

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE FARMÁCIA
TRABALHO DE CONCLUSÃO DE CURSO DE FARMÁCIA

SARSCOVIDB - Banco de dados de infecção pelo SARS-CoV-2: uma nova
plataforma para analisar o impacto molecular da infecção pelo vírus da COVID-19

TUNG SHENG YANG

Porto Alegre

2020

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE FARMÁCIA
TRABALHO DE CONCLUSÃO DE CURSO DE FARMÁCIA

SARSCOVIDB - Banco de dados de infecção pelo SARS-CoV-2: uma nova
plataforma para analisar o impacto molecular da infecção pelo vírus da COVID-19

TUNG SHENG YANG

Prof. Dr. WALTER ORLANDO BEYS DA SILVA
ORIENTADOR

MSc. RAFAEL LOPES DA ROSA
CO-ORIENTADOR

Porto Alegre
2020

RESUMO

A pandemia da COVID-19 provocada pelo novo coronavírus (SARS-COV-2) se tornou uma questão de emergência global para saúde pública devido a rápida propagação e elevado número de óbitos. Esta situação de urgência global levou a uma aceleração na pesquisa relacionada e, conseqüentemente, a um volume sem precedente de dados clínicos e experimentais, nos quais se incluem alterações de expressão gênica resultantes da infecção. O Banco de Dados da Infecção por SARS-COV-2 (SARSCOVIDB – <https://sarscovidb.org/>) foi criado para mitigar as dificuldades relacionadas a este cenário. O SARSCOVIDB é uma plataforma *online* que objetiva integrar todos os dados de expressão gênica diferencial, a nível de RNA mensageiro e proteína, acelerando a pesquisa sobre o impacto molecular da COVID-19. O banco de dados pode ser consultado a partir de diferentes perspectivas experimentais como as diferentes cepas virais utilizadas, hospedeiros, abordagens metodológicas (proteômica ou transcriptômica), genes/proteínas, tipo de amostra (clínica ou experimental). Todas estas informações foram retiradas de todos os 30 artigos de análise de expressão gênica diferencial relacionada cerca de 7000 artigos publicados até o momento e disponibilizados nas duas principais plataformas de busca, o PubMed e o Web of Science. O banco de dados apresenta 10534 genes identificados cuja expressão foi alterada devido a infecção por SARS-COV-2. Assim, o SARSCOVIDB é uma nova ferramenta para apoiar a comunidade científica no entendimento da patogênese e impacto molecular causado pelo SARS-COV-2.

Palavras chaves: SARS-COV-2. COVID-19. Banco de dados. Transcriptômica. Proteômica.

ABSTRACT

The COVID-19 pandemic caused by the new coronavirus (SARS-COV-2) has become a global emergency issue for public health due to the rapid spread and high number of deaths. This global emergency led to an acceleration in related research and, consequently, to an unprecedented volume of clinical and experimental data, which includes changes in gene expression resulting from infection. The SARS-COV-2 Infection Database (SARSCOVIDB - <https://sarscovidb.org/>) was created to mitigate the difficulties related to this scenario. SARSCOVIDB is an online platform that aims to integrate all differential gene expression data, at the level of messenger RNA and protein, accelerating research on the molecular impact of COVID-19. The database can be consulted from different experimental perspectives such as the different viral strains used, hosts, methodological approaches (proteomics or transcriptomics), genes / proteins, type of sample (clinical or experimental). All this information was taken from all 30 articles of analysis of related differential gene expression 7000 published so far and made available on the two main search platforms, PubMed and Web of Science. The database features 10534 identified genes whose expression has been altered due to SARS-COV-2 infection. Thus, SARSCOVIDB is a new tool to support the scientific community in understanding the pathogenesis and molecular impact caused by SARS-COV-2.

Keywords: SARS-COV-2. COVID-19. Database. Transcriptomics. Proteomics.

“A ciência não é apenas uma disciplina de razão, mas também de romance e paixão”

Stephen Hawking

SUMÁRIO

1	INTRODUÇÃO	7
2	OBJETIVOS	9
2.1	Objetivos gerais	9
2.2	Objetivos específicos.....	9
3	Metodologia.....	10
3.1	Seleção dos artigos.....	10
3.2	Dados de expressão gênica e coleta e análise de informações relacionadas	11
3.3	Tabela 1: Informações gerais depositadas no SARSCOVIDB	11
3.4	Construção de página da web, upload de dados e interface do usuário	12
4	Resultados e discussão	13
5	CONCLUSÕES.....	17
6	PERSPECTIVAS.....	18
	REFERÊNCIAS	19

1 INTRODUÇÃO

Em janeiro de 2020, foi reportado uma nova cepa de coronavírus associada a uma síndrome respiratória aguda grave (SARS-CoV-2) (World Health Organization, [2020](#)). Este novo coronavírus é um vírus de RNA fita simples, encapsulado, da família *Coronaviridae* e que apresenta alta virulência gerando impactos globais significativos (WRAPP, Daniel, [2020](#)). A epidemia causada pelo SARS-CoV-2 (COVID-19) teve como epicentro a cidade de Wuhan, China, e, em pouco tempo, tomou proporções mundiais se tornando um grave problema de saúde pública de interesse internacional (World Health Organization, [2020](#)). Atualmente está presente em mais de 200 países do mundo com 37 milhões de casos contabilizados (Covid-19 Dashboard, JHU CSSE, [2020](#)). No Brasil, registrou-se, até o momento, mais de 5 milhões de casos confirmados e 149.639 óbitos desde o início do surto segundo boletim do conselho Nacional de Secretários de Saúde (Conass, [2020](#)).

A crescente proliferação do SARS-CoV-2 e a falta de vacina disponível têm causado preocupação das autoridades públicas e de órgãos internacionais ligados a saúde, tendo como consequência a mobilização da comunidade científica na busca pelo entendimento da doença e no desenvolvimento de tratamentos eficientes para a COVID-19. Este panorama atual gerou mudanças significativas no campo da ciência como um aumento rápido e sem precedente na publicação de artigos científicos acerca de um tópico de pesquisa. Este cenário inclui artigos publicados de forma acelerada através de processos de submissão facilitados ou publicações de *pre-prints*, sem avaliação de pares, disseminando diversos artigos com resultados muito especulativos, como previsões computacionais (HORBACH, [2020](#)).

A compreensão dos aspectos moleculares associados a COVID-19 e a busca pelo entendimento da complexa resposta do hospedeiro após a infecção viral vêm ganhando espaço na pesquisa (BOJKOVA, Denisa et al, [2020](#); OUYANG, Yabo et al, [2020](#); SHEN, BO et al, [2020](#)). Abordagens de estudo que avaliem as infecções virais sob a perspectiva do impacto molecular no hospedeiro contribuem para o entendimento da doença associada, importante para traçar potenciais estratégias antivirais (YANG, Shiping et al, [2019](#)). Estas abordagens de estudo de infecções virais incluem principalmente a análise de genes diferencialmente expressos que permitem uma caracterização “*in silico*” dos desdobramentos moleculares da

resposta/impacto da infecção. Muitos bancos de dados são utilizados nestas pesquisas para o levantamento de dados de expressão, tais como Genbank, ViralZone. No entanto, muitos resultados são obtidos por mineração de texto contendo dados complexos (PRZYBYLA et al. [2016](#)) ou mesmo necessitando de bioinformatas na interpretação de dados brutos, como no caso do The Gene Expression Omnibus (GEO), que utiliza mecanismos avançados de análises (Clough, Emily; Barretth, Tanya, [2020](#)). Este cenário dificulta, por exemplo, o acesso da comunidade médica e cientistas não especializados que possam necessitar destas informações.

No recente surto de Zika vírus, nosso grupo de pesquisa desenvolveu um banco de dados, o *ZIKA Virus Infection Database* (ZIKAVID), que facilitou a análise da expressão diferencial associada à infecção por Zika (ROSA, Rafael L. et al, [2020](#)). Da mesma maneira, este trabalho tem por objetivo disponibilizar o Banco de Dados da Infecção de SARS-CoV-2 (SARSCOVIDB - <https://sarscovidb.org/>), um banco de dados livre de mineração de texto, desenvolvido e acurado manualmente e que contribui para a robustez e a precisão dos dados depositados.

2 OBJETIVOS

2.1 Objetivos gerais

O presente trabalho tem como objetivo criar um banco de dados público reunindo todos os dados moleculares de expressão gênica diferencial, resultante da infecção por SARS-CoV-2, publicados em artigos científicos de revistas indexadas internacionais.

2.2 Objetivos específicos

- Busca de dados de expressão gênica diferencial resultantes de estudos publicados associados a infecção por SARS-CoV-2.
- Compilação e organização de dados coletados.
- Desenvolvimento da plataforma *online*.

3 Metodologia

3.1 Seleção dos artigos

O SARSCOVIDB foi construído por etapas, conforme descrito na figura 1. Inicialmente, foi feita uma busca por mineração de texto de artigos usando termos ‘‘SARS-CoV-2’’ e ‘‘COVID-19’’ nos bancos de dados públicos PubMed, Web of Science, Google Scholar e Science Direct. Esta etapa da pesquisa recuperou cerca de **7000** resultados. Em seguida, através da verificação manual dos títulos e resumos, foram selecionados somente os artigos que continham análises de expressão gênica diferencial, após a infecção por SARS-CoV-2. Esta checagem resultou em 30 artigos selecionados.

Em seguida, os dados dos artigos selecionados foram coletados e organizados em planilhas, contendo as listas de genes identificados, com as proteínas correspondentes, tipos de estudos (*in vivo/ in vitro*, clínico), metodologias aplicadas (proteômica, transcriptômica, qRT-PCR, Immunoblotting, entre outros), cepas virais utilizadas, hospedeiros (amostras clínicas humanas, linhagens celulares, modelos animais), além de dados da expressão (*up* ou *down* regulados, de acordo com a análise realizada).

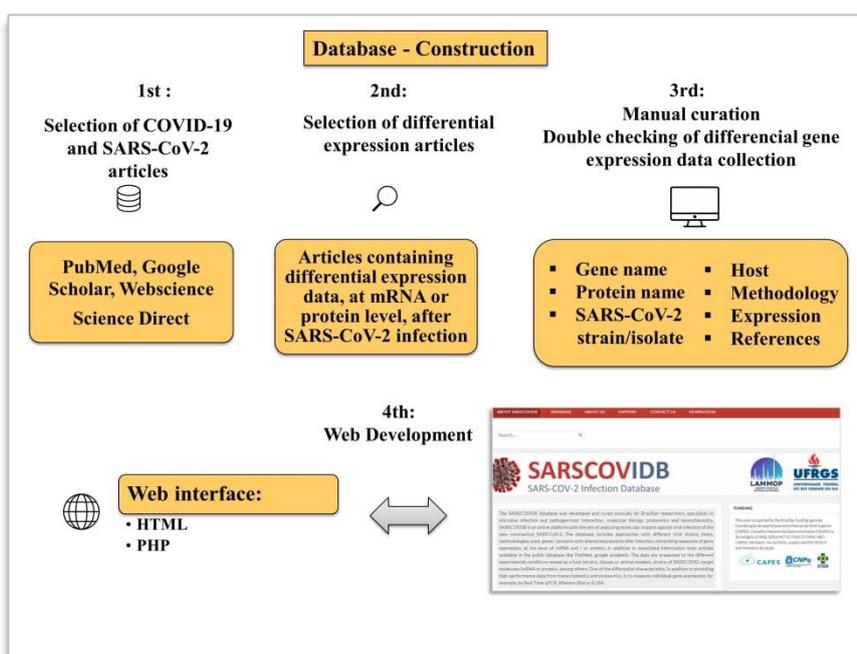


Figura 1: Etapas da construção do banco de dados SARSCOVIDB

3.2 Dados de expressão gênica e coleta e análise de informações relacionadas

Na terceira etapa (figura 1), os dados de expressão gênica diferencial e todas as informações relacionadas de cada artigo foram coletadas, verificadas e organizadas. Compreendem estas informações as listas de genes diferencialmente expressos (DEGs), à nível de RNA e/ou proteína, identificados após a infecção pelo SARS-CoV-2. As informações relacionadas ao gene coletadas foram: o nome do gene, o nome da proteína correspondente, o status da expressão (regulada positivamente ou regulada negativamente, de acordo com os critérios e análise estatística de cada artigo), informações relacionadas à abordagem experimental usada no estudo (se foram utilizadas amostras clínicas, infecção experimental em modelo animal ou linhagem celular), a cepa ou origem do isolado de SARS-CoV-2 usado na infecção, abordagem metodológica (transcriptômica, proteômica, qRT-PCR, entre outros) e a referência do artigo, conforme apresentado na Tabela 1. Se a expressão de qualquer gene foi avaliada com vários métodos, todas as medições foram adicionadas ao banco de dados. Se os dados de expressão gênica forem depositados em um repositório ou banco de dados específico, essa informação é reconhecida na entrada do gene específico no SARSCOVDB.

3.3 Tabela 1: Informações gerais depositadas no SARSCOVDB

SARSCOVDB ID	Identificação de referência da entrada no SARSCOVDB.
Nome do gene	Nomenclatura oficial do gene de cada entrada armazenada.
Nome da proteína	Nome oficial de cada proteína de cada entrada armazenada.
Hospedeiro/amostra	Modelo usado como hospedeiro para infecção por COVID-19 para gerar os dados: in vivo (tecido humano ou órgãos e modelos animais), e in vitro (humanos ou linhagens de células animais ou outras células).
Origem geográfica	Origem geográfica dos isolados virais de SARS-CoV-2
Metodologia	Método usado para medir a expressão gênica em nível de RNA e / ou proteína após a infecção por COVID-19.
Expressão	Informa se um determinado gene aumenta a expressão (Up) e / ou reduz a expressão (Down) em resposta à infecção pelo ZIKV.
Referência	Os PMIDs e o título do artigo usados como fonte de dados.
URL	Link direto para a referência usada como fonte.
Base de dados	Fornecer o banco de dados específico onde os dados são armazenados, como UNIPROT e NCBI.
URL	Link direto para todas as referências usadas como fonte de dados de expressão gênica para construir o banco de dados.

3.4 Construção de página da web, upload de dados e interface do usuário

Na última etapa, MySQL v5.0, PHP v.5.2.99 e HTML10 foram usados para construir o SARSCOVDB. As linguagens de programação PHP v.5.2.99 e HTML10 foram usadas para criar a interface gráfica do usuário (GUI). Os dados são armazenados e mantidos pelo servidor RDBMS, e a linguagem SQL foi usada para gerenciamento de dados. O usuário pode personalizar facilmente as pesquisas; por exemplo, é possível combinar resultados como ao pesquisar proteínas específicas em um determinado modelo animal ou celular ou selecionar um isolado de SARS-CoV-2 específico. Uma lista que contém todos os “X” genes das “X” entradas depositadas no SARSCOVDB é fornecida por um link de download direto (<https://sarscovdb.org/download/>). Além disso, o banco de dados tem links para instruir os usuários sobre como pesquisar proteínas ou genes específicos e também hospedeiros e/ou cepas de SARS-CoV-2 específicos.

4 Resultados e discussão

O acelerado surto de COVID-19 no mundo, somado a falta de tratamentos eficientes comprovados e vacinas regulamentadas, desencadeou grande esforço da comunidade científica e dos governos em pesquisas envolvendo o vírus SARS-CoV-2 e as potenciais comorbidades associadas em humanos (KAMBHAMPATI et al., [2020](#)).

O banco de dados SARSCOVDB reúne DEGs identificados em trabalhos publicados em revistas científicas internacionais indexadas, de plataformas *online* públicos referente a infecção por SARSCOVDB e a resposta do hospedeiro. Os artigos da literatura sobre o tópico COVID e SARS-CoV-2 foram inicialmente coletados por mineração de texto enquanto as buscas subsequentes para a coleta de dados de expressão gênica diferencial foram aferidas manualmente. Para melhorar a experiência da busca e a interface de uso, todos os dados são categorizados e organizados de acordo com recursos ligados as condições experimentais e resultados, tais como modelos de hospedeiros utilizados na infecção, cepas virais utilizadas, metodologia aplicada, status de expressão diferencial e artigo referência. Para o nosso conhecimento, o SARSCOVDB é o primeiro banco de dados que se dedica exclusivamente a reunir todos os genes/proteínas detectados diferencialmente expressos (DEGs), após a infecção pelo SARS-COV-2, sendo assim, uma ferramenta importante a ser explorada pela comunidade científica neste campo de investigação urgente.

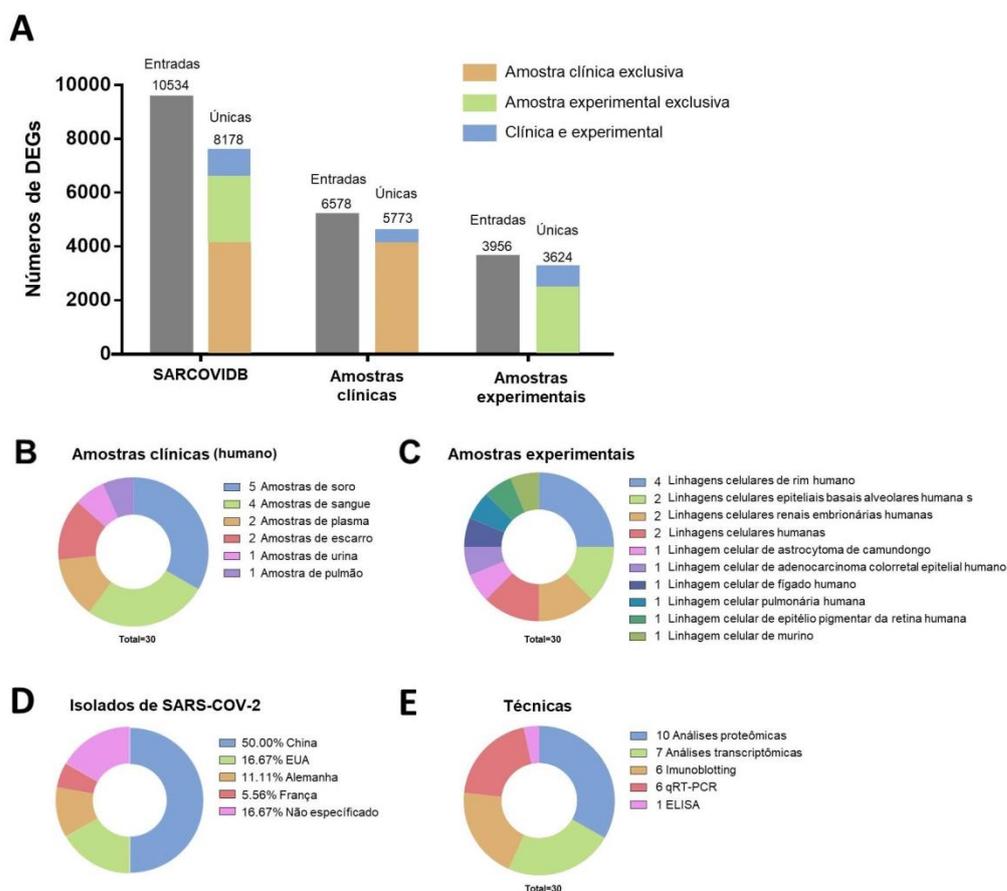


Figura 2: Dados gerais do SARSCOVIDB: **(A)** Total de entradas, em número de DEG, no total, em amostras clínicas e amostras experimentais. **(B)** Total de artigos de amostras clínicas. **(C)** Total de artigos de amostras experimentais. **(D)** Cepa ou origem geográfica do isolado de SARS-CoV-2. **(E)** Técnicas utilizadas para a obtenção dos dados.

Até o momento, o SARSCOVIDB, apresenta um total de **10534** entradas de genes depositadas, as quais equivalem a **8178** genes diferentes, adicionados e revisados manualmente. Esses dados são oriundos de diferentes tipos de experimentos, com distintos objetivos, sendo a maioria das entradas obtidas de análises de amostras clínicas, **6578** no total (Fig. 2A), já os dados obtidos de amostras experimentais totalizaram **3956** entradas (Fig. 2A). Esse número maior de dados clínicos demonstram a dimensão do esforço empregado pela comunidade científica em estudar a COVID-19. Comparando com o banco de dados ZIKAVID, uma quantidade muito menor de amostras clínicas de pacientes infectados por Zika vírus foi analisada (ROSA et al., 2019), mesmo com a epidemia tendo gerado um alerta global pela WHO (PAHO, 2016), causando grande preocupação da sociedade em 2015-2017. Já

em relação as amostras utilizadas, a maioria são amostras de sangue periférico (5 de soro, 4 de sangue e 2 de plasma) e apenas uma amostra de pulmão (Fig. 2B). Por outro lado, as amostras experimentais utilizaram diferentes tipos de células, em quase sua totalidade células de linhagens humanas, de rim, pulmão, neurais, etc (Fig. 2C). É importante destacar esses tipos de modelos e estratégias de estudos que buscam o entendimento molecular da COVID-19. Muitos dados científicos têm sido gerados, porém, ainda não há modelos animais robustos que repliquem fielmente a patogênese da SARS-CoV-2 (NEERUKONDA et al., 2020). Para o nosso conhecimento, inexistem estudos abrangentes comparando diferentes tipos de hospedeiros e suas respostas à infecção. Conhecer esses dados comparativos é importante para o planejamento de futuros experimentos, principalmente na escolha do modelo experimental mais adequado e a geração de dados com maior relevância científica. O SARSCOVIDB compila todos estes dados moleculares de expressão diferencial em diferentes hospedeiros e amostras clínicas. Assim, os usuários podem facilmente consultar os modelos mais frequentemente utilizados, comparar conjuntos de genes específicos, entre outras informações sobre todos os trabalhos publicados com este enfoque. Isto pode facilitar o planejamento de novos experimentos, acelerar o entendimento dos dados gerados e assim contribuir no entendimento mais rápido e necessário da COVID-19.

Os dados presentes no SARSCOVIDB foram gerados com análises realizadas com amostras clínicas ou experimentais com isolados de SARS-CoV-2 de diferentes regiões geográficas (Fig. 2D). Porém, os estudos ainda têm dado pouca importância sobre o impacto que as mutações sofridas pelas cepas, nas diferentes regiões mundiais, possam ter na dinâmica da doença e na virulência, como ocorreu com o vírus da Zika (ESSER-NOBIS, K., et al., 2019). Curiosamente, a maioria dos isolados estudados são de origem chinesa (50%; Fig. 2D), seguidos de isolados americanos (16.67%) e europeus. Em estudos emergentes, isolados virais de SARS-COV-2 podem levar a diferentes impactos patológicos e taxas de mortalidade, a depender da região (ISLAM, M. Rafiul et al, 2020). O exemplo mais recente de um vírus que apresentou um impacto clínico diferencial, a depender da região, foi o vírus da ZIKA, no qual o isolado Brasileiro, diferentemente dos isolados Asiáticos e Africanos, foi fortemente associado a dados neurológicos graves (BEYS-da-Silva et al. 2018). Assim, o SARSCOVIDB contribui para o estudo da patologia considerando a origem

do vírus, pois é possível fazer cruzamento de dados de expressão em modelos similares, utilizando vírus de diferentes origens.

Atualmente existe uma variedade de bancos de dados que reúnem informações de dados genômicos, de transcriptômica e de proteômica do vírus e de seu impacto no hospedeiro (PICKETT, Brett et al, [2020](#); PYLRO et al. [2016](#)). Diversos bancos de dados foram desenvolvidos para obtenção de dados de interação patógeno-hospedeiro como VirHostNet 2.0 (GUIRIMAND et al., [2015](#)). Existem também os que são direcionados para hospedagem de patógenos diferentes no que tange as interações patógeno-humano como HIV-1 human interaction database (AKO-ADJEI, Danso et al, [2014](#)). Da mesma maneira, o Gene Expression Omnibus (GEO), repositório de banco de dados público de expressão de gene de alto rendimento e matrizes de hibridização, chips, microarrays, entre outros recursos avançados de análises e exibição (Clough, Emily; Barreth, Tanya, [2020](#)) ou o “Virus Pathogen Database and Analysis Resource” disponível para uma ampla gama de vírus patogênicos humanos da qual fornece informações de genes e proteínas, dados de hospedeiros, componentes imunológicos baseadas na web (PICKETT, Brett et al, [2020](#)). No entanto, até onde sabemos, o SARSCOVDB, apesar de não ser o primeiro banco de dados com informações públicas contendo informações sobre a resposta molecular de hospedeiros virais, certamente é o primeiro a se dedicar exclusivamente com este enfoque específico em expressão diferencial após infecção por SARS-CoV-2. O banco foi projetado inicialmente por uma necessidade interna do grupo de pesquisa, na busca por marcadores de expressão diferencial em diferentes modelos biológicos. A maioria dos dados gerados por outros bancos de dados são de forma automatizada concentrando informações genéricas o que pode resultar em informações e dados imprecisos (PAPANIKOLAOU et al. [2015](#)).

O SARSCOVDB de propõe a preencher estas lacunas de forma simples, organizada e objetiva. É um banco de dados onde se pode consultar dados sob diferentes perspectivas experimentais como a metodologia utilizada, hospedeiro, cepa do vírus, etc. Além de identificar todos os DEGs envolvidos na infecção por SARS-COV-2. Estes dados e resultados do banco de dados podem servir a comunidade científica na facilitação e aceleração de estudos para entender a patogênese da COVID-19.

5 CONCLUSÕES

O SARSCOVIDB é o primeiro banco de dados a reunir todos os dados de análises de expressão diferencial da infecção por SARS-COV-2. O banco de dados construído reúne dados atualizados e checados manualmente que geram análises mais precisas e rápidas. Os usuários não precisam ter um *background* em bioinformática, pela interface amigável e simples da plataforma. Além disso, é possível explorar todas as possibilidades de recursos de busca, refinando os resultados relacionados aos estudos depositados, como a busca pelas diferentes cepas virais, hospedeiros, metodologias, lista de genes/proteínas. Isto permite que o usuário possa cruzar dados para o melhor entendimento e análise dos dados depositados. Por fim, o SARSCOVIDB é uma ferramenta promissora no apoio de cientistas e profissionais da área médica em pesquisas e análises futuras e na compreensão das abordagens moleculares relacionados à infecção por SARS-COV-2.

6 PERSPECTIVAS

Nosso objetivo futuro é manter o banco de dados SARSCOVIDB atualizado com estudos publicados sobre genes diferencialmente expressos de hospedeiros de SARS-CoV-2, criar ferramentas interativas para facilitar a interpretação dos dados moleculares depositados e compartilhar dados relevantes em redes sociais.

REFERÊNCIAS

AKO-ADJEI, Danso et al. HIV-1, human interaction database: current status and new features. *Nucleic Acids Research*, [S. l.], v. 43, n. D1, p. D566–D570, 2014.

Disponível em: <https://doi.org/10.1093/nar/gku1126>

BEYS-DA-SILVA, Walter O. et al. Zika Virus Infection of Human Mesenchymal Stem Cells Promotes Differential Expression of Proteins Linked to Several Neurological Diseases. *Molecular Neurobiology*, [S. l.], v. 56, n. 7, p. 4708–4717, 2018. Disponível em: <https://doi.org/10.1007/s12035-018-1417-x>

BOJKOVA, Denisa et al. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature*, [S. l.], v. 583, n. 7816, p. 469–472, 2020. Disponível em: <https://doi.org/10.1038/s41586-020-2332-7>

CLOUGH, Emily; BARRETT, Tanya. The Gene Expression Omnibus Database. In: *Methods in Molecular Biology*. [S. l.]: Springer New York, 2016. p. 93–110. E-book. Disponível em: https://doi.org/10.1007/978-1-4939-3578-9_5

CONASS. Conselho Nacional de Secretários de Saúde, 2020. Painel COVID-19. Disponível em: <https://www.conass.org.br/>. Acesso em 21/09/2020.

ELIEZER, Michael et al. Sudden and Complete Olfactory Loss of Function as a Possible Symptom of COVID-19. *JAMA Otolaryngology–Head & Neck Surgery*, [S. l.], v. 146, n. 7, p. 674, 2020. Disponível em: <https://doi.org/10.1001/jamaoto.2020.0832>

Esser-Nobis K, Aarreberg LD, Roby JA, Fairgrieve MR, Green R, Gale M, Jr. 2019. Comparative analysis of African and Asian lineage-derived Zika virus strains reveals differences in activation of and sensitivity to antiviral innate immunity. *J Virol* 93:e00640-19. <https://doi.org/10.1128/JVI.00640-19>.

GORDON, David E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, [S. l.], v. 583, n. 7816, p. 459–468, 2020. Disponível em: <https://doi.org/10.1038/s41586-020-2286-9>

Gurumayum S, Brahma R, Naorem LD, Muthaiyan M, Gopal J, Venkatesan A (2018) ZikaBase: an integrated ZIKV- human interactome map database. *Virology* 514:203–210. <https://doi.org/10.1016/j.virol.2017.11.007>

HORBACH, Serge P. J. M. Pandemic Publishing: Medical journals drastically speed up their publication process for Covid-19. Cold Spring Harbor Laboratory, , 2020. Disponível em: <https://doi.org/10.1101/2020.04.18.045963>

HULO, Chantal et al. ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Research*, [S. l.], v. 39, n. suppl_1, p. D576–D582, 2010. Disponível em: <https://doi.org/10.1093/nar/gkq901>

ISLAM, M. Rafiul et al. Genome-wide analysis of SARS-CoV-2 virus strains circulating worldwide implicates heterogeneity. *Scientific Reports*, [S. l.], v. 10, n. 1, 2020. Disponível em: <https://doi.org/10.1038/s41598-020-70812-6>

JHU CSSE. John Hopkins University, 2020, COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE)]. Disponível em: <https://coronavirus.jhu.edu/map.html>. Acesso em 21/09/2020.

KAMBHAMPATI, Srinivas B. S.; VAISHYA, Raju; VAISH, Abhishek. Unprecedented surge in publications related to COVID-19 in the first three months of pandemic: A bibliometric analytic report. *Journal of Clinical Orthopaedics and Trauma*, [S. l.], v. 11, p. S304–S306, 2020. Disponível em: <https://doi.org/10.1016/j.jcot.2020.04.030>

KWON, Diana. How swamped preprint servers are blocking bad coronavirus research. *Nature*, [S. l.], v. 581, n. 7807, p. 130–131, 2020. Disponível em: <https://doi.org/10.1038/d41586-020-01394-6>

LECHIEN, Jerome R. et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. *European Archives of Oto-Rhino-Laryngology*, [S. l.], v. 277, n. 8, p. 2251–2261, 2020. Disponível em: <https://doi.org/10.1007/s00405-020-05965-1>

NEERUKONDA, Sabari Nath; KATNENI, Upendra. A Review on SARS-CoV-2 Virology, Pathophysiology, Animal Models, and Anti-Viral Interventions. *Pathogens*, [S. l.], v. 9, n. 6, p. 426, 2020. Disponível em: <https://doi.org/10.3390/pathogens9060426>

OUYANG, Yabo et al. Downregulated Gene Expression Spectrum and Immune Responses Changed During the Disease Progression in Patients With COVID-19. *Clinical Infectious Diseases*, [S. l.], 2020. Disponível em: <https://doi.org/10.1093/cid/ciaa462>

PAHO, Pan American Health Organization, 2016. WHO announces a Public Health Emergency of International Concern. Disponível em: https://www.paho.org/hq/index.php?option=com_content&view=article&id=11640:2016-who-statement-on-1st-meeting-ihp-2005-emergency-committee-on-zika-virus&Itemid=135&lang=en Acesso em: 10/10/2020.

PAPANIKOLAOU, Nikolas et al. Protein–protein interaction predictions using text mining methods. *Methods*, [S. l.], v. 74, p. 47–53, 2015. Disponível em: <https://doi.org/10.1016/j.ymeth.2014.10.026>

Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, Liu M, Kumar S, Zaremba S, Gu Z, Zhou L, Larson CN, Dietrich J, Klem EB, Scheuermann RH (2011) ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res* 40:D593–D598. <https://doi.org/10.1093/nar/gkr859>

Przybyła P, Shardlow M, Aubin S, Bossy R, Eckart de Castilho R, Piperidis S, McNaught J, Ananiadou S (2016) Text mining resources for the life sciences. Banco de dados 2016. doi: <https://doi.org/10.1093/database/baw145>

Pylro VS, Oliveira FS, Morais DK, Cuadros-Orellana S, Pais FS-M, Medeiros JD, Geraldo JA, Gilbert J, Volpini AC, Fernandes GR (2016) ZIKV – CDB: a collaborative database to guide research linking snrnas and Zika virus disease symptoms. PLoS Negl Trop Dis 10:e0004817. <https://doi.org/10.1371/journal.pntd.0004817>

ROSA, Rafael L. et al. ZIKAVID—Zika virus infection database: a new platform to analyze the molecular impact of Zika virus infection. Journal of NeuroVirology, [S. l.], v. 26, n. 1, p. 77–83, 2019. Disponível em: <https://doi.org/10.1007/s13365-019-00799-y>

SHEN, Bo et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. Cell, [S. l.], v. 182, n. 1, p. 59- 72.e15, 2020. Disponível em: <https://doi.org/10.1016/j.cell.2020.05.032>

VARGAS, Gabriele et al. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and glial cells: Insights and perspectives. Brain, Behavior, & Immunity - Health, [S. l.], v. 7, p. 100127, 2020. Disponível em: <https://doi.org/10.1016/j.bbih.2020.100127>

WHO. World Health Organization, 2020. Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). Disponível em: [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)). Acesso em 08/07/2020

WHO. World Health Organization, 2020. WHO Statement regarding cluster of pneumonia cases in Wuhan, China. Disponível em: <https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china>. Acesso em: 10/10/2020.

WRAPP, Daniel et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science, [S. l.], v. 367, n. 6483, p. 1260–1263, 2020. Disponível em: <http://dx.doi.org/10.1126/science.abb2507>

YANG, Shiping et al. Understanding Human-Virus Protein-Protein Interactions Using a Human Protein Complex-Based Analysis Framework. mSystems, [S. l.], v. 4, n. 2, 2019. Disponível em: <https://doi.org/10.1128/msystems.00303-18>