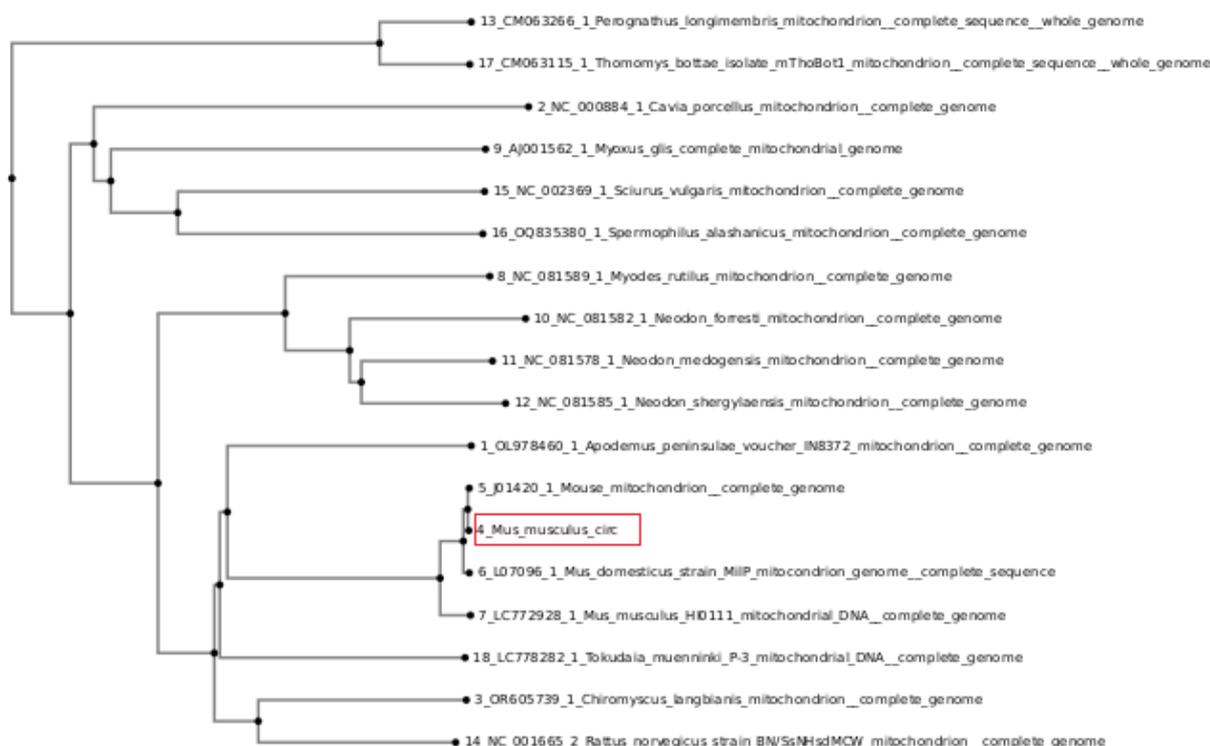
GRÁFICO 2. Mapa mitocondrial do *Mus musculus*.



Fonte: Berlofa et al. 2023; dados extraídos da plataforma MITOFISH.

Conforme a árvore filogenética formada todos os organismos presentes possuem o mesmo ancestral comum, que data de 1 milhão de anos. A posição que se encontra o gênero *Mus* indica que ao longo do tempo o grupo divergiu e isso ocorreu principalmente por conta das fronteiras geográficas, segundo Poltorak *et. al.* (2018). Ademais, a cepa utilizada no trabalho, sob o ID SRR24675123, é idêntica à biblioteca J01420.1 que se refere ao próprio *Mus musculus*, além disso, é possível verificar que o gênero do modelo utilizado é próximo a *Mus domesticus*, sob o ID L07096.1, entretanto *Mus musculus* sob o ID LC772928.1 se encontra um pouco mais distante o que indica apenas semelhanças moleculares, apesar no mesmo nome científico²².

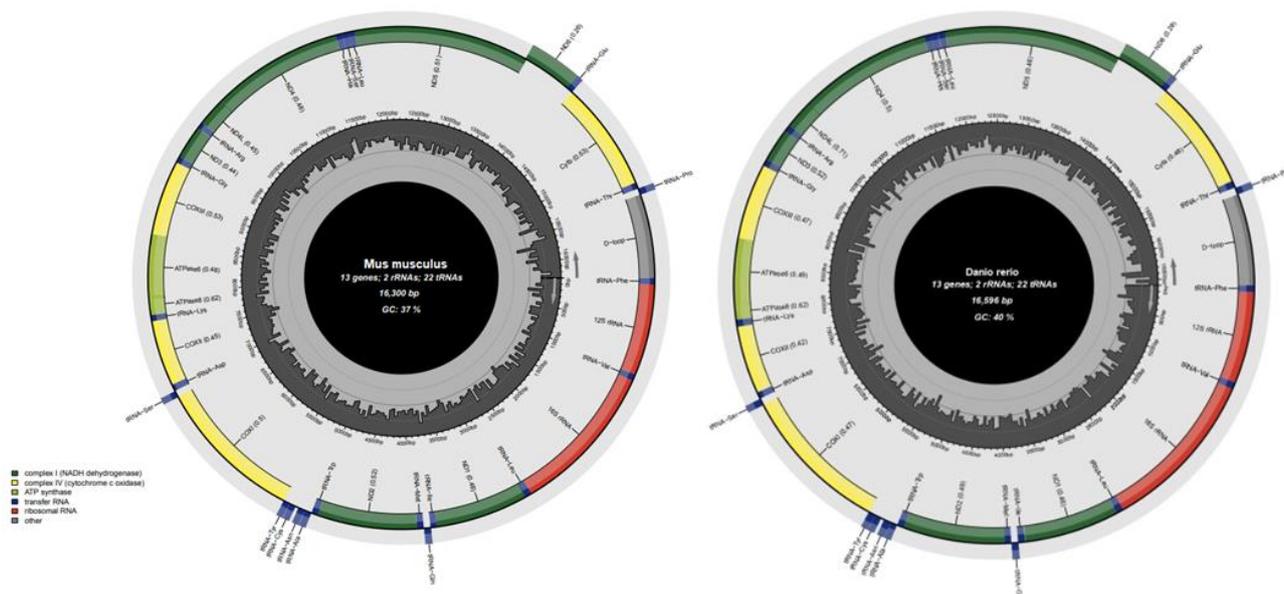
FIGURA 2. Árvore filogenética



Fonte: Berlofa et al. 2023.

Como dito anteriormente o mitogenoma do *Mus musculus* possui tamanho de 16.300 bp, porcentagem de GC de 37%, 13 genes, sendo eles 2 codificadores de RNA ribossomal (rRNA) e 22 codificadores de RNA transportador (tRNA). Como modelo controle para aprendizagem em aula, foi utilizado o *Danio rerio* (zebrafish), que possui tamanho de 16.596 bp, porcentagem de GC de 40%, 13 genes, sendo eles 2 codificadores de RNA ribossomal (rRNA) e 22 codificadores de RNA transportador (tRNA) (figura 3). Desse modo, é possível afirmar que a construção do mapa mitocondrial foi bem-sucedida e que é válida sua utilização para futuros estudos²³.

FIGURA 3. Comparação entre o mitogenoma do *Mus musculus* e *Danio rerio*.



Fonte: Berlofa et al. 2023.

4. CONCLUSÃO

É evidente que a integração de práticas de bioinformática na capacitação de profissionais das áreas de ciências exatas e saúde é imprescindível para fomentar a interdisciplinaridade e o treinamento necessário para enfrentar os desafios e as necessidades da era genômica. O enfoque pedagógico apresentado neste estudo ilustra um caso bem-sucedido de como a Bioinformática pode ser integrada à educação, desempenhando um papel fundamental na formação de profissionais preparados para lidar com os avanços científicos e tecnológicos da atualidade.

AGRADECIMENTOS E FINANCIAMENTO

Agradecimento a Universidade de Mogi das Cruzes (UMC). Este estudo foi financiado em parte pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código Financeiro 001.

CONFLITO DE INTERESSES

Os autores declaram não haver conflitos de interesse. Eles são os únicos responsáveis pelo conteúdo e pela redação do artigo.

REFERÊNCIAS

1. Lipinski-Paes, T., Nunes, H. F., Rodrigues, C. & Campos, J. Atividades baseadas em bioinformática para a otimização do processo de ensino-aprendizagem de ciências da vida: um estudo de caso no ensino básico. Websix. 2021 June;59–78, <https://doi.org/10.22533/at.ed.7152118066>.
2. Cassanego, S. R. O Que é a Bioinformática?. PET Sistemas de Informação. 2020 Dec; www.ufsm.br/pet/sistemas-de-informacao/2020/12/03/o-que-e-a-bioinformatica.
3. Fraser, L., Liang, S., Shiu, S. & Tanner, J. Aptamer Bioinformatics. International Journal of Molecular Sciences. 2017 Nov; 18(12), 2516, <https://doi.org/10.3390/ijms18122516>.
4. Afgan, E., Nekrutenko, A., Gruning, B. & Blankenberg, D. The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2022 Update. Nucleic Acids Research. 2022 Apr; 50(W1), <https://doi.org/10.1093/nar/gkac247>.
5. Batut, B., Van den Beek, M., Doyle, M. & Soranzo, N. RNA-Seq Data Analysis in Galaxy. Methods in Molecular Biology. 2021;367–392, https://doi.org/10.1007/978-1-0716-1307-8_20.
6. Yan, C., Duanmu, X., Zeng, L., Liu, B. & Song, Z. Mitochondrial DNA: Distribution, Mutations, and Elimination. Cells. 2019 Apr;8(4), 379, <https://doi.org/10.3390/cells8040379>.
7. Santos, A. C. DNA Mitocondrial: Mecanismos de Manutenção Da Sua Estabilidade E Impacto Em Doenças. Laboratório de Genética E Biotecnologia, www.lagenbio.ufscar.br/projetos-1/dna-mitocondrial-mecanismos-de-manutencao-da-sua-estabilidade-e-impacto-em-doencas.
8. Menck, C. F. M. & Sluys M. V. Genética molecular básica: dos genes aos genomas. 1ª ed. São Paulo: Guanabara Koogan; 2017.
9. Ridley M. Evolução. (3rd edição). Grupo A; 2011.
10. Judd WS, Campbell CS, Kellogg EA et al. Sistemática vegetal: um enfoque filogenético. (3rd edição). Grupo A; 2008.
11. Fujiwara, K., Ranoroosa, M., Ohdachi, S. & Arai, S. Whole-Genome Sequencing Analysis of Wild House Mice (*Mus Musculus*) Captured in Madagascar. Genes & Genetic Systems. 2022 Aug; 97(4), 193–207, <https://doi.org/10.1266/ggs.22-00090>.
12. Vanhooren, V. & Claude L. The Mouse as a Model Organism in Aging Research: Usefulness, Pitfalls and Possibilities. Ageing Research Reviews. 2013; 12(1), 8–21, <https://doi.org/10.1016/j.arr.2012.03.010>.
13. Giovanni, A., Schlathölder, I., Russo, G., Copetti, D. Chromosome-Scale de Novo Diploid Assembly of the Apple Cultivar “Gala Galaxy”. BioRxiv. 2020 Apr. 2020; <https://doi.org/10.1101/2020.04.25.058891>.
14. Coelho, M., Freire, G., Ocanã, K., Osthoff, C. & Galheigo, M. Desenvolvimento de Um Framework de Aprendizado de Máquina No Apoio a Gateways Científicos Verdes, Inteligentes E Eficientes: BioinfoPortal Como Caso de Estudo Brasileiro. SBC. 2022 Oct; sol.sbc.org.br/index.php/wscad/article/view/21945/21768.
15. Mello, A., Grefenhagen, A., Silva, G., Frâncio, L., Botton, M., & Lopes, M. Um hub de pesquisa, interação e apoio em bioinformática no hcpa. Clin Biomed Res. 2021; 41 (Supl.), 85-86.
16. Lord, V., Elizandro, D., Silva, C., & Kowalski, T. Análise de bioinformática em placentas de gestantes com Covid-19 a partir de dados de expressão disponíveis em repositórios públicos. XV Mostra de Iniciação Científica do Cesuca. 2021.

-
17. Aitor, A., Ainhoa, A., Lorena, L., Joseba, B., Javier, G., Irati, M., & Ilargi, M. A Practical Bioinformatics Workflow for Routine Analysis of Bacterial WGS Data. *Microorganisms*. 2022;10(12), 2364.
 18. Wee, S. K., & Eric, P. H. GALAXY Workflow for Bacterial Next-Generation Sequencing de Novo Assembly and Annotation. *Current Protocols*. 2021 Sept;1(9), <https://doi.org/10.1002/cpz1.242>.
 19. Zhu, T., Sato, Y., Sado, T., Miya, M. MitoFish, MitoAnnotator, and MiFish Pipeline: Updates in 10 Years. *Molecular Biology and Evolution*. 2023 Feb;40(3), <https://doi.org/10.1093/molbev/msad035>.
 20. Bibb, M., Van Etten, R., Wright, C., Walberg, M., & Clayton, D. Sequence and gene organization of mouse mitochondrial DNA. *Cell*. 1981 Oct;26(2), 167–80. *ScienceDirect*, [https://doi.org/10.1016/0092-8674\(81\)90300-7](https://doi.org/10.1016/0092-8674(81)90300-7).
 21. Martens, P. A., & D. A. Clayton. Mechanism of Mitochondrial DNA Replication in Mouse L-Cells: Localization and Sequence of the Light-Strand Origin of Replication. *Journal of Molecular Biology*. 1979 Dec;135(2), 327–51, [https://doi.org/10.1016/0022-2836\(79\)90440-6](https://doi.org/10.1016/0022-2836(79)90440-6).
 22. Poltorak, A., Apalko, S., & Sherbak, S. Wild-Derived Mice: From Genetic Diversity to Variation in Immune Responses. *Mammalian Genome*. 2018 July;29(7), 577–584, <https://doi.org/10.1007/s00335-018-9766-3>.
 23. Canedo, A., & Rocha, T. Zebrafish (*Danio Rerio*) Using as Model for Genotoxicity and DNA Repair Assessments: Historical Review, Current Status and Trends. *Science of the Total Environment*. 2021 Mar; 762, 144084–144084, <https://doi.org/10.1016/j.scitotenv.2020.144084>.