

# Introduction to Bioinformatics

Ernane Rosa Martins  
(Organizador)

 **Atena**  
Editora  
2019



**Ernane Rosa Martins**  
**(Organizador)**

# **Introduction to Bioinformatics**

**Atena Editora**  
**2019**

2019 by Atena Editora

Copyright © da Atena Editora

**Editora Chefe:** Profª Drª Antonella Carvalho de Oliveira

**Diagramação e Edição de Arte:** Geraldo Alves e Natália Sandrini

**Revisão:** Os autores

#### Conselho Editorial

- Prof. Dr. Alan Mario Zuffo – Universidade Federal de Mato Grosso do Sul  
Prof. Dr. Álvaro Augusto de Borba Barreto – Universidade Federal de Pelotas  
Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná  
Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília  
Profª Drª Cristina Gaio – Universidade de Lisboa  
Prof. Dr. Constantino Ribeiro de Oliveira Junior – Universidade Estadual de Ponta Grossa  
Profª Drª Daiane Garabeli Trojan – Universidade Norte do Paraná  
Prof. Dr. Darllan Collins da Cunha e Silva – Universidade Estadual Paulista  
Profª Drª Deusilene Souza Vieira Dall'Acqua – Universidade Federal de Rondônia  
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná  
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul  
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria  
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná  
Profª Drª Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia  
Profª Drª Ivone Goulart Lopes – Istituto Internazionele delle Figlie de Maria Ausiliatrice  
Profª Drª Juliane Sant'Ana Bento – Universidade Federal do Rio Grande do Sul  
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense  
Prof. Dr. Jorge González Aguilera – Universidade Federal de Mato Grosso do Sul  
Profª Drª Lina Maria Gonçalves – Universidade Federal do Tocantins  
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Profª Drª Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa  
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão  
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará  
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista  
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará  
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas  
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande  
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa  
Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

#### Dados Internacionais de Catalogação na Publicação (CIP) (eDOC BRASIL, Belo Horizonte/MG)

I61	Introduction to bioinformatics [recurso eletrônico] / Organizador Ernane Rosa Martins. – Ponta Grossa (PR): Atena Editora, 2019.  Formato: PDF Requisitos de sistema: Adobe Acrobat Reader Modo de acesso: World Wide Web Inclui bibliografia. ISBN 978-85-7247-113-8 DOI 10.22533/at.ed.138191202  1. Bioinformática. 2. Inteligência artificial. I. Martins, Ernane Rosa.  CDD 570.285
-----	---

**Elaborado por Maurício Amormino Júnior – CRB6/2422**

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores.

2019

Permitido o download da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

[www.atenaeditora.com.br](http://www.atenaeditora.com.br)

## APRESENTAÇÃO

A bioinformática é um campo interdisciplinar, que busca analisar, interpretar e processar dados biológicos, com foco na aplicação de técnicas computacionais intensivas, tais como: métodos computacionais, teoria de grafos, inteligência artificial, algoritmos matemáticos, reconhecimento de padrões, mineração de dados, algoritmos de aprendizado de máquina, processamento de imagens e simulação computacional. Como um campo interdisciplinar, a bioinformática combina diversas áreas do conhecimento, como: engenharia, matemática, física, química, estatística, ciência da computação e biologia, entre outras.

A coletânea “*Introduction to bioinformatics*” é um livro composto por 6 capítulos que abordam assuntos atuais, tais como: o adenocarcinoma gástrico que é uma malignidade com elevada incidência e mortalidade no mundo; o vírus zika (VZIK) que é um Arbovirus que pertence à família Flaviviridae; As  $H^+$ -ATPases que são proteínas integrais da membrana plasmática que têm a capacidade de utilizar a energia química da hidrólise de ATP para expulsar os prótons para o ambiente extracelular, atuando na manutenção da homeostase iônica e transporte de solutos; o vírus da família Geminiviridae que tem sido intensamente estudado devido à gravidade das doenças causadas em várias culturas importantes como: feijão, algodão, milho, tomate e mandioca.

Espero que os capítulos deste livro possam contribuir efetivamente na disseminação dos conhecimentos relevantes da bioinformática, proporcionando uma visão ampla sobre este campo de conhecimento.

Assim, desejo a todos uma excelente leitura.

Ernane Rosa Martins

## SUMÁRIO

### **CAPÍTULO 1 ..... 1**

ANÁLISE DE lncRNAs EM NPCs DE HAMSTER GOLDEN SÍRIO (*Mesocricetus auratus*) RECÉM-NASCIDOS INFECTADOS PELO VÍRUS ZIKA

Jardel Fabio Lopes Ferreira  
Samir Mansour Moraes Casseb  
Karla Fabiane Lopes de Melo  
Carlos Alberto Marques de Carvalho  
Gustavo Moraes Holanda  
Paloma Daguer Ewerton dos Santos  
Suellen de Almeida Machado  
Francisco Canindé Ferreira de Luna  
Walter Felix Franco Neto  
Lívia Carício Martins  
Ana Cecília Ribeiro Cruz  
Pedro Fernando da Costa Vasconcelos

**DOI 10.22533/at.ed.1381912021**

### **CAPÍTULO 2 ..... 11**

ANÁLISE *IN SILICO* DA FARMACOCINÉTICA E FARMACODINÂMICA DO COMPOSTO BENZOTIAZÓLICO COM POTENCIAL ANTITUMORAL CONTRA LINHAGEM DE ADENOCARCINOMA GÁSTRICO

Felipe Pantoja Mesquita  
Luina Benevides Lima  
Julio Paulino Daniel  
Adrhyan Jullyanne de Sousa Portilho  
Lais Lacerda Brasil de Oliveira  
Emerson Lucena da Silva  
Eliza de Lucas Chazin  
Thatyana Rocha Alves Vasconcelos  
Maria Elisabete Amaral de Moraes  
Raquel Carvalho Montenegro

**DOI 10.22533/at.ed.1381912022**

### **CAPÍTULO 3 ..... 23**

ANÁLISE PRIMÁRIA DE TRANSCRIPTOMA DE TECIDO TESTICULAR DE HAMSTERS (*MESOCRICETUS AURATUS*) INFECTADOS COM VÍRUS ZIKA

Walter Felix Franco Neto  
Samir Mansour Moraes Casseb  
Karla Fabiane Lopes de Melo  
Wallax Augusto Silva Ferreira  
Ana Paula Sousa Araujo  
Jardel Fabio Lopes Ferreira  
Taiana Andrade Freitas  
Milene Ferreira Silveira  
Livia Carício Martins  
Pedro Fernando da Costa Vasconcelos

**DOI 10.22533/at.ed.1381912023**

<b>CAPÍTULO 4</b> .....	<b>32</b>
CARACTERIZAÇÃO FILOGENÉTICA DA FAMÍLIA MULTIGÊNICA DA H <sup>+</sup> -ATPASE DE MEMBRANA PLASMÁTICA EM MONOCOTILEDÔNEAS DA ORDEM POALES	
Lyndefânia Melo de Sousa Clesivan Pereira dos Santos Thais Andrade Germano Moacíria de Souza Lemos Stelamaris de Oliveira Paula Rafael de Souza Miranda José Helio Costa	
<b>DOI 10.22533/at.ed.1381912024</b>	
<b>CAPÍTULO 5</b> .....	<b>40</b>
CLADISTIC ANALISYS IN GEMINIVIRIDAE: AN EVIDENCE OF MULTISPECIFICITY FOR CULTIVARS HOSTS	
Rafael Trindade Maia Aparecida Yasmim Silva de Azevedo Maria Bartira Chaves de Souza Silva Ana Verônica Silva do Nascimento	
<b>DOI 10.22533/at.ed.1381912025</b>	
<b>CAPÍTULO 6</b> .....	<b>50</b>
DESENVOLVIMENTO DE FRAMEWORK PARA CRIAÇÃO DE MODELOS COMPUTACIONAIS DE CÉLULA COMPLETA	
Frederico Chaves Carvalho Paulo Eduardo Ambrósio	
<b>DOI 10.22533/at.ed.1381912026</b>	
<b>CAPÍTULO 7</b> .....	<b>63</b>
IN-SILICO DETOXIFICATION EVIDENCE OF THE HERBICIDE BISPYRIBAC SODIUM BY A TEORETHICAL MODEL OF GLUTATHIONE S-TRANSFERASE TAU 5 FROM <i>Oryza sativa</i> L.	
Vinícius Costa Amador Ravenna Lins Rodrigues Felipe de Oliveira França Rafael Trindade Maia	
<b>DOI 10.22533/at.ed.1381912027</b>	
<b>CAPÍTULO 8</b> .....	<b>73</b>
INVESTIGAÇÃO IN SILICO DA EFICÁCIA DE FÁRMACOS ANTIVIRAIS NA INIBIÇÃO DA NS5 DO VÍRUS DA ZIKA	
Henriqueta Monalisa Farias Rafael de Lima Medeiros Franklin de Ferreira Farias Nóbrega Rafael Trindade Maia	
<b>DOI 10.22533/at.ed.1381912028</b>	
<b>SOBRE O ORGANIZADOR</b> .....	<b>85</b>

## ANÁLISE DE LNCRNAS EM NPCS DE HAMSTER GOLDEN SÍRIO (*Mesocricetus auratus*) RECÉM-NASCIDOS INFECTADOS PELO VÍRUS ZIKA

**Jardel Fabio Lopes Ferreira**

Instituto Evandro Chagas  
Ananindeua – PA

**Samir Mansour Moraes Casseb**

Instituto Evandro Chagas  
Ananindeua – PA

**Karla Fabiane Lopes de Melo**

Escola Superior da Amazônia  
Belém – PA

**Carlos Alberto Marques de Carvalho**

Centro Universitário Metropolitano da Amazônia –  
UNIFAMAZ – PA

**Gustavo Moraes Holanda**

Universidade do Estado do Pará – UEPA – PA

**Paloma Daguer Ewerton dos Santos**

Instituto Evandro Chagas – IEC – PA

**Suellen de Almeida Machado**

Instituto Evandro Chagas  
Ananindeua – PA

**Francisco Canindé Ferreira de Luna**

Universidade Federal do Pará  
Belém – PA

**Walter Felix Franco Neto**

Instituto Evandro Chagas  
Ananindeua – PA

**Lívia Carício Martins**

Instituto Evandro Chagas  
Ananindeua – PA

**Ana Cecília Ribeiro Cruz**

Instituto Evandro Chagas  
Ananindeua – PA

**Pedro Fernando da Costa Vasconcelos**

Instituto Evandro Chagas  
Ananindeua – PA

**RESUMO:** O Vírus Zika (VZIK) é um arbovírus pertencente à família *Flaviviridae*, gênero *flavivirus* e necessita de vetores artrópodes do gênero *Aedes* (*Ae.aegypti* e *Ae.albopictus*) para manter seu ciclo de replicação viral. Possui RNA de fita simples com polaridade positiva com informações para codificação de proteínas estruturais e não estruturais. O VZIK vem se destacando a partir de epidemias nas Américas e, principalmente, pela capacidade em causar agravos neurológicos severos como a microcefalia em crianças recém-nascidas e síndrome de Guillian-Barré em adultos. Além disso, foi relatada a infertilidade em homens infectados com o vírus, o que evidencia o tropismo por diferentes órgãos. Portanto, tentar elucidar a relação vírus-hospedeiro através de RNAs longos não-codificantes de proteínas (lncRNAs) a partir de transcriptoma dos animais infectados utilizando RNA-seq, é de suma importância para o meio acadêmico científico, uma vez que esses RNAs são reguladores da expressão gênica e estão intimamente relacionados a processos fisiológicos celulares importantes, como apoptose e proliferação celular.

## 1 | INTRODUÇÃO

O Vírus Zika (VZIK) teve seu isolamento na África 1947, em Uganda na floresta de Zika, originando o nome dado a esse arbovírus. O isolamento do VZIK foi a partir de amostras coletadas de macaco do gênero *Rhesus* que servia de estudo para o Vírus da Febre Amarela (VFA) na região.

O VZIK possui genoma completo de aproximadamente 10.5 a 11kb de tamanho, formado por RNA de fita simples com polaridade positiva, responsável por conter informação para codificação de proteínas estruturais e não-estruturais sendo essas últimas responsáveis principalmente pela evasão contra a resposta imune do hospedeiro.

Pertencente à família *Flaviviridae* e gênero *flavivirus*, assim como outros arbovírus importantes para a saúde pública mundial como o Vírus Dengue (VDEN) e VFA, o VZIK necessita de vetores artrópodes do gênero *Aedes* para manter seu ciclo de replicação, sendo *Ae.aegypti* e *Ae.albopictus* seus principais vetores, podendo infectar células animais hospedeiras e ocasionar uma gama de doenças em humanos.

O termo arbovírus provém do inglês *arthropod-borne* que foi designado em 1962, para descrever os arbovírus que infectam animais vertebrados a partir de vetores artrópodes hematófagos. As principais regiões que concentram uma grande variedade de vetores artrópodes são as de climas temperadas e tropicais, possibilitando um ambiente propício para reprodução desses vetores e causando sua disseminação no mundo.

## 2 | FAMÍLIA FLAVIVIRIDAE E GÊNERO FLAVIVIRUS

A família *Flaviviridae* e gênero *flavivirus* contém a maioria dos arbovírus que causam problemas de saúde pública e social em todo o mundo, chamando atenção da Organização da Mundial da Saúde (OMS) para tomar medidas preventivas contra epidemias. Essa classe de família e gênero é formada por vírus que possuem RNA de fita simples com polaridade positiva, chamando atenção para outros arbovírus que além do VZIK, são capazes de ocasionar agravos neurológicos como o Vírus West Nile (VWN) e Vírus da Encefalite Japonesa (VEJ).

O gênero *Flavivirus* possui dois agrupamentos importantes, sendo os vírus transmitidos por carrapatos e os vírus transmitidos por mosquitos. A diversidade de vetores artrópodes de acordo com as características das doenças por eles provocadas e ampla distribuição geográfica, são fatores importantes para considerar que há necessidade de medidas preventivas uma vez que inúmeras populações mundiais residem em áreas endêmicas.

O material genético é lido no sentido 5' e 3', possuindo regiões que não são traduzidas (RNT) de aproximadamente 100 a 500 nucleotídeos. Seu genoma pode variar de 10.5kb a 11kb formando proteínas não-estruturais que auxiliam no processo de replicação e proteínas estruturais, importantes nas estratégias virais de infecção.

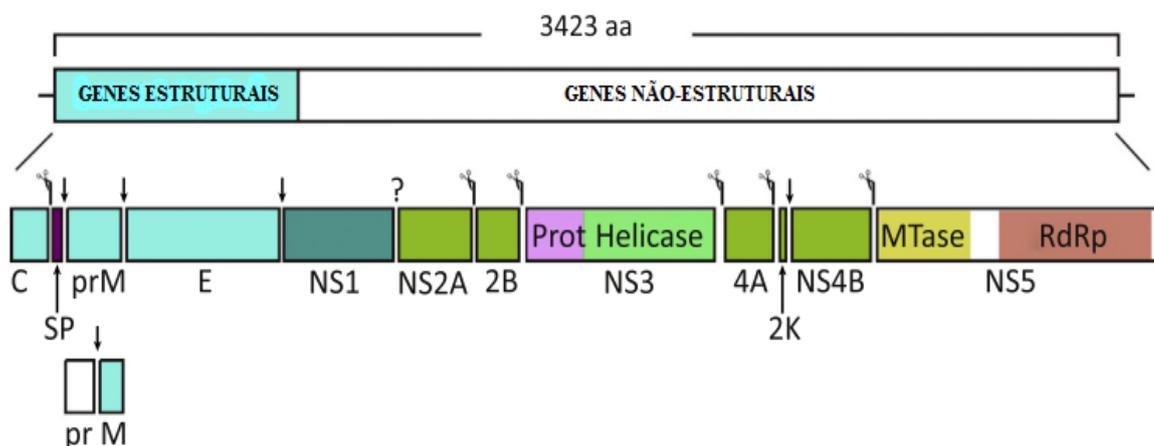
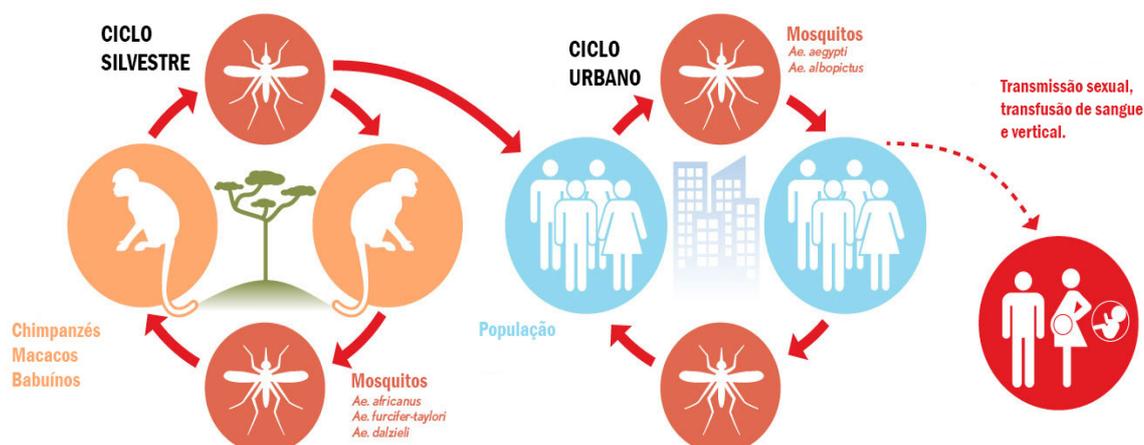


FIGURA 1: Genoma completo de flavivírus, demonstrando a codificação de genes relacionados a morfologia viral e os genes não-estruturais que estão responsáveis por auxiliar o processo de replicação e evasão imunológica da célula hospedeira.

FONTE: Adaptado de SHI and GAO, 2017.

### 3 | DINÂMICA DE TRANSMISSÃO DO VÍRUS ZIKA

Para que ocorra a transmissão do VZIK para seres humanos, ocorre primeiramente o ciclo silvestre e posterior urbano, através das picadas de mosquitos fêmeas infectadas com o vírus, sendo *Aedes aegypti* e *Aedes albopictus*, os principais vetores, assemelhando-se a forma de transmissão dos VDEN e Vírus Chikungunya. Os mosquitos do gênero *Aedes* estão confinados nas regiões tropicais e subtropicais, limitando assim o potencial de transmissão do VZIK.



Sources: CDC, PLOS, Reuters Credits: David Foster, Laurie Garrett, Doug Halsey, Gabriella Meltzer

No entanto, estudos demonstraram transmissão perinatal e potencial risco para a transmissão por transfusão sanguínea. Aliado a isso, a transmissão sexual do VZIK foi sugerida por Faye e colaboradores (2011) e posteriormente o vírus foi detectado em amostras de sêmen e urina de um paciente com o quadro de hematospemia que havia apresentado sintomas da infecção duas semanas antes da coleta.

A presença de VZIK na urina, em títulos mais elevados por um período mais longo que em amostras de soro, sugere sua adequação como amostra para detecção viral e diagnóstico da infecção, sobretudo, depois de intervalos maiores após o início dos sintomas (entre 10 ou mais de 20 dias).

#### 4 | MANIFESTAÇÕES CLÍNICAS

As manifestações clínicas resultantes da infecção pelo VZIK, estão relacionadas com um grande número de pacientes assintomáticos, aproximadamente 80%, tendo como contrastes as manifestações sintomáticas intimamente relacionadas com outras infecções causadas por flavivirus, como VDEN e VCHIK. O VZIK é incubado por cerca de 12 dias, os pacientes sintomáticos apresentam febre, erupção cutânea mialgia, artralgia, conjuntivite, distúrbios gastrointestinais e dor de cabeça.

A apresentação clínica da infecção por VZIK geralmente não é específica e pode ser confundida com outras arboviroses, mais comumente DENV e CHIKV. Em outubro de 2015, casos de microcefalia começaram a ser investigados no nordeste do Brasil para verificar a hipótese de uma associação com o VZIK.

A microcefalia geralmente resulta do desenvolvimento anormal do cérebro. Suas consequências a longo prazo dependem das anomalias fundamentais do cérebro e podem variar de atrasos leves no desenvolvimento a déficit intelectual e motor, como paralisia cerebral. Além das infecções congênitas, a mesma pode resultar de anormalidades cromossômicas, exposição a drogas, álcool ou outras toxinas ambientais, fusão prematura dos ossos do crânio (craniossinostose) e determinados distúrbios metabólicos.

#### 5 | EPIDEMIOLOGIA DO VÍRUS ZIKA

A distribuição geográfica dos arbovírus se estende pelos continentes e por todas as regiões, em especial nas regiões tropicais, devido às condições favoráveis. Apesar de hospedeiros e vetores existirem durante todo o ano, o ciclo de transmissão é interrompido em países de clima temperado na estação de inverno.

O vírus pode ser rapidamente disseminado, uma vez que uma única fêmea

de *Aedes aegypti* pode infectar vários hospedeiros por meio de suas alimentações curtas (CONSOLI & OLIVEIRA, 1998). No início do ano de 2015, métodos moleculares confirmaram a circulação do VZIK na Bahia (CAMPOS; BANDEIRA; SARDI, 2015) e em São Paulo, logo em seguida, nos estados do Rio Grande do Norte, Alagoas, Maranhão, Pará e Rio de Janeiro.

Aliado a isso, durante o surto da infecção pelo VZIK reportado recentemente no nordeste do Brasil no início do ano de 2015, as autoridades em saúde pública receberam relatos de médicos a cerca de um aumento no número de recém-nascidos com microcefalia. Logo em seguida, o Ministério da Saúde do Brasil (MS) confirmou um crescimento na prevalência de nascimentos com microcefalia no nordeste do país, comparado às estimativas registradas anteriormente (cerca de 0,5/10.000 nascidos vivos), que são baseadas na revisão das certidões de nascimento e incluem as descrições das maiores anomalias congênitas. Desde então, o MS estabeleceu um registro de microcefalia no Brasil, e em novembro de 2015, divulgou o aumento no número de casos de microcefalia.

Com frequência, as infecções por VZIK se tornaram uma realidade, com consequências inimagináveis para uma infecção causada por arbovírus. Embora a ocorrência de microcefalia e síndrome de Guillain-Barré seja atribuída ao VZIK, ainda há muito a ser investigado. Em 1º de fevereiro de 2016, a OMS declarou agrupamento de microcefalia e VZIK, principalmente observado nas regiões brasileiras, como “Emergência de Saúde Pública de Interesse Internacional” (PHEIC), isto é indicativo da extensão da emergência global.

## 6 | RNAs LONGOS NÃO-CODIFICANTES (lncRNAs)

Os lncRNAs são RNAs com 200 nucleotídeos ou mais que são sintetizados abundantemente por diversos tipos celulares. Apesar do conhecimento voltado aos lncRNAs ainda ser módico, é elucidado que eles são encarregados de promover a regulação celular, como desenvolvimento, diferenciação e metabolismo.

As alterações fisiopatológicas ocasionadas pela infecção viral nas células do hospedeiro, estão relacionadas com a expressão diferencial de RNAs não-codificantes de proteínas responsáveis pela regulação gênica celular, como os lncRNAs (*long non-coding RNAs*).

Embora a patogênese do VZIK hoje seja bem elucidada, muitos estudos ainda são necessários para uma compreensão de forma mais completa de como ocorre a interação vírus-hospedeiro. Diante disso, uma vasta área de transcritos são alvos importantes para ajudar a compreender tal interação, uma vez que o vírus utiliza de vias diferenciadas que estão relacionadas com importantes mecanismos fisiológicos celulares - afim de evadir da resposta imune e manter seu ciclo de replicação - sendo ativadas a partir da regulação gênica através de seus mediadores, como os lncRNAs

## 7 | MATERIAL E MÉTODOS

### 7.1 Preparação da biblioteca e sequenciamento de RNA

As bibliotecas de RNA-seq foram preparadas de acordo com o protocolo sugerido pelo guia de preparação de amostras TruSeq da Illumina. Resumidamente, 3 a 5 µg de RNA total foram usados para o sequenciamento. O RNA restante foi usado para a construção de ambas as cadeias de cDNA. A biblioteca foi então sequenciada utilizando protocolos e reagentes padrão da Illumina em instrumentação NextSeq 550 (Illumina). Os arquivos Fastq foram gerados pelo software CASAVA 1.8 offline. Análise de controle de qualidade adicional foi realizada usando FastQC (Ref). A qualidade média da sequência por base para todas as amostras foi superior a 32.

### 7.2 Análise dos dados

As leituras foram alinhadas ao genoma de referência do *Mesocricetus auratus* disponível no NCBI usando o programa Tophat (Ref) com os seguintes parâmetros “-G *mesocricetus.gtf* -l 2000 -i 40 --no-coverage-search UCSC/sacCer3/Bowtie2Index/genome”. Limites de transcrição e quantificação foram determinados usando o programa cufflinks (Ref). A visualização de dados foi realizada no programa R Studio (<http://www.R-project.org/>).

Os transcritos foram considerados novos se eles não se sobrepusessem com transcritos anotados ou UTRs das extremidades fossem maiores de 50 pb diferentes dos RNAs não codificantes previamente relatados SUTs entre outros.

## 8 | RESULTADOS E DISCUSSÃO

### 8.1 Identificação dos lncRNAs

Para a identificação de lncRNAs que foram submetidas cerca de 193 milhões de *reads* aos filtros de controle de qualidade para os efeitos aleatórios da variabilidade deixando apenas transcritos que estavam presentes ou eram secundários para RNA-Seq. Para as leituras filtradas, a montagem *de novo* foi realizada, e 193.846 contigs foram gerados com um valor N50 igual a 1,171 e taxa efetiva de leituras mapeadas de 84,66%. Estes contigs foram usados como sequências de referência para a identificação de lncRNA putativos em Hamster.

A dificuldade em caracterizar lncRNA em *Mesocricetus auratus* a falta de um robusto banco de dados lncRNA e a baixa conservação do genoma entre hamsters, exigiram um processamento em vários estágios dos dados com objetivos específicos. A primeira foi a remoção dos transcritos codificadores de proteínas (exclusão de transcritos que codificassem proteínas conhecidas). Como resultado, acima de 70K contigs foram anotados (valor  $E < 10^{-5}$ ) que foram descartados de análise posterior.

Em segundo lugar, um filtro altamente rigoroso foi aplicado para cobertura contig.

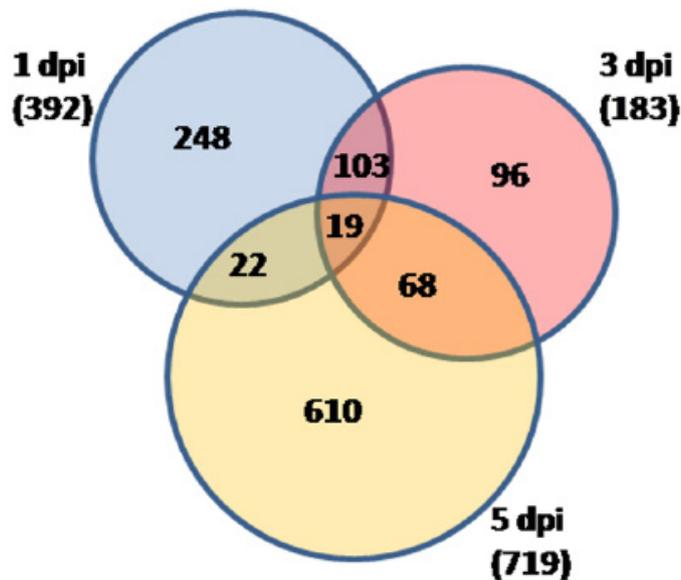


Figura 3: Representação da expressão diferencial de lncRNAs

Este filtro permite minimizar o viés da montagem *de novo* e descarta os contigs com baixa cobertura (12k contigs com alta cobertura <50 reads / pb). Como novos transcritos de codificação de proteína poderiam se associar a novos transcritos não-codificantes no processo de predição, foi utilizado o Calculador do Potencial de Codificação (CPC) entre 28-30 em para avaliar o potencial de codificação de proteínas de novos transcritos e remover transcritos putativos de codificação de proteínas. Em seguida, inspecionamos a presença de quadros de leitura abertos (ORF) > 200 e a presença de domínios conservados de proteína putativa entre todos os quadros de tradução possíveis. Com base no valor de CPC e na previsão de ORF, os contigs não mostraram evidência de codificação de domínios potenciais e putativos conservados dentro das suas sequências de aminoácidos previstas. As transcrições restantes foram mapeadas contra o genoma de referência para garantir a qualidade da montagem. Desse conjunto, 5.635 não apresentaram sobreposição parcial de locos preditos e representam potenciais lncRNAs de *Mesocricetus auratus*.

## 8.2 Caracterização dos lncRNAs

Para examinar se os lncRNAs identificados neste estudo exibem características típicas observadas em outros mamíferos (principalmente Hamsters), analisamos características como potencial de codificação, comprimento de sequência, conteúdo de GC e conservação de sequências com espécies intimamente relacionadas. Como os lncRNAs eram estritamente definidos por sua incapacidade de codificar proteína, determinamos a probabilidade de codificação de nossos lncRNAs recém-identificados e os comparamos com os conhecidos lncRNA, 3'UTR, 5'UTR e mRNA codificador de proteínas. Descobrimos que, à semelhança de outras sequências não codificantes,

como o conhecido lncRNA, 3'UTR e 5'UTR, nossos novos transcritos de lncRNA têm probabilidade de codificação extremamente baixa quando comparados ao mRNA codificador de proteínas. Além disso, descobrimos que os novos e conhecidos lncRNAs (fornecidos por anotação genômica de *Mesocricetus auratus*) eram mais curtos do que os transcritos codificadores de proteínas. O comprimento médio dos novos e conhecidos lncRNAs foi de 815 pb e 137727pb respectivamente, enquanto o mRNA codificador de proteínas tem um comprimento médio de 3220 pb.

Semelhante aos relatos anteriores (LIU; DING, 2017; OUYANG; HU; CHEN, 2016), observamos que os lncRNAs identificados neste estudo apresentavam um conteúdo GC menor que os mRNAs codificadores de proteínas. Por exemplo, o conteúdo médio GC de novos lncRNA e mRNA foi de 38,1% e 45,2%, respectivamente. Por outro lado, o conteúdo GC médio da sequência 5'UTR e 3'UTR foi de 43,1% e 34,6%, respectivamente. No geral, o conteúdo de GC da sequência não codificadora foi relativamente menor que a sequência de codificação.

Ao Determinar os lncRNAs que foram altamente conservados para outras espécies relacionados as espécies de estudo, semelhante ao estudo anterior (CULSHAW; MONGKOLSAPAYA; SCREATON, 2018; OUYANG; HU; CHEN, 2016; WANG et al., 2017), observamos sequencias com alto grau de semelhança. Ao agrupar esses lncRNAs altamente conservados, descobrimos que a maioria deles compartilhava semelhanças de sequência com Hamster domésticos, sugerindo que eles eram presumivelmente específicos do gênero. Embora os lncRNAs do *Mesocricetus auratus* compartilhe alta similaridade, principalmente *Mus musculus*, em comparação com os lncRNAs totais. Apenas 11% (1.258 dos 11.376 lncRNAs) apresentaram alta similaridade com espécies evolutivamente mais relacionados.

## REFERÊNCIAS

BASARAB, M.; BOWMAN, C.; AARONS, E. J.; e CROPLEY, I. Zika Virus. **BMJ** 352:i1049,2016

BUATHONG, R.; HERMANN, L.; THAISOMBOONSUL, B.; RUTVISUTTINUNT, W.; KLUNGTHONG, C.; CHINNAWIROTPISAN, P.; et al. Detection of Zika Virus Infection in Thailand, 2012-2014. **Am. J. Trop. Med. Hyg.** 93: 380-383. 2015

CASSEB, A. R.; CASSEB, L. M. N.; SILVA, S. P. VASCONCELOS, P. F. C. Arbovírus: Importante Zoonose na Amazônia Brasileira. **Veterinária e Zootecnia**, v.20, n.3, p. 9-21, 2013.

CHAMBERS, T. J., HAHN, C. S., GALLER, R.; RICE, C. M. Flavivirus: genome

organization, expression and replication. **Revista de Microbiologia**.44: p. 649-688, 1990.

CULSHAW, A.; MONGKOLSAPAYA, J.; SCREATON, G. The immunology of Zika Virus. **F1000Research**, v. 7, n. 0, p. 203, 2018.

DICK, G. W.; KITCHEN, S. F.; HADDOW, A. J. Zika virus. I. Isolations and serological specificity. **Trans R Soc Trop Med Hyg**, n. 5 v. 46, p. 509-20, 1952.

- GOURINAT, A. C.; O'CONNOR, O.; CALVEZ, E.; GOARANT, C.; DUPONT-ROUZEYROL, M; Detection of zika virus in urine. **Emerging Infectious Diseases**, v. 21, n. 1, p. 84–86, 2015.
- GRITSUN, T.S., TUPLIN, A.K., GOULD, E.A., 2006. Origin, evolution and function of flavivirus RNA in untranslated and coding regions: implications for virus transmission. In: Kalitzky, M., Borowski, P. (Eds.), *Flaviviridae: Pathogenesis, Molecular Biology and Genetics*. **Horizon Scientific Press**, pp. 47–99
- HOLBROOK, R. M. Historical perspectives on Flavivirus Research. **Viruses**. **9**: 97, 2017
- LIU, W.; DING, C. Roles of LncRNAs in Viral Infections. **Frontiers in Cellular and Infection Microbiology**, v. 7, n. May, 26 maio 2017.
- McGRAW, E. A.; O'NEILL, S. L. Beyond insecticides: new thinking on an ancient problem. **Nature Reviews Microbiology**, v.11, n.3, p. 181-193, 2013.
- MERCER, T. R.; MATTICK, J. S. Structure and function of long noncoding RNAs in epigenetic regulation. **Nature Structural & Molecular Biology**, v.20, n.3, p. 300-307, 2013.
- MOTA O. T. M; Moquito-transmitted viruses - the great Brazilian Challenge **Brazilian Journal Of Microbiology**, 2016
- MUSSO, D.; NHAN, T.; ROBIN, E.; ROCHE, C.; BIERLAIRE, D.; ZISOU, K.; SHAN, A. Y.; CAO-LORMEAU, V. M.; BROULT, J.; Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014. **Euro Surveill**, n.14 v.19, p. 1-3, 2014
- NEUFELDT, C. J.; CORTESE, M.; ACOSTA, E. G.; BARTENSCHLAGER, R. Rewiring cellular networks by members of the Flaviviridae Family. **Nature Reviews Microbiology**, v.16, n.3, p. 125-142, 2018.
- OUYANG, J.; HU, J.; CHEN, J.-L. lncRNAs regulate the innate immune response to viral infection. **Wiley Interdisciplinary Reviews: RNA**, v. 7, n. 1, p. 129–143, jan. 2016.
- PAPAIT, R.; KUNDERFRANCO, P.; STIRPARO, G. G.; LATRONICO, M. V. G.; CONDORELLI, G. Long Noncoding RNA: a New Player of Heart Failure?. **Journal of Cardiovascular Translational Research**, v.6, n.6, p. 876-883, 2013.
- REGO, S.; PALÁCIOS, M.; Ética, saúde global e a infecção pelo vírus zika: uma visão a partir do Brasil. **Revista Bioética**, n.3 v.24, p. 430-434, 2016.
- SCHULER-FACCINI, L. et al. Possível associação entre a infecção pelo vírus zika e a microcefalia — Brasil, 2015. **Weekly**, v. 65, n. 3, 2016.
- Strauss, J.H., Strauss, E.G. Evolution of RNA viruses. **Microbiol.** v.42, 657–683, 1988.
- THOMAS, S. M.; OBERMAYR, U.; FISCHER, D.; KREYLLING, J.; BEIERKUHNLEIN, C. Low-temperature threshold for egg survival of a post-diapause and non-diapause European aedine strain, *Aedes albopictus* (Diptera: Culicidae). **Parasites and Vectors**, v.5, n.1, p. 1-7, 2012.
- WANG, Z.; WANG, P.; AN, J. Zika virus and Zika fever. **Virologica Sinica**, v. 31, n. 2, p. 103–109, 2016.
- WANG, A. et al. Zika virus genome biology and molecular pathogenesis. **Emerging Microbes & Infections**, v. 6, n. 3, p. e13, 2017.

ZANLUCA, C.; MELO, V. C. A.; MOSIMANN, A. L. P.; SANTOS, G. I. V.; LUZ, K.; First report of autochthonous transmission of Zika virus in Brazil. **Memorias do Instituto Oswaldo Cruz**, v. 110, n. 4, p. 569–572, 2015

## **SOBRE O ORGANIZADOR**

**ERNANE ROSA MARTINS** Doutorado em andamento em Ciência da Informação com ênfase em Sistemas, Tecnologias e Gestão da Informação, na Universidade Fernando Pessoa, em Porto/Portugal. Mestre em Engenharia de Produção e Sistemas pela PUC-Goiás, possui Pós-Graduação em Tecnologia em Gestão da Informação pela Anhanguera, Graduação em Ciência da Computação pela Anhanguera e Graduação em Sistemas de Informação pela Uni Evangélica. Atualmente é Professor de Informática do Instituto Federal de Educação, Ciência e Tecnologia de Goiás - IFG (Câmpus Luziânia), ministrando disciplinas nas áreas de Engenharia de Software, Desenvolvimento de Sistemas, Linguagens de Programação, Banco de Dados e Gestão em Tecnologia da Informação. Pesquisador do Núcleo de Inovação, Tecnologia e Educação (NITE).

Agência Brasileira do ISBN  
ISBN 978-85-7247-113-8

