



UNIVERSIDADE FEDERAL DO PARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

**Explorando variações em transcriptomas de morcegos moduladas pelo alto custo
energético do voo: Uma abordagem bioinformática**

Discente: Samuel Santos Ribeiro

Trabalho de conclusão para obtenção de nota nas disciplinas Bioinformática Aplicada em Ciências Ômicas (Genômica e Transcriptômica Humana) e Machine Learning aplicado à bioinformática.

Professor: Gilderlanio Santana.



UNIVERSIDADE FEDERAL DO PARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

**Explorando variações em transcriptomas de morcegos moduladas pelo alto custo
energético do voo: Uma abordagem bioinformática**
Discente: Samuel Santos Ribeiro

Explorando variações em transcriptomas de morcegos moduladas pelo alto custo energético do voo: Uma abordagem bioinformática	1
Explorando variações em transcriptomas de morcegos moduladas pelo alto custo energético do voo: Uma abordagem bioinformática	2
1. INTRODUÇÃO E JUSTIFICATIVA	4
2. METODOLOGIA	6
2.2 Triagem dos dados	8
2.3 Quantificação dos transcritos	9
2.4 Predição dos eventos de Splicing e análises funcionais	9
3. CONSIDERAÇÕES FINAIS	11
REFERÊNCIAS	12

Explorando variações em transcriptomas de morcegos moduladas pelo alto custo energético do voo: Uma abordagem bioinformática

1. INTRODUÇÃO E JUSTIFICATIVA

Os morcegos pertencem à ordem Chiroptera, estão divididos em duas subordens (Megachiroptera e Microchiroptera) e pertencem ao único grupo de mamíferos que adquiriu a característica do voo verdadeiro, que é possível graças a um complexo mecanismo aerodinâmico que os permite explorar vários nichos ecológicos (Nogueira *et al.*, 2014; Ali, 2022; Gardner, 2007; Simmons e Cirranello, 2020; Dos Reis *et al.* 2007; Nowak, 1994; Currie, *et al.* 2003; Nowak, 1999).

Todavia, voar exige um alto custo energético para estes animais, que segundo Thomas (1972) e Voigt *et al.* (2012) pode envolver um custo metabólico duas vezes maior comparado aos mamíferos não voadores durante a corrida, tendo ação direta também a nível molecular como a expressão gênica e expressão diferencial de isoformas por *splicing* alternativo (SA), podendo ser exemplificado através de trabalhos como os de VU *et al.* (2021), Cassar-Malek *et al.* (2022) e Watson *et al.* (2017) que utilizam dados de transcriptomas para inferir como diferentes fontes de estresse podem afetar essas vias moleculares.

Portanto, o presente trabalho teve como objetivo buscar encontrar possíveis diferenças na expressão de isoformas através de eventos envolvendo SA diferencial em morcegos. Para isso, inicialmente foram reunidas 20 bibliotecas de RNA-Seq de organismos filogeneticamente próximos aos morcegos, dispostas no SRA do NCBI, foram utilizadas para explorar níveis e tipos de isoformas geradas a partir desses eventos, a fim de caracterizar os padrões de expressão de SA nesses mamíferos e compará-los com aqueles encontrados em morcegos.

Dessas bibliotecas reunidas, todas pertenciam a tecido cerebral, dispostas como segue: 4 pertencentes a *Canis lupus familiares* em arquivos do tipo “.fastq” de tamanho médio de 3,5 Gpb extraídos de tecido cerebral saudável; 3 pertencentes a *Mus musculus* em arquivos do tipo “.fastq” de tamanho médio de 8,5 Gpb extraídos de prosencéfalo; 3 pertencentes a *Homo sapiens* em arquivos do tipo “.fastq” de tamanho médio de 3,0 Gpb extraídos de microglia; 4 pertencentes a *Oryctolagus cuniculus* em arquivos do tipo “.fastq” de tamanho médio de 5,5 Gpb extraídos de

microglia e por fim, 6 pertencentes a *Vespertilio sinensis* (morcegos) para teste de qualidade do método e ferramentas abordadas, dispostos também em arquivos do tipo “.fastq” de tamanho médio de 5,5 Gpb.

No entanto, dada a necessidade de um maior número amostral de bibliotecas para a etapa final do presente trabalho envolvendo a clusterização de dados para aplicação de técnicas envolvendo *Machine learning* (ML), mais bibliotecas foram adquiridas do NCBI, de organismos filogeneticamente próximos aos morcegos. Entretanto, devido a dificuldade na obtenção de amostras seguindo o mesmo filtro das anteriores, as triplicatas foram reunidas de outros tecidos sendo uma triplicata controle e outra triplicata envolvendo uma condição qualquer.

Desse modo, ao final foram reunidas 94 bibliotecas extraídas de 14 organismos, sendo: *Canis lupus familiaris*, *Mus musculus*, *Homo sapiens*, *Vespertilio sinensis*, *Oryctolagus cuniculus*, *Bos taurus*, *Monodelphis domestica*, *Rhinolophus affinis*, *Dasypus novemcinctus*, *Trichechus manatus latirostris*, *Rattus norvegicus*, *Sus scrofa*, *Ovis aries* e *Meriones unguiculatus*. Todavia, devido a uma série de imprevistos quanto ao andamento deste projeto, o que afetou diretamente no tempo de execução, o presente relatório conta com uma análise incompleta e, sendo assim, negativa (por hora) sobre os resultados obtidos até aqui, pois a penúltima etapa condenou um erro grave que só pôde ser notado dentro do ambiente R.

Portanto, neste trabalho mostraremos o decorrer da preparação de dados na etapa inicial e também os *scripts* prontos para as etapas seguintes que não puderam ser executadas devido ao erro supracitado na quantificação das amostras.

2. METODOLOGIA

2.1 Obtenção de bibliotecas

Foram utilizadas bibliotecas de RNA-Seq obtidas do NCBI, geradas a partir de diversos tecidos de 14 mamíferos onde, cada organismo contava com 2 triplicatas sendo controle e condição, que podem ser consultadas utilizando os seguintes códigos através do *BioProject* como demonstrado na Figura 1.

Figura 1: Tabela demonstrando os códigos de identificação por cada organismo.

PRJNA671877	Canis lupus familiaris
PRJNA1047137	Mus musculus
PRJNA945424	Homo sapiens
PRJNA636635	Vespertilio sinensis
PRJNA757533	Oryctolagus cuniculus
PRJNA629585	Bos taurus
PRJNA193216	Monodelphis domestica
PRJNA644044	Rhinolophus affinis
PRJNA892621	Dasypus novemcinctus
PRJNA342504	Trichechus manatus latirostris
PRJNA825500	Rattus norvegicus
PRJNA798047	Sus scrofa
PRJNA773775	Ovis aries
PRJNA509196	Meriones unguiculatus

Fonte: National Center for Biotechnology Information (NCBI)

Tendo acesso aos *BioProject's* de cada uma das amostras, foram coletadas as bibliotecas de cada organismo, selecionando quais eram os controles e quais eram as condições, como demonstrado na Figura 2. Desse modo, os códigos de cada biblioteca que iniciam por “SRR” foram adicionados a um *looping* para *download* automático de todas as amostras diretamente do Bash.

Figura 2: Tabela demonstrando as bibliotecas baixadas onde, a primeira coluna refere-se aos controles (1) e condições (2), a segunda coluna refere-se aos códigos SRR de cada biblioteca e a terceira coluna refere-se aos respectivos organismos.

CONDIÇÃO	ORGANISMO				
1	SRR12899120	Canis lupus familiaris	1	SRR776573	Monodelphis domestica
1	SRR12899121	Canis lupus familiaris	1	SRR776577	Monodelphis domestica
1	SRR12899122	Canis lupus familiaris	1	SRR776569	Monodelphis domestica
1	SRR12899123	Canis lupus familiaris	2	SRR776557	Monodelphis domestica
2	SRR12899124	Canis lupus familiaris	2	SRR776561	Monodelphis domestica
2	SRR12899125	Canis lupus familiaris	2	SRR776553	Monodelphis domestica
2	SRR12899126	Canis lupus familiaris	1	SRR12145330	Rhinolophus affinis
2	SRR12899127	Canis lupus familiaris	1	SRR12145331	Rhinolophus affinis
1	SRR27004681	Mus musculus	1	SRR12145332	Rhinolophus affinis
1	SRR27004682	Mus musculus	2	SRR12145334	Rhinolophus affinis
1	SRR27004683	Mus musculus	2	SRR12145337	Rhinolophus affinis
2	SRR27004678	Mus musculus	2	SRR12145338	Rhinolophus affinis
2	SRR27004679	Mus musculus	1	SRR21987044	Dasyus novemcinctus
2	SRR27004680	Mus musculus	1	SRR21987045	Dasyus novemcinctus
2	SRR23881850	Homo sapiens	1	SRR21987046	Dasyus novemcinctus
2	SRR23881851	Homo sapiens	2	SRR21987038	Dasyus novemcinctus
2	SRR23881852	Homo sapiens	2	SRR21987039	Dasyus novemcinctus
1	SRR23881844	Homo sapiens	2	SRR21987040	Dasyus novemcinctus
1	SRR23881845	Homo sapiens	1	SRR4228541	Trichechus manatus latirostris
1	SRR23881846	Homo sapiens	1	SRR4228542	Trichechus manatus latirostris
1	SRR11906443	Vespertilio sinensis	1	SRR4228543	Trichechus manatus latirostris
1	SRR11906444	Vespertilio sinensis	2	SRR4228545	Trichechus manatus latirostris
1	SRR11906445	Vespertilio sinensis	2	SRR4228546	Trichechus manatus latirostris
2	SRR11906450	Vespertilio sinensis	2	SRR4228547	Trichechus manatus latirostris
2	SRR11906451	Vespertilio sinensis	1	SRR18713679	Rattus norvegicus
2	SRR11906452	Vespertilio sinensis	1	SRR18713680	Rattus norvegicus
1	SRR11906446	Vespertilio sinensis	1	SRR18713681	Rattus norvegicus
1	SRR11906447	Vespertilio sinensis	2	SRR18713675	Rattus norvegicus
1	SRR11906448	Vespertilio sinensis	2	SRR18713676	Rattus norvegicus
2	SRR11906453	Vespertilio sinensis	2	SRR18713678	Rattus norvegicus
2	SRR11906454	Vespertilio sinensis	1	SRR17646267	Sus scrofa
2	SRR11906455	Vespertilio sinensis	1	SRR17646268	Sus scrofa
1	SRR15602500	Oryctolagus cuniculus	1	SRR17646269	Sus scrofa
1	SRR15602501	Oryctolagus cuniculus	2	SRR17646264	Sus scrofa
1	SRR15602502	Oryctolagus cuniculus	2	SRR17646265	Sus scrofa
1	SRR15602503	Oryctolagus cuniculus	2	SRR17646266	Sus scrofa
2	SRR15602496	Oryctolagus cuniculus	1	SRR16538453	Ovis aries
2	SRR15602497	Oryctolagus cuniculus	1	SRR16538454	Ovis aries
2	SRR15602498	Oryctolagus cuniculus	1	SRR16538455	Ovis aries
2	SRR15602499	Oryctolagus cuniculus	2	SRR16538444	Ovis aries
1	SRR11657441	Bos taurus	2	SRR16538445	Ovis aries
1	SRR11657442	Bos taurus	2	SRR16538446	Ovis aries
1	SRR11657443	Bos taurus	1	SRR8309415	Meriones unguiculatus
2	SRR11657429	Bos taurus	1	SRR8309416	Meriones unguiculatus
2	SRR11657430	Bos taurus	1	SRR8309417	Meriones unguiculatus
2	SRR11657431	Bos taurus	2	SRR8309418	Meriones unguiculatus
			2	SRR8309419	Meriones unguiculatus
			2	SRR8309420	Meriones unguiculatus

Fonte: National Center for Biotechnology Information (NCBI)

Sendo assim, foi utilizado um dos servidores do Laboratório de Genômica e Biotecnologia (LGBIOTEC) para *download* e preparação inicial dos dados durante toda a primeira etapa, pois o número e o tamanho das bibliotecas tornava inviável a utilização de *notebook* com capacidade de armazenamento inferior a 1TB de HD livre. De igual modo, a necessidade de utilização de um servidor também se deu pela disponibilidade de internet no prédio em período integral, que permitiria que os arquivos fossem baixados sem interrupções.

No entanto, alguns imprevistos ainda foram enfrentados durante essa etapa. Isso porque a universidade passou por um período de instabilidade de internet e de energia, que tornou o processo mais demorado que o esperado e por diversas vezes o *looping* para

download precisou ser reiniciado com aquelas que ainda não haviam sido baixadas, o que demorou alguns dias.

Sendo assim, após a obtenção de todas as bibliotecas, os arquivos foram devidamente preparados para seguir o fluxo de trabalho nos seguintes passos: filtragem, quantificação, predição dos eventos de AS e clusterização de amostras para ML.

2.2 Triagem dos dados

As bibliotecas foram submetidas às ferramentas FASTQC (Andrews, 2010) e PRINSEQ (Schmieder; Edwards, 2011) para uma avaliação de qualidade e filtragem/eliminação das *reads* com padrões de qualidade baixos ($Q < 20$). Desse modo, os resultados obtidos foram tabulados (Figura 3) e foram usados para as etapas seguintes.

No entanto, como evidenciado ainda na figura abaixo (Figura 3), apenas as primeiras bibliotecas baixadas foram submetidas à filtragem de *reads* de baixa qualidade, pois esta é uma etapa importante mas opcional, uma vez que não havia mais tempo de filtragem para todas as 94 bibliotecas, garantindo a entrega com êxito dos resultados finais tendo passado por todo o passo a passo. Porém, vale ressaltar que esta etapa (filtragem) não inviabiliza as etapas seguintes de quantificação e predição, uma vez que apenas torna as leituras mais confiáveis por eliminar o aquilo que possui qualidade baixa.

Figura 3: Tabela de dados avaliados e filtrados, onde são demonstrados: Pré tratamento e pós tratamento seguindo as etapas FASTQC e PRINSEQ.

			PRÉ TRATAMENTO (seq. brutas)	PÓS TRATAMENTO (Prinseq)
Fastq-dump	SRR12899120	Canis lupus familiaris	36191043	36155906
Fastq-dump	SRR12899121	Canis lupus familiaris	36690194	36490561
Fastq-dump	SRR12899122	Canis lupus familiaris	71591871	70496835
Fastq-dump	SRR12899123	Canis lupus familiaris	26186699	25252484
Fastq-dump	SRR27004683	Mus musculus	30905692	30903194
Fastq-dump	SRR27004682	Mus musculus	25680992	25679457
Fastq-dump	SRR27004681	Mus musculus	25879520	25877911
Fastq-dump	SRR23881846	Homo sapiens	25988090	25988007
Fastq-dump	SRR23881845	Homo sapiens	25483297	25483193
Fastq-dump	SRR23881844	Homo sapiens	29503464	29503347
Fastq-dump	SRR11906443	Vespertilio sinensis	25838506	25835486
Fastq-dump	SRR11906444	Vespertilio sinensis	21404396	21402281
Fastq-dump	SRR11906445	Vespertilio sinensis	25872185	25867808
Fastq-dump	SRR11906446	Vespertilio sinensis	20812890	20809559
Fastq-dump	SRR11906447	Vespertilio sinensis	24191327	24186305
Fastq-dump	SRR11906448	Vespertilio sinensis	22074489	22072041
Fastq-dump	SRR15602503	Oryctolagus cuniculus	12812954	12811246
Fastq-dump	SRR15602502	Oryctolagus cuniculus	22366659	22363750
Fastq-dump	SRR15602501	Oryctolagus cuniculus	23509696	23507426
Fastq-dump	SRR15602500	Oryctolagus cuniculus	31138980	31136155

Fonte: O autor

2.3 Quantificação dos transcritos

Anteriormente, ao entregar a segunda etapa do trabalho, foi relatado um equívoco durante a utilização da ferramenta Salmon (Patro *et al.*, 2017) para a quantificação das amostras, todavia, foi constatado que a forma como as bibliotecas haviam sido quantificadas pela primeira vez estava correta. Desse modo, o processo foi repetido para cada biblioteca individualmente, utilizando o comando demonstrado na Figura 4, o qual gerou arquivos de output em formato “quant.sf” que seriam usados futuramente na predição.

Foi gerado também um índice a partir de um genoma humano, adquirido de um arquivo “.fasta” diretamente do “Ensembl genome browser 112”, etapa em que o arquivo “index”, necessário para a etapa Salmon, foi gerado normalmente, e adaptado ao código usado para quantificar os transcritos por biblioteca de forma individual.

Figura 4: Código em um arquivo de apoio “Workflow.sh” criado no Sublime Text, onde a primeira linha demonstra a criação do índice a partir de um Genoma humano e a segunda linha demonstra a etapa de quantificação por amostra.

```
#Run data - Salmon (antes de fazer a predição é necessária a criação de um 'index' com um genoma de referência que pode ser baixado do Ensembl em formato .fasta)
salmon index -t Homo_sapiens.fa -i Homo_index #para criação do index que gera o diretório Homo_index com todos os arquivos necessários para a próx etapa
salmon quant -i caminho_do_index_criado -l A -r /caminho/da/triplicata/01.sf -o salmon_output
```

Fonte: O autor

Sendo assim, a quantificação também foi custosa, pois apenas 1 servidor estava disponível para uso, mas era compartilhado por mais 2 alunos. Por esse motivo, foram utilizados apenas 8 *threads* por análise (mínimo exigido pelo Salmon) para todas as quantificações.

2.4 Predição dos eventos de *Splicing* e análises funcionais

Com os arquivos resultantes da etapa anterior e aqueles novos que não passaram pela filtragem ,já quantificados pelo Salmon, subdiretórios foram criados (Sample_1, Sample_2, Sample_3,...) para comportar cada amostra de modo que, durante o carregamento em RStudio, fossem fornecidos os caminhos do Diretório Base e dos subdiretórios individualmente (método utilizado segundo outras análises envolvendo predição de isoformas).

No entanto, antes de preparar o ambiente R com as amostras carregadas, foi usado

o *header* para avaliação de cada biblioteca, onde foi visto que quase todas estavam com valores de Transcritos Por Milhão (TPM) extremamente baixos ou zerados. Desse modo, com a maior parte das bibliotecas possuindo os valores de TPM zerados, foi possível constatar que a etapa de quantificação no Salmon gerou um erro grave que se repetiu para todas as amostras e que pode ter sido por conta do Index escolhido ou pela execução da própria ferramenta.

Sendo assim, as poucas bibliotecas que demonstraram algum valor de TPM, foram preenchidas apenas 4 ou 5 linhas de 512, como demonstrado na tabela reduzida da figura 5, de uma das amostras de *Vespertilio*.

Figura 5: Tabela demonstrando a baixa frequência dos valores de TPM de uma das amostras dos 14 organismos.

	Name	Length	EffectiveLength	TPM	NumReads
1	HG76_PATCH	6367528	6367278	0.00	0
2	HG2365_PATCH	5500449	5500199	0.00	0
3	HSCHR15_4_CTG8	5161414	5161164	0.00	0
4	HSCHR6_MHC_SSTO_CTG1	4929269	4929019	0.00	0
5	HSCHR6_MHC_MCF_CTG1	4827813	4827563	0.00	0
6	HSCHR6_MHC_COX_CTG1	4795265	4795015	0.00	0
7	HSCHR6_MHC_MANN_CTG1	4677643	4677393	0.00	0
8	HSCHR6_MHC_APD_CTG1	4672374	4672124	0.00	0
9	HSCHR6_MHC_QBL_CTG1	4606388	4606138	0.00	0
10	HSCHR6_MHC_DBB_CTG1	4604811	4604561	10547.49	4
11	HSCHR17_7_CTG4	2877074	2876824	4220.50	1
12	HSCHR16_1_CTG1	2659700	2659450	0.00	0
13	HSCHR15_6_CTG8	2365364	2365114	0.00	0
14	HG926_PATCH	1927115	1926865	0.00	0
15	HSCHR17_1_CTG5	1821992	1821742	0.00	0
16	HSCHR5_2_CTG1_1	1612928	1612678	0.00	0
17	HG1343_HG173_HG459_PATCH	1572686	1572436	0.00	0
18	HSCHR14_7_CTG1	1511111	1510861	0.00	0
19	HSCHR17_2_CTG5	1423190	1422940	0.00	0
20	HSCHR14_3_CTG1	1351393	1351143	0.00	0
21	HG2280_PATCH	1154574	1154324	0.00	0
22	HSCHR5_1_CTG1_1	1144418	1144168	0.00	0
23	HSCHR7_2_CTG6	1111570	1111320	0.00	0

Fonte: O autor

Desse modo, como já mencionado, os passos seguintes incluindo a utilização da ferramenta IsoformSwitchAnalyzer (VITTING-SEERUP; SANDELIN, 2019), e a etapa de ML não puderam ser executados em tempo hábil para a entrega do relatório final em formato de projeto com os resultados bem explicados e finalizados pela falta de um número mínimo de quantidade de dados. Por isso, espera-se para o trabalho em questão, reunir uma maior quantidade de organismos com base nos mesmos filtros utilizados para a escolha das bibliotecas já baixadas, a fim de inturgescer o número amostral, criar uma matriz bem mais robusta e identificar com clareza o erro gerado pelo Salmon, para que ao final possamos ter uma matriz grande o suficiente para a criação de um modelo confiável baseado em aprendizado de máquina capaz de identificar padrões de expressão de isoformas nesses animais.

Ainda assim, todos os scripts para as etapas seguintes foram preparados e serão utilizados no decorrer do projeto original, que também envolve a quantificação de isoformas em morcegos da família Phyllostomidae, que foi escolhida devido ao tamanho corporal e carga alar dos animais, o que pode afetar diretamente em seu consumo energético e indicar bons resultados baseados na hipótese do presente trabalho.

3. CONSIDERAÇÕES FINAIS

Mesmo com a apresentação parcial dos dados no presente estudo, que contempla as etapas de Bioinformática aplicada às ômicas mas não a etapa envolvendo ML, asseguro que o projeto é totalmente executável e que será finalizado, utilizando não apenas os dados obtidos durante o decorrer desta pesquisa, mas também os conhecimentos repassados durante toda a disciplinas.

Também reitero que devido a extensão do trabalho, o que envolve o *download* de bibliotecas, filtragem e quantificação, e aos imprevistos com relação a utilização de servidores e indisponibilidade de internet no campus, toda a entrega final de uma matriz foi comprometida para entrega em tempo hábil. Isso porque dentro do tempo planejado, não foram consideradas todas as variáveis que envolviam possíveis contratempos e um plano B para a entrega final.

Sendo assim, enfatizo que muito embora o desfecho deste projeto não tenha sido o esperado, os conhecimentos repassados em sala de aula por meio das atividades de fixação e do conteúdo foram de grande valia, uma vez que poderão acrescentar novas visões para o decorrer do trabalho final e torná-lo mais robusto e atrativo em termos bio-informáticos.

REFERÊNCIAS

- ANDREWS, Simon. **FastQC: A Quality Control Tool for High Throughput Sequence Data**. Brabraham Bioinformatics, 2010. Disponível em: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- ALI, Azad. Species diversity of bats (Mammalia: Chiroptera) in Assam, Northeast India. **Journal of Wildlife and Biodiversity**, v. 6, n. 3, p. 115-125, 2022.
- CASSAR-MALEK, Isabelle et al. Transcriptome profiling reveals stress-responsive gene networks in cattle muscles. **PeerJ**, v. 10, p. e13150, 2022.
- DOS REIS, Nelio R. et al. (Ed.). **Morcegos do brasil**. Universidade Estadual de Londrina, 2007.
- GARDNER, Alfred L. (Ed.). **Mammals of South America, volume 1: marsupials, xenarthrans, shrews, and bats**. University of Chicago Press, 2007.
- NOGUEIRA, Marcelo Rodrigues et al. Checklist of Brazilian bats, with comments on original records. **Volume 10, Número 4, Pags. 808-821**, 2014.
- NOWAK, Ronald M. **Walker's Mammals of the World**. JHU press, 1999.
- PATRO, Rob et al. Salmon provides fast and bias-aware quantification of transcript expression. **Nature Methods**, [S.l.], v. 14, p. 417-419, 2017. DOI: 10.1038/nmeth.4197.
- SIMMONS, N. B.; CIRRANELLO, A. L. Bat species of the world: a taxonomic and geographic database. Recurso online (<https://batnames.org>), acessado em 19 de março de 2023.
- THOMAS, Steven P.; SUTHERS, Roderick A. The physiology and energetics of bat flight. **Journal of Experimental Biology**, v. 57, n. 2, p. 317-335, 1972.
- VITTING-SEERUP, Kristoffer; SANDELIN, Albin. IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. **Bioinformatics**, [S. l.], v. 35, n. 21, p. 4469-4421, 2019. DOI: 10.1093/bioinformatics/btz247.
- VOIGT, C. C.; WINTER, Y. Energetic cost of hovering flight in nectar-feeding bats (Phyllostomidae: Glossophaginae) and its scaling in moths, birds and bats. **Journal of Comparative Physiology B**, v. 169, p. 38-48, 1999.
- VU, Trieu-Duc et al. Alternative splicing plays key roles in response to stress across

different stages of fighting in the fish *Betta splendens*. **BMC genomics**, v. 22, n. 5, p. 1-12, 2021.

WATSON, Hannah et al. Transcriptome analysis of a wild bird reveals physiological responses to the urban environment. **Scientific Reports**, v. 7, n. 1, p. 44180, 2017.