

UNIVERSIDADE ESTADUAL PAULISTA
"Júlio de Mesquita Filho"
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU

MATHEUS RODRIGUES SAUDA

Análise de redes complexas do transcriptoma de células
epiteliais do tecido brônquico primário com Fibrose Cística e
associação com Rinovírus

UNIVERSIDADE ESTADUAL PAULISTA
"Júlio de Mesquita Filho"
INSTITUTO DE BIOCIÊNCIAS DE BOTUCATU

Análise de redes complexas do transcriptoma de células
epiteliais do tecido brônquico primário com Fibrose Cística e
associação com Rinovírus

MATHEUS RODRIGUES SAUDA

ORIENTADOR

GUILHERME TARGINO VALENTE

Trabalho de Conclusão de Curso apresentado ao
Instituto de Biociências, Campus de Botucatu, UNESP,
para obtenção de Bacharel em Ciências Biomédicas.

BOTUCATU – SP

2023

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP

BIBLIOTECÁRIA RESPONSÁVEL: MARIA CAROLINA A. CRUZ E SANTOS-CRB 8/10188

Sauda, Matheus Rodrigues.

Análise de redes complexas do transcriptoma de células epiteliais do tecido brônquico primário com Fibrose Cística e associação com Rinovírus / Matheus Rodrigues Sauda. - Botucatu, 2023

Trabalho de conclusão de curso (bacharelado - Ciências Biomédicas) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu
Orientador: Guilherme Targino Valente
Capes: 20200005

1. Bioinformática. 2. Fibrose cística. 3. Transcriptoma.

Palavras-chave: Bioinformática; Redes; Transcriptoma.

AGRADECIMENTOS

Exponho aqui meus sinceros agradecimentos.

Gostaria de agradecer primeiramente aos meus pais, irmã e avó: Roberto dos Santos Sauda, Elizangela Rodrigues Sauda, Juliana Rodrigues Sauda e Luzia Rodrigues de Moura, respectivamente, por me proporcionarem as condições necessárias para concluir o curso, tanto no âmbito emocional com o incentivo de seguir minhas escolhas, quanto no âmbito financeiro com a viabilização de realizar a faculdade em outra cidade, arcando com os custos. Amo vocês.

Agradeço a minha namorada, Maria Letícia de Oliveira Lyra, por estar sempre ao meu lado em momentos bons e momentos ruins, me apoiando durante todo o percurso do bacharelado e, fornecendo o apoio de uma verdadeira companheira, Amo você.

Também gostaria de agradecer ao meu orientador, Professor Doutor Guilherme Targino Valente, pelas inúmeras correções em trabalhos e scripts, compartilhamento de conhecimento, paciência ao ensinar e também pela possibilidade de aprender sobre o mundo de Bioinformática e me permitir ingressar nessa área tão interessante e vasta. Obrigado!

Gostaria de agradecer a amigos próximos ou distantes que de alguma forma puderam contribuir para a minha formação, seja de forma direta na faculdade, ou de forma indireta, proporcionando momentos de risada e felicidade.

Por fim, sou grato a UNESP de Botucatu pela experiência adquirida durante o curso de Ciências Biomédicas, obrigado.

SUMÁRIO

1. RESUMO	1
1. ABSTRACT	3
2. INTRODUÇÃO	5
3. OBJETIVO	9
4. METODOLOGIA	10
5. RESULTADOS	16
6. DISCUSSÃO	21
7. CONCLUSÃO	23
8. REFERÊNCIAS BIBLIOGRÁFICAS	25

1. RESUMO

A Fibrose Cística é uma doença autossômica recessiva que afeta cerca de 70.000 pessoas por ano no mundo e 1 a cada 7.576 nativos no Brasil. A mutação altera o gene responsável por codificar a proteína responsável pela condutância de cloro nas células epiteliais e há diferenças no tipo de mutação ocorrida no gene, sendo caracterizada em alterações que podem variar desde a síntese de proteínas incompletas ou truncadas até a não expressão proteica. O quadro clínico de pacientes com Fibrose Cística inclui a produção em excesso de secreções nas vias aéreas e episódios de exacerbação, que abrangem tosse e dispneia. A associação comum entre vírus oportunistas e Fibrose Cística conduziu ao estudo realizar transcriptomas de células do epitélio brônquico primário em pacientes infantis com a doença e associação ou não de Rinovírus.

O objetivo do projeto foi o de reanalisar os transcriptomas obtidos pelo artigo “RNA sequencing of primary bronchial airway epithelial cells from young children with and without CF, including those with and without rhinovirus infection in vitro” e também conduzir análises em redes complexas. A reanálise foi realizada utilizando linguagens de programação como R e Python e uso de banco de dados como o BioGRID e Uniprot, obtendo um comparativo com os valores de contagens de genes dos grupos estabelecidos: Controle - Pacientes sem Fibrose Cística e sem Rinovírus (A); Pacientes sem Fibrose Cística e com Rinovírus (B); Pacientes com Fibrose Cística e sem Rinovírus (C); Pacientes com Fibrose Cística e com Rinovírus (D).

Os resultados revelaram vias enriquecidas como a de morte celular, regulação do metabolismo de proteínas e biossíntese de RNA quando há a presença do Rinovírus nas amostras. Cabe salientar que a reanálise obteve uma via enriquecida de

regulação positiva da sinalização de proteínas, não presente de forma direta em literaturas que envolvem a presença de Rinovírus em pacientes com a doença e em amostras de pacientes com doenças respiratórias, que pode servir de auxílio em pesquisas futuras envolvendo os genes enriquecidos na comparação. Além disso, vias de reparo do DNA no ciclo mitótico também foram enriquecidas e mostram-se, na literatura, como um desafio para a tecnologia de reparo do material genético.

Palavras chave: Bioinformática, Transcriptoma, RNA-Seq, Fibrose Cística

1. ABSTRACT

Cystic Fibrosis is an autosomal recessive disease that affects around 70,000 people per year worldwide and 1 in every 7,576 natives in Brazil. The mutation alters the gene responsible for coding the protein responsible for chlorine conductance in epithelial cells and there are differences in the type of mutation that occurs in the gene, being characterized in changes that can range from the synthesis of incomplete or truncated proteins to non-expression of the protein. The clinical picture of patients with Cystic Fibrosis includes excessive production of secretions in the airways and episodes of exacerbation, which include coughing and dyspnea. The common association between opportunistic viruses and Cystic Fibrosis led to the study carrying out transcriptomes of primary bronchial epithelial cells in pediatric patients with the disease and whether or not it was associated with Rhinovirus.

The objective of the project was to reanalyze the transcriptomes obtained by the article "RNA sequencing of primary bronchial airway epithelial cells from young children with and without CF, including those with and without rhinovirus infection in vitro" and also to conduct complex network analyses. The reanalysis was carried out using programming languages such as R and Python and use of databases such as BioGRID and Uniprot, obtaining a comparison with the gene count values of the established groups: Control - Patients without Cystic Fibrosis and without Rhinovirus (A) ; Patients without Cystic Fibrosis and with Rhinovirus (B); Patients with Cystic Fibrosis and without Rhinovirus (C); Patients with Cystic Fibrosis and Rhinovirus (D).

The results revealed enriched pathways such as cell death, regulation of protein metabolism and RNA biosynthesis when the Rhinovirus is present in the samples. It is worth noting that the reanalysis obtained an enriched pathway for positive

regulation of protein signaling, not present directly in literature involving the presence of Rhinovirus in patients with the disease and in samples from patients with respiratory diseases, which can serve as an aid in future research involving the genes enriched in the comparison. Furthermore, DNA repair pathways in the mitotic cycle were also enriched and are shown in the literature to be a challenge for genetic material repair technology.

Keywords: Bioinformatics, Transcriptome, RNA-Seq, Cystic Fibrosis

2. INTRODUÇÃO

A Fibrose Cística (FC) é caracterizada como uma doença genética autossômica recessiva que afeta cerca de 70.000 pessoas por ano no mundo (CUTTING, 2015). No Brasil a incidência é estimada em 1 a cada 7.576 nativos com diferenças regionais, ademais a estimativa de vida é de 43,8 anos (VENDRUSCULO; FAGUNDES DONADIO; ARAÚJO PINTO, 2021). Cabe salientar que apesar da cobertura para triagem neonatal da Fibrose Cística ter melhorado o diagnóstico precoce de pacientes, a estimativa de vida mediana de idade de óbito ainda é de 18,4 anos (VENDRUSCULO; FAGUNDES DONADIO; ARAÚJO PINTO, 2021).

A mutação causadora da doença, de herança autossômica recessiva (BUENO; ROSA; TREVIZANI, [s.d.]), ocorre no cromossomo responsável pela codificação da proteína Reguladora de Condutância Transmembrana em Fibrose Cística (CFTR) (GIBSON; BURNS; RAMSEY, 2003), responsável pela passagem do cloro no núcleo de células epiteliais (BUENO; ROSA; TREVIZANI, [s.d.]). A mutação de CFTR culmina em alterações que podem variar desde a síntese de proteínas incompletas ou truncadas até a não expressão proteica (GIBSON; BURNS; RAMSEY, 2003).

As possíveis mutações no gene CFTR podem ser classificadas em 6 classes, assumindo como base o efeito funcional causado na célula pela alteração de CFTR:

- Classe I: Mutações que resultam na terminação de sinais de transcrição que resultam em instável, truncada ou nula expressão de proteína (GIBSON; BURNS; RAMSEY, 2003);
- Classe II: Mutações $\Delta 508$, que afetam o processamento da proteína, impedindo-a de exercer sua função e levando a sua degradação prematura (GIBSON; BURNS; RAMSEY, 2003);

- Classe III: Mutações localizadas nos locais dois locais de domínio de ligação, acarretando em uma menor atividade do canal de cloro (GIBSON; BURNS; RAMSEY, 2003);
- Classe IV: Mutações localizadas primariamente nos domínios que abrangem a membrana, resultando em uma menor condutância de cloro pelo canal (GIBSON; BURNS; RAMSEY, 2003);
- Classe V: Mutações que caracterizam uma menor quantidade de proteínas funcionais expressas causadas pelo *splicing* anormal (GIBSON; BURNS; RAMSEY, 2003);
- Classe VI: Mutações que resultam na expressão de proteínas instáveis na membrana apical da célula, sendo prontamente removidas e degradadas (BUENO; ROSA; TREVIZANI, [s.d.]).

O quadro clínico em pacientes pode ser identificado pela liberação exacerbada de secreções nas vias aéreas e também no ducto pancreático, acarretando em obstruções e inflamações nas regiões citadas anteriormente (CUTTING, 2015). Outros órgãos também podem ser afetados pela doença, tais como glândulas sudoríparas, ducto biliar, intestino e o trato reprodutor masculino (CUTTING, 2015).

Episódios de Exacerbação Pulmonar, definido como crises de tosse, dispneia e diminuição da função pulmonar, são observados em pacientes infantis, causando perda de função irreversível dos pulmões (LING et al., 2020). Ademais, a relação entre episódios de Exacerbação e infecções por vírus oportunistas, mesmo não compreendida completamente, já foi relatada (ASNER et al., 2012). Segundo pesquisa feita com 280 pacientes com CF que passaram por ocorrências de Exacerbação, foi possível identificar que, em mais de 85% das amostras

processadas, foi detectada a presença de vírus como o Rinovírus (ASNER et al., 2012).

O Rinovírus pertence a família dos *Picornaviridae* e é um dos maiores causadores de resfriado comum. Em doenças pulmonares, como o caso da Fibrose Cística, o patógeno mencionado é responsável por acentuar complicações respiratórias em pacientes doentes. A resposta imune exacerbada em pacientes com doenças pulmonares intensifica a resposta anti-inflamatória e anti-viral em infecções por Rinovírus, podendo levar a uma inflamação persistente devido ao recrutamento constante de células de defesa, como macrófagos e neutrófilos (GANJIAN et al., 2020).

Compreendido como um patógeno oportunista, o Rinovírus é o vírus mais detectado nas vias aéreas de adultos e crianças com Fibrose Cística e seu impacto clínico pode variar desde a hospitalização até a redução da capacidade funcional dos pulmões (LING et al., 2020).

Dada a relevância do Rinovírus em Pacientes com Fibrose Cística, o artigo *Rhinovirus Infection Drives Complex Host Airway Molecular Responses in Children With Cystic Fibrosis* (LING et al., 2020) recrutou 24 pacientes infantis. Dentre eles, é possível citar criança sem Fibrose Cística que iriam passar por procedimentos não relacionados com a via aérea e crianças com Fibrose Cística através de avaliações de visitas anuais. Os grupos foram divididos em:

- Grupo 1: 5 pacientes sem Fibrose Cística controles e 7 pacientes com Fibrose Cística.
- Grupo 2: 6 pacientes sem Fibrose Cística controles e 6 pacientes com Fibrose Cística.

O grupo 1 foi selecionado para a realização de sequenciamento de RNA e o grupo 2 foi criado com o intuito de validar os experimentos (LING et al., 2020). As amostras foram coletadas da mucosa da traqueia dos pacientes utilizando uma escova de citologia e tratadas em um meio de cultura próprio para posterior contaminação pelo Rinovírus RV1b. O RNA obtido foi extraído utilizando um kit *PureLink RNA* para posterior análise de sequenciamento (LING et al., 2020). Foram então realizados sequenciamentos de RNA (RNA-seq) utilizando a plataforma da Illumina HiSeq 2500. Os arquivos FASTQ, obtidos do RNA-seq foram tratados e foi construído uma tabela contendo o número de contagens dos genes para posterior análise (LING et al., 2020).

Cabe salientar que a expressão diferencial também foi realizada com a finalidade de testar a hipótese inicial proposta pelo estudo, a qual o vírus afetaria a resposta antiviral do organismo (LING et al., 2020).

De modo geral, o estudo tentou fornecer uma compreensão mais detalhada sobre a resposta analisada nos transcriptomas das células do epitélio brônquico primário em casos de CF com associação a Rinovírus, identificando possíveis alvos terapêuticos que podem reduzir o impacto do vírus nos organismos (LING et al., 2020).

3. OBJETIVO

O objetivo do presente trabalho consiste na produção de reanálise do artigo “RNA sequencing of primary bronchial Airwar epitelial cells from young children with and without CF, including those with and without rhinovirus infection in vitro” utilizando os transcriptomas fornecidos pelo estudo para criar redes regulatórias com a finalidade de comparar as amostras dos pacientes participantes do estudo. Ressalta-se que a análise de redes não foi realizada no artigo escolhido para esse estudo.

4. METODOLOGIA

O estudo foi iniciado com a pesquisa de artigos para revisão no banco de dados *Gene Expression Omnibus* (GEO), que contém artigos com todos os materiais utilizados pelo artigo armazenados no site, facilitando a busca e uso dos dados para pesquisas na área de Bioinformática. Após a pesquisa, o artigo “RNA sequencing of primary bronchial airway epithelial cells from young children with and without CF, including those with and without rhinovirus infection in vitro” foi escolhido para reanálise, após revisão da bibliografia básica, materiais fornecidos pelo estudo.

Após ter escolhido o tema a ser trabalhado, foi realizada a instalação do sistema operacional Ubuntu baseado em Linux, instalação das linguagens de programação Python e R em conjunto da interface PyCharm Community e R Studio, respectivamente e instalação do programa Cytoscape, o qual auxiliou na visualização da rede de interações no processo de reanálise dos dados de RNA-Seq fornecidos pelo artigo selecionado.

Em seguida foi realizada uma pesquisa por repositórios de interação de proteínas e genes e, o repositório escolhido foi o *Biological General Repository for Interaction Datasets* (BioGRID). Com a escolha foi possível construir uma rede de interações entre proteínas. Além disso, o banco de dados utilizado para converter a interação entre proteínas e genes em uma linguagem que pudesse ser usada posteriormente foi *Universal Protein Resource* (UniProt). Dessa forma, um *script in house* utilizando o Python foi criado com o objetivo de filtrar o repositório BioGRID para criar uma rede de interações, apenas, entre genes de humanos, também realizando a troca da nomenclatura utilizada pelo BioGRID para a nomenclatura

utilizada no UniProt, com o objetivo final de facilitar a criação de posteriores *scripts* em R, Python e no Cytoscape.

Para efetuar o teste de validação necessário após a construção de uma rede de interações, um *script in house* em R foi criado, utilizando o pacote *igraph*, com a finalidade de criar um gráfico de *Degree Distribution*.

A variável Degree Distribution é representada como $P(k)$ e pode ser definida como a probabilidade de que um nó selecionado ao acaso terá exatamente “k” arestas ou ligações com outros nós (HEIN et al., 2013). Além disso, o uso da variável contribui para comprovar que a rede está apta para o uso por seguir o padrão de distribuição Poisson, o qual indica que o maior número dos nós são caracterizados, em sua maioria, pela mesma quantidade de arestas e a chance de encontrar um nó com um número alto de arestas é baixa (HEIN et al., 2013).

É possível observar que no gráfico de Degree Distribution (FIGURA 1) a rede de interações filtrada para seres humanos utilizando o banco de dados BioGRID segue de fato a distribuição de Poisson, comportando baixos número de nós com muitas conexões, sendo possível concluir que a rede possui uma estabilidade alta a alterações randômicas.

Como próximo passo, a contagem de genes realizada por Ling et al. (2020) foi filtrada e organizada para formar 4 grupos dos 24 transcriptomas de células do tecido brônquico primário fornecidos pelo artigo, esses foram definidos em:

- Controle - Pacientes sem Fibrose Cística e sem Rinovírus (A);
- Pacientes sem Fibrose Cística e com Rinovírus (B);
- Pacientes com Fibrose Cística e sem Rinovírus (C);
- Pacientes com Fibrose Cística e com Rinovírus (D).

Após organizado, o arquivo de contagem de genes passou por uma alteração na nomenclatura dos genes, trocando de EsemblID para o vocabulário do UniProt, com o objetivo de manter a mesma terminologia durante a reanálise. Com a separação dos grupos completa, a inclusão da contagem dos genes para cada grupo foi feita; ou seja, cada grupo foi separado por sua quantidade de amostras, respectivas contagens para cada gene e integrado à rede de interações criada anteriormente, obtendo-se uma rede com 108.919 interações entre genes ao final do código.

Com a separação dos grupos, foi realizada uma média de cada gene dentro dos grupos separados, ou seja, uma média das contagens de cada gene foi calculada com base em todos os valores de contagens dos genes das amostras daquele grupo, para exemplificar é possível utilizar o gene Q9NSG2 do grupo A, que possuía os valores de contagem: 24, 26, 18, 47 e 34. Calculando a média de contagens do gene para esse grupo obtemos o valor 29,8. Para elucidar de forma mais clara é possível utilizar como exemplo uma interação do grupo A como a interação dos genes P49795 e P19086 que receberam as médias 50 e 1, respectivamente. Em outras palavras, é possível afirmar que o gene P49795 obteve a média 50 após o cálculo da média de todos os valores para aquele gene no grupo de Pacientes Sem Fibrose Cística e sem Rinovírus (grupo A). O procedimento foi repetido para todos os outros grupos e suas respectivas interações, obtendo uma tabela de interações e a média dos pesos de cada interação separadas nos grupos A, B, C e D.

Após o cálculo das médias de contagens cada gene, ou seja, pesos das arestas, na interação, foi calculado o valor do peso das arestas de cada interação nos grupos. Seguindo o exemplo anterior, o cálculo do peso da aresta da interação

entre os genes P49795 e P19086 seria dada por **50 x 1**, logo o valor do peso da aresta entre essa interação seria de **50**. Posteriormente foi feito o cálculo de log 2 *fold change* (Log2FC, **log 2(X – Y) / Y**), definido pelo valor quantitativo de mudança entre um valor inicial e um valor final, de cada interação nos grupos, comparando-os na seguinte ordem:

- 1° Comparação: Grupo B comparado com o grupo A;
- 2° Comparação: Grupo C comparado com o grupo A;
- 3° Comparação: Grupo D comparado com o grupo A;
- 4° Comparação: Grupo D comparado com o grupo C.

As interações foram então filtradas utilizando um valor da média + dois desvios padrão dos resultados obtidos pelo cálculo do valor de Log2FC conforme é exibido no gráfico (FIGURA 2). Para exemplificar o cálculo de Log2FC é possível retomar os mesmos genes citados no grupo A e utilizar a primeira comparação, que consiste no cálculo comparativo entre a aresta da interação entre os genes P49795 e P19086 grupo B e do grupo A:

$$\text{Log2FC} = \text{Log2} (\text{Aresta de B} - \text{Aresta de A}) / \text{Aresta de A}$$

Os valores foram atribuídos no seguinte modo:

- 1° Comparação: Log2FC ≥ 1,0
- 2° Comparação: Log2FC ≥ 1,34
- 3° Comparação: Log2FC ≥ 1,5
- 4° Comparação: Log2FC ≥ 1,4

Após a obtenção dos valores de Log2FC, a filtragem dos valores e a criação dos grupos que serão comparados, as tabelas de comparação foram submetidas à

análise no programa Cytoscape, em conjunto da extensão SAFE, com o intuito de realizar o enriquecimento por Ontologias Gênicas de processos biológicos e a transformação das comparações em gráficos, fornecendo uma observação detalhada das expressões diferenciais entre os grupos comparados.

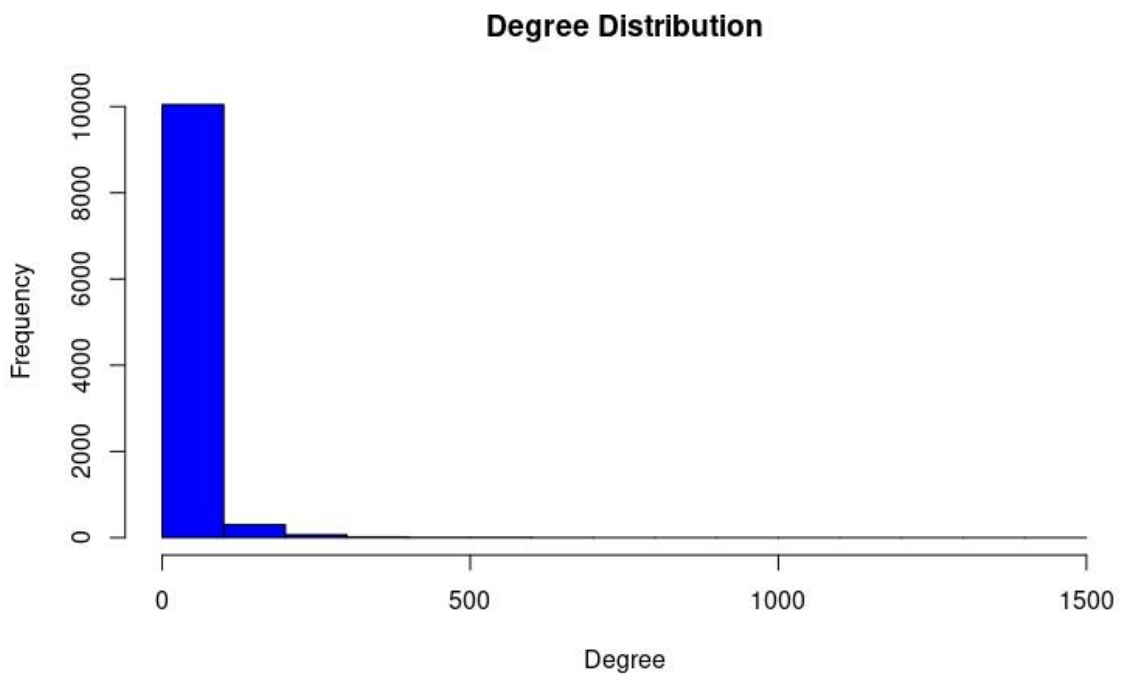


FIGURA 1: Gráfico de Degree Distribution criado no R.

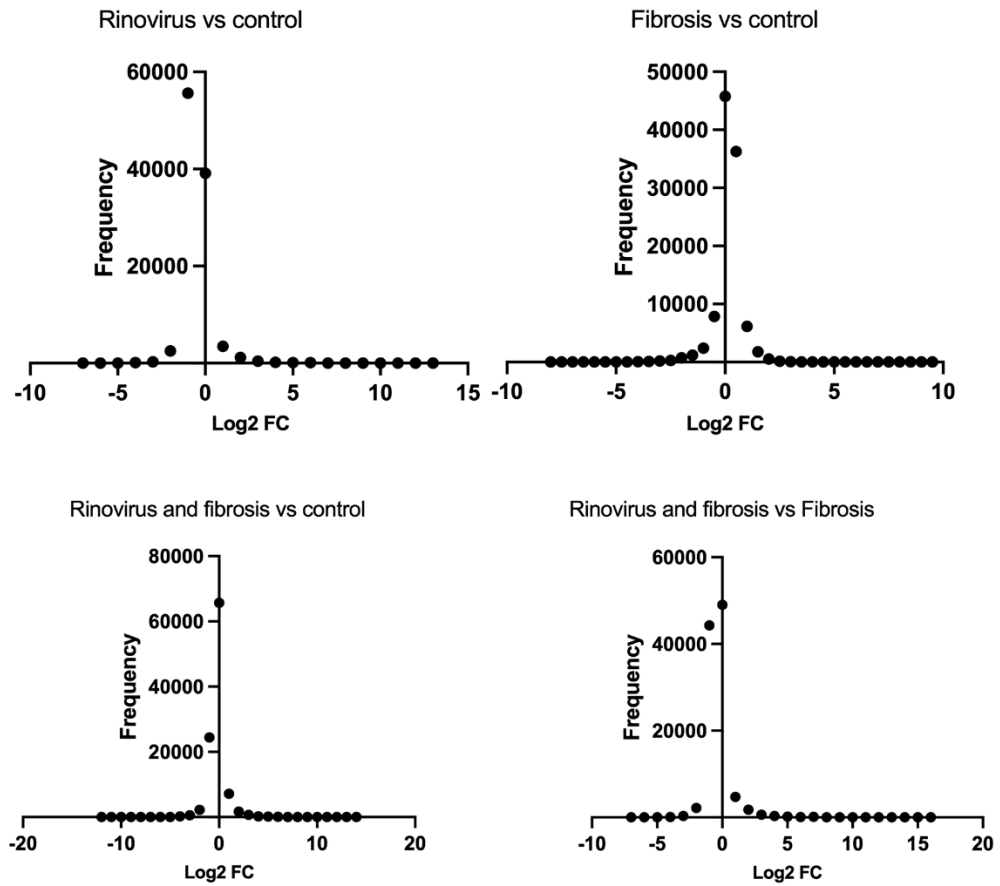


FIGURA 2: Valores normalizados, da comparação entre os grupos, utilizando um valor da média + dois desvios padrão dos resultados obtidos pelo cálculo do valor de Log2FC para filtragem das conexões.

5. RESULTADOS

Os resultados foram criados utilizando a ferramenta SAFE alocada no aplicativo Cytoscape, o qual é responsável pela análise de enriquecimento da comparação entre as redes de interações e possui o objetivo de indicar vias que estão em destaque na comparação entre os grupos. A primeira comparação é definida entre o grupo de Pacientes sem Fibrose Cística e com Rinovírus (B) e o grupo de Pacientes sem Fibrose Cística e sem Rinovírus (A). No mapa de enriquecimento por Ontologia Gênica de processos biológicos (FIGURA 3), há um destaque das seguintes vias: Via de regulação positiva da sinalização de proteínas; Via de regulação metabólica biosintética de RNA; Via de regulação da morte celular e Via de regulação da reposta peptídica celular.

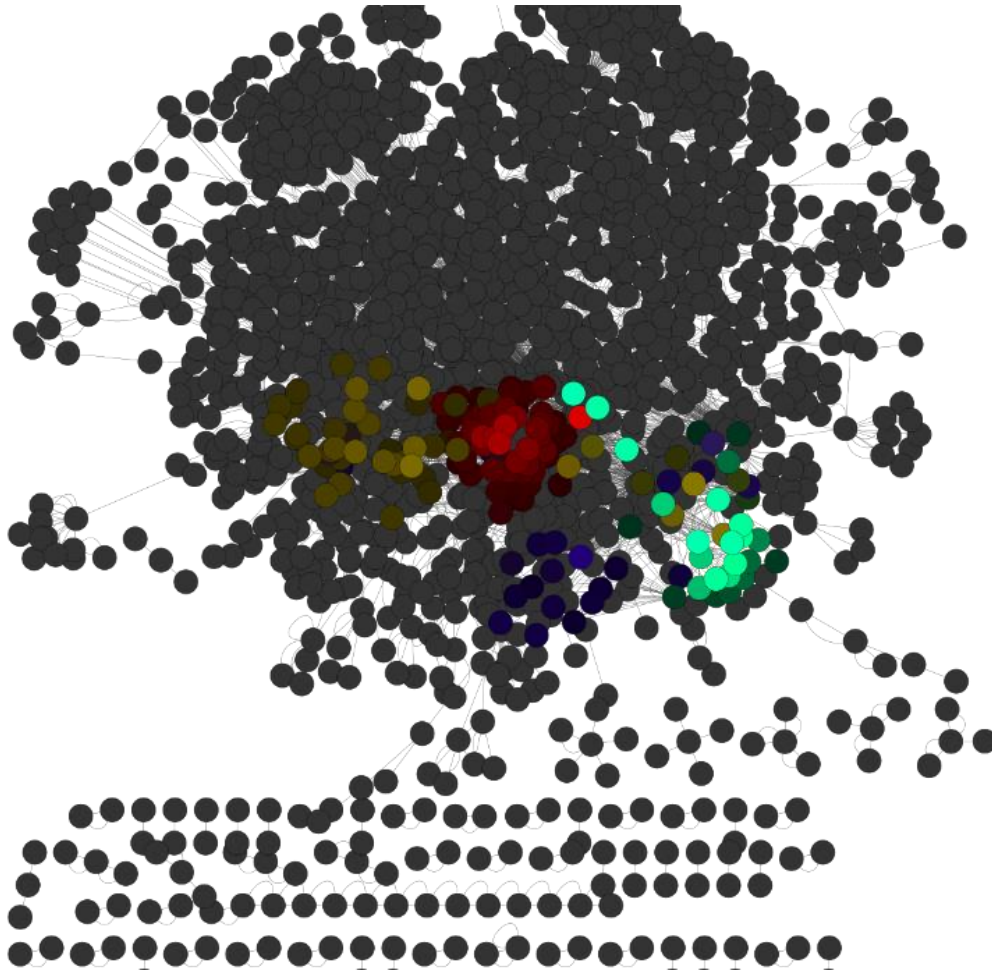


FIGURA 3: As principais vias enriquecidas estão em destaque. A cor verde representa a via de regulação positiva da sinalização de proteínas; A cor vermelha representa a via de regulação metabólica biosintética de RNA; a cor amarela representa a Via de regulação da morte celular; a cor roxa representa a Via de regulação da reposta peptídica celular.

A seguir foi feita a segunda comparação, que foi determinada entre o grupo de Pacientes com Fibrose Cística e sem Rinovírus (C) e o grupo de Pacientes sem Fibrose Cística e sem Rinovírus (A). No mapa de enriquecimento da comparação (FIGURA 4) há destaque para as vias: Via de replicação do DNA no ciclo mitótico e Via de resposta no reparo da quebra do DNA.

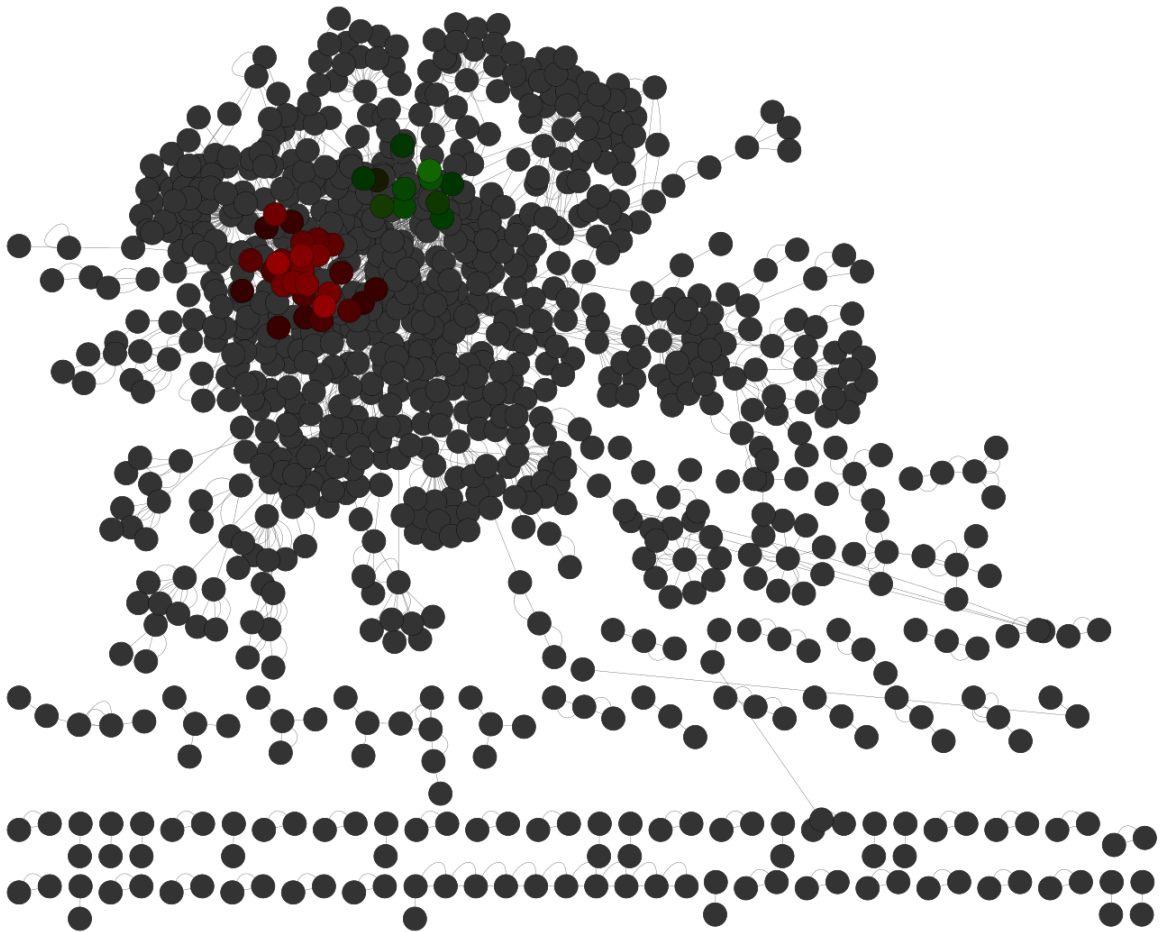


FIGURA 4: As principais vias enriquecidas estão em destaque. A cor vermelha representa a Via de replicação do DNA no ciclo mitótico; A cor verde representa a Via de resposta no reparo da quebra do DNA.

Já a terceira comparação, que foi indicada entre o Pacientes com Fibrose Cística e com Rinovírus (D) e o grupo de Pacientes sem Fibrose Cística e sem Rinovírus (A). No mapa de enriquecimento da comparação (FIGURA 5) há destaque para as seguintes vias: Via de regulação catabólica de sinalização em proteínas; Via de regulação metabólica Biosintética de RNA celular e Via de regulação sinalizadora de morte celular I – Kappa B.

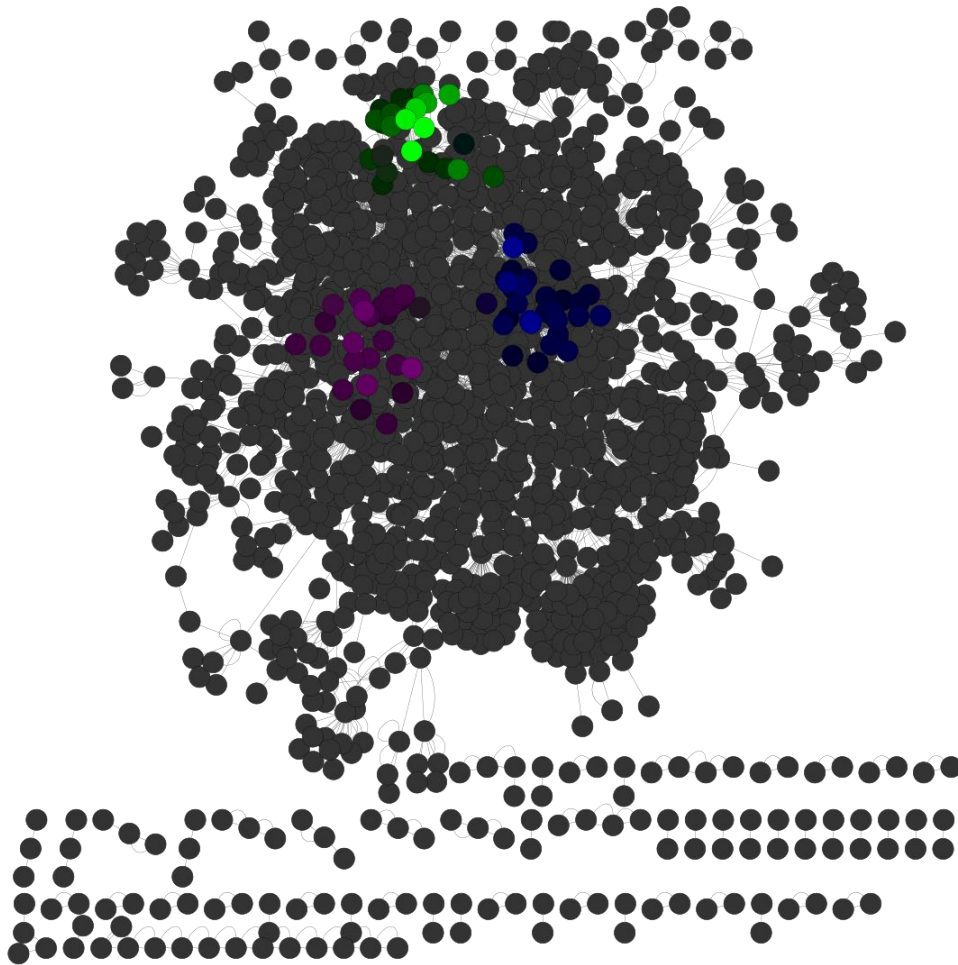


FIGURA 5: As principais vias enriquecidas estão em destaque. A cor verde representa a Via de regulação catabólica de sinalização em proteínas; a cor azul representa a Via d regulação metabólica Biosintética de RNA celular e a cor rosa representa a Via de regulação sinalizadora de morte celular I – Kappa B.

A quarta e última comparação, entre o grupo de Pacientes com Fibrose Cística e com Rinovírus (D) e o grupo de Pacientes com Fibrose Cística e sem Rinovírus (C). Observando o mapa de enriquecimento da comparação (FIGURA 6), é possível apreciar as seguintes vias em destaque: Via de regulação da sinalização de proteínas na célula; Via de regulação metabólica biosintética de RNA e Via de regulação de morte celular.

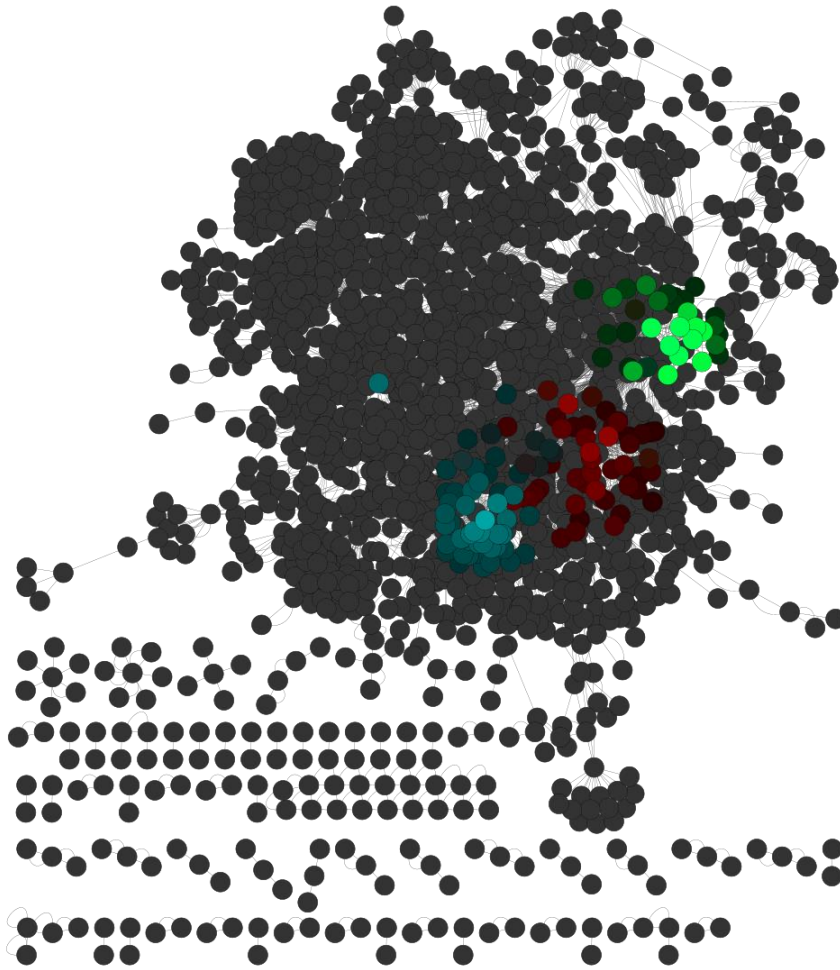


FIGURA 6: As principais vias enriquecidas estão em destaque. A cor verde representa a Via de regulação da sinalização de proteínas na célula; A cor azul representa a Via de regulação metabólica biosintética de RNA e a cor vermelha representa a Via de regulação de morte celular.

6. DISCUSSÃO

Os resultados de enriquecimentos das comparações são relevantes para ressaltar vias de regulação que estão em realce. Como primeiro destaque no enriquecimento, é possível citar a via de morte celular ou apoptose, que está presente na 1º comparação de Pacientes sem Fibrose Cística e com Rinovírus (B) e de Pacientes sem Fibrose Cística e sem Rinovírus (grupo A ou Controle), na 3º comparação de Pacientes com Fibrose Cística (D) e com Rinovírus e o grupo (A) de pacientes controle e, também na 4º comparação, composta por Pacientes com Fibrose Cística e com Rinovírus (D) e de Pacientes com Fibrose Cística e sem Rinovírus (C). É possível pontuar que a presença do Rinovírus nas comparações, tanto entre pacientes que só possuíam o vírus, quanto entre pacientes que possuíam Fibrose Cística, enriqueceram a via de morte celular. Na literatura é possível encontrar resultados parecidos quando amostras de pacientes com doenças respiratórias e sem a doença são expostos ao patógeno, enriquecendo as vias regulação negativa de morte celular ou apoptose (BOCHKOV et al., 2010).

O enriquecimento de vias como a de regulação metabólica biosintética, de metabolismo (regulação catabólica e sinalização) de proteínas encontrados nas análises de enriquecimento das comparações 1, 2 e 4 também se afirmam na literatura com a análise de enriquecimento de Ontologia Gênica em amostras de pacientes com Fibrose Cística e sem Fibrose Cística e seu comportamento quando expostos ao vírus, obtendo resultados com genes responsáveis pelas mesmas vias (LING et al., 2020). A via de regulação da resposta peptídica celular encontrada nos resultados de análise de enriquecimento da 1º comparação também pode ser parcialmente na literatura com o estudo de Ling et al. (2020) que obteve

enriquecimento nas vias relacionadas a metabolismo de proteínas. Cabe salientar que a 1ª comparação obteve uma via enriquecida de regulação positiva da sinalização de proteínas, sugerindo diferença se comparados com as literaturas envolvendo células do tecido brônquico primário, as quais possuem uma maior ligação com sistema de produção de citocinas (IFN beta) induzido pelo Rinovírus em amostras com Fibrose Cística e sem Fibrose Cística (LING et al., 2020) ou expõem que a maior parte (90%) dos achados na expressão diferencial dos genes encontrados são de expressão reduzida em amostras de células de pacientes com doenças respiratórias como a asma (BOCHKOV et al., 2010).

Já 2ª comparação, composta por Pacientes com Fibrose Cística e sem Rinovírus (C) e por Pacientes do grupo A (controle), possui as seguintes vias enriquecidas: via de replicação do DNA no ciclo mitótico e de resposta no reparo da quebra do DNA. As vias enriquecidas encontradas nos resultados da comparação corroboram para uma possível tentativa de reparo do material genético mutado da célula em seu ciclo celular. O tema complexo é abordado pelo *Cystic Fibrosis Foundation* (“Gene Editing for Cystic Fibrosis | Cystic Fibrosis Foundation”, [s.d.]) que descreve a tentativa de edição genética como uma solução para o reparo não sucedido da célula na doença, contudo a tentativa de reparo ainda contém muitos erros e se mostra muito seletiva para aplicação em pacientes.

7. CONCLUSÃO

O presente trabalho abordou a doença Fibrose Cística e sua associação com o Rinovírus no tecido epitelial brônquico primário. A reanálise do artigo “RNA sequencing of primary bronchial Airway epithelial cells from young children with and without CF, including those with and without rhinovirus infection in vitro” buscou como objetivo, utilizar transcriptomas de pacientes infantis obtidos pelo estudo original para efetuar uma análise de enriquecimento dos dados, obtendo novas comparações e vias em evidência.

Para melhores resultados, os dados utilizados foram filtrados através de *scripts in house* em variadas linguagens de programação como Python e R, utilizando, no processo, banco de dados confiáveis com o objetivo final de obter vias em destaque na comparação dos grupos criados, sendo esses: Controle - Pacientes sem Fibrose Cística e sem Rinovírus (A); Pacientes sem Fibrose Cística e com Rinovírus (B); Pacientes com Fibrose Cística e sem Rinovírus (C); Pacientes com Fibrose Cística e com Rinovírus (D). Os grupos passaram por uma normalização das médias de cada gene em cada grupo, obtendo assim o peso das arestas com o cálculo de $\text{Log}_2 \text{Fold Change}$. Esse valor foi aplicado ao programa Cytoscape e os grupos foram separados em 4 comparativos para observar melhor as vias enriquecidas entre as comparações. No geral, as vias de morte celular e sinalização de proteínas foram as mais observadas na construção do enriquecimento por Ontologias Gênicas e também observadas na literatura com estudos de análise de enriquecimento com amostras de pacientes com Fibrose Cística e sem a doença e sua exposição ao Rinovírus e também de doenças do trato respiratório e sua interação com o patógeno. É possível ressaltar que o trabalho trouxe uma possível

nova abordagem com a via de regulação positiva de sinalização de proteínas, criando a possibilidade de novos estudos que envolvem os genes enriquecidos por essa via, já que a literatura atual não possui uma abordagem direta e especializada sobre essa via destacada. Há que se comentar sobre as vias de reparo do DNA evidenciadas pela análise de enriquecimento na segunda via e a possibilidade de futuras pesquisas envolvendo os genes enriquecidos, já que as soluções de reparo do material genético ainda estão em pesquisa pela alta seletividade do processo de edição nas células com o gene mutado. Os dados se mostraram relevantes no que tange Fibrose Cística e associação com rinovírus e podem servir como base para futuras pesquisas nessa interação entre a doença e o patógeno.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ASNER, S. et al. Role of respiratory viruses in pulmonary exacerbations in children with cystic fibrosis. **Journal of Cystic Fibrosis**, v. 11, n. 5, p. 433–439, 1 set. 2012.

BOCHKOV, Y. et al. Rhinovirus-induced modulation of gene expression in bronchial epithelial cells from subjects with asthma. **Mucosal immunology**, v. 3, n. 1, p. 69–80, jan. 2010.

BUENO, E. C.; ROSA, G. C.; TREVIZANI, L. FIBROSE CÍSTICA: ASPECTOS GENÉTICOS E PRINCIPAIS CRITÉRIOS DIAGNÓSTICOS. [s.d.].

CUTTING, G. R. Cystic fibrosis genetics: from molecular understanding to clinical application. **Nature Reviews Genetics**, v. 16, n. 1, p. 45–56, jan. 2015.

GANJIAN, H. et al. Rhinovirus and Innate Immune Function of Airway Epithelium. **Frontiers in Cellular and Infection Microbiology**, v. 10, 2020.

Gene Editing for Cystic Fibrosis | Cystic Fibrosis Foundation. Disponível em: <<https://www.cff.org/research-clinical-trials/gene-editing-cystic-fibrosis>>. Acesso em: 22 nov. 2023.

GIBSON, R. L.; BURNS, J. L.; RAMSEY, B. W. Pathophysiology and Management of Pulmonary Infections in Cystic Fibrosis. **American Journal of Respiratory and Critical Care Medicine**, v. 168, n. 8, p. 918–951, 15 out. 2003.

HEIN, M. Y. et al. Proteomic Analysis of Cellular Systems. Em: **Handbook of Systems Biology: Concepts and Insights**. [s.l.] Academic Press, 2013. p. 3–25.

LING, K.-M. et al. Rhinovirus Infection Drives Complex Host Airway Molecular Responses in Children With Cystic Fibrosis. **Frontiers in Immunology**, v. 11, 2020.

VENDRUSCULO, F. M.; FAGUNDES DONADIO, M. V.; ARAÚJO PINTO, L. Cystic fibrosis in Brazil: achievements in survival. **Jornal Brasileiro de Pneumologia**, p. e20210140, 30 abr. 2021.