UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE MICROBIOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA

GABRIEL AUGUSTO PIRES DE SOUZA

INTEGRANDO TRANSCRIPTOMAS COMO FERRAMENTA DE ENTENDIMENTO DE PATOGÊNESE E SELEÇÃO DE ALVOS TERAPÊUTICOS EM INFECÇÕES VIRAIS EMERGENTES E DESENVOLVIMENTO DE UM MATERIAL DIDÁTICO PARA AULAS PRÁTICAS DE VIROLOGIA

BELO HORIZONTE
2023

GABRIEL AUGUSTO PIRES DE SOUZA

INTEGRANDO TRANSCRIPTOMAS COMO FERRAMENTA DE ENTENDIMENTO DE PATOGÊNESE E SELEÇÃO DE ALVOS TERAPÊUTICOS EM INFECÇÕES VIRAIS EMERGENTES E DESENVOLVIMENTO DE UM MATERIAL DIDÁTICO PARA AULAS PRÁTICAS DE VIROLOGIA

Versão final

Tese apresentada ao Programa de Pós-Graduação em Microbiologia da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Microbiologia.

Orientador: Prof. Luiz Felipe Leomil Coelho

Coorientador: Prof. Jônatas Santos Abrahão

BELO HORIZONTE
2023

043 Souza, Gabriel Augusto Pires de.
Integrando transcriptomas como ferramenta de entendimento de patogênese e seleção de alvos terapêuticos em infecções virais emergentes e desenvolvimento de um material didático para aulas práticas de virologia [manuscrito] / Gabriel Augusto Pires de Souza. – 2023.
91 f. : il. ; 29,5 cm.

Orientador: Prof. Luiz Felipe Leomil Coelho. Coorientador: Prof. Jônatas Santos Abrahão.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Microbiologia.

1. Microbiologia. 2. Transcriptoma. 3. Virologia. 4. Materiais de Ensino. I. Coelho, Luiz Felipe Leomil. II. Abrahão, Jônatas Santos. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título..

CDU: 579

Ficha Catalográfica elaborada pela bibliotecária Fabiane C M Reis CRB 6 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PÓS-GRADUAÇÃO EM MICROBIOLOGIA

ATA DE DEFESA DE TESE

ATA DA DEFESA DE TESE DE **GABRIEL AUGUSTO PIRES DE SOUZA**

Nº REGISTRO: 2019703879

Às 13:30 horas do dia **13 de março de 2023**, reuniu-se, por via remota, a Comissão Examinadora composta pelos Drs. Betânia Paiva Drumond (Departamento de Microbiologia/ICB/UFMG), Gabriel Magno de Freitas Almeida (Arctic University of Norway), Rafael Kroon Campos (University of Texas Medical Branch), Gabriel Gerber Hornink (Universidade Federal de Alfenas (UNIFAL-MG) e o Prof. Dr. Luiz Felipe Leomil Coelho (Orientador), para julgar o trabalho final **"Integrando transcriptomas como ferramenta de entendimento de patogênese e seleção de alvos terapêuticos em infecções virais emergentes e desenvolvimento de um material didático para aulas práticas de virologia"** do aluno **Gabriel Augusto Pires de Souza**, requisito final para a obtenção do Grau de **DOUTOR EM CIÊNCIAS BIOLÓGICAS: MICROBIOLOGIA**. Abrindo a sessão, o Presidente da Comissão, Prof. Dr. Luiz Felipe Leomil Coelho, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para a apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. O candidato foi considerado **APROVADO**. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ata, que será assinada por todos os membros participantes da Comissão Examinadora. O candidato tem 60 (sessenta) dias, a partir desta data, para entregar a versão final da tese ao Programa de Pós-graduação em Microbiologia da UFMG e requerer seu diploma.

Belo Horizonte, 13 de março de 2023

Membros da Banca:

Profa. Dra. Betânia Paiva Drumond

Prof. Dr. Gabriel Magno de Freitas Almeida

Prof. Dr. Rafael Kroon Campos

Prof. Dr. Gabriel Gerber Hornink

De acordo:

Prof. Dr. Luiz Felipe Leomil Coelho

(Orientador)

Profa. Dra. Daniele da Glória de Souza

(Coordenadora do Programa de Pós-graduação

em Microbiologia)

Documento assinado eletronicamente por **Luiz Felipe Leomil Coelho**, **Usuário Externo**, em 15/03/2023, às 07:07, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Gabriel Magno de Freitas Almeida**, **Usuário Externo**, em 15/03/2023, às 08:41, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Betania Paiva Drumond**, **Professora do Magistério Superior**, em 15/03/2023, às 08:46, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Daniele da Gloria de Souza**, **Coordenador(a) de curso de pós-graduação**, em 15/03/2023, às 14:52, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Gabriel Gerber Hornink**, **Usuário Externo**, em 16/03/2023, às 09:16, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rafael Kroon Campos**, **Usuário Externo**, em 29/03/2023, às 09:14, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2093818** e o código CRC **3EDB15E5**.

AGRADECIMENTOS

Primeiramente, agradeço à minha família, mãe, pai e irmão, por todo o apoio, confiança e pelos valores que me foram ensinados desde a infância, mas acima de tudo pela confiança nas minhas escolhas e por caminharem ao meu lado na busca dos meus objetivos. Sei que mesmo distante, nunca estarei sozinho.

À Universidade Federal de Minas Gerais (UFMG) e ao Departamento de Microbiologia, pela oportunidade de adquirir novos conhecimentos e experiências enriquecedoras ao longo da minha formação acadêmica. Agradeço a todos os professores que contribuíram direta ou indiretamente para a minha formação, especialmente pelo seu papel na minha jornada acadêmica. Especialmente aos professores do Laboratório de Vírus, pelo seu trabalho árduo e dedicação ao ensino. Sou grato por todo o apoio e orientação que me forneceram e continuo a valorizar as lições valiosas que aprendi com vocês, incluidno os puxões de orelha, mas também muito dos conselhos que sigo passando a diante.

Ao meu orientador, Prof. Dr. Luiz Felipe Leomil Coelho, gostaria de expressar minha profunda gratidão pela paciência, dedicação, competência e profissionalismo. Seu exemplo de orientação e a confiança que depositou em mim foram essenciais para meu crescimento acadêmico e pessoal. Obrigado por estar sempre disponível e por toda a ajuda que me ofereceu ao longo desses anos.

Ao meu co-orientador, Jônatas Santos Abrahão, minha eterna gratidão por sua orientação e exemplos de pesquisa ambição para uma pesquisa de alta qualidade. Agradeço também por abrir caminhos para uma carreira internacional e pelo belíssimo projeto que me inseriu com os vírus gigantes.

Gostaria de expressar minha gratidão aos meus colegas do Laboratório de Vírus, em especial às garotas do Grupo de Estudo e Prospecção de Vírus Gigantes (GEPVIG), Ana Paula Crispim, Bruna Azevedo, Isabella Aquino, Talita Machado e Victoria Queiroz, pela acolhida e amizade. O trabalho em equipe e as experiências compartilhadas com vocês foram inesquecíveis, mas mais do que isso o laço que criamos nas adversidades e que se mantém mesmo distante. Estarei com vocês nas lutas e nas muitas conquistas que certamente estão por vir.

Também agradeço aos novos alunos do meu antigo laboratório, O Laboratorio de vacinas, em Alfenas, nas, pela colaboração e empenho em conjunto, sem os quais muito deste trabalho não teriam sido viáveis. Especialmente meu amigo, Renato Fróes Goulart de Castro (*in memoriam*) que tanto fez para me auxiliar e infelizmente não me deu a oportunidade ou tempo de retribuir, mas que se foi sabendo da minha gratidao.

Por fim, agradeço à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro e suporte à pesquisa, o que tornou possível a realização deste trabalho.

“A Mãe Natureza está ao telefone: O que você tem feito? Não se esqueça que eu sou sua casa! Os vírus vêm, o fogo queima, até que os seres humanos aprendam com cada desastre; você não é meu mestre" (Marina Diamandis, Purge The Poison, 2021)

RESUMO

O aumento na incidência de eventos de doenças emergentes virais tem sido observado desde a década de 1940. Recentemente temos observado vários virus sendo introduzidos em novas regiões do globo, como em 2015 com a disseminação do Zika vírus pela América do Sul. Ainda mais recentemente, presenciamos a emergência de um novo corona vírus, o SARS-CoV-2, na China que rapidamente atingiu níveis pandêmicos. Esse tipo de evento, além de trazer preocupações no que diz respeito à saúde pública e economia, exige uma rápida resposta da comunidade cientifica, seja para o entendimento dos mecanismos que levam à doença, seja na busca de tratamentos contra a doença em questão. As técnicas de proteômica, transcriptômica e genômica mostram-se grandes aliadas após a emergência de novos doenças virais, pois são amplamente difundidas e é comum que sejam empregadas na compreensão do ciclo de replicação do vírus, a patogênese viral e para a definição de alvos terapêuticos para novos vírus. Estudos de transcriptômica geram grandes quantidades de dados que têm aplicações potenciais muito além dos objetivos originais de um experimento, uma vez que são armazenados em bancos de dados, normalmente de acesso público. Através destes dados disponíveis em banco de dados buscamos desenvolver um protocolo de análise integrativa de transcriptomas, utilizando como base para a validação deste protocolo usando dados de células neurais infectadas com ZIKV. Foram identificados genes diferencialmente expressos compartilhados. Alguns desses genes, vias e funções biológicas foram ligadas à neurogênese e/ou apoptose. A análise *in silico* integrada capaz de prever e identificar biomarcadores putativos a partir de diferentes dados de transcriptoma, que seriam úteis para o entendimento da patogênese da doença viral e para aplicação na identificação de candidatos antivirais. Com a emergência do SARS-CoV-2, este protocolo foi revistado e atualizado, com uma estratégia de *deep learning* para identificar alvos em células infectadas que possam ser inibidos por moléculas já aprovadas disponíveis em bibliotecas, sugerindo formas de controle da infecção por SARS-CoV-2. Este tipo de análise só é possível devido ao grande volume de informação que é gerado a partir dos estudos de transcriptômica e de virologia em geral. Em sala de aula muitas vezes mostra-se desafiador transferir todo esse conhecimento aos alunos. Um dos fatores identificados é a dificuldade de se elaborar

aulas práticas de virologia. Portanto, buscamos desenvolver um material didático que permitisse a visualização de partículas de vírus gigantes de amebas e efeitos citopáticos de vírus em células animais, com objetivo de afastar a virologia do campo abstrato da aprendizagem. Incluindo também estes vírus gigantes como modelos para o entendimento da coevolução ao revisar seus mecanismos de invasão nas células hospedeiras.

Palavras-chaves: Análise integrativa de transcritora, vírus emergentes, Materiais de Ensino.

ABSTRACT

The increase of the incidence of viral emerging disease events has been observed since the 1940s. Recently we have observed several viruses being introduced in new regions of the globe, for example, the spread of the Zika virus in South America in 2015. Even more recently, we witnessed the emergence of a new coronavirus in China that quickly reached pandemic levels. In addition to raising concerns about public health and the economy, this type of event requires a quick response from the scientific community, either to understand the mechanisms that lead to the disease or in the search for treatments against the disease in question. Proteomics, transcriptomics, and genomics techniques are great allies after the emergence of new viral diseases, as they are widely disseminated, and it is common for them to be used to understand the virus replication cycle and viral pathogenesis and to define therapeutic targets for new viruses. Transcriptomics studies generate large amounts of data that have potential applications far beyond the original objectives of an experiment since they are stored in databases, usually publicly accessible. Through the data available in the database, we sought to develop a protocol for the integrative analysis of transcriptomes, using it as a basis for validating this protocol using data from neural cells infected with ZIKV. Shared differentially expressed genes were identified. Some of these genes, pathways and biological functions have been linked to neurogenesis and apoptosis and integrated into silico analysis capable of predicting and identifying putative biomarkers from different transcriptome data that would be useful for understanding viral disease pathogenesis and for application in identifying antiviral candidates. With the emergence of SARS-CoV-2, this protocol was revised and updated, with a deep learning strategy to identify targets in infected cells that can be inhibited by already approved molecules available in libraries, suggesting controlling mechanisms for SARS-CoV-2 infection. This type of analysis is only possible due to the large volume of general information generated from transcriptomics and virology studies. In the classroom, transferring all this knowledge to students is often challenging. One of the factors identified is the difficulty of preparing for practical virology classes. Therefore, we sought to develop didactic material that allowed the visualization of Amoeba's giant virus particles and cytopathic effects of viruses in animal cells, to move virology away from the abstract field of learning, also including

these giant viruses as models for understanding coevolution by reviewing their invasion mechanisms in host cells.

Keywords: Integrative transcriptome analysis, emerging viruses, Teaching Materials.

SUMÁRIO

1 INTRODUÇÃO

Além dos grandes impérios, guerras e avanços tecnológicos ao longo do tempo, a história também é definida pelas grandes enfermidades que afetaram uma variedade de povos. As doenças têm uma história e cada época tem suas doenças (HEGENBERG, 1998). Podemos citar: a Peste de Atenas, que ocorreu no século V a.C. durante a Guerra do Peloponeso e afetou gravemente a cidade de Atenas, matando uma grande parte da população, incluindo o famoso líder Péricles. A Peste Negra dizimou cerca de um terço da população europeia na Idade Média, a varíola, que dizimou populações inteiras nas Américas após a chegada dos europeus, a gripe espanhola, que matou cerca de 50 milhões de pessoas em todo o mundo no início do século XX, e mais recentemente, a pandemia de COVID-19, que vem afetando todo o mundo desde o final de 2019 (COOK, 1998; HAYS, 2005). Cada uma dessas enfermidades teve impactos profundos na história, moldando culturalmente a sociedade.

Embora as doenças acompanhem a humanidade desde o seu surgimento, o entendimento do que é uma doença se altera conforme ao longo da história e das sociedades. Na Grécia Antiga, Hipócrates postulou a saúde estava relacionada ao equilíbrio entre os quatro humores do corpo (bile amarela, bile negra) (SCLIAR, 2007; TSOUCALAS *et al.*, 2018). Durante a Idade Média, a doença era frequentemente vista como uma punição divina ou resultado da má disposição dos astros. No Iluminismo, surgiram teorias mais científicas sobre a doença, na qual passa a se observar a influência do ambiente e da dieta na saúde dos indivíduos (KARAMANOU *et al.*, 2012). Finalmente, a teoria microbiana da doença, desenvolvida no século XIX, revolucionou a compreensão das doenças infecciosas (KARAMANOU *et al.*, 2012), demonstrando que muitas delas são causadas por microrganismos, como fungos, bactérias e, pelos vírus.

Os vírus são considerados como as entidades mais abundantes na Terra e têm grande impacto sobre a vida e sobre a manutenção da vida no planeta, não só por atuarem como importantes forças de controle populacional, mas também por, dessa forma, influenciar ciclos importantes, como no ciclo do carbono nos ecossistemas aquáticos (MD TAIMUR KHAN *et al.*, 2022; PAEZ-ESPINO *et al.*, 2016; ROUX *et al.*, 2016; VILLARREAL; WITZANY, 2010). Estes parasitas intracelulares obrigatórios,

independentemente da discussão se possuem vida ou não, são, sem dúvidas, entidades biológicas essenciais na biosfera (PAEZ-ESPINO *et al.*, 2016).

Uma vez descobertos, os vírus inicialmente eram considerados agentes filtráveis, e foi natural que despertassem o fascínio humano e se tornassem alvos de estudos. No começo da virologia, o estudo dos vírus limitava-se aos vírus de plantas e, posteriormente, aos vírus bacterianos (bacteriófagos), pois o cultivo de seus hospedeiros era relativamente fácil (BOS, 1981; CLOKIE *et al.*, 2011; SIMON, 1912; VAN HELVOORT, 1994).

Simultaneamente a descoberta dos vírus de animais, foi apresentada sua face patogênica, por estarem frequentemente associados a diversas doenças (KEILIN, 1959; LECOQ, 2001; PAEZ-ESPINO *et al.*, 2016). Na década de 1940, com o aperfeiçoamento das técnicas de cultura celular, o estudo dos vírus animais teve avanços significativos, quando muitos vírus passaram a ser isolados em laboratório (FAISST, 1999; LELAND; GINOCCHIO, 2007).

A virosfera, reúne a diversidade dos vírus, foi entendida para além de diversa, muito vasta, afinal, estima-se que para cada organismo vivo na terra, exista ao menos um vírus que possa infectá-lo (ENQUIST, 2009; MUSHEGIAN, 2020). Contudo, ao despertar da virologia, muitos estudos foram direcionados principalmente para agentes patogênicos ao homem ou aos agentes que causassem perdas econômicas na agropecuária (RODRIGUES *et al.*, 2017). Ainda assim, independente do hospedeiro que infecte, todos os vírus devem atingir o espaço intracelular para se reproduzirem.

No momento em que o vírus toma controle da maquinaria celular hospedeira, ele passa a regular também a expressão gênicas daquela célula (STERN-GINOSSAR *et al.*, 2019). É comum que vírus silencie genes ligados a fatores de resistência e aumente a expressão de genes que favoreçam a sua replicação. Como exemplo, a expressão aumentada do gene PRPF8 promovida pelo vírus da Influenza A, desencadeia uma maior produção viral (YANG *et al.*, 2017). Portanto, entender quais genes são diferencialmente expressos ao longo da infecção da célula hospedeira em comparação a uma célula não-infectada nos fornece informações valiosas sobre o ciclo de replicação daquele vírus.

O conteúdo da informação de um organismo é registrado no DNA, seu genoma, com exceção dos vírus com genoma de RNA. O conteúdo da informação armazenado no genoma deve ser posteriormente expresso por meio da transcrição. O transcriptoma de um organismo representa a soma de todos os seus transcritos de RNA (LOWE *et al.*, 2017). Avaliar os níveis a expressão dos genes de um organismo em diferentes condições, por exemplo frente a infecções e em diferentes tempos de infecção, fornece informações sobre como os genes são regulados por aquele patógeno e revela detalhes da biologia do organismo em estudo.

Este tipo de avaliação só foi possível com o desenvolvimento de técnicas de hibridização ou sequenciamento de RNA, que se tornaram particularmente populares em meados dos anos de 1990 e 2000, inaugurando a era da transcriptômica (LOWE *et al.*, 2017). Publicada pela primeira vez em 1995, a técnica de *microarrays*, mede a abundância de um conjunto definido de transcritos por meio da hibridação destes transcritos com uma matriz de sondas complementares (NELSON, 2001; WANG, ZHONG; GERSTEIN; SNYDER, 2009). RNA-Seq, por outro lado, refere-se ao sequenciamento de cDNAs transcritos, em que a abundância é derivada do número de contagens de cada transcrito (MOROZOVA; HIRST; MARRA, 2009; WANG, ZHONG; GERSTEIN; SNYDER, 2009).

A análise transcriptômica permitiu o estudo de como a expressão gênica e é fundamental para a compreensão de doenças humanas, incluindo doenças emergentes (LOWE *et al.*, 2017). Uma doença infecciosa emergente é caracterizada pelo aumento de sua incidência recentemente (nos últimos 20 anos) e seu potencial aumento de incidência no futuro próximo (FALLIS, 2013; WOOLHOUSE, MARK *et al.*, 2012; WOOLHOUSE, MARK E.J.; GOWTAGE-SEQUERIA, 2005).

Estima-se que as infecções emergentes representem pelo menos 12% de todos os patógenos humanos (TAYLOR; LATHAM; WOOLHOUSE, 2001). Elas podem ser causadas por microrganismos recentemente identificados, como o caso mais recente do coronavírus causador da síndrome respiratória aguda grave 2 (SARS-CoV-2) ou pela emergência de um patógeno conhecido, como ocorre com novas cepas de influenza (FAUCI, 2005; MASCOLA; GRAHAM; FAUCI, 2021). Podem também ser resultado da disseminação de uma doença existente para uma nova

população em uma região geográfica diferente, como ocorreu com o surto *Zika vírus* (ZIKV) especialmente no nordeste do Brasil, mas não restrito a ele, no ano de 2015.

O aumento na incidência de eventos de doenças emergentes virais tem sido observado desde a década de 1940, muitas vezes associadas a zoonoses relacionadas à vida selvagem. A perda de biodiversidade causada pela atividade humana é reconhecida como um dos principais mecanismos correlacionado a este aumento, bem como as mudanças climáticas (CARLSON *et al.*, 2022; KEESING *et al.*, 2010; MAHMOOD; GUINTO, 2022). A emergência de um novo patógeno viral traz uma série de preocupações, pois além do impacto na saúde pública, pode levar a perdas econômicas, bem como consequências sociais.

Portanto, ao se detectar um novo vírus emergente, é comum que a comunidade cientifica direcione seus esforços a compreender o patógeno em si, assim como a doença causada por ele, a fim de atender à uma demanda da sociedade (LUNA, 2002). Nos eventos mais recentes de emergência e reemergência de vírus, as técnicas de transcriptômica já estavam amplamente difundidas e é comum que sejam empregadas na compreensão do ciclo de replicação do vírus, a patogênese viral e para a definição de alvos terapêuticos para novos vírus (ALQUTAMI; SENOK; HACHIM, 2021; DAAMEN *et al.*, 2021; SUN *et al.*, 2021).

O entendimento sobre os vírus e seus mecanismos é que fornece ferramentas que serão aplicadas mais tarde pela biotecnologia. No caso especifico dos estudos de transcriptômica geram grandes quantidades de dados que têm aplicações potenciais muito além dos objetivos originais de um experimento (LOWE *et al.*, 2017). Estes dados são normalmente armazenados em bancos de dados específicos, como por exemplo o Gene Expression Omnibus, e ficam disponíveis para serem explorados por outros pesquisadores (LOWE *et al.*, 2017).

É inegável que há um grande fluxo de informações sendo gerados a partir do estudo dos vírus. Contudo, apesar de todo o conhecimento que é gerado pela virologia, é fácil identificar uma dificuldade na transferência desse conhecimento que para os alunos em sala de aula e, para a sociedade de um modo geral. Um fenômeno foi evidenciada ao longo da pandemia de COVID-19, a disseminação de informações falsas e desinformação, conhecida como "fake news", foi um grande desafio enfrentado pela virologia e pela saúde pública em geral (APUKE; OMAR, 2021;

GREENE; MURPHY, 2021). A proliferação de teorias da conspiração, tratamentos falsos e informações enganosas na internet e nas redes sociais levou a uma maior desconfiança do público em relação às autoridades de saúde e às vacinas, bem como a um aumento da hesitação em se vacinar (ROCHA *et al.*, 2021). Isso tornou ainda mais difícil para os virologistas e profissionais de saúde disseminarem informações precisas e confiáveis sobre a doença e suas formas de prevenção e tratamento. A disseminação de fake news pode ter consequências graves para a saúde pública, tornando o controle de epidemias e pandemias mais difícil e prolongado.

A lacuna da transferência de conhecimento para os alunos está associada a inúmeros fatores, entre eles, a dificuldade de se realizar aulas práticas em disciplinas de virologia, devido aos riscos de biossegurança, aos altos custos e ao fato de muitas vezes a virologia parecer um tanto abstrata, visto que a representação dos vírus muitas vezes está restrita a figuras esquemáticas, ilustrações e imagens de microscopia eletrônicas (AKASHI *et al.*, 2019).

Essa última limitação, contudo, parece ter sido deixada para trás com a descoberta dos vírus gigantes de amebas (LA SCOLA *et al.*, 2003). Estes vírus, que infectam amebas de vida livre, e, portanto, não são patogênicos aos seres humanos. Eles recebem essa denominação pois são particularmente grandes se comparados a maioria dos vírus conhecidos que infectam outros organismos (especialmente animais). Os vírus gigantes frequentemente deixam a escala dos nanômetros para a escala de micrômetros, podendo ser visualizados em microscópios óticos comuns (COLSON *et al.*, 2017; LA SCOLA *et al.*, 2003). Portanto, apresentam um grande potencial no campo educacional, pois possibilitam que temas da virologia sejam trazidos para os laboratórios aulas práticas de forma segura.

Esta tese foi elaborada para permear entre diferentes áreas de virologia, retratando a trajetória de meu trabalho na UFMG. Ela se inicia com uma seção de dois capítulos, apresentados no formato de artigos científicos publicados, em que explora o potencial da análise integrativa de transcriptoma para determinar genes diferencialmente expressos em decorrência de uma infecção, seja com a intenção de estudar a patogênese desta novo agente viral emergente ou para encontrar alvos terapêuticos para a doença. O primeiro capítulo apresenta um protocolo de análise integrativa de transcriptoma, utilizando dados disponíveis em banco de dados de

transcriptômica e os integrando uma plataforma virtual amigável, gerando um protocolo acessível e útil para identificar genes-chave e vias úteis que permitirão o melhor entendimento da patogênese de doenças virais emergentes. Esse protocolo foi validado com dados de células neurais infectadas com ZIKV, definido como organismo alvo, devido a sua recente emergência no Brasil. No capítulo seguinte este protocolo foi revistado e atualizado para ser empregado na identificação de alvos terapêuticos para nova doença emergente, a COVID-19. Publicado no periódico *Brazilian Journal of Microbiology* no final de 2022, este trabalho foi conduzido remotamente ao longo da pandemia de COVID-19, no período de confinamento e em que se iniciou também uma trajetória de estudos na Aix-Marseille Université, em Marseille na França.

Na seção seguinte são abordadas outras oportunidades exploradas ao longo do doutorado e conduzidas junto ao Grupo de Estudos e Prospecção de Vírus Gigantes (GEPVIG), explorando os aspectos revolucionários dos vírus gigantes e suas aplicações na educação. Nesta seção esta apresentada uma ferramenta que visa trazer para dentro da sala de aula descobertas recentes do campo da virologia, chamada de “Virus Goes Viral” e publicada na revista *Virology Journal.* Este trabalho também foi desenvolvido no primeiro ano de doutorado e pensado como o desenvolvimento de um produto para a sala de aula, e por essa razão também foi agregado a tese. Cada um dos capítulos é apresentado individualmente por uma breve introdução, justificativa e os objetivos. Ao final desta tese, está apresentada a lista de publicações associadas ao doutorado, incluindo colaborações e outros trabalhos não estão apresentados neste documento.

2. ARTIGO I: An *in silico* integrative protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis

Este é um artigo original publicado no periódico "Virus Research" em julho de 2020.

2.1 Introdução

A reemergência de arbovírus tem sido uma preocupação dos órgãos de saúde, especialmente por afetarem países em desenvolvimento nas zonas tropicais e subtropicais do globo. O Zika vírus foi descoberto em campanhas de identificação de vigilância da febre amarela e outros de novos agentes virais em meados da década de 1940 (WEAVER *et al.*, 2016; YUN; LEE, 2017). Ele foi isolado a partir de uma amostra de sangue de um macaco *Rhesus* febril, na floresta de Zika em Uganda (JAMALI MOGHADAM *et al.*, 2016). Posteriormente o vírus também foi isolado de um paciente humano, sendo então identificado como um patógeno humano.

O ZIKV pertence à família *Flaviviridae* e ao gênero *Flavivirus*, no qual estão inclusos também o Dengue virus (DENV), o West Nile virus (WNV) e Yellow Fever virus (YFV) (CHAN *et al.*, 2016). Assim como os demais membros desse gênero, o ZIKV é um vírus envelopado e com genoma de RNA simples fita senso positiva (SHI; GAO, 2017). Assim como ocorre com o DENV, o principal vetor de transmissão conhecido do ZIKV é o *Aedes aegypti*, assim como o *A. albopictus*, que são mosquitos hematófagos amplamente distribuídos nas zonas tropicais e subtropicais do globo (SONG *et al.*, 2017).

Esse arbovírus permaneceu por muito tempo restrito a surtos em regiões endêmicas da África e Ásia e ganhou destaque após sua emergir na América do Sul, por volta do ano de 2015, quando desordens neurológicas em recém nascidos foram associadas ao surto de ZIKV no nordeste do Brasil (PETERSEN *et al.*, 2016; YUN; LEE, 2017). Em adultos, as preocupações giravam em torno da associação da infecção pelo ZIKV e a síndrome de Guillain-Barré, uma desordem neuroimunológica em que o sistema imunológico ataca parte do sistema nervoso periférico, causando formigamento, fraqueza muscular, paralisia e até mesmo a morte.

Esses casos graves associados a doença trouxeram preocupação aos órgãos de saúde brasileiros e internacionais e o combate ao ZIKV foi iniciado, dando a uma

série de estudos que possibilitassem uma melhor compreensão da patogênese viral (CHAN *et al.*, 2016; SONG *et al.*, 2017; ZANLUCA; DOS SANTOS, 2016). Desvendar a patogênese de um vírus emergente é particularmente desafiador, devido aos poucos dados que estão disponíveis inicialmente. Nesse contexto, as abordagens de transcriptômica podem trazer pistas valiosas para compreensão dos mecanismos da infecção e do desenvolvimento da doença.

Com o acúmulo de dados gerados em diferentes condições de infecção e em diferentes tipos celulares, compilar estes resultados e buscar por genes diferencialmente expressos compartilhados por diferentes células em diferentes condições podem guiar estudos sobre a patogênese daquele vírus emergente (LIAO *et al.*, 2017).

A análise integrativa do transcriptoma, que tem se mostrado uma ferramenta valiosa para a compreensão de eventos biológicos em doenças complexas. A abordagem integrativa do transcriptoma tem sido empregada, por exemplo, para determinar genes associados doenças como, por exemplo, pré-eclâmpsia (MOSLEHI *et al.*, 2013), HIV (LIAO *et al.*, 2017), cancer de pulmão (CINEGAGLIA *et al.*, 2016) e carcinoma hepatocelular (HOSHIDA *et al.*, 2009).

Neste trabalho, nós apresentamos um protocolo de análise integrativa do transcriptômica utilizando uma plataforma virtual amigável útil para identificar genes-chave e vias úteis que permitirão o melhor entendimento da patogênese de doenças virais emergentes, utilizando dados disponíveis dos bancos de dados de transcriptoma. Para validação do protocolo, o ZIKV foi definido como organismo alvo, uma vez que era um vírus emergente com um surto recente no Brasil. Através deste protocolo usando dados de células neurais infectadas com ZIKV foram identificados genes diferencialmente expressos compartilhados. Alguns desses genes, vias e funções biológicas foram ligadas à neurogênese e/ou apoptose. A análise *in silico* integrada capaz de prever e identificar biomarcadores putativos a partir de diferentes dados de transcriptoma, que seriam úteis para o entendimento da patogênese da doença viral e para aplicação na identificação de candidatos antivirais.

2.2 Justificativa

O Zika vírus é um arbovírus transmitido principalmente por mosquitos do gênero *Aedes*. Embora tenha sido descoberto há bastante tempo, em 1947, em Uganda, passou por muitas décadas sendo associado a surtos doença febril em países endêmicos e sem despertar grandes preocupações da comunidade cientifica e dos órgãos de saúde mundiais. Esse cenário se reverteu com uma rápida disseminação do vírus para outros continentes, como a Oceania e as Américas, onde passou a ser associado com a malformação de recém-nascidos e por desordens neuro-imunológicas em adultos. Este fato alterou a percepção em torno deste vírus e fez com que fosse iniciada a busca por formas de tratamento de prevenção da doença, assim como ressaltou a necessidade do entendimento dos mecanismos de patogênese viral. Portanto, determinar genes que são diferencialmente expressos após a infecção pelo ZIKV é uma estratégia valiosa na compreensão da patogênese. Um protocolo que integra estes genes diferencialmente expressos em diferentes tecidos e células em diferentes condições, pode possibilitar a identificação de genes frequentemente alterados no curso da infecção e assim, determinam-se rotas metabólicas e processos biológicos que são afetados pelo vírus em controle da célula e como agente causador de uma doença.

2.3 Objetivos

2.3.1 Objetivo Geral

Elaborar um protocolo integrativo para identificar os principais genes e vias potencialmente associados a patogênese de um vírus emergente em células humanas.

2.3.2 Objetivos Específicos

- Identificar genes que são frequentemente diferencialmente expressos em diferentes tipos células neurais infectadas por ZIKV.
- Determinar as vias metabólicas e funções biológicas associadas a esses genes e o impacto de sua expressão diferencial na célula.

ANEXO

Artigo I

An *in silico* integrative protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis

DOI: 10.1016/j.virusres.2020.197986

Virus Research 284 (2020) 197986

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

An *in silico* integrative protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis

Gabriel Augusto Pires de Souza[a,b], Ezequiel Aparecido Salvador[a], Fernanda Roza de Oliveira[a], Luiz Cosme Cotta Malaquias[a], Jonatas Santos Abrahão[b], Luiz Felipe Leomil Coelho[a,*]

[a] Laboratório de Vacinas, Instituto de Ciências Biomédicas, Departamento de Microbiologia e Imunologia, Universidade Federal de Alfenas, Alfenas, Minas Gerais, Brazil
[b] Laboratório de Vírus, Instituto de Ciências Biológicas, Departamento de Microbiologia, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

ARTICLE INFO

Keywords:
Integrative bioinformatics
Emerging virus disease
Pathogenesis

ABSTRACT

The pathogenesis of an emerging virus disease is a difficult task due to lack of scientific data about the emerging virus during outbreak threats. Several biological aspects should be studied faster, such as virus replication and dissemination, immune responses to this emerging virus on susceptible host and specially the virus pathogenesis. Integrative *in silico* transcriptome analysis is a promising approach for understanding biological events in complex diseases. In this study, we propose an *in silico* protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis. To validate our protocol, the emerging arbovirus *Zika virus* (ZIKV) was chosen as a target micro-organism. First, an integrative transcriptome data from neural cells infected with ZIKV was used to identify shared differentially expressed genes (DEGs). The DEGs were used to identify the potential candidate genes and pathways in ZIKV pathogenesis through gene enrichment analysis and protein-protein interaction network construction. Thirty DEGs (24 upregulated and 6 downregulated) were identified in all ZIKV-infected cells, primarily associated with endoplasmic reticulum stress and DNA replication pathways. Some of these genes and pathways had biological functions linked to neurogenesis and/or apoptosis, confirming the potential of this protocol to find key genes and pathways involved on disease pathogenesis. Moreover, the proposed *in silico* protocol performed anintegrated analysis that is able to predict and identify putative biomarkers from different transcriptome data. These biomarkers could be useful to understand virus disease pathogenesis and also help the identification of candidate antiviral drugs.

1. Introduction

Outbreaks of emerging and/or re-emerging viral infections are common threats to human health. The emerging viral pathogens are defined as novel viruses that have been recently introduced in a population. In the last years, several medical important infectious outbreaks have occurred worldwide, including West Nile virus, Chikungunya virus, Zika virus, SARS, MERS, influenza and nCov-2019 (Ellwanger and Chies, 2016; Olival et al., 2017). Nevertheless, the knowledge and understanding of virus pathogenesis is a difficult task due to the lack of scientific data about the emerging virus during outbreak threats. Several biological aspects should be studied faster, such as virus replication and dissemination, immune responses to this emerging virus on a susceptible host and especially virus pathogenesis. A rapid and regular focus on basic research about virus pathogenesis is imperative and could help researches to better understand the pathogenesis of these viruses and also to develop therapeutics and vaccines and to control the disease (Afrough et al., 2019).

Changes in gene expression is a hallmark of virus infection and the identification of the altered Gene expression profiling is very important to understand the virus pathogenesis and also to the development of new antivirals drugs and vaccines for the emerging/re-emerging viruses. Recent developments in high-throughput sequencing technologies result in an intense accumulation of *omics* data from cells, tissues and patients infected with different viruses. Integrative *in silico* transcriptome analysis (ITA) is a promising approach for understanding biological events in complex diseases. ITA also contribute to understandings of disease pathogenesis and to find new drugs to treat the disease. For example, ITA was used to found genes with significant associations with Alzheimer's disease (Jin et al., 2018), preeclampsia (Moslehi et al., 2013), HIV (Liao et al., 2017), lung cancer (Cinegaglia et al., 2016) and hepatocellular carcinoma (Hoshida et al., 2009).

* Corresponding author at: Laboratório de Vacinas, Instituto de Ciências Biomédicas, Universidade Federal de Alfenas, Rua Gabriel Monteiro, 700 Centro, CEP: 37130-000, Alfenas, Minas Gerais, Brazil.
E-mail address: luiz.coelho@unifal-mg.edu.br (L.F. Leomil Coelho).

https://doi.org/10.1016/j.virusres.2020.197986
Received 2 March 2020; Received in revised form 21 April 2020; Accepted 21 April 2020
Available online 24 April 2020

In this study, we propose an *in silico* protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis. To validate our protocol, the emerging arbovirus *Zika virus* (ZIKV) was chosen as a target micro-organism. ZIKV is an enveloped virus with approximately 50 nm in diameter with a single-stranded positive RNA genome. It belongs to the family *Flaviviridae* and genus *Flavivirus*. ZIKV was first isolated in the Zika Forest, Uganda, in 1947 (Shi and Gao, 2017; Yun and Lee, 2017). The virus is transmitted to humans by hematophagous mosquitoes, mainly those belonging to the *Aedes* genus (Petersen et al., 2016). The virus circulation was reportedly restricted in endemic areas of Africa and Asia for several decades. Individuals infected by ZIKV in these areas exhibited a self-limited febrile disease. ZIKV started to gain attention from international public health communities in 2013 when an outbreak occurred in French Polynesia (Musso et al., 2015). In Brazil, the high rate of ZIKV infection was established as one of the leading causes of an increase in cases of newborns with microcephaly in the Brazilian Northeast. The severity of the reported neurological deficiency in the ZIKV cases led the World Health Organization (WHO) to propose measures to intensify ZIKV prevention. They also suggested the surveillance and investigation of the correlation between it and microcephaly in neonates and Guillain-Barré syndrome generally in adults (Larocca et al., 2016; Laura et al., 2018).

Our proposed integrative transcriptome analysis protocol using data from neural cells infected with ZIKV was able to identify shared differentially expressed genes (DEGs). The DEGs were used to identify the potential candidate genes and pathways in ZIKV pathogenesis through gene enrichment analysis and protein-protein interaction network construction. Some of the genes, pathways and biological functions were linked to neurogenesis and/or apoptosis, confirming the potential of this protocol to find key genes and pathways involved in disease pathogenesis. Moreover, the proposed *in silico* protocol performed an integrated analysis that is able to predict and identify putative biomarkers from different transcriptome data. These biomarkers could be useful to understand virus disease pathogenesis and also help the identification of candidate antiviral drugs.

2. Materials and methods

2.1. Transcriptome data of neural cells infected with Zika virus

Gene expression profiling data from ZIKV-infected human neural cells were retrieved from RNA-sequencing data from the *Gene Expression Omnibus* platform (https://www.ncbi.nlm.nih.gov/gds). The platform was extensively searched for datasets of any neural-origin cell type that was infected with ZIKV. Information about the number of samples used in each experimental condition, cell type, ZIKV strain and time of infection were collected in an Excel spreadsheet. The number of each dataset was recorded, and the raw sequence data were accessed using the online Galaxy platform (https://usegalaxy.org/). After the sequencing data was uploaded on the Galaxy platform, the "Tuxedo suite" protocol was applied to identify differentially expressed genes among uninfected and infected cells (Amrit and Ghazi, 2017). This protocol uses a set of tools to analyze the quality and variety of RNA-sequencing data, including short-readout mapping, splicing junction identification and differential expression detection of transcripts and isoforms, as well as tools for data visualization and control metrics of sequencing quality. After the protocol was completed, a list of all DEGs from each study with log fold change greater than 1.5 and p value $\leq$ 0.05 was collected in an Excel spreadsheet. Only datasets that were performed with at least two replicates for each experimental sample and that had a minimum of 100 DEGs were used for the subsequent analysis. Therefore, once a successful Galaxy run is completed we used the final DEG list identified from each study to perform an integrative analysis. This integrative analysis aims find genes and pathways which were coincident among all DEG lists and therefore more likely to be involved in the etiology.

2.2. Network construction

All networks presented in this work were built using Gephi version 0.9.2 (Bastian and Heymann, 2009). The cells and genes were listed in a comma-separated values (.csv) spreadsheet for each graph, and this file was imported into the software. Another. csv spreadsheet with the connections between the cells and genes was also imported to generate the network graph. In all networks, the node diameter is directly proportional to the edge degree. The layout was generated using algorithms based on the force of attraction and repulsion of the nodes (Force Atlas 2). The nodes were submitted to local rearrangement for better visualization of the connections between nodes.

2.3. Pathway and Gene Ontology (GO) enrichment analyses of DEGs

The Enrichr platform (http://amp.pharm.mssm.edu/Enrichr/) was used to study pathway enrichment analyses of DEGs (Kuleshov et al., 2016). The Enrichr online tool for GO (http://www.geneontology.org/; Harris et al., 2004) was used to determine the function of DEGs. Data from cellular components, biological processes and molecular functions were recorded from each set of genes. An adjusted $p \leq 0.05$ was considered statistically significant for all analyses. The lists of significant GO terms were submitted to the Reduce + Visualize Gene Ontology tool (REVIGO; Supek et al., 2011) and GO terms were visualized in an interactive graph tool. Results were exported into Cytoscape software to create graph-based visualization of the identified terms for each GO category.

2.4. Protein-protein interaction (PPI) network construction

The candidate DEGs were searched in the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database, version 11 (stringdb.org). A combined score > 0.4 was used as the criteria to establish the PPI network. The isolated nodes were deleted from the network, and the final network was downloaded as a simple tabular text file. This network file was imported into Cytoscape software (version 3.7.1) for visualization.

3. Results

3.1. Differential expressed genes on ZIKV infected cells

As illustrated in Fig. 1, a six-step immunoinformatic approach was used to identify identifying key genes and pathways useful to understand emerging virus disease pathogenesis. To validate our protocol, the emerging arbovirus ZIKV was chosen as a target micro-organism. An extensive search using public databases and indexed publications was performed to find transcriptomic studies from ZIKV-infected neural cells. The search resulted in the selection of seven studies with transcriptome data from different human neural cell lines or cerebral organoids (CO) infected at different multiplicity of infection (MOI) and different times of infection. Experimental parameters, including ZIKV strains, MOI, number of replicates and next-generation sequencing (NGS) platform, used in each study are described in Table 1. The predominant cell type used for ZIKV infection was human cortical neural progenitors (HCNP). However, transcriptome data from other ZIKV-infected cells were also identified including microglia (MCG), glioblastoma stem cells (GSC), human neural crest cells (HNCC), human peripheral neurons (PN) and CO. Different ZIKV strains (MR766, Mex-1, PRVABC59 and Dakar 41,519) and a prolonged time of infection (69.8 ± 29.96 h) were used by these studies.

To analyze changes in cellular gene expression after ZIKV infection, only datasets that had at least duplicates of the experimental conditions were used. Therefore, the dataset from ZIKV-infected microglia (Tiwari et al., 2017) was excluded from the subsequent analysis. The Tuxedo suite protocol was performed using the transcriptome, and the

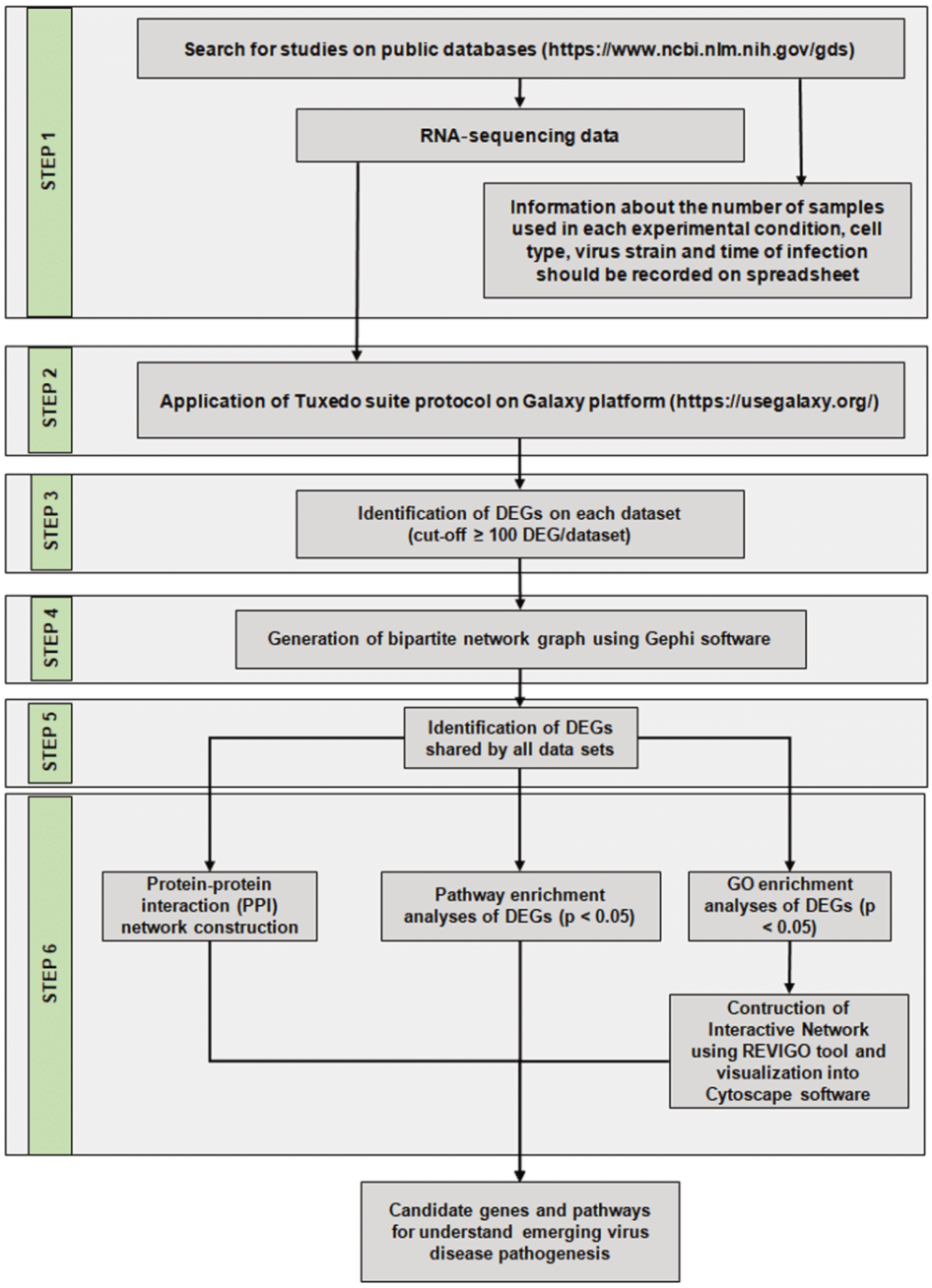


Fig. 1. Schematic representation of entire *in silico* approach for identification of key genes and pathways useful to understand emerging virus disease pathogenesis.

statistically significant DEGs were recorded. Applying the Tuxedo protocol on datasets generated two sets of results. The first one comprised the results from HCNP and GSC datasets there was a high number of DEGs in these cells after ZIKV infection. The second group was related to HNCC, PN and CO datasets, and it was characterized by a low number of DEGs after ZIKV infection (Supplementary Table 1).

To construct a final list with the DEGs in HCNP after ZIKV infection, the lists of downregulated and upregulated genes obtained in each dataset (GSE78711, GSE80434 and GSE93385) were compared to use only the genes that were considered differentially expressed in at least two studies. For ZIKV-infected CO, only the dataset derived from 3-day infection was used, because the dataset from 5-day infected CO had a low number of DEGs ($\leq$ 100). The final list with the identified DEGs (n = 11,049) was used to build a bipartite network graph to better represent the effect of ZIKV infection on each cell type (Table 2). The network graph for upregulated genes was composed of 5311 nodes, with 5306 genes and 5 cells all connected by 6424 edges with the same weight (w) = [1]. A network graph for downregulated genes was composed of 3688 nodes, with 3683 genes and 5 cells all connected by 4611 edges with the same weight (w) = [1]. The network of upregulated and downregulated genes showed a very similar pattern (Fig. 2A). Network analysis revealed a marked change in GSC and HCNP gene expression once these cells exhibited a high number of DEGs. There were also changes in HNCC, PN and CO gene expression, but the number of DEGs in these cells after ZIKV infection was relatively low. Furthermore, the network analysis revealed a high number of DEGs shared by GSC and HCNPC, but a low number of common DEGs among all cells. Data analysis also demonstrated that the majority of upregulated DEGs were expressed in one cell type (4338/ 81.71 %). Only 851 genes (16.02 %) were upregulated in two cell types, 94 genes (1.77 %)

Table 1
Summary of datasets and sample details.

Study	GEO Acces	Nº of samples		Cell Type	ZIKV Strain	MOI	Time of infection	NGS Platform
		Control Cells	Infected cells					
Tang et al. (2017)	GSE78711	3	3	Human cortical neural progenitors	MR766	< 0.1	56 h	GPL18573 IlluminaNextSeq 500 GPL15520 IlluminaMiSeq
Zhang et al. (2016)	GSE80434	4	4 2	Human cortical neural progenitors	MR766 FSS13025	0.02 0.04	64 h 64 h	GPL11154 IlluminaHiSeq 2000 GPL15520 IlluminaMiSeq GPL18573 IlluminaNextSeq 500
McGrath et al (2017)	GSE93385	9	9	Human cortical neural progenitors	Mex1 – 7	10	5 days	GPL18460 IlluminaHiSeq 1500
Tiwari et al. (2017)	GSE93870	1	1	Microglia	MR766	1	24 h	GPL18573 IlluminaNextSeq 500
Zhu et al. (2017)	GSE102924	3	3	Glioblastoma stem cells	Dakar 41,519	5	36 – 48 h	GPL21290 IlluminaHiSeq 3000
Oh et al. (2018)	GSE87750	3	4	Human neural crest cells	PRVABC59	0.4	65 h	GPL16791 IlluminaHiSeq 2500 GPL21290 IlluminaHiSeq 3000
	GSE87750	2	4	Human Peripheral neurons	PRVABC59	0.4	65 h	GPL16791 IlluminaHiSeq 2500 GPL21290 IlluminaHiSeq 3000
Watanabe et al. (2017)	SE104279	3	3	Cerebral organoids	PRVABC59	0.3	3 or 5 days	GPL21290 IlluminaHiSeq 3000

Table 2
Number of Differentially expressed genes used to construct interaction networks among ZIKV infected cells.

Cell type	Up	Dow	Total
Human cortical neural progenitors	3586	1821	5407
Glioblastoma stem cells	2487	2537	5024
Human neural crest cells	188	137	325
Peripheral neurons	99	44	143
Cerebral organoids	76	74	150

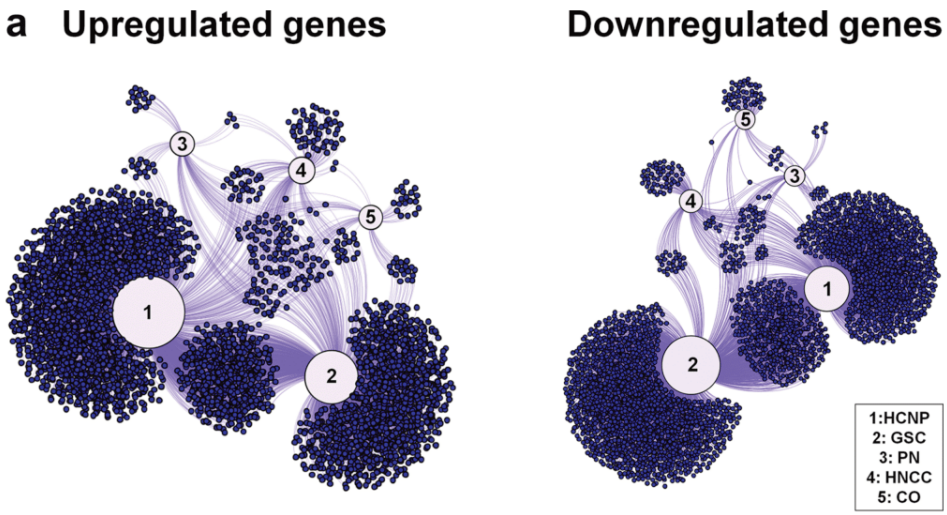
in three, 23 genes (0.43 %) in four and 3 genes (0.05 %) in all cell types. Regarding downregulated DEGs, 3207 genes (82.59 %) were expressed in only one cell type, 630 genes (16.22 %) in two, 40 genes (1.03 %) in three and 6 genes (0.15 %) in four (Fig. 2B).

3.2. Protein-Protein interaction network using the DEGs shared by all cells types after ZIKV infection

The results showed that the differential effect of ZIKV on cell transcriptome could be related to the differentiation status of the cell. Whereas the CO is composed of several cell types, using DEGs from this type of tissue can induce an error in the integrative analysis of the transcriptome of ZIKV-infected cells. Therefore, a new analysis was performed using only the DEGs identified in HCNP, GSC, HNCC and PN cells after ZIKV infection. Using this approach, we identified DEGs expressed in all cells (Fig. 3A; Table 3; Supplementary Table 2). The PPI network was constructed using the DEGs shared by all cells. The PPI network of upregulated DEGs was composed of 14 nodes and 28 edges, and the PPI network of downregulated genes was composed of 4 nodes and 6 edges. DEGs that encode proteins without any interaction were excluded from the PPI networks (upregulated: CLK1, EIF1B, GOT1, MXD1, SEC11C, SEC24D, SLC25A25, STMN4, TSPYL2 and TUFT1; downregulated: PRTG and IGFBPL1). To investigate the expression level of the target genes of ZIKV infected host cells, the RNA-seq transcriptomic expression data sets were analyzed and the fold change values were calculated comparing the mock and ZIKV treated cells. The DEGs shared by all cells were highly differentially modulated compared with the control condition (Fig. 3B).

3.3. Pathway and Gene Ontology enrichment analysis using the DEGs shared by all cells types after ZIKV infection

Pathway enrichment analysis using the upregulated DEGs indicated that they were involved in protein processing in the endoplasmic reticulum (ER), and downregulated DEGs were involved in DNA replication and cell cycle (Fig. 4A). Using the upregulated DEGs shared by all cells, it was possible to identify several biological processes related to the response to ER stress (GO:1905898, GO:1902235, GO:0034976, GO:1902043, GO:0036498, GO:1900102 and GO:0036499; Fig. 4B). There were also biological processes related to type I interferon signaling pathway (GO:0060337 and GO:0071357) and positive regulation of transcription from RNA polymerase II promoter (GO:0045944). Molecular functions related to upregulated DEGs were core promoter binding (GO:0000987, GO:0000978, GO:0001046, GO:0001047 and GO:1990837) and regulation of transcription (GO:0044212, GO:0000976 and GO:0000982). For downregulated DEGs, DNA replication (GO:0006260) and mitotic cell cycle phase transition were the identified biological process (GO:0000082, GO:0044843, GO:0044772). For downregulated DEGs, the identified molecular functions were related to DNA helicase activity (GO:0003678 and GO:0004003). The upregulated DEG list did not produce any significant cellular component ($p \geq 0.05$). However, the downregulated DEG list indicated chromosome telomeric region (GO:0000781), nuclear chromosome-telomeric region (GO:0000784) and nuclear chromosome part (GO:0044454) as cellular components related to these genes (Fig. 4B).



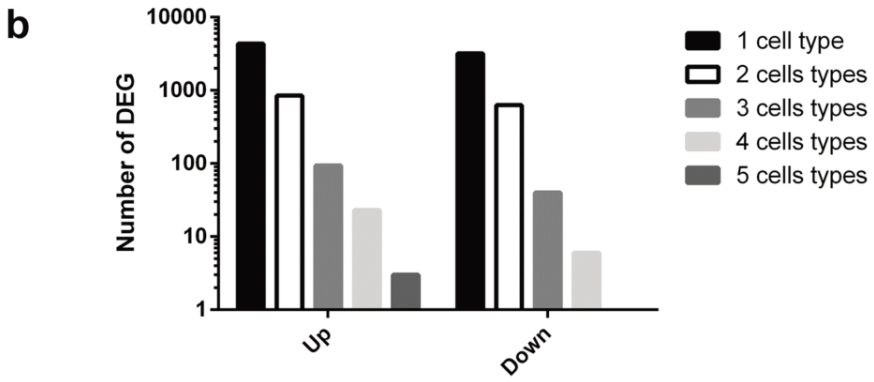


Fig. 2. The network of up- and downregulated genes after ZIKV infection of different cells types. (A) The bipartite network graph shows a spatially connected network among differentially expressed genes (DEGs) and cells types after ZIKV infection. Each node represents a gene or cell type. The larger nodes represent cell types and the connected smaller nodes represent DEG. The layout was generated using a force-based algorithm followed by manual rearrangement to better visualize the connections. A total of 11,049 DEGs (6436 upregulated and 4613 downregulated) and five cell types are represented. (B) Number of DEGs expressed in ZIKV-infected cells.

3.4. Search for published studies on Pubmed

To verify if the identified candidate genes and pathways had a correlation with ZIKV pathogenesis, a bibliographic survey was conducted in the PubMed (https://www.ncbi.nlm.nih.gov/pubmed). The search strategy for original articles was composed of two components: a) the gene, pathway or molecular function name and b) *Zika virus*. To identify relevant indexed studies the search filter was structured by combining the descriptors selected by boolean operators AND and NOR, as well as the Title / Abstract [TIAB] algorithm. Reviews are excluded from this analysis. Among the upregulated DEGs, it was identified 26 published studies that investigated the correlation of 14 upregulated DEGs with ZIKV pathogenesis (Supplementary Table 3). ATF3 and STAT1 genes exhibit the highest number of hits. For downregulated DEGs, it was identified 8 published studies, being the CENPF the most frequent gene (Fig. 5A). Regarding the pathways and biological processes, all of them had at least one published study showing the modulation of these pathways or biological processes after ZIKV infection (Fig. 5B). In a special way, 7 of 15 identified published studies showed a correlation of the inhibition of the cell cycle pathway with ZIKV pathogenesis. The ER stress has the highest number of hits among the significant biological process related to DEGs.

4. Discussion

Outbreaks of emerging and/or re-emerging viral infections are common threats to human health. When a new virus was discovery and associated with human disease, it is necessary to understand virus pathogenesis as soon as possible to develop countermeasures to the disease such as vaccines and antiviral drugs (Afrough et al., 2019). The *omics* technologies, especially the transcriptome studies could generate thousands of information about the gene regulation on cells after virus infection and therefore, it can help researches to gain insights about the pathogen-host interaction (Berkhout and Coombs, 2013; Jean Beltran et al., 2017). In this way, we propose a novel and friendly *in silico* integrative protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis. The main advantage of this protocol is the use of galaxy platform to identify DEGs from different transcriptome studies. Galaxy is a web open-source, web-based platform for intensive biomedical research. Galaxy is used by several researches, because this platform has thousands of different tools for many different scientific fields. It can allow the analysis of large sequencing datasets by researchers without programming skills (Afgan et al., 2018). In this way, our proposed *in silico* protocol could permit the researcher DEGs from different datasets and also integrate the data. Data integration is vital to connect all the different data sources to exploit the value of insights because it reduces data complexity and also increases the value of data through unified systems. Once all the DEGs are available in a single place in real-time, researches would be able to use this protocol to integrate different transcriptome data and also perform the identification of candidate genes and pathways involved in virus pathogenesis.

To validate our protocol, the emerging arbovirus ZIKV was chosen as a target micro-organism. Since ZIKV is considered to be a risk factor for microcephaly development in humans, several studies attempted to understand the effect of ZIKV infection on different cell types. It is

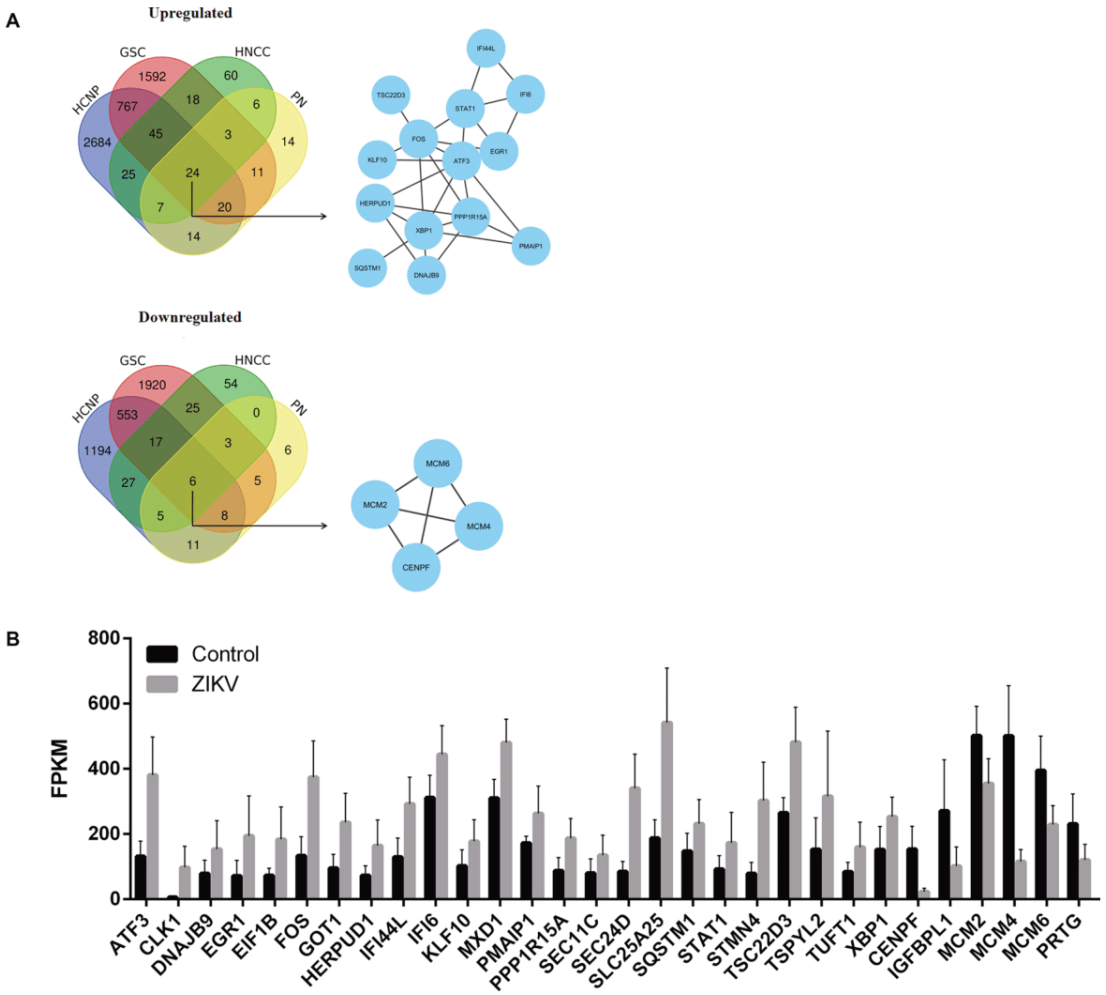


Fig. 3. Integrative transcriptome analysis of ZIKV-infected neural cells. (A) The Venn diagram shows the degree of overlap of the upregulated and downregulated genes in ZIKV-infected cells. The Venn diagrams were constructed using the DEGs identified in each cell type using the Calculate and draw custom Venn diagrams tool available at http://bioinformatics.psb.ugent.be/webtools/Venn/. The protein-protein interaction (PPI) networks for the common upregulated and downregulates genes were constructed using the STRING online database. DEG-encoded proteins without any interaction were excluded from the network. The network file was visualized with Cytoscape software. Abbreviations – HCNP: human cortical neural progenitors; GSC: glioblastoma stem cells, HNCC: human neural crest cells; PN: human peripheral neurons. (B) Expression levels (FPKM) of DEGs on ZIKV infected cells.

Table 3
Number of Differentially expressed genes indentified by integrative analysis.

Group	Up	Dow	Total
All cells	24	6	30
Undifferentiated cells (HCNP and GSC)	856	584	1440
Differentiated cells (HNCC and PN)	40	14	54

crucial to characterize the gene expression profile in ZIKV-infected cells, including neural-origin cells. Therefore, some studies published the transcriptome of ZIKV-infected cells to identify genes that could be involved in the development of neurological disorders induced by ZIKV. A search of the literature identified some transcriptome studies from ZIKV-infected neural cells, particularly neural progenitor cells (Table 1). This finding may be associated with the damage that ZIKV infection induces in this cell type during cell differentiation and in the development of the cerebral compartment.

Many studies presented evidence of disrupted brain development after infection of neural stem cells and neuroprogenitors by ZIKV infectious particles (Russso and Beltrão-Braga, 2017). However, other studies also showed the effect of ZIKV infection on GSC, HNCC, PN and CO transcriptomes (Table 1). In these studies, ZIKV can infect the target cells, but the outcomes are different. Data from many studies showed more severe cell death and higher viral load in undifferentiated cells. ZIKV infection and apoptosis rates are significantly higher in GSC compared to proliferating tumor or differentiated cells (Zhu et al., 2017). In the same context, ZIKV infection and cell death occur preferentially in neural progenitors present in CO (Watanabe et al., 2017). Zhang et al. (2016) reported higher viral load and cell death in ZIKV-infected HCNP. Regarding the effect of ZIKV infection on differentiated cells, the virus replicates approximately 10-fold more in HNCC compared to PN. There is also a significantly higher rate of apoptosis in HNCC compared to PN (Oh et al., 2018).

Applying the Tuxedo protocol on the identified datasets allowed the identification of DEGs in these cells after ZIKV infection (Table 2). However, the integrative analysis of the transcriptomes showed a differential effect on gene expression depending on the infected cell type. HCNP and GSC showed more DEGs when compared to other cells (Fig. 2A). Additionally, as HCNP and GSC shared many DEGs, data suggest a similar effect of ZIKV infection on the transcriptome of these cells. The data presented by Kaid et al. (2018) showed novel *in vitro* and *in vivo* evidence about the use of a Brazilian ZIKV strain as an oncolytic therapy to treat aggressive human embryonal tumors of the central nervous system. Therefore, the most prominent oncolytic effects of

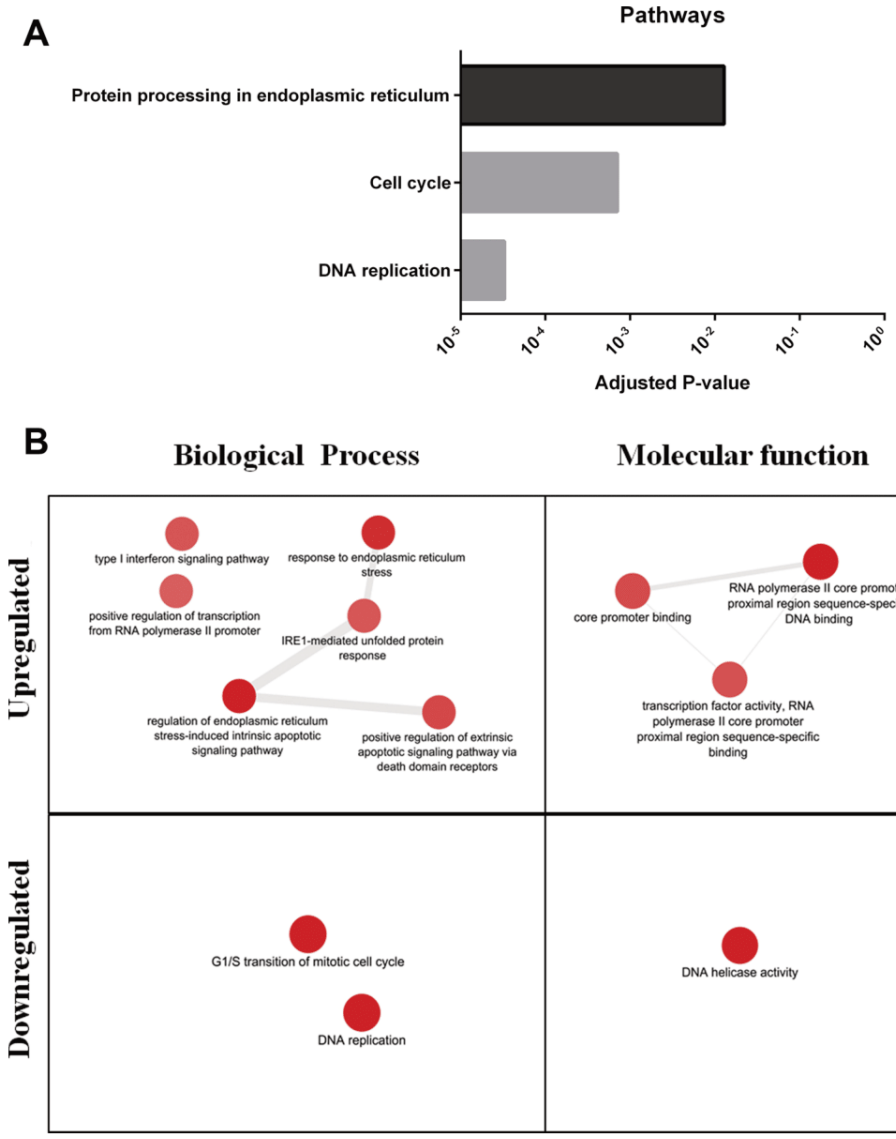


Fig. 4. Enriched Kyoto Encyclopedia of Genes and Genome (KEGG) pathways and Gene Ontology (GO) terms identified in integrative analysis of ZIKV-infected cells. A) DEG-encoded proteins identified from the protein-protein interaction (PPI) network in each cell type after ZIKV infection were used to identify significant pathways (adjusted $p < 0.05$) using the Enrichr platform (http://amp.pharm.mssm.edu/Enrichr/). B) Visualization of the biological process and molecular function GO annotations in ZIKV-infected cells using REVIGO. Bubble color indicates the p-value; bubble size indicates the frequency of the GO term. Highly similar GO terms are linked by edges in the graph, where the line width indicates the degree of similarity.

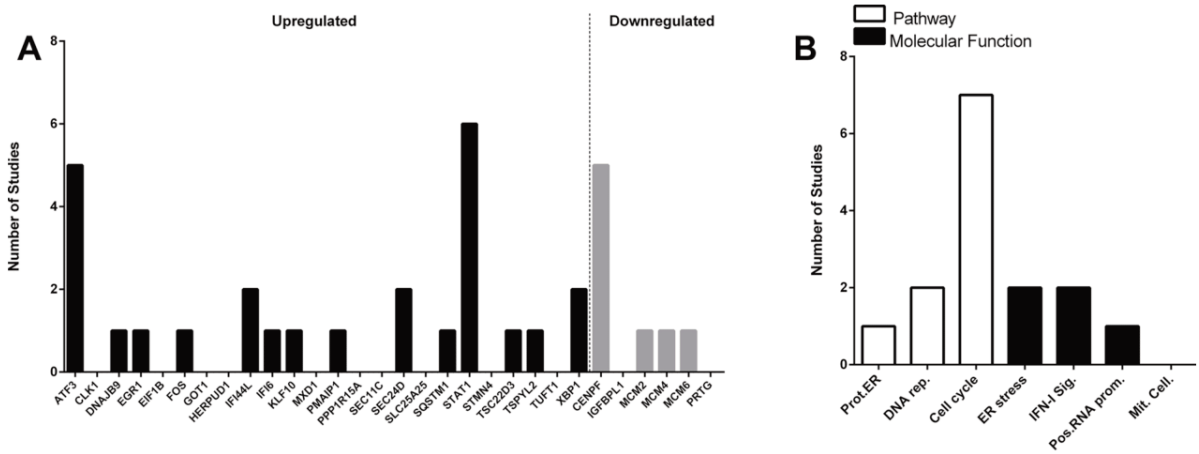


Fig. 5. Search hits for terms related to DEGs, pathways, gene ontology terms and ZIKV identified in PubMed. (A) Search for upregulated and downregulated DEGs. (B) Search for pathways and gene ontology terms. Legend: *Prot.ER*: protein processing in the endoplasmic reticulum, *DNA rep.*: DNA replication; *ER stress*: endoplasmic reticulum stress; *IFN-I sig.*: type I interferon signaling pathway; *Mit. Cell.*: mitotic cell cycle phase.

ZIKV infection was observed in tumors generated by cell lines with highly similar characteristics to early neural stem and neuroprogenitor cells. The few identified DEGs shared by all cell types could be attributed to the differentiation status of these cells. HCNP and GSC are undifferentiated, while HNCC and PN are differentiated cells. Therefore, the number of shared DEGs between cells under the same differentiation status should be greater than the DEG shared by undifferentiated and differentiated cells (Table 3). There were few DEGs shared by the four cell types. Consequently, few pathways, molecular functions and biological processes were enriched using this set of genes (24 upregulated and 6 downregulated genes).

The searches for published studies that investigate the expression of DEGs on ZIKV infected cells also confirm the potential of this *in silico* protocol to gain insights about virus pathogenesis. Eighteen of thirty identified DEGs (60 %) had at least one published study (Fig. 5A). Some of these DEGs have a direct correlation with antiviral response, neurogenesis and apoptosis. STAT1 and ATF3 were the upregulated DEGs with the highest number of hits. The increase of STAT1 expression on infected cells is a common antiviral response on virus-infected cells. Like other flaviviruses, ZIKV can antagonize the type I interferon pathway through interference with STAT-1 signaling. This immune evasion could be responsible to increase viral replication and apoptosis on infected cells of neural or non-neural origin (Chen et al., 2018). The ATF3 is a member of the activation transcription factor/cAMP-responsive element-binding protein family of transcription factors in the mammalians. Changes in ATF3 expression was observed on ZIKV infected cells (Moni and Lio, 2017; Zanini et al., 2018). This transcription marker is activated on injured peripheral neurons and therefore it is linked to neural cell survival (Lindå et al., 2011; Mahar and Cavall, 2018). CENPF is the downregulated DEG with the highest number of hits. The CENPF encoded protein associates with the centromere-kinetochore complex and its nuclear localization suggests that it play a role in chromosome segregation during mitosis (Varis et al., 2006). Several studies had demonstrated the decrease of CENPF on ZIKV neural infected cells, showing the correlation of CENPF downregulation with microcephaly (Dang et al., 2019; Moni and Lio, 2017; Paul et al., 2018; Zhang et al., 2016b).

Protein processing in the ER was the only significant pathway identified using the upregulated DEGs. This data is in agreement with several molecular functions related to the response to ER stress. Our literature search was able to identify several published studies that investigate the role of ER stress on ZIKV pathogenesis (Fig. 5B). The ZIKV can induce the expression of ER stress sensors in mice and human neural cells (Singh et al., 2018; Tan et al., 2018). Like other flaviviruses, ZIKV depends on the ER for its translation, replication and packaging. The biological events related to ZIKV replication in ER membranes induce modifications that could trigger ER stress. Therefore, the unfolded protein response should be activated to reduce stress on ZIKV-infected neural cells. This event could be crucial to the microcephaly development in newborns from ZIKV-infected mothers (Tan et al., 2018). Enrichment analysis using the downregulated DEGs identified cell cycle and DNA replication as important affected pathways with deregulated biological processes in all ZIKV-infected cells. Hammack et al. (2019), showed that ZIKV can impair the cell cycle by inducing DNA double-strand breaks in the host genome. This mechanism could be especially important for HNCP, because the growth arrest on these cells could impair brain development. Our data strengthen the idea that cell cycle arrest and DNA replication could be an important mechanism in all neural ZIKV-infected cells, independent of the differentiation status but with a more detrimental effect on the undifferentiated cells. Among the six downregulated DEGs identified by our analysis, four are related to the regulation of early events of DNA replication (MCM2, MCM4 and MCM6) on neocortex development. The proteins encoded by these genes are part of the minichromosomal maintenance complex (MMC). This complex consists of six subunits (MCM2 through MCM7) with ATP-dependent helicase activity. The main function of this complex is to attach to double-stranded DNA and induce the binding of multiple replication factors to the replisome. Defective DNA replication is involved in the pathophysiology of some cortical malformations like microcephaly. Therefore, decreased MCM2, MCM4 and MCM6 expression could disrupt the MMC in ZIKV-infected cell and thus affect the cell cycle. Some works demonstrated a reduction of proliferation and apoptosis in ZIKV-infected neural progenitor cells (Li et al., 2016; Tang et al., 2017). These data in connection with the downregulation of CENPF strengths the importance of maintaining the cell cycle to inhibit the deleterious effects caused by the multiplication of ZIKV in neural progenitor cells.

Using ZIKV as a model to validate our protocol we were able to perform an integrative transcriptome analysis of different ZIKV-infected cells. This analysis allowed us the identification of gene expression profiles in each cell type. The results showed that each cell type exhibited a particular response to ZIKV infection. Indeed, most identified DEGs were present in only one cell type. The common DEGs are related to two main critical pathways (DNA replication and ER stress) that could be involved in defects in brain development in ZIKV-infected newborns. Therefore, these DEGs can be subsequently used as potential targets for the development of new drugs that allow effective treatment of the disease. In conclusion, the proposed *in silico* protocol performed an integrated analysis that is able to predict and identify putative biomarkers from different transcriptome data. These biomarkers could be useful to understand virus disease pathogenesis and also help the identification of candidate antiviral drugs.

Funding

This work was supported in part by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) (grants APQ-01165–16 and PPM-00399–18). G.A.P.S. and E.A.S. received fellowships from FAPEMIG. F.R.O. received fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - PIBIC). L.F.L.C. and J.S.A. received PQ fellowship from CNPq. This study was financed in part by the Coordenação deAperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES). Finance Code 001.

CRediT authorship contribution statement

G.A.P.S. and L.F.L.C. conceived and planned the experiments. G.A.P.S., E.A.S., F.R.O. carried out the experiments. G.A.P.S., E.A.S., F.R.O., L.C.C.M., J.S.A., L.F.L.C. contributed to the interpretation of the results. G.A.P.S. and L.F.L.C took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2020.197986.

References

Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., Blankenberg, D., 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 46, W537–W544. https://doi.org/10.1093/nar/gky379.

Afrough, B., Dowall, S., Hewson, R., 2019. Emerging viruses and current strategies for vaccine intervention. Clin. Exp. Immunol. 196, 157–166. https://doi.org/10.1111/cei.13295.

Amrit, F.R.G., Ghazi, A., 2017. Transcriptomic analysis of C. Elegans RNA sequencing data through the Tuxedo suite on the galaxy project. J. Vis. Exp. 122, 1–12. https://doi.org/10.3791/55473.

Bastian, M., Heymann, S., 2009. Gephi : an open source software for exploring and manipulating networks. Fond. Maison des Sci. l'Homme.

Berkhout, B., Coombs, K.M., 2013. Quantitative omics and its application to study virus-host interactions-a new frontier. Front. Microbiol. 4 (2012–2013). https://doi.org/10.3389/fmicb.2013.00031.

Chen, Jian, Yang, Y.F., Yang, Y., Zou, P., Chen, Jun, He, Y., Shui, S.L., Cui, Y.R., Bai, R., Liang, Y.J., Hu, Y., Jiang, B., Lu, L., Zhang, X., Liu, J., Xu, J., 2018. AXL promotes Zika virus infection in astrocytes by antagonizing type i interferon signalling. Nat. Microbiol. 3, 302–309. https://doi.org/10.1038/s41564-017-0092-4.

Cinegaglia, N.C., Andrade, S.C.S., Tokar, T., Pinheiro, M., Severino, F.E., Oliveira, R.A., Hasimoto, E.N., Cataneo, D.C., Cataneo, A.J.M., Defaveri, J., Souza, C.P., Márques, M.M.C., Carvalho, R.F., Coutinho, L.L., Gross, J.L., Rogatto, S.R., Lam, W.L., Jurisica, I., Reis, P.P., 2016. Integrative transcriptome analysis identifies deregulated microRNA-transcription factor networks in lung adenocarcinoma. Oncotarget 7, 28920–28934. https://doi.org/10.18632/oncotarget.8713.

Dang, J.W., Tiwari, S.K., Qin, Y., Rana, T.M., 2019. Genome-wide integrative analysis of zika-virus-Infected neuronal stem cells reveals roles for MicroRNAs in cell cycle and stemness. Cell Rep. 27, 3618–3628.e5. https://doi.org/10.1016/j.celrep.2019.05.059.

Ellwanger, J.H., Chies, J.A.B., 2016. Emergent diseases in emergent countries: we must study viral ecology to prevent new epidemics. Braz. J. Infect. Dis. 20, 403–404. https://doi.org/10.1016/j.bjid.2016.02.003.

Hammack, Christy, Ogden, S.C., Madden Jr., J.C., Medina, A., Xu, C., Philips, E., Son, Y., Cone, A., Giovinazzi, S., Didier, R.A., Gilbert, D., 2019. Zika virus infection induces DNA damage response in human neural progenitors that enhaces viral replication. J. Virol. https://doi.org/10.1128/JVI.00638-19.

Harris, M.A., et al., 2004. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res. 32, 258–261. https://doi.org/10.1093/nar/gkh036.

Hoshida, Y., Nijman, S.M.B., Kobayashi, M., Chan, J.A., Brunet, J., Chiang, D.Y., Villanueva, A., Newell, P., Hashimoto, M., Watanabe, G., Gabriel, S., Friedman, S.L., 2009. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res. 69, 7385–7392. https://doi.org/10.1158/0008-5472.CAN-09-1089.Integrative.

Jean Beltran, P.M., Federspiel, J.D., Sheng, X., Cristea, I.M., 2017. Proteomics and integrative omic approaches for understanding host–pathogen interactions and infectious diseases. Mol. Syst. Biol. 13, 922. https://doi.org/10.15252/msb.20167062.

Jin, C., Kang, H., Ryu, J.R., Kim, S., Zhang, Y., Lee, Y., Kim, Y., Han, K., 2018. Integrative brain transcriptome analysis reveals region-specific and broad molecular changes in Shank3-Overexpressing mice. Front. Mol. Neurosci. 11, 1–14. https://doi.org/10.3389/fnmol.2018.00250.

Kaid, C., Goulart, E., Caires-j, L.C., Ventini, D.C., Astray, R.M., Assoni, A.F., 2018. Zika virus selectively kills aggressive human embryonal CNS tumor cells in vitro and in vivo. Cancer Res. 26, 3363–3375. https://doi.org/10.1158/0008-5472.CAN-17-3201.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., Mcdermott, M.G., Monteiro, C.D., Gundersen, W., Ma, A., 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44, 90–97. https://doi.org/10.1093/nar/gkw377.

Larocca, R.A., Abbink, P., Peron, J.P.S., Zanotto, P.M.D.A., Iampietro, M.J., Badamchi-zadeh, A., Boyd, M., Ng, D., Kirilova, M., Nityanandam, R., Mercado, N.B., Li, Z., Moseley, E.T., Bricault, C.A., Borducchi, E.N., Giglio, P.B., Jetton, D., Neubauer, G., Nkolola, J.P., Maxfield, L.F., Barrera, R.A.D.La, Jarman, R.G., Eckels, K.H., Michael, N.L., Thomas, S.J., Barouch, D.H., Badamchi-zadeh, A., Boyd, M., Ng, D., Kirilova, M., Nityanandam, R., Mercado, N.B., Maxfield, L.F., Barrera, R.A.D.La, Jarman, R.G., Eckels, K.H., Neubauer, G., Joseph, P., Michael, N.L., Thomas, S.J., Barouch, D.H., 2016. Vaccine protection agains Zika virus from Brazil. Nature 000, 1–19. https://doi.org/10.1038/nature18952.

Laura, A., Victoria, T., María, L., Morales, A., Groisman, B., Baricalla, A., Fabbri, C., Sinchi, A., Alonso, A., Gonzalez, C., Ledesma, B., Masi, P., Silva, M., Israilev, A., Rocha, M., Quaglia, M., Paz, M., Rosa, B., Pablo, L., 2018. Surveillance of microcephaly and selected brain anomalies in Argentina : relationship with Zika virus and other congenital infections. Birth Defects Res. 110, 1016–1026. https://doi.org/10.1002/bdr2.1347.

Li, H., Saucedo-cuevas, L., Regla-nava, J.A., Chai, G., Tang, W., Terskikh, A.V., Shresta, S., Gleeson, J.G., 2016. Zika Virus Infects Neural Progenitors in the Adult Mouse Brain and Alters Proliferation 19. pp. 593–598.

Liao, Q., Wang, J., Pei, Z., Xu, J., Zhang, X., 2017. Identification of miRNA-mRNA crosstalk in CD4+ T cells during HIV-1 infection by integrating transcriptome analyses. J. Transl. Med. 15, 1–11. https://doi.org/10.1186/s12967-017-1130-y.

Lindå, H., Sköld, M.K., Ochsmann, T., 2011. Activating transcription factor 3, a useful marker for regenerative response after nerve root injury. Front. Neurol. (MAY), 1–6. https://doi.org/10.3389/fneur.2011.00030.

Mahar, Marcus, Cavall, Valeria, 2018. Intrinsic mechanisms of neuronal axon regeneration. Nat. Rev. Neurosci. 19, 323–337. https://doi.org/10.1038/s41583-018-0001-8.

Moni, M.A., Lio, P., 2017. Genetic profiling and comorbidities of zika infection. J. Infect. Dis. 216, 703–712. https://doi.org/10.1093/infdis/jix327.

Moslehi, R., Mills, J.L., Signore, C., Kumar, A., Ambroggio, X., Dzutsev, A., 2013. Integrative transcriptome analysis reveals dysregulation of canonical cancer molecular pathways in placenta leading to preeclampsia. Sci. Rep. 3, 1–11. https://doi.org/10.1038/srep02407.

Musso, D., Cao-Lormeau, V.M., Gubler, D.J., 2015. Zika virus: following the path of dengue and chikungunya? Lancet 386, 243–244. https://doi.org/10.1016/S0140-6736(15)61273-9.

Oh, Y., Zhang, F., Wang, Y., Lee, E.M., Choi, I.Y., Lim, H., Mirakhori, F., Li, R., Huang, L., Xu, T., Wu, H., Li, C., Qin, F., Wen, Z., Wu, Q., Tang, H., Xu, Z., Song, H., Ming, G., Lee, G., Malvin, H., Orleans, N., Sciences, B., Biology, C., 2018. Zika virus directly infects peripheral neurons and induces cell death. Nat. Neurosci. 20, 1209–1212. https://doi.org/10.1038/nn.4612.Zika.

Olival, K.J., Hosseini, P.R., Zambrana-Torrelio, C., Ross, N., Bogich, T.L., Daszak, P., 2017. Host and viral traits predict zoonotic spillover from mammals. Nature 546, 646–650. https://doi.org/10.1038/nature22975.

Paul, A.M., Acharya, D., Neupane, B., Ashely Thompson, E., Gonzalez-Fernandez, G., Copeland, K.M., Garrett, M., Liu, H., Lopez, M.E., Cruz, Mde, Flynt, A., Liao, J., Guo, Y.L., Gonzalez-Fernandez, F., Vig, P.J.S., Bai, F., 2018. Congenital Zika virus infection in immunocompetent mice causes postnatal growth impediment and neurobehavioral deficits. Front. Microbiol. 9, 1–12. https://doi.org/10.3389/fmicb.2018.02028.

Petersen, E., Wilson, M.E., Touch, S., McCloskey, B., Mwaba, P., Bates, M., Dar, O., Mattes, F., Kidd, M., Ippolito, G., Azhar, E.I., Zumla, A., 2016. Rapid spread of zika virus in the Americas - implications for public health preparedness for mass gatherings at the 2016 Brazil olympic games. Int. J. Infect. Dis. 44, 11–15. https://doi.org/10.1016/j.ijid.2016.02.001.

Russso, B.R., Beltrão-Braga, P., 2017. Biochemical and Biophysical Research Communications the impact of Zika virus in the brain. Biochem. Biophys. Res. Commun. 492, 603–607. https://doi.org/10.1016/j.bbrc.2017.01.074.

Shi, Y., Gao, G.F., 2017. Structural biology of the zika virus. Trends Biochem. Sci. 1–14. https://doi.org/10.1016/j.tibs.2017.02.009. xx.

Singh, P.K., Khatri, I., Jha, A., Pretto, C.D., Spindler, K.R., Arumugaswami, V., Giri, S., Kumar, A., Bhasin, M.K., 2018. Determination of system level alterations in host transcriptome due to Zika virus (ZIKV) Infection in retinal pigment epithelium. Sci. Rep. 8, 1–16. https://doi.org/10.1038/s41598-018-29329-2.

Supek, F., Bosnjak, M., Skunca, N.T.S., 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS Negl. Trop. Dis. 6, 1–9. https://doi.org/10.1371/journal.pone.0021800.

Tan, Z., Zhang, W., Sun, J., Fu, Z., Ke, X., Zheng, C., Zhang, Y., Li, P., Liu, Y., Hu, Q., Wang, H., Zheng, Z., 2018. ZIKV infection activates the IRE1-XBP1 and ATF6 pathways of unfolded protein response in neural cells. J. Neuroinflammation 1–16.

Tang, H., Hammack, C., Ogden, S.C., Wen, Z., Qian, X., 2017. Zika virus infects human cortical neural precursors and attenuates their growth. Cell Stem Cell 18, 587–590. https://doi.org/10.1016/j.stem.2016.02.016.Zika.

Tiwari, S.K., Dang, J., Qin, Y., Lichinchi, G., Bansal, V., Rana, T.M., 2017. Zika virus infection reprograms global transcription of host cells to allow sustained infection. Emerg. Microbes Infect. 6, e24–10. https://doi.org/10.1038/emi.2017.9.

Varis, A., Salmela, A.L., Kallio, M.J., 2006. Cenp-F (mitosin) is more than a mitotic marker. Chromosoma 115, 288–295. https://doi.org/10.1007/s00412-005-0046-0.

Watanabe, M., Buth, J.E., Vishlaghi, N., Taxidis, J., Khakh, B., Coppola, G., Pearson, C.A., Gong, D., Dai, X., Damoiseaux, R., Aliyari, R., Liebscher, S., Schenke-layland, K., Caneda, C., Huang, E.J., Zhang, Y., Cheng, G., Geschwind, D.H., Golshani, P., Novitch, B.G., Disabilities, D., Sciences, B., Behavior, H., Pharmacology, M., Genetics, M., 2017. Self-organized cerebral organoids with human specific features predict effective drugs to combat Zika virus infection. Cell Rep. 21, 517–532. https://doi.org/10.1016/j.celrep.2017.09.047.Self-organized.

Yun, S.-I., Lee, Y.-M., 2017. Zika virus: an emerging flavivirus. J. Microbiol. 55, 204–219. https://doi.org/10.1007/s12275-017-7063-6.

Zanini, F., Pu, S.Y., Bekerman, E., Einav, S., Quake, S.R., 2018. Single-cell transcriptional dynamics of flavivirus infection. Elife 7, 1–21. https://doi.org/10.7554/eLife.32942.

Zhang, F., Hammack, C., Ogden, S.C., Cheng, Y., Lee, E.M., Wen, Z., Qian, X., Nguyen, H.N., Li, Y., Yao, B., Xu, Miao, Xu, T., Chen, L., Wang, Z., Feng, H., Huang, W., Yoon, K., Shan, C., Huang, L., Qin, Z., Christian, K.M., Shi, P., Xu, Mingjiang, Xia, M., Zheng, W., Wu, H., Song, H., Tang, H., 2016a. Molecular signatures associated with ZIKV exposure in human cortical neural progenitors. Nucleic Acids Res. 44, 8610–8620. https://doi.org/10.1093/nar/gkw765.

Zhang, N., Zhang, N., Qin, C.F., Liu, X., Shi, L., Xu, Z., 2016b. Zika virus disrupts neural progenitor development and leads to microcephaly in mice. Cell Stem Cell 19, 120–126. https://doi.org/10.1016/j.stem.2016.04.017.

Zhu, Z., Gorman, M.J., Mckenzie, L.D., Chai, J.N., Hubert, C.G., Prager, C., Fernandez, E., Richner, J.M., Zhang, R., Shan, C., Tycksen, E., Wang, X., Shi, P.Y., Diamond, M.S., Rich, J.N., Chheda, M.G., 2017. Zika virus has oncolytic activity against glioblastoma stem cells. J. Exp. Med. 214, 2843–2857.

3. ARTIGO II: Integrative transcriptome analysis of SARS-CoV-2 human-infected cells combined with deep learning algorithms identifies two potential cellular targets for the treatment of coronavirus disease

Este é um artigo original publicado no periódico "Brazilian Journal of Microbiology" em novembro de 2022.

3.1 Introdução

Assim como o surto de ZIKV em 2015 trouxe a atenção da comunidade cientifica para um vírus emergente, devido as sequelas neurológicas da doença em adultos e recém-nascidos, em dezembro de 2019 o mundo passou a observar com preocupação a emergência de uma doença respiratório na China (GORBALENYA *et al.*, 2020; HUI *et al.*, 2020; WU *et al.*, 2020). O agente etiológico dessa doença foi posteriormente identificado como um novo coronavírus capaz de infectar humanos (LU *et al.*, 2020; ZHU *et al.*, 2020).

Uma vez que se tratava de um novo corona vírus emergente, denominou como a doença do coronavírus 2019 (COVID-19). Enquanto o agente etiológico foi inicialmente denominado 2019-nCoV, mas posteriormente renomeado como coronavírus da síndrome respiratória aguda grave 2 (SARS-CoV-2) pelo Comitê Internacional de Taxonomia de vírus (ICTV). Esta alteração da nomenclatura se deu uma vez que as analise filogenéticas o correlacionaram com o SARS-CoV-1, o vírus responsável pelo surto de síndrome respiratória aguda grave no continente asiático em 2002 (LAI *et al.*, 2020; LUDWIG; ZARBOCK, 2020; SHEREEN *et al.*, 2020).

Tanto o SARS-CoV-1 quanto o SARS-CoV-2 estão relacionados a grupo de vírus de RNA fita simples com sentido positivo (ssRNA+) associados infecções nas vias respiratórias de mamíferos e aves (BONILAURI; RUGNA, 2021). Esse grupo em particular constitui a subfamília *Orthocoronavirinae*, na família *Coronaviridae*, ordem *Nidovirales* e reino *Riboviria* (ICTV, [S.d.]; ZHOU; QIU; GE, 2021). A subfamília *Orthocoronavirinae*abriga quatro gêneros: *Alphacoronavirus, Betacoronavirus, Gammacoronavirus* e *Deltacoronavirus* (ZHOU; QIU; GE, 2021). Mas apenas coronavírus dos gêneros *Alphacoronavirus* e *Betacoronavirus* costumam ser capazes de infectar e causar doenças com um amplo espectro de sintomas respiratórios

(MURGOLO *et al.*, 2021). Ambos, SARS-CoV-1 e SARS-CoV-2, pertencem ao gênero *Betacoronavirus*.

As partículas de coronavírus são caracterizadas por projeções em forma de espiculas que se projetam da superfície viral. Estas espiculas criam uma imagem que lembra a coroa solar em micrografias eletrônicas, aspectos dos quais deriva o nome corona virus (ALMEIDA; TYRRELL, 1967). Eles são vírus envelopados com um genoma de RNA de fita simples de sentido positivo (26-32 kilobases) e um nucleocapsídeo de simetria helicoidal (CHERRY *et al.*, 2019).

O tamanho do genoma do SARS-CoV-2 é de aproximadamente 29,9 kb (BRANT *et al.*, 2021; LU *et al.*, 2020) e codifica dezesseis proteínas não estruturais (nsp1-16) e quatro proteínas estruturais (WANG, QIHUI *et al.*, 2020): a proteína do nucleocapsídeo (N), que é responsável por ligar o RNA genômico dentro do vírion e empacotá-lo no complexo ribonucleoproteína (RNP) que é ainda empacotado por um envelope lipídico associado às outras três proteínas estruturais, a proteína de membrana (M), a proteína spike (S) e proteína do envelope (E) (HARDENBROOK; ZHANG, 2022; WANG, QIHUI *et al.*, 2020).

Por ser um vírus emergente, pouco se sabia inicialmente sobre sua biologia ou citopatologia, e, mesmo que fosse possível de se inferir muito a partir dos vírus filogeneticamente próximos e estudados anteriormente, muitas questões ainda precisam ser respondidas em relação ao SARS-CoV-2. As tecnologias ômicas, incluindo as de transcriptômica, podem gerar milhares de informações sobre o processo de regulação gênica em células hospedeiras após a infecção pelo SARS-CoV-2 e, portanto, podem auxiliar no entendimento dos mecanismos moleculares da patogênese da infecção por SARS-CoV-2 e vírus-hospedeiro (BERKHOUT; COOMBS, 2013; JEAN BELTRAN *et al.*, 2017).

Aqui, revisitamos o protocolo de análise integrativa de transcriptoma utilizado anteriormente nos nossos estudos com o ZIKV e o atualizamos para o SARS-CoV-2, incluindo no protocolo uma nova etapa, utilizando algoritmos de *deep learning* para, mais do que gerar dados a respeito da patogênese viral, também identificar na célula hospedeira infectada por SARS-CoV-2 potenciais alvos celulares para o tratamento da doença de coronavírus. Aplicando este novo protocolo de análise integrativa de transcriptoma, foram identificados dois alvos, um transportador de frutose (GLUT5) e

um componente do proteassoma 26s. Esses alvos foram posteriormente modelados molecularmente e através do *docking* molecular identificou-se potenciais inibidores para ambas as estruturas. A inibição de estruturas que têm a expressão aumentada pelo vírus pode representar uma estratégia para reduzir a replicação viral por meio da seleção de células infectadas. Associar essas ferramentas de bioinformática representa uma estratégia promissora de triagem de moléculas para novos usos. Abrindo espaço para uma abordagem que garante economia de recursos financeiros, tempo e fazer uma triagem personalizada para cada doença infecciosa emergente.

3.2 Justificativa

Com a emergência de um novo coronavírus, o SARS-CoV-2, no final de 2019 e sua rápida dispersão pelo globo ainda no início ano seguinte, o desenvolvimento de estratégias de identificação de fármacos já aprovados para outras doenças que poderiam ser reposicionados se tornou uma emergência para o controle da COVID-19. As ferramentas de bioinformática representam uma estratégia valiosa na seleção de fármacos conhecidos e aprovados contra doenças emergentes, poupando recursos financeiros e tempo de experimentos *in vitro* e *in vivo*. Através da triagem *in sílico*, as análises posteriores podem ser direcionadas e personalizadas para cada doença e/ou agente infecioso. Entre os muitos recursos da bioinformática, a análise integrativa de transcriptoma tem sido utilizada para identificar genes chave do processo de patogênese de diversas doenças, incluindo doenças infeciosas. Esses genes chave tornam-se potenciais alvos terapêuticos para essas doenças, que depois de modelados molecularmente, podem ser utilizados contra bibliotecas de moléculas em busca de fármacos que tenham potencial inibir tais alvos e consequentemente a replicação do agente viral em questão.

3.3 Objetivos

3.3.1 Objetivo geral

Determinar potenciais alvos celulares terapêuticos e seus respectivos inibidores pela identificação de genes diferencialmente expressos em células infectadas por SARS-CoV-2.

3.3.2 Objetivos específicos

- Identificar genes alvos diferencialmente expressos em células após a infecção por SARS-CoV-2.
- Determinar genes alvos com potenciais drogáveis e realizar a modelagem *in sílico* das proteínas codificadas por esses genes.
- Identificar moléculas com potencial inibidor dos alvos previamente determinados

ANEXO

Artigo II

Integrative transcriptome analysis of SARS-CoV-2 human-infected cells combined with deep learning algorithms identifies two potential cellular targets for the treatment of coronavirus disease

DOI: 10.1007/s42770-022-00875-2

Brazilian Journal of Microbiology
https://doi.org/10.1007/s42770-022-00875-2

BACTERIAL, FUNGAL AND VIRUS MOLECULAR BIOLOGY - RESEARCH PAPER

Integrative transcriptome analysis of SARS-CoV-2 human-infected cells combined with deep learning algorithms identifies two potential cellular targets for the treatment of coronavirus disease

Ricardo Lemes Gonçalves[1] · Gabriel Augusto Pires de Souza[2,3] · Mateus de Souza Terceti[2] · Renato Fróes Goulart de Castro[2] · Breno de Mello Silva[1] · Romulo Dias Novaes[4] · Luiz Cosme Cotta Malaquias[2] · Luiz Felipe Leomil Coelho[2]

Received: 29 January 2022 / Accepted: 14 November 2022

Abstract
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) quickly spread worldwide, leading coronavirus disease 2019 (COVID-19) to hit pandemic level less than 4 months after the first official cases. Hence, the search for drugs and vaccines that could prevent or treat infections by SARS-CoV-2 began, intending to reduce a possible collapse of health systems. After 2 years, efforts to find therapies to treat COVID-19 continue. However, there is still much to be understood about the virus' pathology. Tools such as transcriptomics have been used to understand the impact of SARS-CoV-2 on different cells isolated from various tissues, leaving datasets in the databases that integrate genes and differentially expressed pathways during SARS-CoV-2 infection. After retrieving transcriptome datasets from different human cells infected with SARS-CoV-2 available in the database, we performed an integrative analysis associated with deep learning algorithms to determine differentially expressed targets mainly after infection. The targets found represented a fructose transporter (GLUT5) and a component of proteasome 26s. These targets were then molecularly modeled, followed by molecular docking that identified potential inhibitors for both structures. Once the inhibition of structures that have the expression increased by the virus can represent a strategy for reducing the viral replication by selecting infected cells, associating these bioinformatics tools, therefore, can be helpful in the screening of molecules being tested for new uses, saving financial resources, time, and making a personalized screening for each infectious disease.

Keywords Integrative bioinformatic · Integrative transcriptome analysis · SARS-CoV-2 · COVID-19 · Deep learning · Emerging virus disease

Introduction

Coronaviruses belong to a family of enveloped RNA viruses of medical and veterinary importance, which have recently induced several outbreaks worldwide and caused fatal respiratory diseases in humans [1, 2]. The severe acute respiratory

Ricardo Lemes Gonçalves and Gabriel Augusto Pires de Souza contributed equally to this work and shared authorship.

Responsible Editor: Luis Nero

✉ Luiz Felipe Leomil Coelho
luiz.coelho@unifal-mg.edu.br

[1] Núcleo de Pesquisas em Ciências Biológicas, NUPEB, Universidade Federal de Ouro Preto, Ouro Preto 35400-000, Brazil

[2] Laboratório de Vacinas, Departamento de Microbiologia e Imunologia, Instituto de Ciências Biomédicas, Universidade Federal de Alfenas, Rua Gabriel Monteiro da Silva, 700, 37130-001 Alfenas, Brazil

[3] Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Belo Horizonte 31270-901, Brazil

[4] Instituto de Ciências Biomédicas, Departamento de Biologia Estrutural, Universidade Federal de Alfenas, Alfenas, Minas Gerais, Brazil

Published online: 26 November 2022

coronavirus syndrome (SARS-CoV) that emerged in 2002 [3] and the Middle East respiratory syndrome coronavirus (MERS-CoV) of 2012 [4, 5] are examples of two highly pathogenic coronaviruses with zoonotic origin.

In late 2019, a new highly transmissible coronavirus called SARS-CoV-2 emerged in China, beginning an outbreak of viral pneumonia known as coronavirus disease 2019 (COVID-19) that spread rapidly around the world configuring itself as a pandemic [6–8]. SARS-CoV-2 overwhelmingly surpassed SARS and MERS in terms of the number of infected people and the spatial breadth of epidemic areas. As a new beta-coronavirus SARS-CoV-2 shares 79% genome sequence identity with SARS-CoV and 50% with MERS-CoV [9].

The World Health Organization (WHO) has confirmed more than 630 million cases of COVID worldwide, including about 6.5 million deaths [8]. Thus, the ongoing outbreak of COVID-19 represents an extraordinary threat to the world's public health [10–13]. Although several studies have shown that certain drugs may bring some benefits in patients' subpopulations, there are few effective therapies for COVID-19 or antivirals against SARS-CoV-2.

The drugs tested so far can act by preventing the entry of SARS-CoV-2 into the cell [14–19], directly inactivating viral replication [20–24], or acting as immunomodulating agents to reduce the excessive inflammatory response [25–30]. However, despite the anti-virus drugs and immunomodulators tested so far, it is clear that vaccination is the most effective long-term strategy for prevention and control [2].

Furthermore, the molecular mechanisms of pathogenesis of SARS-CoV-2 and virus–host interactions infection remain primarily unclear [2]. In this sense, omics technologies, especially transcriptome studies, can generate thousands of information about the gene regulation process in host cells after infection by a virus and, therefore, can help obtain relevant data on pathogenesis and pathogen–host interaction[31, 32].

Thus, the in silico integrative transcriptome analysis (ITA) has proven to be a promising approach for understanding biological events in complex diseases [33]. It can connect different data sources, reduce complexity, and increase data's predictive value through unified systems. Therefore, identifying the altered expression profile by ITA can contribute to understanding the virus pathogenesis and the development of new antiviral drugs, as performing joint and integrated analysis of several transcriptomes [34]. In addition, ITA has been used to find genes with significant associations for neurodegenerative disease (Alzheimer), cancer, and emerging viral disease (Zika) [33–40].

Therefore, this study aimed to perform an integrative analysis of transcriptomes from SARS-Cov-2-infected cells to identify genes and biological pathways essential for understanding the pathogenesis and identifying therapeutic strategies for COVID-19. Here, we show that a combination of ITA and highly efficient structural modeling through the breakthrough advance of deep learning algorithms can identify shared differentially expressed genes (DEGs) between SARS-CoV-2-infected cells. Furthermore, structural bioinformatics analysis confirms the potential drug–gene interaction of two DEGs (SLC2A5 and PSMD2). Therefore, the proposed in silico protocol performed can predict and identify putative cellular proteins that can be useful to guide the identification of candidate antiviral drugs or repurposing drugs to treat SARS-CoV-2.

Methods

The transcriptome of cells infected with SARS-CoV-2

Gene expression profiling data from SARS-CoV-2-infected human cells were retrieved from RNA sequencing data from the Gene Expression Omnibus platform (https://www.ncbi.nlm.nih.gov/gds). The platform was extensively searched for datasets of any human cell infected with SARS-COV-2. Information about the cell type used, the multiplicity of infection (MOI), time of infection, and sequencing platform were represented in Table 1. The raw sequence data were accessed using the online Galaxy platform (https://usegalaxy.org/), and the identification of differentially expressed genes (DEG) were obtained using the "Tuxedo suite" protocol as described in Pires de Souza and colleagues (2020). Only datasets that were performed with at least two replicates for each experimental sample were considered for the analysis. The final DEG list was identified from each study to fulfill an integrative analysis.

Network construction

All networks were built using the Gephi software version 0.9.2 [41]. The cells and genes were listed in a comma-separated values (.csv) spreadsheet for each graph, and this file was imported into the software. Another.csv spreadsheet with the connections between the cells and genes was also imported to generate the network graphs. The node diameter is directly proportional to the edge in all networks. The layout was generated using algorithms based on the force of attraction and repulsion of the nodes (Force Atlas 2). Finally, the nodes were submitted to local rearrangement to better visualize the connections between nodes.

Molecular modeling and validation of SLC2A5 and PSMD2 isoform structures

The three-dimensional models of isoform 1 of SLC2A5 (id: P22732) and PSMD2 (ID: Q13200) were generated from the AlphaFold database (https://alphafold.ebi.ac.uk/)

Table 1 Summary of datasets and sample details

GEO access	Cell type	SARS-COV-2 strain	MOI	Time of infection	NGS platform
GSE150392	Cardiomyocyte	USA-WA1/2020	0.1	72 h	Illumina NextSeq 500
GSE148729	Caco2	NR	0.3	24 h	Illumina NextSeq 500 Illumina HiSeq 4000
GSE148729	Calu3	NR	0.3	24 h	Illumina NextSeq 500 Illumina HiSeq 4000
GSE148729	HT199	NR	0.3	24 h	Illumina NextSeq 500 Illumina HiSeq 4000
GSE153970	Primary human airway epithelial cell (hEC)	USA-WA1/2020	0.25	48 h	Illumina NovaSeq 6000
GSE154613	A459 cells expressing human ACE2	USA-WA1/2020	0.2	24 h	Illumina NextSeq 500
GSE150728	Peripheral blood mononuclear cells*	—	—	—	Illumina NovaSeq 6000

NR, not reported

ACE2, angiotensin I-converting enzyme-2

NGS, next-generation sequencing

*Cells isolated from COVID-19 patients

partnership between DeepMind and EMBL-EBI, which set out to solve all the main protein structures encoded by the human genome through the new tool based on deep learning, AlphaFold2 [42], responsible for a "breakthrough advance" of protein modeling in CASP14. However, the structure of PSMD2 isoform 1 was considered unfeasible for the present study due to the low quality of the sizeable N-terminal portion of the polypeptide chain (data not shown). Furthermore, the AlphaFold Database does not present any other known isoforms for the modeled proteins. Therefore, isoform 2 (ID: P22732-2) of SLC2A5 and isoforms 1 and 2 (ID: Q13200-2) and 3 (ID: Q13200-3) of PSMD2 were generated through the Robetta Server (https://robetta.bakerlab. org/) using the RobettaAFold algorithm [43], which is based on the same deep learning method used by AlphaFold2. From the models generated for each structure, those with the best quality were selected and verified through the MolProbity score [44].

Those generated models were submitted to structural refinement simulation through the GalaxyRefine [45] server (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE), which provided different refined models, being the structures with the highest quality selected for the job. Finally, the final models' quality was validated through the Prosa-web [46] servers (https://prosa.services.came.sbg.ac.at/prosa.php) which compares the quality of the targets as a function of their total energy distributions. Target structure with the protein structures experimentally resolved by X-ray diffraction and nuclear magnetic resonance. Also used to check the quality of individual models, the MolProbity server (http://molprobity.biochem.duke.edu/) provided MolProbity score that combines different structural parameters to provide a normalized score, favoring the comparison of models with experimentally solved structures and, lastly, the individual construction of Ramachandran plot.

Preparation of receptors and ligands

The amino acid residues of SLC2A5 and PSMD2 isoforms received charges and polar hydrogens. Later, the models were aligned and had their structural centers determined. A grid box was delimited from these, completely covering each MGLTools (Morris et al., 2009) model. The coordinates and box sizes were *X*, 13,153, 64; *Y*, 33,099, 76; and *Z*, – 18,273, 54, for the structures of SLC2A5 and *X*, 1703, 80; *Y*, 43,775, 92; and *Z*, 38,746, 108, for PSMD2. Potential ligands for SLC2A5 had their structures obtained through PubChem: glufosfamide (CID: 123,628) and streptozocin (CID29327). In addition, D-fructose (ChEBI: 37,721) was used as the substrate of the isoforms of SLC2A5, and fludeoxyglucose (18F) (CID: 68,614) was used as an inhibitor. Potential ligands for PSMD2 were also obtained from PubChem (bortezomib (CID: 387,447), carfilzomib (CID: 11,556,711), ixazomib citrate (CID: 56,844,015), and oprozomib (CID: 25,067,547)). All composite structures had their torsions and charges, and hydrogens were added by Open Babel. The blind molecular docking's independent simulations were performed using Vina software [47] using exhaustiveness 32. Finally, the result of the simulations was analyzed and visualized using BIOVIA Discovery Studio® and Pymol software [48].

Results

Differentially expressed genes on SARS-CoV-2-infected cells

As illustrated in Fig. 1, a seven-step bioinformatic approach was used to identify candidate genes that could be targeted for antiviral development compounds against SARS-CoV-2. The search for transcriptome studies of human cells

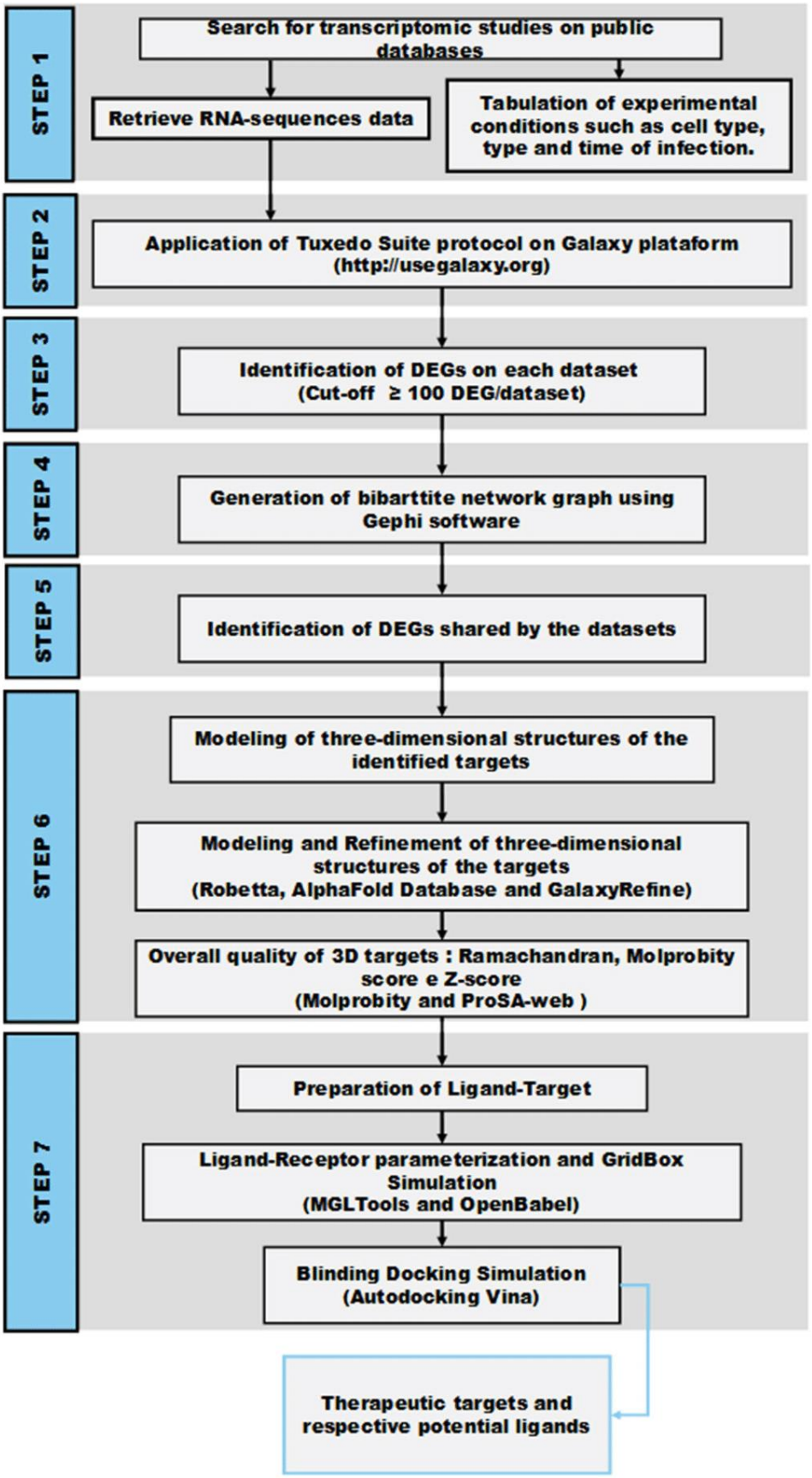


Fig. 1 Schematic representation of the entire in silico approach for identifying potential targets for the treatment of coronavirus disease

infected with SARS-CoV-2 resulted in the selection of 7 datasets. All experimental data, such as time of infection, cell type, SARS-CoV-2 strain, the multiplicity of infection (MOI), GEO number access, and sequencing platform, were described in Table 1. Several cell types were used in these experiments, such as cells derived from cardiovascular (cardiomyocyte), epidermal (HT199 melanoma cells), intestinal (Caco2), pulmonary (Calu3, primary human airway epithelial cells, A459), and immune systems (peripheral blood mononuclear cells). Tuxedo suite protocol was performed based on the transcriptome data from each study to understand better the changes in gene expression in each cell type after SARS-CoV-2 infection. Table 2 shows the number of differentially expressed genes (DEG) with statistical significance identified in each study.

Interaction networks between DEGs and cells were built from the identification of differentially expressed genes in each study. As shown in Fig. 2, there is a cell-specific gene expression profile since few differentially expressed genes were common to different cell types. Few genes were shared between two distinct cells (339/4.18%) and one of three cell types (7/0.08%). No DEG was shared between four or more cell types (Fig. 3). Table 3 represents the DEGs that were shared between three cell types. DEGs that had their expression increased after SARS-COV-2 infection are chosen for molecular docking analysis. The increase in their expression may be related to a dependence on the molecular function of the proteins encoded by these genes in the replication cycle of SARS-CoV-2.

Once the differentially expressed genes in three cell types were identified, the search for drugs that could directly interact with the proteins encoded by these genes began, using the DGIdb 3.0 platform (Wagner ET AL., 2018). The results obtained demonstrate that only the SLC2A5, PSMD2, and GNB3 genes presented drugs with the potential to interact with the proteins encoded by these genes (Table 4). Therefore, analyzing the potential gene–drug interactions obtained and the category analysis identified for each gene, the drugs for interaction with SLC2A5 (glufosfamide and streptozocin) and with PSMD2 (carfilzomib, bortezomib, ixazomib citrate, and oprozomib) are candidates for in vitro assays (Table 5).

Table 2 Number of differentially expressed genes used to construct interaction networks among SARS-CoV-2-infected cells

GSE access	Cell type	Up	Down
GSE150392	Cardiomyocyte	1404	1543
GSE148729	Caco2	50	33
GSE148729	Calu3	39	21
GSE148729	HT199	2240	1458
GSE153970	Primary human airway epithelial cell (hEC)	463	450
GSE154613	A459 cells expressing human ACE2	204	187
GSE150728	Peripheral blood mononuclear cells	119	70
	Total	4329	3762

Molecular modeling and validation of SLC2A5 and PSMD2 isoforms

The general information for the construction and validation of the three-dimensional models are shown in Table S1. The two SLC2A5 isoforms had excellent qualities, with the MolProbity score being evaluated at the 100th percentile compared to the best three-dimensionally resolved structures. The same can be observed for the three models generated for the PSMD2 isoforms, where the individual percentiles reached were 100, 100, and 99, respectively. The comparison of the models with structures solved using the Z-score and the Ramachandran plot is presented in Fig. S1. All isoforms of SLC2A5 and PSMD2 are distributed within the expected Z-score area (Fig. S1a, c, e, g, i) for proteins of the same size, again providing evidence of the high quality of the models, even for those containing > 900 amino acid residues (PSMD2 isoform 1) (Fig. S1e). Additionally, the SLC2A5 isoform 2 model obtained (Fig. S1c) is compared to the best structures resolved by nuclear magnetic resonance, and the SLC2A5 isoform 1 model (Fig. S1a) stands out for being among the best-resolved structures with the same characteristics. The individual Ramachandran plots are shown in Fig. S1b, d, f, h, and j, where, respectively, 99.4%, 99.6%, 99.8%, 99.6%, and 99.4% are distributed in allowable regions and > 97% of the residues distributed in favorable regions for all models of SLC2A5 and PSMD2, respectively. The superposition of all cured isoforms models can be seen in Fig. S2a.

Molecular docking

The ligands used in the docking simulations (Fig. S2b) had their interaction details with the receptors shown in Table 4. Most of the pockets involved in the ligand–receptor interactions from all structures had promisingly measured druggability scores > 0.8. On the other hand, P7 of isoform 1 and P8 of isoform 2, both from PSMD2, were identified in the interaction interface with carfilzomib.

SLCA2A5

The pocket P0 of SLCA2A5 isoform 1 was responsible for the binding of all docked compounds (Table 5), which represents the open core of the transmembrane channel of the GLUT5 receptor (extracellular face) with a volume of 1899Å^3 (Fig. 3). At P0, the ligand candidates presented the

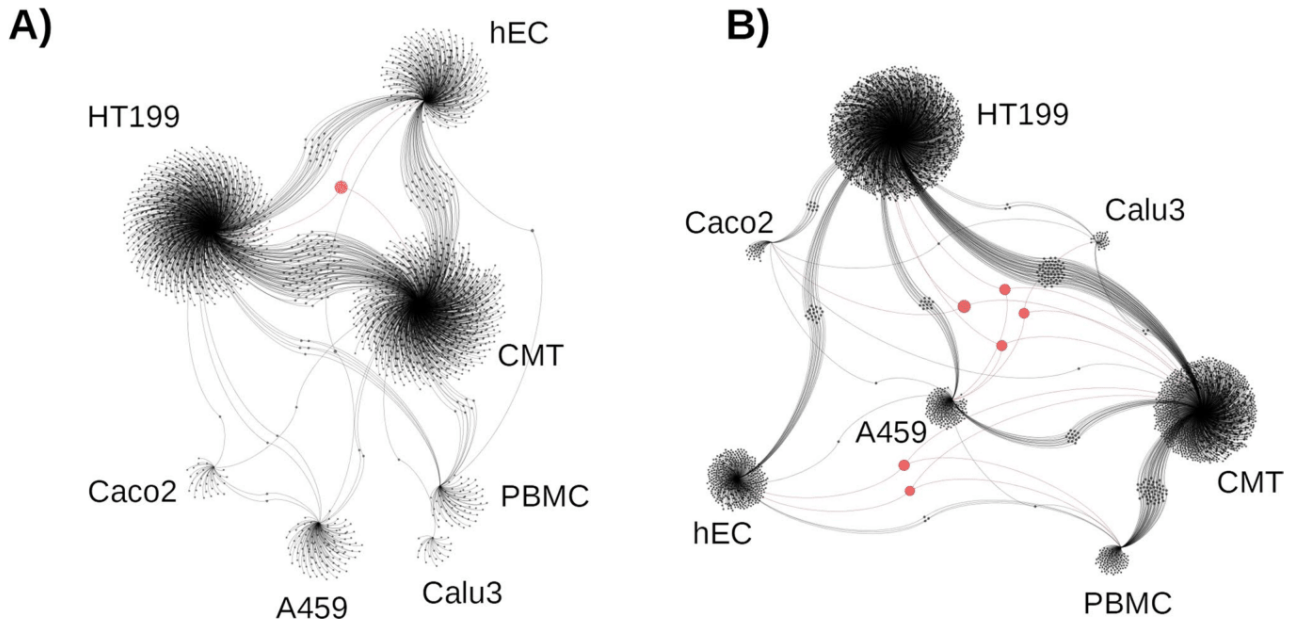


Fig. 2 Network of upregulated and downregulated genes after SARS-COV-2 infection of different human cells types. The bipartite network plot shows a spatially connected network between differentially expressed genes (DEGs) and cell types after SARS-COV-2 infection. Each node represents a gene or cell type. The layout was generated using a force-based algorithm followed by manual rearrangement to visualize the connections better. A total of 8091 DEGs (4329 DEGs with increased expression and 3762 DEGs that had decreased) and seven cell types are represented. Genes shared between three cell types are represented in red. Abbreviations: hEC, human respiratory tract epithelial cells; PBMC, peripheral blood mononuclear cells

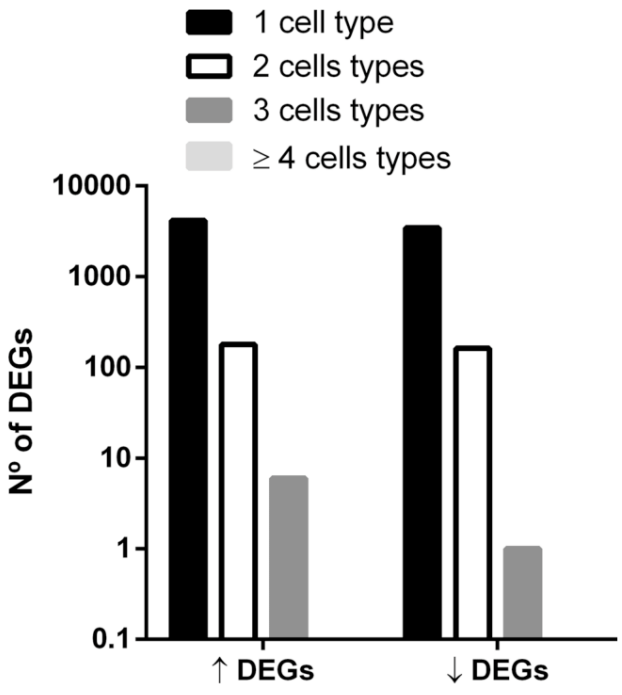


Fig. 3 Number of differentially expressed genes in SARS-COV-2 human-infected cells

same interaction interface when compared to the positive ligand control (D-fructose) and the positive inhibition control (fludeoxyglucose (18F)) (Fig. 3)—indicating potential competitive binding/inhibition property of selected compounds docked in isoform 1. On the other hand, isoform 2 presented distinct interaction pockets (Fig. S3), where the positive controls remained in P0 but at this model with a volume of 1206Å^3. However, the binding candidates were located at P1 (867Å^3), still in the open core of the GLUT5 channel (intracellular face).

Promisingly higher binding energies and higher interface areas (buried surface area (BSA)) ligand–receptor of glufosfamide and streptozocin were observed compared to both binding/inhibition controls (Table 5). In particular, streptozocin presented energy − 6.2 in pocket P0 of isoform1 with BSA of 629Å^2 and − 6.9 in P1 of isoform2 with BSA of 569Å^2. At the observed ligand–receptor interfaces, hydrogen bonds are the main interaction mechanisms presented at the core binding sites of both transmembrane channels (Fig. 4), which is consistent with a large number of hydroxyl groups and amine groups present in the compounds, in addition to an ample supply of acceptors and hydrogen donors in the recipient's residues.

Table 3 List of the genes that were differentially expressed in three cell types after the infection with SARS-COV-2

Gene	Cells type	Expression
IFI44	Cardiomyocyte/human respiratory tract epithelial cells/PBMC	Up
SLC2A5	Cardiomyocyte/human respiratory tract epithelial cells/PBMC	Up
PSMD2	Cardiomyocyte/A549 + hACE2/HT199	Up
PLEKHG6	Cardiomyocyte/A549 + hACE2/HT199	Up
PPP1R15A	Cardiomyocyte/Calu3/A459 + hACE2	Up
CPM	Cardiomyocyte/Caco2/HT199	Up
GNB3	Cardiomyocyte/human respiratory tract epithelial cells/HT199	Down

Table 4 Analysis of potential drugs able to interact with the proteins encoded by the identified genes

		Interaction analysis			
Gene	Category analysis	Drug	Type of interaction	Value	Interaction's value
IFI44	-	U	-	-	-
SLC2A5	Druggable genome transporter	Glufosfamide	U	0.54	3.55
		Streptozocin	U	0.19	1.21
PSMD2	Kinase; clinical action	Carfilzomib	Inhibitory	0.31	1.00
		Bortezomib	Inhibitory	0.29	0.95
		Ixazomib citrate	Inhibitory	0.23	0.75
		Oprozomib	Inhibitory	0.22	0.73
PLEKHG6		U	-	-	-
PPP1R15A		U	-	-	-
CPM	Druggable genome; protease cell surface; enzyme	U	-	-	-
GNB3		Sibutramine	U	1.45	1.45
		Clonidine	U	1.01	1.01
		Torsemide	U	0.73	0.73
		Sildenafil	U	0.67	0.67
		Furosemide	U	0.62	0.62
		Bumetanide	U	0.58	0.58
		Telmisartan	U	0.54	0.55
		Sumatriptan	U	0.48	0.49
		Sertraline	U	0.36	0.36
		Nortriptyline	U	0.36	0.36

U, unidentified

PSMD2

The 3 isoforms of the monomeric 26s proteasome subunits have remarkable structural similarities shown by the overlap (Fig. S2a). However, the binding sites in the receptors were distributed in different ways, with 4, 2, and 1 pockets being observed, respectively (Figs. 5 and S4). Noteworthy is the pocket P3 of isoform 3 with a volume of 983Å^3 and estimated druggability of 0.81 and volume of 2191Å^3. The sharing between the mentioned regions is observed due to the similar interfaces of bortezomib, which is present in P1 of isoforms 1–2, and oprozomib, which is present in P5 of isoform1, partially incorporates the residues of P1 of isoform2 (Figs. 5 and S4)—indicating potential specificity of these compounds to more than one of the monomeric subunit isoforms of the 26 s proteasome. The binding energies varied between − 7.1 and − 8.5 (Table 5), emphasizing the energies and BSA's observed in P3 of isoform 3, slightly higher than in the other isoforms. A great variety of natures of interactions presented by ligand–receptor interaction was observed. In addition to conventional hydrogen bonds, carfilzomib has several pi–alkyl interactions caused by the presence of its two aromatic rings (Figs. 5 and S4).

Table 5 Docking ligand–receptor interaction

	Ligands	Accession (ID)	N° of atoms	Vina score (Kcal/mol)			Pocket (ID & drug score)				Receptor–ligand interface (Å^2)			
SLC2A5	Glufosfamide	123,628	43	ISO1	ISO2		ISO1		ISO2		ISO1		ISO2	
	Streptozocin	29,327	33	− 5.8	− 5.7		P0/0.81		P1/0.82		778.3		671.5	
	D-Fructose	37,721	12	− 6.2	− 6.9				P1/0.82		629.4		569.9	
	Fludeoxyglucose (18F)	68,614	13	− 5.6	− 4.9				P0/0.80		490.8		454.2	
				− 5.6	-4.9				P0/0.80		510.7		486.7	
PSMD2	Bortezomib	387,447	53	ISO1	ISO2	ISO3	ISO1	ISO2		ISO3	ISO1	ISO2		ISO3
	Carfilzomib	11,556,711	109	− 7.3	− 7.7	− 7.7	P1/0.8	P1/0.81		P3/0.84	909.4	837.4		1132.3
	Ixazomib Citrate	56,844,015	57	− 7.5	− 8.3	− 8.2	P8/0.45	P7/0.4			1217.6	1478.8		1546.7
	Oprozomib	25,067,547	69	− 7.1	− 7.9	− 8.5	P2/0.82	P0/0.80			908.2	944.3		1148
				− 7.0	− 7.0	− 7.9	P5/0.84	P0/0.80			1082	1064.5		1249.6

Similarly, the influence of bortezomib rings in the presence of pi–cation/anion and pi–alkyl bonds can also be observed, mainly due to the presence of the pyrazine group (Figs. 5 and S4). Furthermore, many hydroxyl groups present in Ixazomib citrate give it a remarkable affinity to the receptor's hydrogen acceptors, with several hydrogen bonds being observed as the primary binding mechanism (Figs. 5 and S4). Finally, oprozomib has, in addition to a large number of hydrogen bonds, the inherent ability to create pi–sulfur interactions due to the existence of sulfur from the thiazole group (Figs. 5 and S4).

Discussion

In late 2019, SARS-CoV-2 emerged in China, and this new coronavirus spread rapidly around the world, configuring itself as one of the most pandemic ever experienced by humankind [6–8]. Due to high transmissibility, the high rate of hospitalization, and death, the development of vaccines and effective antiviral treatment is a priority of many research groups.

Although several studies have identified some effective drugs against SARS-CoV-2, to date, there are few approved effective therapies for COVID-19. Most of them are monoclonal antibodies (e.g., regdanvimab, tocilizumab, sotrovimab, and casirivimab/imdevimab), but there is also one antiviral drug (remdesivir) [49]. Therefore, there is an urgent need to identify new viral and cellular targets that could be useful to identify new antiviral compounds or be useful for repurposing drugs.

In this study, we used an in silico protocol to integrate different transcriptome data combined with deep learning algorithms to identify candidate genes important to the development of antiviral treatment of COVID-19. By this approach, it was verified that there is a specific transcriptional response of cells to SARS-CoV-2 infection since few DEGS are shared between the cells (Figs. 2 and 3). This specific transcriptional response for each cell type can also reflect the different times of infection, which ranged from 24 to 72 h, in an asynchrony condition (started with MOI < 0.3) (Table 1). Consequently, most cells of each cell model analyzed could be at different times of the replication cycle and justify the different expression patterns between these cell models. After all, the replication kinetics of SARS-CoV-2 is influenced by the host cell, and the virus's modulation of host gene expression will also depend on the replication and transcription step of SARS-CoV-2 itself in that cell [50–53], and this could explain why few DEGS are shared between the cell types. ITA was able to identify some potential targets for the treatment of coronavirus disease (Table 3), with some of the potential to be a target for the development of new treatment strategies for COVID-19 (Table 4).

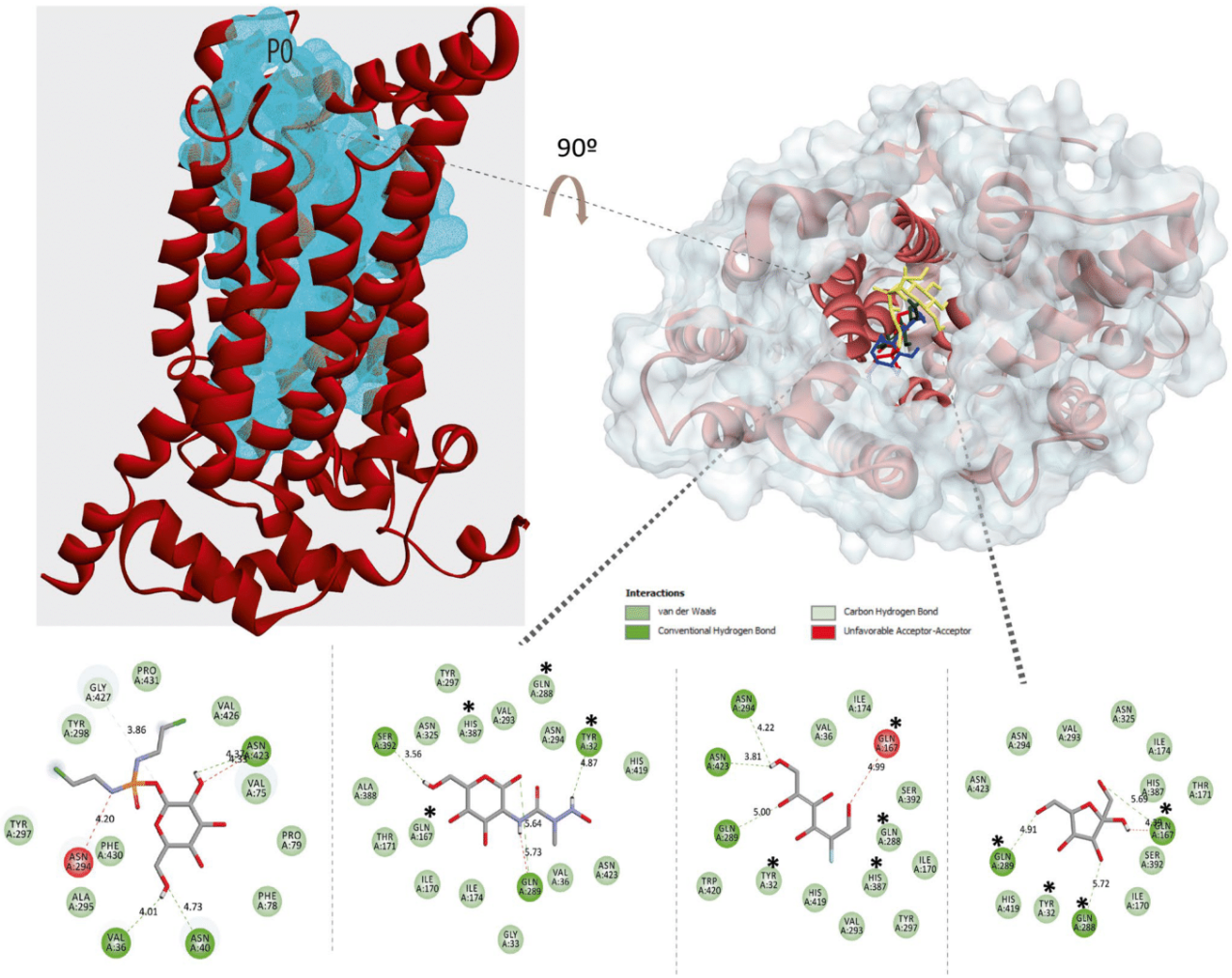


Fig. 4 SLC25A binding site isoform 1. Isoform 1 was represented as cartoons on the top left with their respective surface binding pockets. Isoform 1 in the cartoon with a transparent surface on the right, showing the open channel with all the ligands in pocket P0. At the bottom, the 2D maps of the interactions are represented, containing from left to right: glufosfamide, streptozocin, fludeoxyglucose (18F), and D-fructose are represented in stick, with all the amino acid residues involved in the interaction described by balloons. The hydrogen donor/acceptor interactions are characterized by a dotted line with their respective distances. Shaded areas on atoms from balloons indicate solvent exposure. The asterisk points out the key residues binding to the natural substrate (D-fructose)

Among these DEGs, two upregulated DEGs encoded proteins that can be used as antiviral targets. The SLC2A5 gene encodes a fructose transporter responsible for fructose uptake and therefore for energetic metabolism [54, 55]. The PDSM2 gene encodes a component of the 26S proteasome, a multiprotein complex involved in the ATP-dependent degradation of ubiquitinated proteins [56]. These two proteins have recently attracted attention for their druggable and therapeutic potential [57, 58].

Our analysis identified SLC2A5 as an upregulated DEG in three SARS-CoV-2 infected cells (cardiomyocyte, human respiratory tract epithelial cells, and PBMC). The SLC2A5 gene was reported as one of many genes differentially expressed related to a severe neurological progression in COVID-19 patients [59]. It also demonstrated high levels of fructose in PBMC [60] and serum metabolome of SARS-CoV-2 patients [61]. The high levels of this metabolite can also induce glycolysis in SARS-CoV-2-infected human monocytes [62]. These data suggest an increasing need for energy production in SARS-CoV-2-infected cells since fructose could be an efficient carbon and energy source. Therefore, higher levels of fructose would require greater transport efficiency into the cell, which could explain the higher expression of the SLC2A5 gene to supply the energy needs of the infected cells.

Other viruses are known to manipulate other glucose transporters to increase the acquisition of their target carbohydrate. For example, the human immunodeficiency virus led

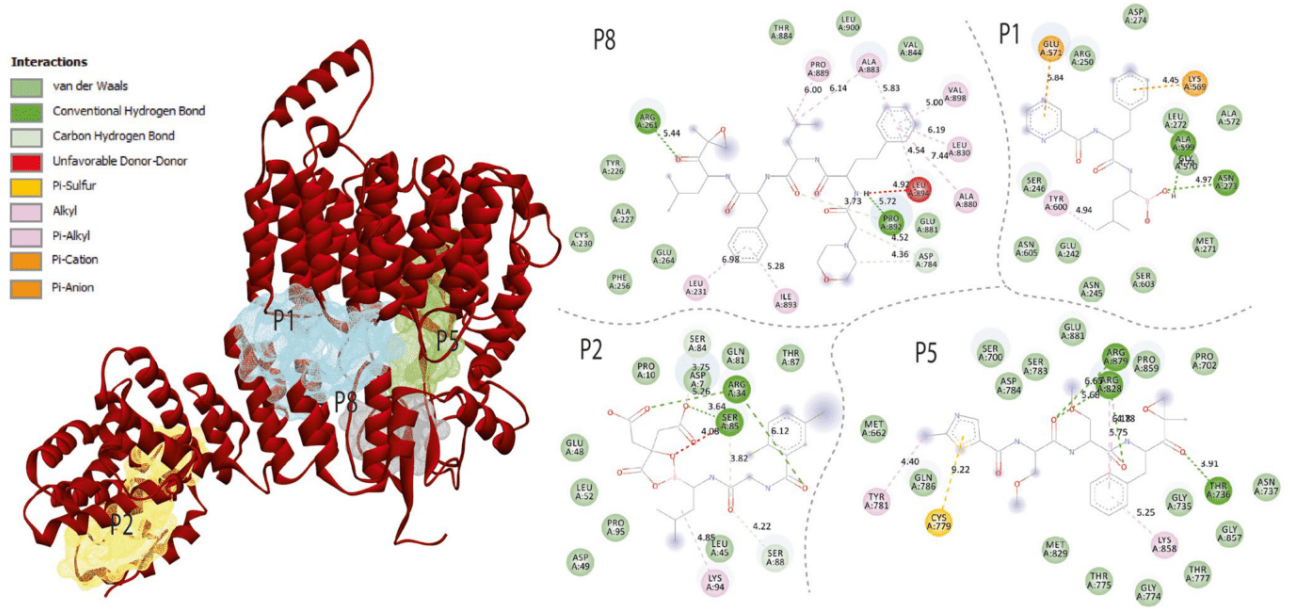


Fig. 5 PSMD2 binding site isoform 1. On the left, isoform 1 was represented as cartoons with their respective surface binding pockets. On the right, the ligands are represented in 2D interaction maps. The first pair of ligands is represented by lines, from left to right: carfilzomib and bortezomib and, the second pair, ixazomib and oprozomib. Balloons represent the amino acid residues involved in the interaction. The hydrogen donor/acceptor interactions, pi–sulfur, pi–cation, and pi–anion are characterized by a dotted line with their respective distances. Shaded areas on atoms and balloons indicate solvent exposure

to an increased expression of GLUT3 and increased glucose transport in H9 lymphocytic infected cells [63]. In addition, the human cytomegalovirus replaces GLUT1 with GLUT4, abundant in adipose tissue, and increases glucose transport capacity [64]. In this particular case, this change in GLUTs appears essential in cytomegalovirus infection of fibroblasts; once the function is specifically inhibited, the result is not only less glucose uptake but also dramatically inhibits the formation of infectious viral infection progeny [64, 65].

Therefore, an increase of the expression of SLC2A5 in SARS-CoV-2-infected cells derived from heart and lung tissues could be important to the pathophysiology of COVID-19 once these organs are two important systems affected COVID-19 patients. It is also suggested that the increase in glycolysis induced by SARS-CoV-2 infection may also be an important factor in cytokine regulation [62]. An increased expression of SLC2A5 by SARS-CoV-2 may represent a vital mechanism of infection, and inhibiting it could lead to a significant reduction in viral production. Previously, the inhibition of glycolysis by 2-deoxy-D-glucose prevents conversion to fructose, and this event resulted in inhibition of SARS-CoV-2 replication in a colon adenocarcinoma cell (Caco-2) line and monocytes [62, 66, 67].

Once it was understood that the SLC2A5 is a potential target for the control of SARS-CoV-2 replication, this enzyme was modeled to be later used to search for potential inhibitors of the target and, consequently, viral replication. The molecular modeling of GLUT-5 included the three-dimensional models of its isoforms (Table S1) derived from the alternative-splicing demonstrated in accessions P22732 and Q13200 (Uniprot ID). Even though interactions with the lipid bilayer make it particularly difficult to achieve high-quality resolution of transmembrane transporters [68], the complete structure of GLUT-5 presented here offers quality comparable to the best structures ever solved in PDB (Fig. S1), providing a unique path to search for potential drugs from its three-dimensional model.

The molecular modeling allowed determining the druggable regions classified in different pockets and starting the search for putative ligands. D-Fructose was used as an empirical reference of potential new ligands. Fludeoxyglucose (18F), an experimentally known inhibitor of the sugar transport function [69], was also included in the analysis. It was also considered the "open core of the channel," which has the hotspot binding points characterized for the amino acid residues 32TYR, 167GLN, 288GLN, and 387HIS responsible for the binding site of the transmembrane receptor SLC2A5 to D-fructose [70]. All these benchmarks were considered for the new putative ligand/inhibitor appointment.

The accuracy of the receptor–ligand binding is supported by the result of the docking simulation with D-fructose which presented an interaction interface with its respective natural binding site, constituting interactions with all

key residues and observed for the fludeoxyglucose (18F) (Fig. 3). Furthermore, both molecules were observed in pocket P1 with the same − 5.6 kcal/mol of Vina score and with a high similar BSA score (Table 4). These data suggest that a competitive inhibition mechanism of SLC2A5 in this region could be important to find new potential drugs for inhibiting SLC2A5 activity.

Through the binding and inhibition references presented, we promisingly point to the putative inhibition of SLC2A5 by streptozocin through comparison with the fludeoxyglucose (18F) binding site in pocket P1. Streptozocin presents a binding interface with all the key residues observed in Fig. 3. Additionally, streptozocin had a higher Vina (− 6.2 kcal/mol) and BSA (629.4 $Å^2$) score than the known inhibitor, making it a candidate for inhibiting sugar transport also through competition. Streptozocin is particularly an eclectic drug; it was first discovered as an antibiotic against Gram-negative bacteria once it inhibits the synthesis of DNA in microorganisms and mammalian cells [71, 72]. It is also used as a chemotherapeutic drug for treating certain pancreatic cancers and was approved to be used in the USA in 1982. The anticancer activity is due to the inhibition of glucose transporter 2 (GLUT2) expressed at pancreatic beta cells [71, 73]. These data strengthen the potential of streptozocin to be used in in vitro and in vivo assays, and therefore, new experiments should be performed to determine if the inhibition of GLUT-5 activity could reduce viral replication.

Our analysis also indicates that the ubiquitin–proteasome system is a potential target to develop antivirals against SARS-CoV-2. This intracellular system has a crucial role in the degradation of most cellular proteins, including short-lived, regulatory, and misfolded/denatured proteins [74, 75]. While viruses evolved, exploiting the cellular machinery and taking control of a protein degradation system are undoubtedly an adaptive advantage. Therefore, it is not surprising that viruses induce proteasome regulation after cellular infection. The proteasome can play an antiviral role in response to many infections, through degradation of viral proteins (e.g., West Nile virus and hepatitis C virus) or suppressing viral protein activity by ISGylation (e.g., influenza A virus and human papillomavirus) [76–78], although many other pro-viral functions are reported [79].

Viruses such as adenovirus and human immunodeficiency viruses 1 and 2 use proteasome to regulate cellular protein degradation [80, 81]. However, others may employ this system to maintain proper levels of viral proteins (e.g., human papillomavirus and hepatitis C virus) [82, 83]. It was described that SARS-CoV might use the proteasome for its benefit, counteracting the post-translational modification of signaling molecules involved in innate immunity. The papain-like protease of SARS-CoV can promote deubiquitinating (and deISGylating) that are proposed to counteract the post-translational modification of signaling molecules that activate the innate immune response [79, 84]. At least two other SARS-CoV proteins interact with this system. The E protein interacts with the non-structural protein 3 to be ubiquitinated and the N protein with the host cell proteasome subunit p42, a 26S proteasome subunit [85, 86].

Since the proteasome system plays an important role during the SARS-CoV infection cycle, proteasome inhibition appears to be a promising strategy in combating SARS-CoV-2 infection. Proteasome inhibition leads to the accumulation of polyubiquitinated proteins in the cytoplasm [87]. This unbalanced protein homeostasis leads to a cellular stress response that can be responsible for the induction of cell death [87, 88].

The inhibition of the proteasome with different compounds, such as MG132 and epoxomicin, leads to impaired viral entry and a decrease of RNA synthesis and protein expression of different coronaviruses [89]. On the other hand, a study demonstrated that the inhibition of SARS-CoV replication by MG132 occurs independently of inhibition of the proteasome system [90]. Although this potential of proteasome inhibitors has been suggested to combat SARS-CoV-2 infection, we do not know that they have been used in clinical trials.

Notwithstanding, a proteasome inhibitor (α-keto phenylamide) derivative compound demonstrated an anti-SARS-CoV-2 activity. However, the authors associated the viral progeny reduction with an interaction with the virus's protease [91]. Anyway, it seems plausible to consider proteasome inhibitors as a target for the search for drugs that inhibit viral replication, especially if we consider the expression profile of the proteasome observed by our integrative analysis and its roles in the control of coronaviruses replication.

Large polypeptide chains are difficult to model without high-quality homologous templates. The three-dimensional structures of the 26 s proteasome subunits are known from the 6MSK crystal (PDB ID). However, dozens of monomers are incomplete and mostly resolved at low resolutions. Therefore, it was relevant to obtain high-quality three-dimensional models for the monomeric chain of the 26S proteasome (Fig. S1e and S1j) before starting the search for its putative inhibitors.

The 26 s proteasome showed four high druggability different pockets of ligand–receptor interaction, each containing distinct and promising interaction profiles (Fig. 4). Besides, recent studies reported curcumin, LDN192960, and oprozomib as 26 s proteasome inhibitors, the pathways responsible for its inhibition are poorly understood. Therefore, there are no specific ligands/inhibitors that can be used as controls in docking studies using the 26 s proteasome [92–94].

Nevertheless, four ligands have been identified, and they had similar binding energy scores (from − 7.0 to − 7.5). The compound ixazomib citrate is promisingly docked to P2 in the N-terminal portion of the 26 s proteasome (Fig. 4). The N-terminal portion is understood as an attractive drug target once it constitutes a regulatory region of the 26 s proteasome [95]. The ligands bortezomib and oprozomib interact in P1 and P5, respectively. These two regions are both considered promising druggable pockets (Table 4) located around the dozens of helices in the structural center of the monomer. On the other hand, carfilzomib has the largest binding interface area and Vina energy score (Table 4). However, it is located at P8, with a low estimated druggability (0.45).

These compounds are already known as proteasome inhibitors and are used/suggested for anti-myeloma therapy [96]. Bortezomib reversibly binds with high affinity to the 26S proteasome β-subunit [97]. This drug has been recommended in treating adult patients with multiple myeloma cell lymphoma by the Food and Drug Administration (FDA) since 2003 [98]. Ixazomib binds and inhibits the β5 subunit of the 20S proteasome and has been approved by the FDA in 2015 for use in the treatment of patients with multiple myeloma [99]. Unlike bortezomib, the carfilzomib, approved in 2012 by the FDA, binds irreversibly and selectively to the chymotrypsin-like activity region of the 20S proteasome [100, 101]. Oprozomib is a second-generation proteasome inhibitor, which the US FDA has also approved for clinical treatment of recurrent multiple myeloma [102, 103].

Identifying new uses outside the original indication for approved or investigational drugs has become a very desired strategy, especially with the need to treat COVID-19 [104–106]. This strategy, known as drug repurposing, offers various advantages over developing an entirely new drug. For example, drug repurposing has a lower failure risk and reduced drug development and approval [105, 107]. In this study, we used several bioinformatic tools to find new drugs that could repurpose drugs for different COVID-19. Through the initial integrative analysis, genes that are differentially expressed during the infection in cells from different tissues are determined, resulting in an almost global pattern of infection. This previously determined pattern will guide the subsequent modeling of the targets selected by this integrative transcriptomic analysis of machine learning structural modeling. Furthermore, molecular docking would allow identifying available drugs that interact with the targets and propose laboratory and clinical essays for proof of concept. This represents a bioinformatics tool for screening molecules tested for new uses, saving financial resources time, and making a personalized screening for each infectious disease.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s42770-022-00875-2.

Acknowledgements A special acknowledgement to Renato Fróes Goulart de Castro (in memoriam) for all his time working with us, for his dedication and friendship. We regret that he could not have seen more of this published work.

Author contribution All authors contributed to the study's conception and design. Material preparation, data collection, and analysis were performed by Ricardo Lemes Gonçalves, Gabriel Augusto Pires de Souza, Renato Fróes Goulart de Castro, and Luiz Felipe Leomil Coelho. The first draft of the manuscript was written by Ricardo Lemes Gonçalves, Gabriel Augusto Pires de Souza, and Luiz Felipe Leomil Coelho. All authors commented on previous versions of the manuscript. Finally, all authors read and approved the final manuscript.

Funding This work was funded by the Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG) (grants number APQ-01165–16 and PPM-00399–18); the Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior, Brazil (CAPES) (Finance Code 001); and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Availability of data and material Not applicable.

Code availability Not applicable.

Declarations

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

1. Cui J, Li F, Shi ZL (2019) Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol 17:181–192. https://doi.org/10.1038/s41579-018-0118-9
2. Hu B, Guo H, Zhou P (2020) Shi ZL (2020) Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 193(19):141–154. https://doi.org/10.1038/S41579-020-00459-7
3. Zhong NS, Zheng BJ, Li YM et al (2003) Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. Lancet (London, England) 362:1353–1358. https://doi.org/10.1016/S0140-6736(03)14630-2
4. Zaki AM, van Boheemen S, Bestebroer TM et al (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814–1820. https://doi.org/10.1056/NEJMOA1211721
5. van Boheemen S, de Graaf M, Lauber C et al (2012) Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. MBio 3(6):e473–12. https://doi.org/10.1128/MBIO.00473-12
6. Wu F, Zhao S, Yu B et al (2020) A new coronavirus associated with human respiratory disease in China. Nat 5797798(579):265–269. https://doi.org/10.1038/S41586-020-2008-3
7. Hui DS, I Azhar E, Madani TA et al (2020) The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — the latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 91:264–266.https://doi.org/10.1016/j.ijid.2020.01.009

8. WHO (2020) Coronavirus disease 2019 (COVID-19): situation report, 61. https://apps.who.int/iris/handle/10665/331605. Accessed 24 Jan 2022
9. Lu R, Zhao X, Li J et al (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 395:565–574. https://doi.org/10.1016/S0140-6736(20)30251-8/ATTACHMENT/1686D69B-3F2C-44B2-A8A5-C648EF9D1C2B/MMC1.PDF
10. Chan JFW, Yuan S, Kok KH et al (2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 395:514–523. https://doi.org/10.1016/S0140-6736(20)30154-9
11. Deng SQ, Peng HJ (2020) Characteristics of and public health responses to the coronavirus disease 2019 outbreak in China. J Clin Med 9(2):575. https://doi.org/10.3390/JCM9020575
12. Chen R, Liang W, Jiang M et al (2020) Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in China. Chest 158:97–105. https://doi.org/10.1016/j.chest.2020.04.010
13. Han Q, Lin Q, Jin S, You L (2020) Coronavirus 2019-nCoV: a brief perspective from the front line. J Infect 80:373–377. https://doi.org/10.1016/J.JINF.2020.02.010
14. Li Y, Xie Z, Lin W et al (2020) Efficacy and safety of lopinavir/ritonavir or Arbidol in adult patients with mild/moderate COVID-19: an exploratory randomized controlled trial. Med (New York, NY) 1:105-113.e4. https://doi.org/10.1016/J.MEDJ.2020.04.001
15. Lian N, Xie H, Lin S et al (2020) Umifenovir treatment is not associated with improved outcomes in patients with coronavirus disease 2019: a retrospective study. Clin Microbiol Infect 26:917–921. https://doi.org/10.1016/J.CMI.2020.04.026
16. Wang M, Cao R, Zhang L et al (2020) Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res 30:269–271. https://doi.org/10.1038/S41422-020-0282-0
17. Rosenberg ES, Dufort EM, Udo T et al (2020) Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York state. JAMA 323:2493–2502. https://doi.org/10.1001/JAMA.2020.8630
18. Monteil V, Kwon H, Prado P et al (2020) Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. Cell 181:905-913.e7. https://doi.org/10.1016/j.cell.2020.04.004
19. Tian X, Li C, Huang A et al (2020) Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. Emerg Microbes Infect 9:382–385. https://doi.org/10.1080/22221751.2020.1729069
20. Tahir ul Qamar M, Alqahtani SM, Alamri MA, Chen LL (2020) Structural basis of SARS-CoV-2 3CL pro and anti-COVID-19 drug discovery from medicinal plants. J Pharm Anal 10:313–319.https://doi.org/10.1016/J.JPHA.2020.03.009
21. Williamson BN, Feldmann F, Schwarz B et al (2020) (2020) Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. Nat 5857824(585):273–276. https://doi.org/10.1038/S41586-020-2423-5
22. Cai Q, Yang M, Liu D et al (2020) Experimental treatment with favipiravir for COVID-19: an open-label control study. Eng (Beijing, China) 6:1192–1198. https://doi.org/10.1016/J.ENG.2020.03.007
23. Cao B, Wang Y, Wen D et al (2020) A trial of lopinavir–ritonavir in adults hospitalized with severe Covid-19. N Engl J Med 382:1787–1799. https://doi.org/10.1056/NEJMOA2001282/SUPPL_FILE/NEJMOA2001282_DATA-SHARING.PDF
24. Hung IFN, Lung KC, Tso EYK et al (2020) Triple combination of interferon beta-1b, lopinavir–ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. Lancet 395:1695–1704. https://doi.org/10.1016/S0140-6736(20)31042-4/ATTACHMENT/4753030C-5333-4ABE-BE66-F099C5E96D3A/MMC1.PDF
25. Xu X, Han M, Li T et al (2020) Effective treatment of severe COVID-19 patients with tocilizumab. Proc Natl Acad Sci U S A 117:10970–10975. https://doi.org/10.1073/PNAS.2005615117
26. Diurno F, Numis FG, Porta G et al (2020) Eculizumab treatment in patients with COVID-19: preliminary results from real life ASL Napoli 2 Nord experience. Eur Rev Med Pharmacol Sci 24:4040–4047. https://doi.org/10.26355/EURREV_202004_20875
27. Duan K, Liu B, Li C et al (2020) Effectiveness of convalescent plasma therapy in severe COVID-19 patients. Proc Natl Acad Sci U S A 117:9490–9496. https://doi.org/10.1073/PNAS.2004168117
28. Zost SJ, Gilchuk P, Case JB et al (2020) (2020) Potently neutralizing and protective human antibodies against SARS-CoV-2. Nat 5847821(584):443–449. https://doi.org/10.1038/S41586-020-2548-6
29. Park A, Iwasaki A (2020) Type I and type III interferons - induction, signaling, evasion, and application to combat COVID-19. Cell Host Microbe 27:870–878. https://doi.org/10.1016/J.CHOM.2020.05.008
30. Shen C, Wang Z, Zhao F et al (2020) Treatment of 5 critically Ill patients with COVID-19 with convalescent plasma. JAMA 323:1582–1589. https://doi.org/10.1001/JAMA.2020.4783
31. Berkhout B, Coombs KM (2013) Quantitative omics and its application to study virus-host interactions-a new frontier. Front Microbiol 4:31. https://doi.org/10.3389/FMICB.2013.00031/BIBTEX
32. Jean Beltran PM, Federspiel JD, Sheng X, Cristea IM (2017) Proteomics and integrative omic approaches for understanding host–pathogen interactions and infectious diseases. Mol Syst Biol 13(3):922. https://doi.org/10.15252/MSB.20167062
33. Tulsyan S, Hussain S, Mittal B et al (2020) A systematic review with in silico analysis on transcriptomic profile of gallbladder carcinoma. Semin Oncol 47:398–408. https://doi.org/10.1053/J.SEMINONCOL.2020.02.012
34. Pires de Souza GA, Salvador EA, de Oliveira FR et al (2020) An in silico integrative protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis. Virus Res 284:197986. https://doi.org/10.1016/j.virusres.2020.197986
35. Cinegaglia NC, Andrade SCS, Tokar T et al (2016) Integrative transcriptome analysis identifies deregulated microRNA-transcription factor networks in lung adenocarcinoma. Oncotarget 7:28920–28934. https://doi.org/10.18632/ONCOTARGET.8713
36. Hoshida Y, Nijman SMB, Kobayashi M et al (2009) Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res 69:7385–7392. https://doi.org/10.1158/0008-5472.CAN-09-1089
37. Liao Q, Wang J, Pei Z et al (2017) Identification of miRNA-mRNA crosstalk in CD4+ T cells during HIV-1 infection by integrating transcriptome analyses. J Transl Med 15:1–11. https://doi.org/10.1186/S12967-017-1130-Y/FIGURES/2
38. Moslehi R, Mills JL, Signore C et al (2013) (2013) Integrative transcriptome analysis reveals dysregulation of canonical cancer molecular pathways in placenta leading to preeclampsia. Sci Reports 31(3):1–11. https://doi.org/10.1038/SREP02407
39. Raj T, Li YI, Wong G et al (2018) (2018) Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. Nat Genet 5011(50):1584–1592. https://doi.org/10.1038/S41588-018-0238-1

40. Vavougios GD, Nday C, Pelidou SH et al (2020) Double hit viral parasitism, polymicrobial CNS residency and perturbed proteostasis in Alzheimer's disease: a data driven, in silico analysis of gene expression data. Mol Immunol 127:124–135. https://doi.org/10.1016/J.MOLIMM.2020.08.021
41. Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. Proc Int AAAI Conf Web Soc Media 3(1):361–362. https://doi.org/10.1609/ICWSM.V3I1.13937
42. Jumper J, Evans R, Pritzel A et al (2021) (2021) Highly accurate protein structure prediction with AlphaFold. Nat 5967873(596):583–589. https://doi.org/10.1038/S41586-021-03819-2
43. Baek M, DiMaio F, Anishchenko I et al (2021) Accurate prediction of protein structures and interactions using a three-track neural network. Science (80)373:871–876. https://doi.org/10.1126/science.abj8754
44. Williams CJ, Headd JJ, Moriarty NW et al (2018) MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci 27:293–315. https://doi.org/10.1002/PRO.3330
45. Heo L, Lee H, Seok C (2016) GalaxyRefineComplex: refinement of protein-protein complex model structures driven by interface repacking. Sci Reports 6(1):1–10. https://doi.org/10.1038/srep32153
46. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407–W410. https://doi.org/10.1093/NAR/GKM290
47. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 31:455. https://doi.org/10.1002/JCC.21334
48. Yuan S, Chan HCS, Hu Z (2017) Using PyMOL as a platform for computational drug design. Wiley Interdiscip Rev Comput Mol Sci 7:e1298. https://doi.org/10.1002/WCMS.1298
49. European Medicines Agency (2021) COVID-19 treatments: authorised. https://www.ema.europa.eu/en/human-regulatory/overview/public-health-threats/coronavirus-disease-covid-19/treatments-vaccines/treatments-covid-19/covid-19-treatments-authorised. Accessed 24 Dec 2021
50. Malone B, Urakova N, Snijder EJ, Campbell EA (2022) Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. Nat Rev Mol Cell Biol 23:21–39. https://doi.org/10.1038/s41580-021-00432-z
51. Mautner L, Hoyos M, Dangel A et al (2022) Replication kinetics and infectivity of SARS-CoV-2 variants of concern in common cell culture models. Virol J 19:1–11. https://doi.org/10.1186/s12985-022-01802-5
52. V'kovski P, Kratzel A, Steiner S et al (2021) Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol 19:155–170. https://doi.org/10.1038/s41579-020-00468-6
53. Pires De Souza GA, Le Bideau M, Boschi C et al (2022) Choosing a cellular model to study SARS-CoV-2. Front Cell Infect Microbiol 12:1583. https://doi.org/10.3389/FCIMB.2022.1003608/BIBTEX
54. Kayano T, Burant CF, Fukumoto H et al (1990) Human facilitative glucose transporters. Isolation, functional characterization, and gene localization of cDNAs encoding an isoform (GLUT5) expressed in small intestine, kidney, muscle, and adipose tissue and an unusual glucose transporter pseudogene-like. J Biol Chem 265:13276–13282. https://doi.org/10.1016/s0021-9258(19)38295-x
55. Zwarts I, van Zutphen T, Kruit JK et al (2019) Identification of the fructose transporter GLUT5 (SLC2A5) as a novel target of nuclear receptor LXR. Sci Rep 9:1–10. https://doi.org/10.1038/s41598-019-45803-x
56. Kanayama H-o, Tamura T, Ugai S et al (1992) Demonstration that a human 26S proteolytic complex consists of a proteasome and multiple associated protein components and hydrolyzes ATP and ubiquitin-ligated proteins by closely linked mechanisms. Eur J Biochem 206:567–578. https://doi.org/10.1111/j.1432-1033.1992.tb16961.x
57. Longhitano L, Tibullo D, Giallongo C et al (2020) Proteasome inhibitors as a possible therapy for SARS-CoV-2. Int J Mol Sci 21:1–11. https://doi.org/10.3390/ijms21103622
58. Ferreira RS, Pons JL, Labesse G (2019) Insights into substrate and inhibitor selectivity among human GLUT transporters through comparative modeling and molecular docking. ACS Omega 4:4748–4760. https://doi.org/10.1021/ACSOMEGA.8B03447
59. Rahman MH, Rana HK, Peng S et al (2021) Bioinformatics and system biology approaches to identify pathophysiological impact of COVID-19 to the progression and severity of neurological diseases. Comput Biol Med 138:104859. https://doi.org/10.1016/J.COMPBIOMED.2021.104859
60. Singh Y, Trautwein C, Fendel R et al (2021) SARS-CoV-2 infection paralyzes cytotoxic and metabolic functions of the immune cells. Heliyon 7:e07147. https://doi.org/10.1016/J.HELIYON.2021.E07147
61. Shi D, Yan R, Lv L et al (2021) The serum metabolome of COVID-19 patients is distinctive and predictive. Metabolism 118:154739. https://doi.org/10.1016/J.METABOL.2021.154739
62. Codo AC, Davanzo GG, de Brito Monteiro L et al (2020) Elevated glucose levels favor SARS-CoV-2 infection and monocyte response through a HIF-1α/glycolysis-dependent axis. Cell Metab 32(3):437–446.e5. https://doi.org/10.1016/j.cmet.2020.07.007
63. Sorbara LR, Maldarelli F, Chamoun G et al (1996) Human immunodeficiency virus type 1 infection of H9 cells induces increased glucose transporter expression. J Virol 70:7275–7279. https://doi.org/10.1128/jvi.70.10.7275-7279.1996
64. Yu Y, Maguire TG, Alwine JC (2011) Human cytomegalovirus activates glucose transporter 4 expression to increase glucose uptake during infection. J Virol 85:1573–1580. https://doi.org/10.1128/jvi.01967-10
65. Yu Y, Clippinger AJ, Alwine JC (2011) Viral effects on metabolism: changes in glucose and glutamine utilization during human cytomegalovirus infection. Trends Microbiol 19:360–367. https://doi.org/10.1016/j.tim.2011.04.002
66. Bojkova D, Klann K, Koch B et al (2020) Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. Nature 583:469–472. https://doi.org/10.1038/s41586-020-2332-7
67. Bojkova D, Costa R, Reus P et al (2021) Targeting the pentose phosphate pathway for SARS-CoV-2 therapy. Metab. https://doi.org/10.3390/metabo11100699
68. Kermani AA, Kermani AA (2021) A guide to membrane protein X-ray crystallography. FEBS J 288:5788–5804. https://doi.org/10.1111/FEBS.15676

69. Trayner BJ, Grant TN, West FG, Cheeseman CI (2009) Synthesis and characterization of 6-deoxy-6-fluoro-d-fructose as a potential compound for imaging breast cancer with PET. Bioorganic Med Chem 17:5488–5495. https://doi.org/10.1016/j.bmc.2009.06.034
70. Nomura N, Verdon G, Kang HJ et al (2015) Structure and mechanism of the mammalian fructose transporter GLUT5. Nature 526:397. https://doi.org/10.1038/NATURE14909
71. Abdollahi M, Hosseini A (2014) Streptozotocin. Encycl Toxicol (3rd edn.) 402–404. https://doi.org/10.1016/B978-0-12-386454-3.01170-2
72. Aoki T, Kokudo N, Komoto I et al (2015) Streptozocin chemotherapy for advanced/metastatic well-differentiated neuroendocrine tumors: an analysis of a multi-center survey in Japan. J Gastroenterol 50:769–775. https://doi.org/10.1007/s00535-014-1006-3
73. Rosati M (2017) Saunders handbook of veterinary drugs: small and large animals, 4th edn. Can Vet J 58(7):728. https://doi.org/10.1016/B978-0-323-70957-6.01001-3
74. Gong B, Radulovic M, Figueiredo-Pereira ME, Cardozo C (2016) The ubiquitin-proteasome system: potential therapeutic targets for Alzheimer's disease and spinal cord injury. Front Mol Neurosci 9:1–16. https://doi.org/10.3389/fnmol.2016.00004
75. Schwartz AL, Ciechanover A (2009) Targeting proteins for destruction by the ubiquitin system: implications for human pathobiology. Annu Rev Pharmacol Toxicol 49:73–96. https://doi.org/10.1146/annurev.pharmtox.051208.165340
76. Durfee LA, Lyon N, Seo K, Huibregtse JM (2010) The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. Mol Cell 38:722–732. https://doi.org/10.1016/j.molcel.2010.05.002
77. Ko A, Lee E-W, Yeh J-Y et al (2010) MKRN1 induces degradation of west nile virus capsid protein by functioning as an E3 ligase. J Virol 84:426–436. https://doi.org/10.1128/jvi.00725-09
78. Zhao C, Hsiang TY, Kuo RL, Krug RM (2010) ISG15 conjugation system targets the viral NS1 protein in influenza A virus-infected cells. Proc Natl Acad Sci U S A 107:2253–2258. https://doi.org/10.1073/pnas.0909144107
79. Luo H (2016) Interplay between the virus and the ubiquitin-proteasome system: molecular mechanism of viral pathogenesis. Curr Opin Virol 17:1–10. https://doi.org/10.1016/j.coviro.2015.09.005
80. Querido E, Blanchette P, Yan Q et al (2001) Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. Genes Dev 15:3104–3117. https://doi.org/10.1101/gad.926401
81. Schwefel D, Boucherit VC, Christodoulou E et al (2015) Molecular determinants for recognition of divergent SAMHD1 proteins by the lentiviral accessory protein Vpx. Cell Host Microbe 17:489–499. https://doi.org/10.1016/j.chom.2015.03.004
82. Franck N, Le Seyec J, Guguen-Guillouzo C, Erdtmann L (2005) Hepatitis C virus NS2 protein is phosphorylated by the protein kinase CK2 and targeted for degradation to the proteasome. J Virol 79:2700–2708. https://doi.org/10.1128/jvi.79.5.2700-2708.2005
83. Kamio M, Yoshida T, Ogata H et al (2004) SOC1 inhibits HPV-E7-mediated transformation by inducing degradation of E7 protein. Oncogene 23:3107–3115. https://doi.org/10.1038/sj.onc.1207453
84. Ratia K, Kilianski A, Baez-Santos YM, Baker SC, Mesecar A (2014) Structural basis for the ubiquitin-linkage specificity and deISGylating activity of SARS-CoV papain-like protease. PLOS Pathog 10(5):e1004113. https://doi.org/10.1371/JOURNAL.PPAT.1004113
85. Álvarez E, DeDiego ML, Nieto-Torres JL et al (2010) The envelope protein of severe acute respiratory syndrome coronavirus interacts with the non-structural protein 3 and is ubiquitinated. Virology 402:281–291. https://doi.org/10.1016/j.virol.2010.03.015
86. Wang Q, Li C, Zhang Q et al (2010) Interactions of SARS coronavirus nucleocapsid protein with the host cell proteasome subunit p42. Virol J 7:1–8. https://doi.org/10.1186/1743-422X-7-99
87. Sojka D, Šnebergerová P, Robbertse L (2021) Protease inhibition—an established strategy to combat infectious diseases. Int J Mol Sci 22(11):5762. https://doi.org/10.3390/IJMS22115762
88. Sterz J, von Metzler I, Hahne JC et al (2008) The potential of proteasome inhibitors in cancer therapy. Expert Opin Investig Drugs 17:879–895. https://doi.org/10.1517/13543784.17.6.879
89. Raaben M, Posthuma CC, Verheije MH et al (2010) The ubiquitin-proteasome system plays an important role during various stages of the coronavirus infection cycle. J Virol 84:7869–7879. https://doi.org/10.1128/jvi.00485-10
90. Schneider M, Ackermann K, Stuart M et al (2012) Severe acute respiratory syndrome coronavirus replication is severely impaired by MG132 due to proteasome-independent inhibition of M-calpain. J Virol 86:10112–10122. https://doi.org/10.1128/jvi.01001-12
91. Wang J, Liang B, Chen Y et al (2021) A new class of α-ketoamide derivatives with potent anticancer and anti-SARS-CoV-2 activities. Eur J Med Chem 215:113267. https://doi.org/10.1016/J.EJMECH.2021.113267
92. Banerjee S, Wei T, Wang J et al (2019) Inhibition of dual-specificity tyrosine phosphorylation-regulated kinase 2 perturbs 26S proteasome-addicted neoplastic progression. Proc Natl Acad Sci U S A 116:24881–24891. https://doi.org/10.1073/PNAS.1912033116/-/DCSUPPLEMENTAL
93. Banerjee S, Ji C, Mayfield JE et al (2018) Ancient drug curcumin impedes 26S proteasome activity by direct inhibition of dual-specificity tyrosine-regulated kinase 2. Proc Natl Acad Sci U S A 115:8155–8160. https://doi.org/10.1073/PNAS.1806797115/-/DCSUPPLEMENTAL
94. Caputi FF, Di Cesare Mannelli L, Rullo L, et al (2020) The active second-generation proteasome inhibitor oprozomib reverts the oxaliplatin-induced neuropathy symptoms. Biochem Pharmacol 182:114255. https://doi.org/10.1016/J.BCP.2020.114255
95. Nguyen KT, Mun SH, Lee CS (2018) Hwang CS (2018) Control of protein degradation by N-terminal acetylation and the N-end rule pathway. Exp Mol Med 507(50):1–8. https://doi.org/10.1038/s12276-018-0097-y
96. Leleu X, Martin TG, Einsele H et al (2019) Role of proteasome inhibitors in relapsed and/or refractory multiple myeloma. Clin Lymphoma, Myeloma Leuk 19:9–22. https://doi.org/10.1016/j.clml.2018.08.016
97. Curran MP, McKeage K (2009) Bortezomib: a review of its use in patients with multiple myeloma. Drugs 69:859–888. https://doi.org/10.2165/00003495-200969070-00006
98. FDA (2003) VELCADE® (bortezomib) for injection, for subcutaneous or intravenous use. In: Accessdata.fda.gov. https://www.accessdata.fda.gov/drugsatfda_docs/label/2008/021602s015lbl.pdf
99. Shirley M (2016) Ixazomib: first global approval. Drugs 76:405–411. https://doi.org/10.1007/s40265-016-0548-5
100. Demo SD, Kirk CJ, Aujay MA et al (2007) Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome.

Cancer Res 67:6383–6391. https://doi.org/10.1158/0008-5472.CAN-06-4086
101. Groen K, Van de Donk NWCJ, Stege CAM et al (2019) Carfilzomib for relapsed and refractory multiple myeloma. Cancer Manag Res 11:2663–2675. https://doi.org/10.2147/CMAR.S150653
102. Obrist F, Manic G, Kroemer G, Vitale I, Galluzzi L (2015) Trial watch: proteasomal inhibitors for anticancer therapy. Mol Cell Oncol 2(2):e974463–1. https://doi.org/10.4161/23723556.2014.974463
103. Zhu H, Wang T, Xin Z et al (2019) An oral second-generation proteasome inhibitor oprozomib significantly inhibits lung cancer in a p53 independent manner in vitro. Acta Biochim Biophys Sin (Shanghai) 51:1034–1040. https://doi.org/10.1093/abbs/gmz093
104. Ng YL, Salim CK, Chu JJH (2021) Drug repurposing for COVID-19: approaches, challenges and promising candidates. Pharmacol Ther 228:107930. https://doi.org/10.1016/j.pharmthera.2021.107930
105. Pushpakom S, Iorio F, Eyers PA et al (2018) Drug repurposing: progress, challenges and recommendations. Nat Publ Gr. https://doi.org/10.1038/nrd.2018.168
106. Serafin MB, Bottega A, Foletto VS et al (2020) Drug repositioning is an alternative for the treatment of coronavirus COVID-19. Int J Antimicrob Agents 55:1–4. https://doi.org/10.1016/j.ijantimicag.2020.105969
107. Harris MJCAIJLMARFKM (2004) The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 32:258–261. https://doi.org/10.1093/nar/gkh036

4. ARTIGO III: *Virus goes viral: an educational kit for virology classes*

Este é um artigo original publicado no periódico "Virology Journal" em janeiro de 2020.

4.1 Introdução

A disseminação de noticiais falsas na pandemia de COVID-19 tem uma relação direta com a dificuldade de se transmitir conhecimento sobre vírus e doenças para a população em geral. A complexidade dos conceitos científicos envolvidos, a linguagem técnica e a falta de conhecimento prévio de grande parte da população sobre o assunto são alguns dos fatores que dificultam a comunicação clara e eficaz sobre a doença e suas formas de prevenção e tratamento. Isso abre espaço para a disseminação de informações falsas e desinformação, que muitas vezes são mais simples e apelativas, mas que podem ter consequências graves para a saúde pública. Nesse sentido, é importante que os virologistas e profissionais de saúde não apenas gerem conhecimento, mas também trabalhem na disseminação de informações precisas e acessíveis para o público em geral, de forma a combater a desinformação e as notícias falsas.

Na sala de aula, essa lacuna na transferência de conhecimento é ainda mais grave, pois falha no instrumento de instrução de parte da população. Nesta lacuna identificam-se inúmeros fatores, como a dificuldade de se realizar aulas práticas em disciplinas de virologia, devido aos riscos de biossegurança e aos altos custos envolvidos nesse tipo de aula (AKASHI *et al.*, 2019). Outro fator apontado é que são difíceis as observações praticas dos vírus devido ao seu tamanho, o que faz com que os alunos se restrinjam a figuras esquemáticas, ilustrações e imagens de microscopia eletrônica, o que torna o conhecimento gerado um tanto abstrato.

Contudo, desde o início deste século, com a descoberta dos vírus gigantes, essa limitação de tamanho tornou-se obsoleta na virologia, mas ainda é pouco explorada em sala de aula. Os vírus gigantes, tipicamente infectam amebas e são caracterizados por possuírem grandes partículas virais associadas a extensos genomas e foram responsáveis pela quebra de diversos paradigmas da virologia clássica, além de por apresentar genes jamais identificados anteriormente na vasta virosfera (BORATTO *et al.*, 2022).

O tamanho colossal de suas partículas dentro da virosfera que torna estes vírus tão especiais para se desenvolver ferramentas para aulas práticas de virologia, afinal, eles podem ser visualizados por microscopia óptica, assim como as bactérias e os fungos, que são tradicionalmente apresentados aos alunos através de microscópios comuns. Pensando nisso, desenvolveu-se um kit de lâminas, que fosse inédito e de baixo custo, para ser utilizado em aulas práticas de virologia. Esse kit conta com lâminas para a visualização de partículas de vírus gigantes e suas fábricas virais em

células de amebas, assim como lâminas que apresentam o efeito citopático de vírus animais em monocamadas de células. O *Virus Goes Viral* deve ser trabalhado para explorar os aspectos peculiares e inspiradores dos vírus.

4.2 Justificativa

Nas salas de aula de virologia, seja para alunos de graduação ou de ensino médio, existe uma grande dificuldade em se tornar os vírus conceitos menos abstratos, vistos apenas por figuras em livros didáticos. Aulas práticas de virologia são particularmente desafiadoras de serem elaboradas, pois geralmente envolvem altos custos e uma preocupação com a biossegurança. Contudo, romper essas dificuldades é uma necessidade para garantir que os alunos alcancem a melhor compreensão possível nos estudos dos vírus. Um potencial pouco explorado em salas de aulas são o dos vírus gigantes, que são uma descoberta relativamente recente, e por serem vírus de amebas podem ser utilizados em salas de aula sem grandes riscos de biossegurança e devido a suas partículas visíveis a microscopia ótica comum, podem ser visualizados assim como é feito com bactérias de tamanho similar, dando oportunidade de que os alunos pudessem visualizar na prática sua morfologia e tamanho, correlacionando estas características com a grande diversidade presente na virosfera. O desenvolvimento de um kit que permita que o educador melhore a sua forma de transferir o seu conhecimento aos alunos, dando a eles a oportunidade de fazer ciência no lugar de apenas estudar sobre ela pode transformar o ambiente de aula em um lugar inspirador.

5.3 Objetivos

4.3.1 Objetivo Geral

Construir uma ferramenta de uso seguro e de baixo custo para aulas práticas de virologia.

4.3.2 Objetivos Específicos

- Confeccionar lâminas de microscopia ótica que permitam a visualização de partículas de vírus gigantes, destacando sua morfologia e tamanho.
- Confeccionar lâminas que facilitem a visualização de alterações morfológicas em células infectadas por virus.
- Fundamentar os docentes com conceitos básicos de virologia e instruções para o uso do material confeccionado.

ANEXO A

Artigo III

Virus goes viral: an educational kit for virology classes

DOI: 10.1186/s12985-020-1291-9

Souza *et al. Virology Journal* (2020) 17:13
https://doi.org/10.1186/s12985-020-1291-9

Virology Journal

RESEARCH Open Access

Virus goes viral: an educational kit for virology classes

Gabriel Augusto Pires de Souza[1†], Victória Fulgêncio Queiroz[1†], Maurício Teixeira Lima[1†], Erik Vinicius de Sousa Reis[1], Luiz Felipe Leomil Coelho[2] and Jônatas Santos Abrahão[1*]

Abstract

Background: Viruses are the most numerous entities on Earth and have also been central to many episodes in the history of humankind. As the study of viruses progresses further and further, there are several limitations in transferring this knowledge to undergraduate and high school students. This deficiency is due to the difficulty in designing hands-on lessons that allow students to better absorb content, given limited financial resources and facilities, as well as the difficulty of exploiting viral particles, due to their small dimensions. The development of tools for teaching virology is important to encourage educators to expand on the covered topics and connect them to recent findings. Discoveries, such as giant DNA viruses, have provided an opportunity to explore aspects of viral particles in ways never seen before. Coupling these novel findings with techniques already explored by classical virology, including visualization of cytopathic effects on permissive cells, may represent a new way for teaching virology. This work aimed to develop a slide microscope kit that explores giant virus particles and some aspects of animal virus interaction with cell lines, with the goal of providing an innovative approach to virology teaching.

Methods: Slides were produced by staining, with crystal violet, purified giant viruses and BSC-40 and Vero cells infected with viruses of the genera *Orthopoxvirus*, *Flavivirus*, and *Alphavirus*. Slides with amoebae infected with different species of giant viruses and stained with hemacolor reagents were also produced.

Results: Staining of the giant viruses allowed better visualization of the viral particles, and this technique highlights the diversity in morphology and sizes among them. Hemacolor staining enabled visualization of viral factories in amoebae, and the staining of infected BSC-40 and Vero cell monolayers with crystal violet highlights plaque-forming units.

Conclusions: This kit was used in practical virology classes for the Biological Sciences course (UFMG, Brazil), and it will soon be made available at a low-cost for elementary school teachers in institutions that have microscopes. We hope this tool will foster an inspiring learning environment.

Keywords: Virology, Virology education, Microbiology education, biology education, Didactic material, Giant viruses, Mimivirus, Poxvirus, Flavivirus, Alphavirus

* Correspondence: jonatas.abrahao@gmail.com
†Gabriel Augusto Pires de Souza, Victória Fulgêncio Queiroz and Maurício Teixeira Lima contributed equally to this work.
[1]Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Belo Horizonte 31270-901, Brazil
Full list of author information is available at the end of the article

Background

Viruses are the most numerous entities on Earth, and they are found in the majority of ecosystems [1]. Over a century after their discovery, viruses are often recognized by the population as pathogens associated with diseases and epidemics. They generate fear and fascination once they directly influence human life [2]. The study of viruses is known as virology; this subject is often considered to be a part of microbiology. In the early years, strong background in this discipline was essential for studying medicine and biology. Virology pedagogy presents several limitations and challenges, including a high monetary cost for materials and biosafety laboratory requirements. Additionally, students work with biohazardous materials that can endanger themselves and their colleagues [3].

The small size of viruses is a major obstacle that limits the study of virology. Therefore, learning viral morphology is often limited to electron microscopy figures and schematic illustrations of viruses [2]. This obstacle, however, became obsolete after the discovery of the *Acanthamoeba polyphaga* mimivirus (APMV), the first described amoeba giant virus [4–6]. APMV is the prototype species of the genus *Mimivirus* (4); it belongs to the nucleocytoplasmic large DNA viruses (NCLDV) [7–9]. Following the discovery of APMV, other giant viruses were described and characterized, including cedratvirus (CVG), pandoravirus (PDV), kaumoebavirus (KAUV), Orpheovirusbrasiliensis (OBRV), faustovirus (FSTV), and tupanvirus (TPV) [10–17]. Additionally, with the discovery of giant viruses, one of the paradigms of virology was broken, namely that viruses are considered to be "filterable organisms". This shift led researchers to think that a portion of the virosphere is trapped in filters and begin research on prospection and characterization of giant viruses [18].

Another key point in virology is the observation of cytopathic effects (CPEs) [19]. CPEs refer to structural changes in host cells that are caused by the viral invasion. Some viruses cause characteristic CPEs, and observation of these effects is an important tool for virologists concerned with isolating and identifying viruses [19, 20]. From an educational perspective, myriad CPEs are visible by students under the optical microscope, such as changes in cell morphology, inclusion bodies, and lysis plaques [2, 19, 20]. Many medically important animal viruses show these effects over the course of infection [20]. In this study, we present the CPEs of poxviruses (vaccinia virus [VACV] and cowpox virus [CPXV] and arboviruses (yellow fever virus [YFV], chikungunya virus [CHIKV], and mayaro virus [MAYV]) on materials that are safe for classroom application. Besides, we used a wide range of giant virus preparations where particles and other viral structures as well as CPEs can be visualized.

Practical activities in biology provide opportunities for students to perform science rather than only learn about it. This modality allows the educator to expand topics in the classroom and connect them to recent discoveries [21, 22]. Some studies highlight the importance of combining laboratory and theoretical courses to attain a deeper understanding and greater satisfaction [23, 24]. Our goal in the present work was to develop an innovative way to address the concrete aspects of viral particle and host cell effects. This method, from an educational perspective, combines hands-on experience with a classroom virology course to allow the instructor to expand the students (and her/his own) knowledge. Thus, the purpose of our "Virus Goes Viral" kit was to explore the revolutionary aspect of giant viruses and also the classically explored aspects of animal virus interactions with cell lines and integrate them into the study of virology.

Methods

Virus and cells

TPV, CVG, PDV, and Niemeyer virus (NYMV) were propagated individually using *Acanthamoeba castellanii* (American Type Culture Collection [ATCC] 30010) cultured at 32 °C in 175 cm^2 cell culture flasks with 50 mL peptone-yeast extract with glucose (PYG) medium supplemented with 25 mg/ml amphotericin B (Fungizone; Cristalia, São Paulo, Brazil) and 500 U/ml penicillin (Schering-Plough, Brazil). The amoebae were infected with at a multiplicity of infection (MOI) of 0.01. Cells were incubated until the expected CPEs appeared in culture, and the media was collected by two centrifugation steps. *A. castellanii* cells and cellular debris were first removed by centrifugation (400 x g for 10 min), and the particles were purified by centrifugation (36,000 x g for 1 h) through a sucrose cushion (40–50%), suspended in PAS, and stored at − 80 °C. KAUV, OBRV, and FTSV were propagated individually in *Vermamoeba vermiformis* (ATCC50237) cultivated at 32 °C in PYG medium with a MOI of 0.01. After the appearance of CPEs, the cells and supernatants were collected, with sterile serological pipettes, stored in sterile conical tubes, and the viruses were purified through ultracentrifugation with a sucrose cushion.

African green monkey kidney BSC-40 cells (ATCC CRL-2761) and Vero cells (ATCC CCL-81) were maintained in an atmosphere with 5% CO_2 at 37 °C in Eagle's minimum essential medium (MEM; Gibco BRL, Invitrogen, Carlsbad, CA, United States), supplemented with 5% fetal bovine serum (FBS; Cultilab, Brazil), 2.5 μg/mL amphotericin B, 500 U/mL penicillin (Cristalia), and 50 μg/mL streptomycin (Schering-Plough, São Paulo, Brazil).

YFV (vaccine strain 17DD), MAYV (Strain BeAr20290), CHIKV (Genotype ECSA – Strain BHI3762/H804917, kindly provided by Dr. Maurício Lacerda Nogueira),

VACV group I (Isolate Caragola eye I), VACV group II (Isolate Caragola eye II), and CPXV (strain Brighton Red) were individually multiplied in Vero cells in MEM that contained 1–2.5% FBS, 0.25 μg/mL amphotericin B, 100 U/mL penicillin 100 U/mL (Schering-Plough, Brazil), and 100 μg/mL streptomycin at 37 °C in an atmosphere with 5% CO_2 in a large 175 cm^2 bottle. Subsequently, for YFV, MAYV, and CHIKV, the cell infection supernatant was transferred to tubes and centrifuged at 3,000 x g for 5 min. The clarified supernatants from these centrifugations were stored at − 70 °C. VACV and CPXV were purified on a sucrose gradient as previously described [25] and stored at − 70 °C.

Slide preparation for microscopic visualization

Viral particles

Aliquots of purified virus were diluted 1:10 (CVG, OBRV, PdV, and TPV) or 1:20 (NYMV), and 10 μL of the appropriate dilution was placed on the center of glass slide. The liquid was spread using circular movements to obtain an approximately 1-cm diameter drop. The slide was kept at room temperature until the liquid dried on the surface. Subsequently, the virus was fixed by adding 200 μL methanol over the center of the slide and stained with crystal violet for 15 min. After staining, the slides were washed in running water and dried at room temperature. Once dried, the stained region was covered by a 13 mm glass coverslip and affixed to the slide with Canada balsam (Synth, Brazil).

Viral factories

Approximately 1 million *A. castellanii* cells in PYG medium were cultured in cell culture flasks (25 cm^2). After the amoebae adhered to the flask, NYMV was inoculated with a MOI of 1. After 12-h infection, cells were removed from the flask and centrifuged at 885 x g for 10 min and stained with hemacolor reagents (Renylab, Brazil). Once dried, the stained region was covered with a 13 mm glass coverslip using Canada balsam.

Plaque-forming units

Two cell lines were used for visualization of virus plaque-forming units. For MAYV, CHIKV, and YFV, Vero cells (2×10^5 cells/well) were plated on sterile 13 mm coverslips, grown in MEM with 5% FBS for 24 h, and maintained at 37 °C in an atmosphere with 5% CO_2. The next day, the cells were infected with viruses and observed until the characteristic CPE appeared (lysis plaques). Once CPEs were detected, the coverslips were fixed in formaldehyde (3.7% v/v) for 2 h and stained with crystal violet. After staining, the coverslips were washed, dried, and affixed to the glass slides using Canada balsam. For viruses from both VACV groups and CPXV, lysis plaque visualization was performed by infecting BSC-40 cells. The same procedure described above was followed to prepare this slide.

Inclusion bodies

BSC-40 cells (2×10^5 cells/well) were plated on sterile 13 mm coverslips, grown in MEM with 5% FBS for 24 h, and maintained at 37 °C in an atmosphere that contained 5% CO_2. The cells were infected with CPXV (MOI: 10), and once CPEs were detected, the coverslips were fixed in formaldehyde (3.7% v/v) for 2 h and stained with the solution rich in eosin from the hemacolor kit for 10 min. After staining, the coverslips were washed, dried, and affixed to glass slides using Canada balsam.

Results

Viral particle visualization slides

The slides produced with the purified viral particles stained with crystal violet allowed a simple evaluation of the distinct morphology and particle sizes for each isolate (Fig. 1). CVG, OBRV, and PDV particles were ovoid and measured ~ 1 μm in size and ~ 0.5 μm in diameter (Fig. 1). NYMV particles were ~ 0.6 μm and appeared spherical when viewed with optical microscopy. TPV particles were large and strongly stained with crystal violet. They appeared to have a spherical head and a cylindrical tail under optical microscopy (Fig. 1).

Viral factory visualization slides

Viral factories were visualized after staining the cells with the hemacolor kit. The *Mimivirus* viral factory, such as for NYMV, is very identifiable, and was represented by a light halo with a strongly marked dark purple center located in the cell cytoplasm (Fig. 1, blue arrow). The nucleus was stained dark purple (Fig. 1, yellow arrow).

CPE visualization slides

Two cell lines were infected with different viruses to produce CPE slides (Figs. 2 and 3). BSC-40 cells assays provided visibly distinct lysis plaques after infection with different viruses. CPXV generated large circular plaques of clear destruction with very stained round cell remains (Fig. 2). The presence of A-type eosinophilic inclusion bodies was also observed in CPXV-infected BSC-40 cells (Fig. 2). These inclusion bodies appeared as large pink circles in the host cell cytoplasm (Fig. 2, black arrow). Both VACV groups presented a distinct plaque phenotype; group I showed small lysis plaques, while group II showed large plaques with very stained and stretched cells that formed a web (Fig. 2). There were lysis plaques generated by arbovirus infection in the Vero cell monolayers. YFV generated largely undefined edge lysis plaques that included numerous stretched cells (Fig. 3).

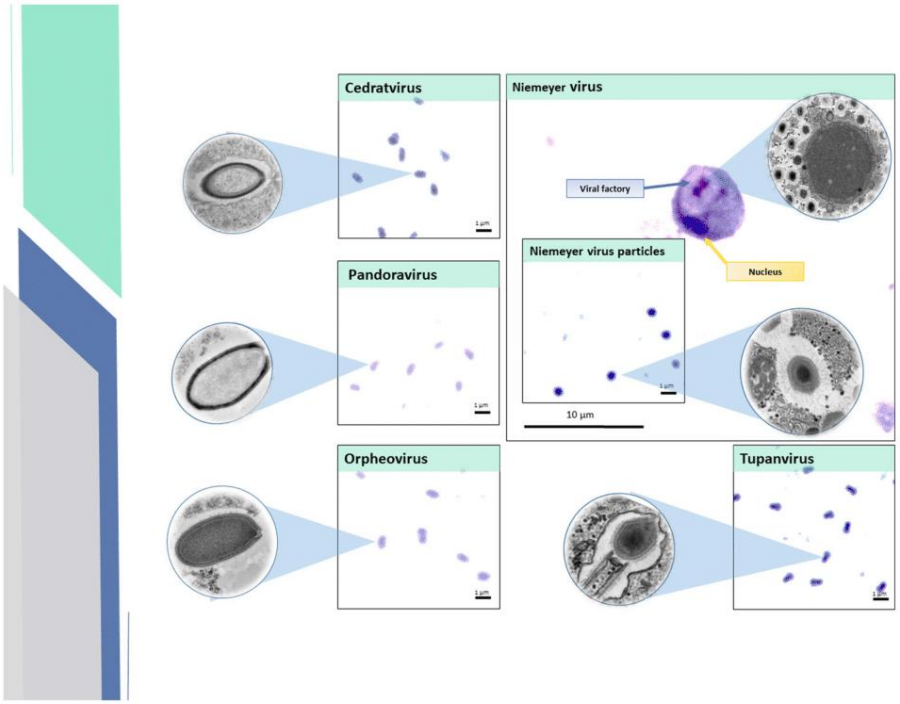


Fig. 1 Presentation of different giant virus particles (cedratvirus, pandoravirus, orpheovirus, Niemeyer virus and tupanvirus) and viral factories of Niemeyer virus under light microscopy. Visualization and comparison of optical and electron micrographs for visualization of viral morphology and factories. Total magnification is 1000x

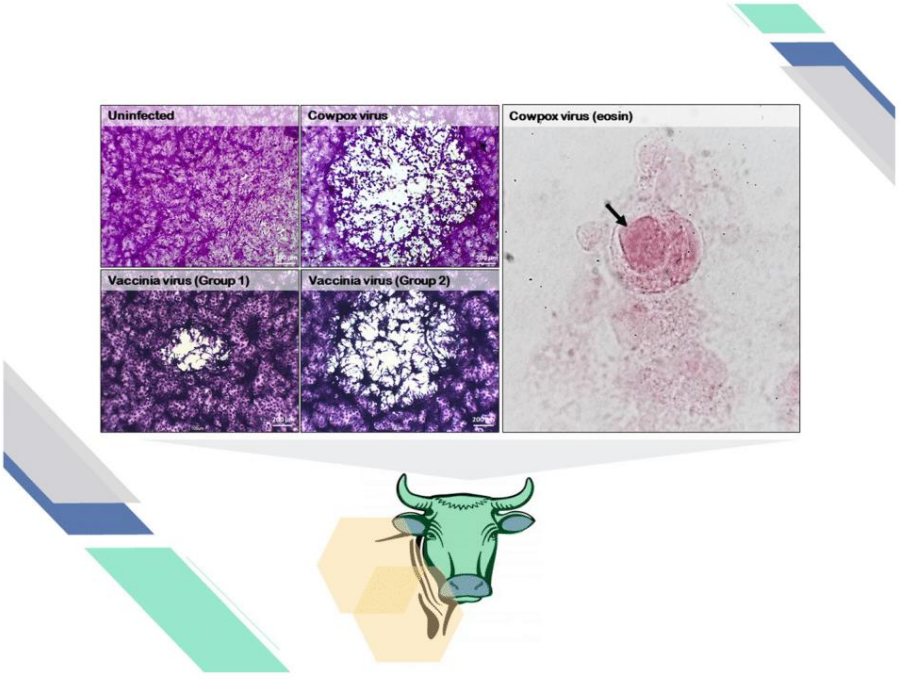


Fig. 2 Cytopathic effects of orthopoxvirus in BSC-40 cells. Monolayers of cells infected or not infected with orthopoxvirus and stained with crystal violet. Visualization of inclusion corpuscle in an eosin-stained infected cell is also shown. Total magnification is 100x

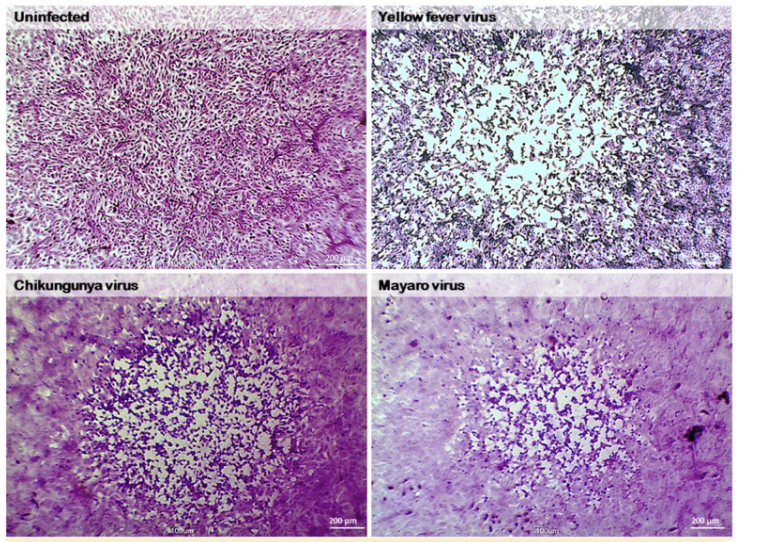


Fig. 3 Cytopathic effects of arboviruses in Vero cells. Monolayers of cells infected or not infected with arboviruses (flavivirus and alphavirus) and stained with crystal violet. Total magnification is 100x

CHIKV presented large circular plaques, while MAYV showed small undefined edge lysis plaques (Fig. 3).

"Virus Goes Viral" kit and associated teaching materials

The slides generated in previous sections were used to compose the "Virus Goes Viral" kit. The kit has a total of 15 labeled slides: five viral particle slides, nine CPE slides, and one viral factory slide (Fig. 4). The kit includes associated teaching materials on a CD-ROM, including an infographic material, which aims to assist in understanding what is being observed (Additional file 3) and also high-resolution model images of each slide and CPE images from several giant viruses in *A. castellanii* and *V. vermiformis* that were not fixed and stained in slide form (Additional file 1: Figure S1 and Additional file 2: Figure S2). These CPEs include: cell rounding for CVG, Marseillevirus, PDV, TPV, and KAUV; cell stretching for OBRV; cell bunches for TPV; lysis plaques for FSTV (Additional file 1: Figure S1 and Additional file 2: Figure S2). The "Virus Goes Viral" kit and its attached material encompass 13 different viral species and 17 miscellaneous effects in host cells. This kit has been used in practical virology classes for the Biological Sciences course (Universidade Federal de Minas Gerais [UFMG], Brazil). The students seemed very receptive to the proposal and many were excited about the practice. Some expressed interest in starting research with giant viruses and initiated contact with our laboratory.

Discussion

The field of virology has a long and strong history of educational innovations. The actual challenge is to develop specific initiatives that can be broadly implemented early in the curriculum ([26], see Table 1). Virology is a very important subject for several undergraduate courses. Students from medical and biological courses need to learn virology; however, costs and adequate facilities, including biological hoods in aseptic rooms, are frequently obstacles [3]. Although digital learning and other approaches to traditional texts or lectures increase student receptiveness, standard educational materials, such as microscope reference slides, are invaluable pedagogical resources [27, 28]. Two main goals led us to develop the "Virus Goes Viral" kit. The first goal was to foster better understanding among students about the basic concepts of virology. Achieving this goal requires only simple installations with one light microscope or computer, and these measures are congruent to the Brazilian reality.

The second goal was to introduce and spread knowledge about giant viruses to high school and undergraduate

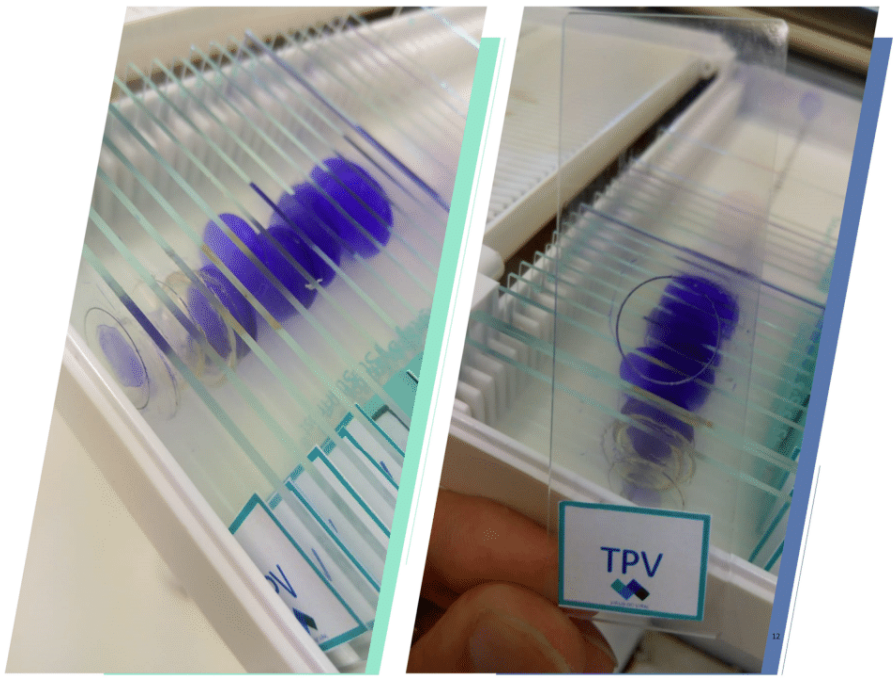


Fig. 4 "Virus Goes Viral" kit. The stock of kit blades that highlights the individual labels

students. Our group has been dedicated to the exploration of giant viruses and isolates from different samples, including water, soil, sewage, and clinical samples, as well as in extreme environments, for example, permafrost and soda lakes [12, 17, 29, 30]. Giant viruses led biologists to rethink the nature of life. In recent years, giant viruses have been widely publicized in magazines, newspapers, blogs, television channels, and video sharing platforms; this coverage is an important example of scientific dissemination.

A common misconception in the general population is that all microorganisms are associated with human diseases [31]. Integrating giant viruses into the classroom represents a way to expand knowledge and dispel the notion that viruses are solely human pathogens. Indeed, giant viruses highlight that all organisms can be infected with viruses. Besides, giant viruses can represent tools to help the general public understand the importance of microorganisms, their evolutionary biology and ecology in the ecosystem, and their ancestry. With this material, we believe the educator can explore the revolutionary aspects of these viruses and create an inspiring learning environment. This is intended to be a simple but useful material for teaching virology. It can and should be associated with other educational strategies, including other hands-on lessons, according to the educator's plan and institution resources.

Our kit covers 13 viral species, including vertebrate viruses, protozoan viruses, and arboviruses. This range of viral models illustrates for students the diversity of the virosphere. Besides, the CPEs covered by the kit are an important discussion point with students about specific interactions between virus and host. An interesting example of virus-host interactions is the mechanism of TPV and amoebas [32]. Virus-induced infected cells can aggregate with uninfected cells to form bunches that can increase TPV fitness [32]. These bunches are present in the "Virus Goes Viral" attached material. Another example is the VACV isolates that show a biological diversity that allows separating the viruses into different groups [33]. The VACV strain shows plaque phenotypes that differ in size; for example, the VACV group II has larger plaques than the VACV group I [33]. Our kit contains slides with VACV differentiation to make it clear to students that even within the same viral species there is variation.

Conclusion

We designed a slide microscope kit named "Virus Goes Viral" that explores giant virus particles and some aspects of animal virus interactions with cells, with the aim to provide an innovative approach in virology teaching. Slides were produced by staining, with crystal violet, purified giant viruses and BSC-40 and Vero cells infected with viruses of the genera *Orthopoxvirus*, *Flavivirus*, and *Alphavirus*. The kit contains a slide for viral factory visualization in amoebas stained with the hemacolor kit as well as associated teaching materials on a CD-ROM, including high-resolution model images of each slide

and CPE images of several giant viruses in *A. castellanii* and *V. vermiformis*. This kit has been used in practical virology classes for the Biological Sciences course (UFMG, Brazil), and it will soon be made available at a low cost to elementary school teachers in institutions that have microscopes. We hope this tool will foster an inspiring learning environment about virology.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12985-020-1291-9.

Additional file 1: Figure S1. The cytopathic effects caused by different viruses on *Acanthamoeba castellanii*.

Additional file 2: Figure S2. The cytopathic effects caused by different viruses on *Vermamoeba vermiformis*.

Additional file 3. Infographics related to *Virus Goes Viral* educational kit.

Abbreviations

APMV: *Acanthamoeba polyphaga* mimivirus; CHIKV: Chikungunya virus; CPEs: Cytopathic effects; CPXV: Cowpoxvirus; CVG: Cedratvirus; DNA: Deoxyribonucleic acid; FBS: Fetal bovine serum; FSTV: Faustovirus; MAYV: Mayaro virus; MOI: Multiplicity of infection; NCLDV: Nucleocytoplasmic large DNA viroses; NYMV: Niemeyer virus; OBRV: Orpheovirus brasiliensis; PAS: Phophate saline; PDV: Pandora virus; PYG: Peptone-yeast extract with glucose; TPV: Tupanvirus; VACV: Vaccinia virus; YFV: Yellow fever virus

Acknowledgements

We are grateful to our colleagues from the Laboratório de Vírus of Universidade Federal de Minas Gerais. Additionally, we thank the Laboratório de Taxomomia, Biodversidade e Biotecnologia de Fungos for lending us their microscope.

Authors' contributions

GS, VQ, ML and ER produced the kit material. GS, VQ and ML captured the images. G.S. worked in the design of images for the attached material. GS, VQ, ML and JA wrote the manuscript. All authors read the final version of the manuscript.

Funding

We thank Conselho Nacional de Desenvolvimento Científco e Tecnológico (CNPq), Ministério da Saúde (MS), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do estado de Minas Gerais (FAPEMIG), and Pro-Reitoria de Pesquisa of UFMG (PRPq-UFMG). J.S.A. and L.C. are CNPq researchers. E.R. and M.L. received fellowships from CAPES-PROEX. G.S. received fellowships from FAPEMIG and V.Q. received fellowships from CNPq (PIBIC). This study was financed in part by the CAPES (Finance Code 001).

Availability of data and materials

All relevant information is provided in this current manuscript. If required, the data presented in this work can be shared by e-mail.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Author details

[1]Laboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Belo Horizonte 31270-901, Brazil. [2]Laboratório de Vacinas, Departamento de Microbiologia e Imunologia, Instituto de Ciências Biomédicas, Universidade Federal de Alfenas, Rua Gabriel Monteiro da Silva, 700, Alfenas 37130-001, Minas Gerais, Brazil.

Received: 27 September 2019 Accepted: 22 January 2020
Published online: 31 January 2020

References

1. Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, et al. Uncovering Earth's virome. Nature. 1996;536(7617):425–30.
2. Akashi M, Fukaya S, Uchiyama C, Aoki K, Takemura M. Visualization of giant virus particles and development of "VIRAMOS" for high school and university biology course. Biochem Mol Biol Educ. 2019;47:426–31.
3. Matza-Porges S, Nathan D. A biosafety level 2 virology lab for biotechnology undergraduates. Biochem Mol Biol Educ. 2017;45:537–43.
4. La Scola B, Audic S, Robert C, Jungang L, De Lamballerie X, Drancourt M, et al. A giant virus in amoebae. Science. 2003;299:2033.
5. Raoult D, Scola BL, Birtles R. The discovery and characterization of Mimivirus, the largest known virus and putative pneumonia agent. Clin Infect Dis. 2007;45:95–102.
6. Sharma V, Colson P, Pontarotti P, Raoult D. Mimivirus inaugurated in the 21st century the beginning of a reclassification of viruses. Curr Opin Microbiol. 2016;31:16–24.
7. Van Etten JL, Lane LC, Dunigan DD. DNA viruses: the really big ones (giruses). Annu Rev Microbiol. 2010;64:83–99.
8. Colson P, De Lamballerie X, Fournous G, Raoult D. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. Intervirology. 2012;55:321–32.
9. Colsonf P, Yutinf N, Shabalina SA, Robert C, Fournous G, La Scola B, et al. Viruses with more than 1,000 genes: Mamavirus, a new Acanthamoeba polyphaga mimivirus strain, and reannotation of mimivirus genes. Genome Biol Evol. 2011;3:737–42.
10. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, et al. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci U S A. 2009;106:21848–53.
11. Philippe N, Legendre M, Doutre G, Couté Y, Poirot O, Lescot M, et al. Pandoraviruses: Amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. Science. 2013;341:281–6.
12. Boratto PVM, Arantes TS, Silva LCF, Assis FL, Kroon EG, La Scola B, et al. Niemeyer virus: a new mimivirus group a isolate harboring a set of duplicated aminoacyl-tRNA synthetase genes. Front Microbiol. 2015;6:1256.
13. Reteno DG, Benamar S, Khalil JB, Andreani J, Armstrong N, Klose T, et al. Faustovirus, an Asfarvirus-related new lineage of giant viruses infecting amoebae. J Virol. 2015;89:6858–594.
14. Andreani J, Aherfi S, Khalil JYB, Di Pinto F, Bitam I, Raoult D, et al. Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. Viruses. 2016;3:E300.
15. Bajrai LH, Benamar S, Azhar EI, Robert C, Levasseur A, Raoult D, et al. Kaumoebavirus, a new virus that clusters with Faustoviruses and Asfarviridae. Viruses. 2016;28:E278.
16. Andreani J, Khalil JYB, Baptiste E, Hasni I, Michelle C, Raoult D, et al. Orpheovirus IHUMI-LCC2: a new virus among the giant viruses. Front Microbiol. 2018;8:2643.
17. Abrahão J, Silva L, Silva LS, Khalil JYB, Rodrigues R, Arantes T, et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. Nat Commun. 2018;9:749.
18. Rodrigues RAL, Andrade ACDSP, Boratto PVM, Trindade GS, Kroon EG, Abrahão JS. An anthropocentric view of the virosphere-host relationship. Front Microbiol. 2017;8:1673.
19. Knipe DM, Howley P. Fields Virology. 6th ed. Philadephia: Lippincott Williams & Wilkins; 2013.
20. Netherton C, Moffat K, Brooks E, Wileman T. A guide to viral inclusions, membrane rearrangements, factories, and viroplasm produced during virus replication. Adv Virus Res. 2007;70:101–82.
21. Çimer A. Effective teaching in science: a review of literature. J Turkish Sci Educ. 2007;4:20–44.
22. Laursen S, Liston C, Thiry H, Graf J. What good is a scientist in the classroom? Participant outcomes and program design features for a short-duration science outreach intervention in K-12 classrooms. CBE Life Sci Educ. 2007;6:49–64.

23. Aronson BD, Silveira LA. From genes to proteins to behavior: a laboratory project that enhances student understanding in cell and molecular biology. CBE Life Sci Educ. 2009;8:291–308.
24. Kipnis M, Hofstein A. The inquiry laboratory as a source for development of metacognitive skills. Int J Sci Math Educ. 2008;6:601–27.
25. Joklik WK. The purification of four strains of poxvirus. Virology. 1962;18:9–18.
26. Hatfull GF. Innovations in undergraduate science education: going viral. J Virol. 2015;89:8111–3.
27. Evans C. The effectiveness of m-learning in the form of podcast revision lectures in higher education. Comput Educ. 2008;50:491–8.
28. Michels KK, Michels ZD, Hotchkiss SC. Advantages of live microscope video for laboratory and teaching applications. Nat Sci Educ. 2016;45:1.
29. Andrade ACDSP, Arantes TS, Rodrigues RAL, Machado TB, Dornas FP, Landell MF, et al. Ubiquitous giants: a plethora of giant viruses found in Brazil and Antarctica. Virol J. 2018;15:22.
30. Silva LKDS, Andrade ACDSP, Dornas FP, Rodrigues RAL, Arantes T, Kroon EG, et al. Cedratvirus getuliensis replication cycle: an in-depth morphological analysis. Sci Rep. 2018;8:4000.
31. de Fraga FBFF. Towards an evolutionary perspective in teaching and popularizing microbiology. J Microbiol Biol Educ. 2018;19:19.1.45.
32. Oliveira G, Silva L, Leão T, Mougari S, da Fonseca FG, Kroon EG, et al. Tupanvirus-infected amoebas are induced to aggregate with uninfected cells promoting viral dissemination. Sci Rep. 2019;9:183.
33. Kroon EG, Abrahão JS, de Souza Trindade G, Oliveira GP, Luiz APMF, Costa GB, et al. Natural Vaccinia virus infection: diagnosis, isolation, and characterization. Curr Protoc Microbiol. 2016;42:14A.5.1–14A.5.43.

Publisher's Note

ANEXO B

Material Suplementar

Virus goes viral: Manual do Usuário

Virus Goes Viral

Manual do Usuário

Abundância

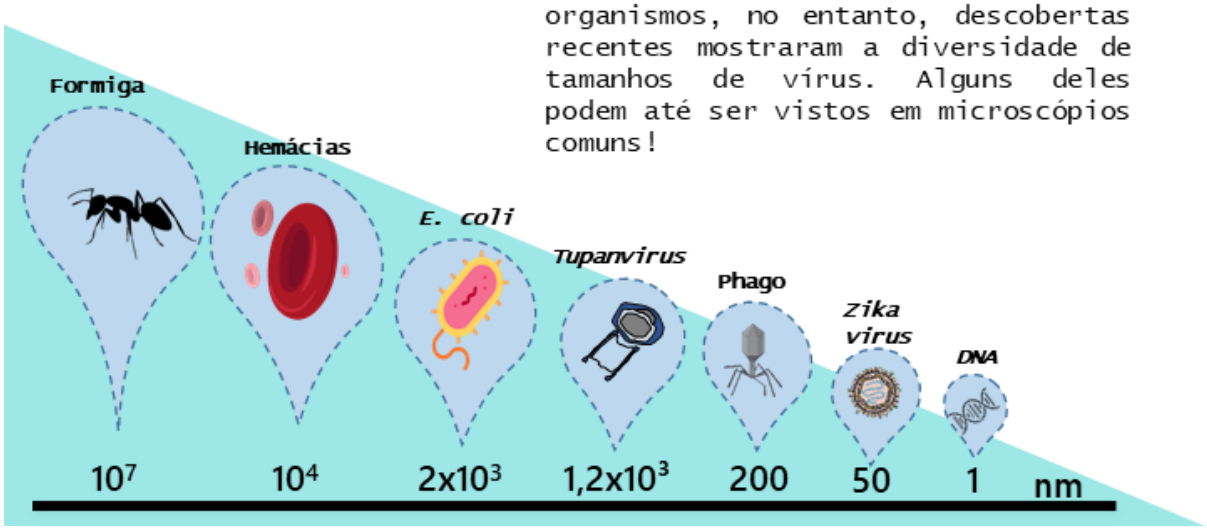
Os vírus são as entidades mais numerosas da Terra e são encontrados na maioria dos ecossistemas.

Patogênicos?

Embora os vírus estejam frequentemente associados à doença, nem todos são patogênicos. Eles representam fortes forças de seleção natural e controle populacional para todos os organismos vivos.

Menor dos menores?

Anteriormente, os vírus eram considerados o menor grupo de organismos, no entanto, descobertas recentes mostraram a diversidade de tamanhos de vírus. Alguns deles podem até ser vistos em microscópios comuns!

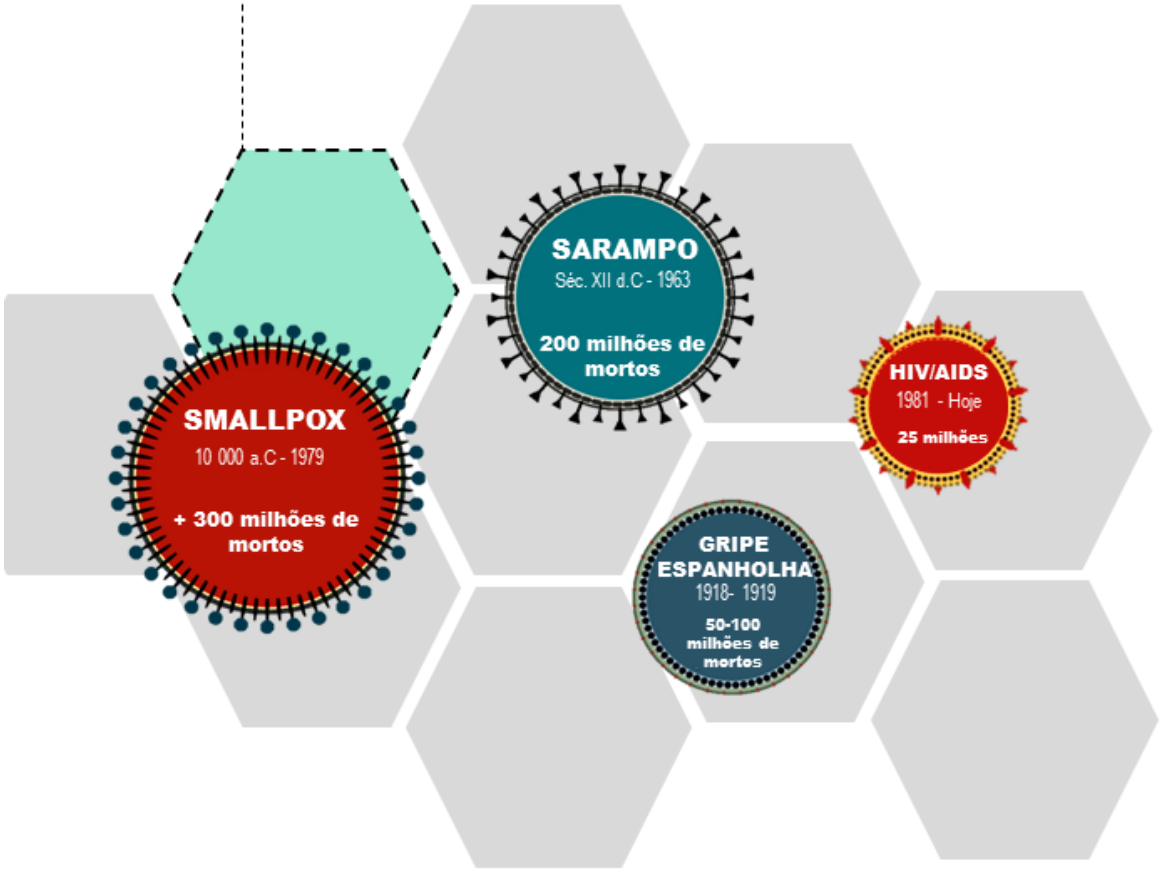


Virus Goes Viral

Manual do Usuário

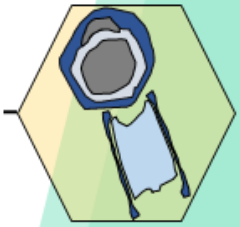
Protagonistas

Os vírus vêm sendo protagonistas de inúmeros episódios na história da humanidade. Viroses, como varíola e sarampo, causaram a morte de milhões de seres humanos somente no século passado.

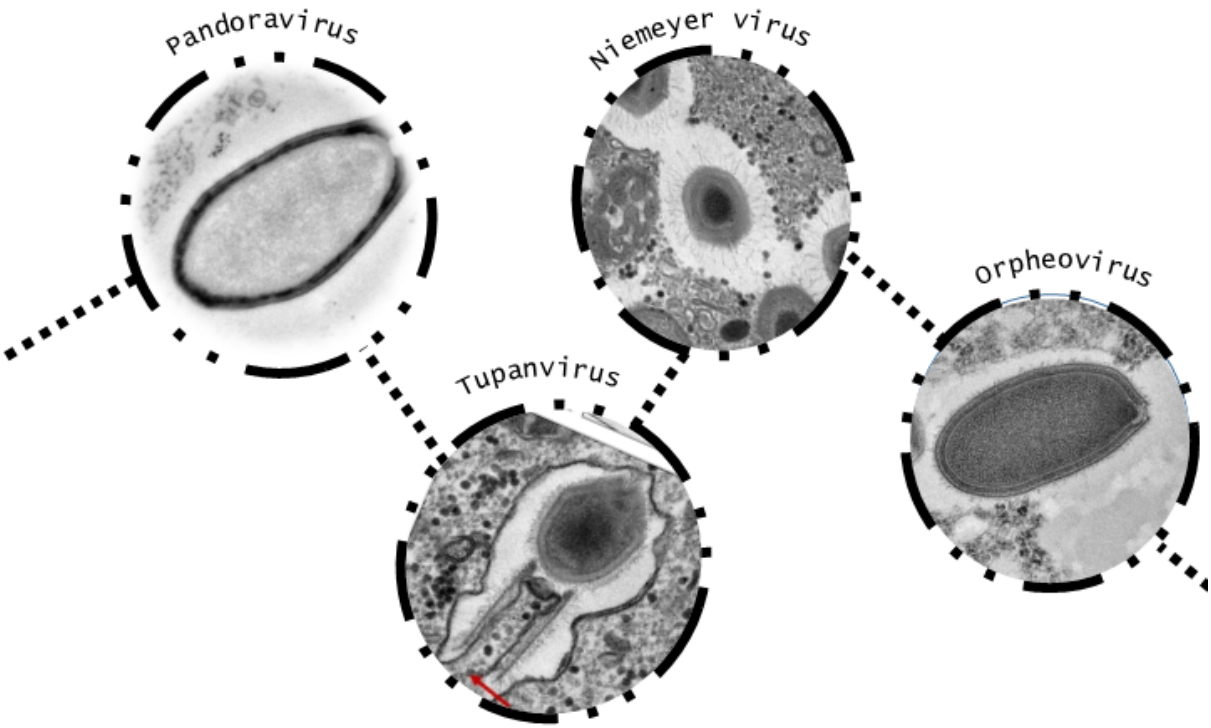


Virus Goes Viral
Manual do Usuário

Vírus gigantes



Como nome já indica os vírus gigantes possuem partículas de grandes dimensões, o que possibilita a visualização de alguns em microscópio óptico, além de genomas extensos e complexos, quando comparados a outros vírus. Eles fazem parte do grupo de vírus grandes núcleo-citoplasmáticos de DNA, que possuem representantes capazes de infectar hospedeiros vertebrados e invertebrados. Alguns exemplos de vírus gigantes que infectam amebas são: os mimivírus como os Tupanvírus, os Orpheovírus e os Pandoravírus

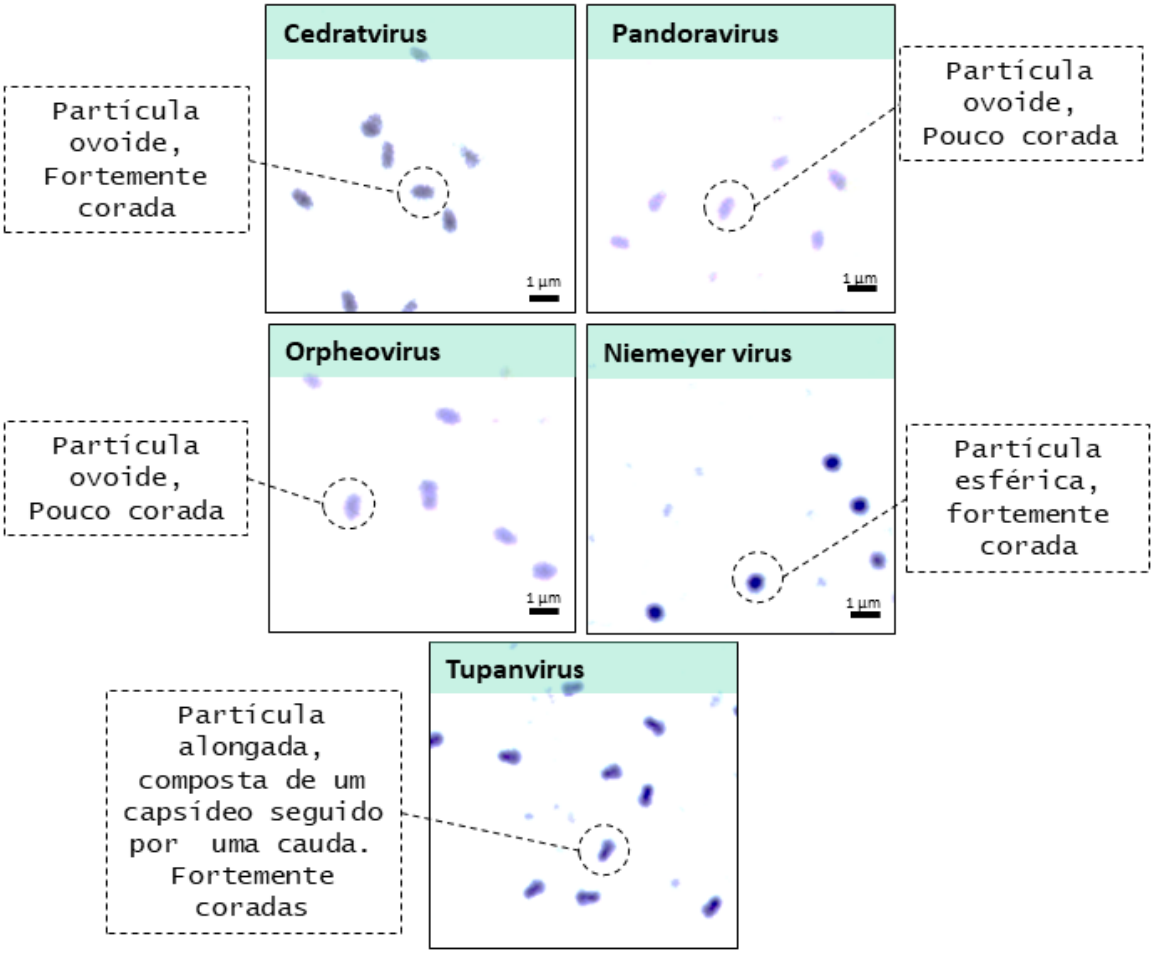


Virus Goes Viral
Manual do Usuário

- **Observando partículas virais**

Aumento: 100x

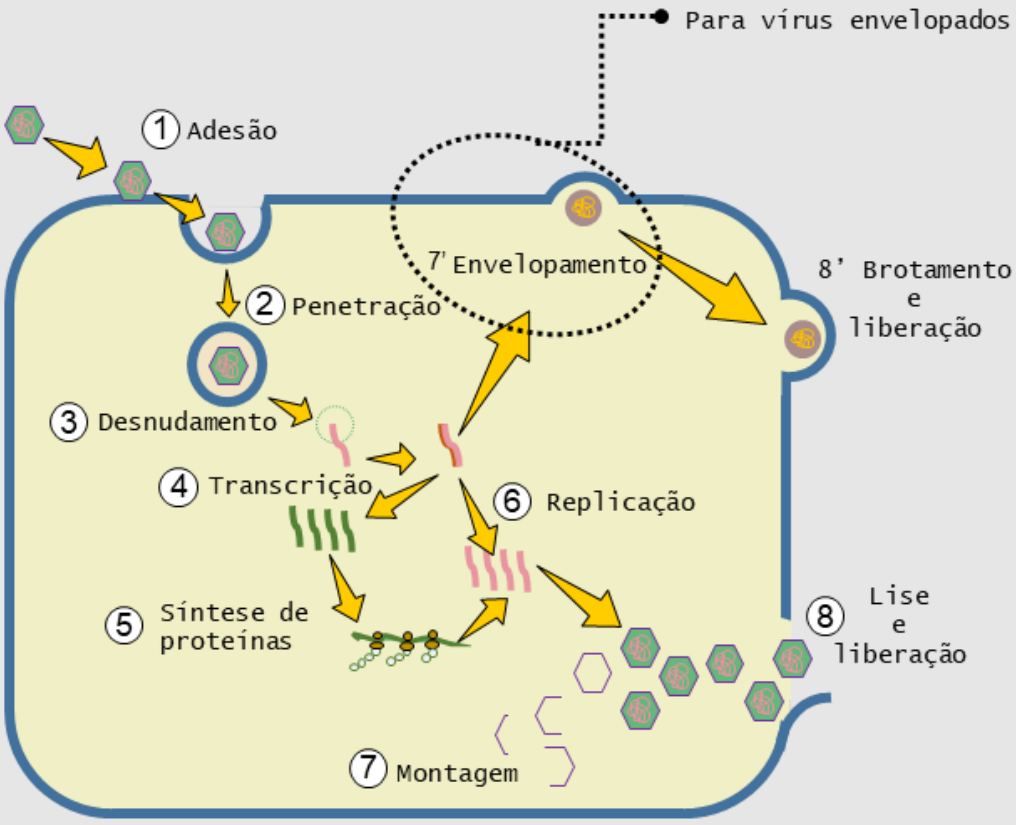
Óleo de imersão



Virus Goes Viral

Manual do Usuário

- Ciclo geral de replicação



Parasitas Intracelulares

Os vírus são parasitas intracelulares obrigatórios, e, portanto, utilizam o mecanismo das células infectadas para se multiplicar. Embora cada vírus possa ter suas próprias peculiaridades durante seu ciclo de replicação, geralmente segue as etapas descritas acima.

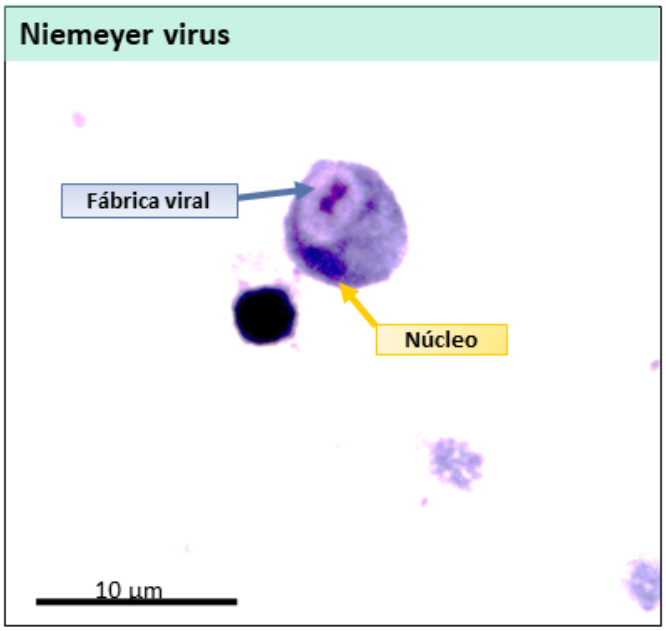
Virus Goes Viral
Manual do Usuário

- Fabricas Virais

As fábricas virais são compartimentos intracelulares nos quais ocorre, com alta eficiência, a montagem da partícula viral e também pode representar uma proteção contra as defesas do hospedeiro. As fábricas virais podem ser citoplasmáticas ou nucleares e geralmente surgem de um rearranjo extensivo dos compartimentos do citoesqueleto e/ou da membrana celular das células hospedeiras.

- Observando a fábrica viral

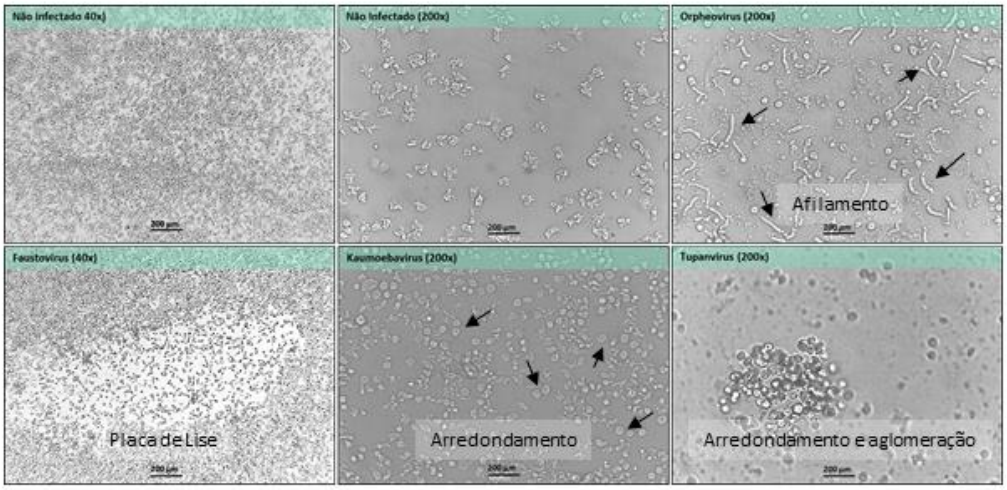
Aumento: 100x Óleo de imersão



Virus Goes Viral
Manual do Usuário

- **Interação vírus-célula: Efeito citopático**

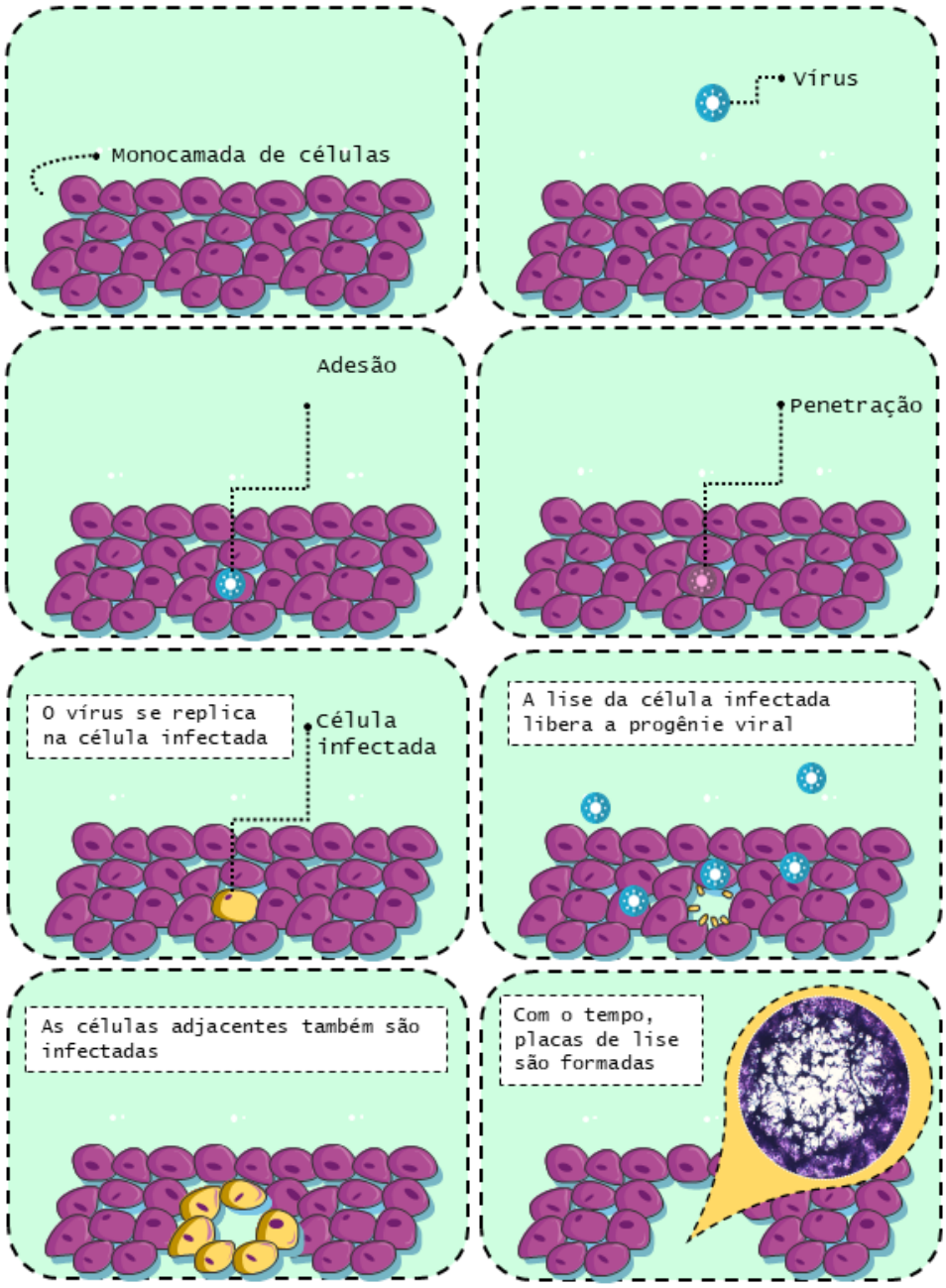
Um ponto chave na virologia é a observação de efeitos citopáticos (ECPs). ECPs referem-se a alterações estruturais nas células hospedeiras causadas pela invasão viral. Alguns vírus causam ECPs característicos, e a observação desses efeitos é uma ferramenta importante para os virologistas que buscam isolar e identificar vírus.



Efeitos citopáticos de diferentes vírus em *Vermamoeba vermiformis* (Ameba de vida livre)

Virus Goes Viral
Manual do Usuário

- Interação vírus-célula: como se formam as placas de lise?

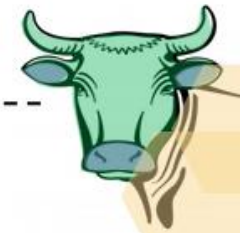


Virus Goes Viral

Manual do Usuário

Orthopoxvirus

São vírus que infectam muitos animais diferentes como bovinos, equinos e humanos e estão presentes em praticamente todo o planeta. Em humanos causam mal-estar, febre e lesões na pele, com alta capacidade de contágio; em cultura de células formam placas de lise. Alguns representantes são: o vaccínia vírus, o monkeypox vírus, o cowpox e o vírus da varíola.



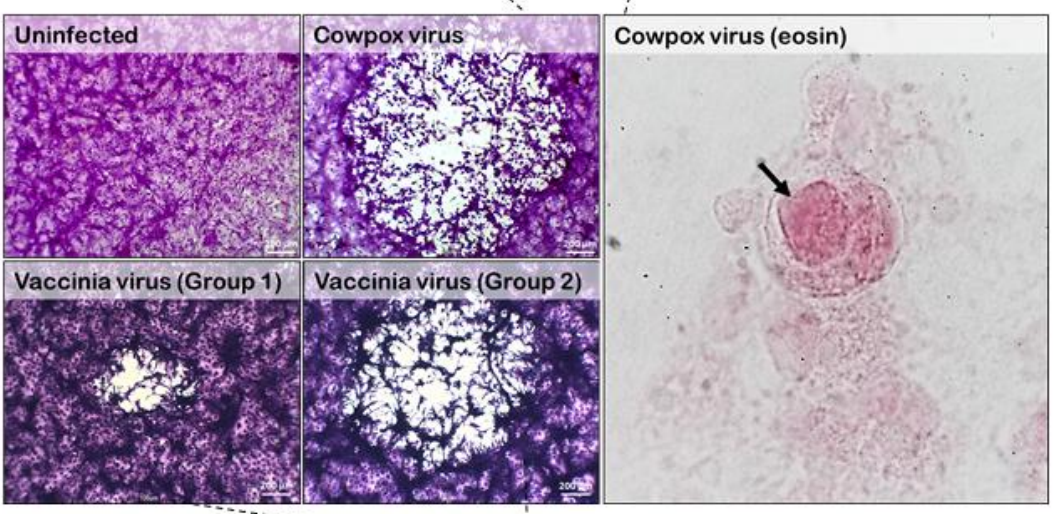
• Observando efeito citopático

Aumento: 40x

Linhagem celular: BSC-40 (Rim de Macaco)

Grandes **placas de lise** (parte clara) com restos celulares bem corados

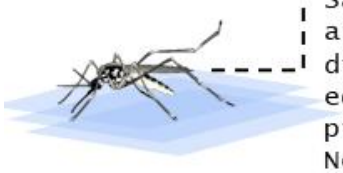
Corpúsculo de inclusão: grandes círculos cor-de-rosa no citoplasma da célula hospedeira



Os grupos VACV apresentam um fenótipo distinto de placa; o grupo I apresenta pequenas placas de lise, enquanto o grupo II apresenta grandes placas com células muito coradas, esticadas e que formaram uma teia

Virus Goes Viral

Manual do Usuário



• Arbovírus

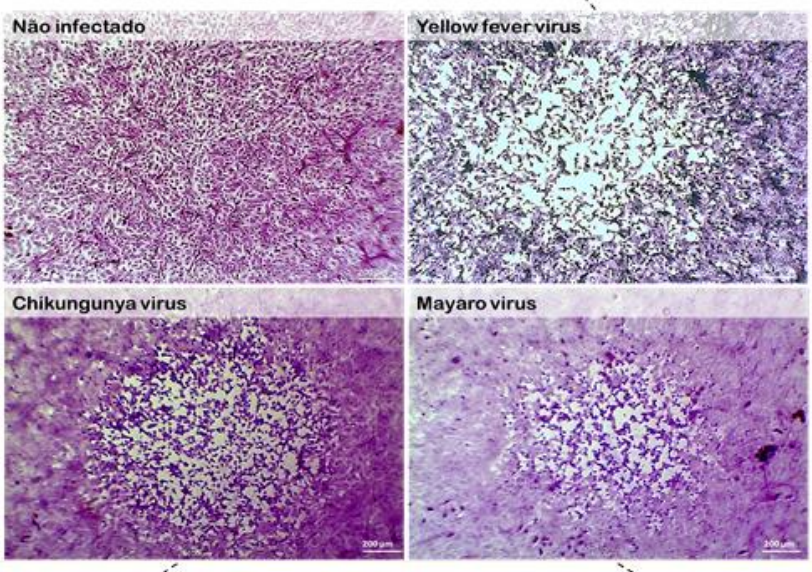
São um grupo de vírus que utilizam como vetores os artrópodes, como carrapatos e mosquitos. Atingem diversas regiões do planeta, em especial regiões equatoriais. O ciclo de transmissão ocorre após a picada de um artrópode infectado em um humano. Neste, os vírus podem causar sintomas inespecíficos, como febre, exantemas e dores articulares, podendo levar à morte; em cultura de células formam placas de lise. Exemplos clássicos: o dengue, o chikungunya, o febre amarela e o Zika.

• Observando efeito citopático

Aumento: 40x

Linhagem celular: Vero (Rim de Macaco)

Placas de lise grandes com bordas indefinidas que incluem numerosas células alongadas



Placas de lise grandes e circulares

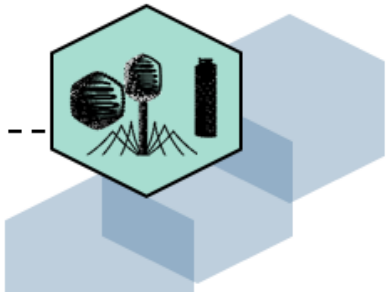
Placas de lise pequenas e com bordas indefinidas

- Lâminas fornecidas a partir dos seguintes vírus:

Vírus	Família	Gênero	Genoma	Tamanho do genoma	Tamanho da partícula	Hospedeiro principal	ECP
Cedratvirus	*Unassigned*	*Unassigned*	dsDNA	~590 kbp	1 µm	Amebas de vida livre	Arredondamento celular e perda de núcleo
Chikungunya virus	*Togaviridae*	*Alphavirus*	ssRNA	11.8 kb	70 nm	Primatas não humanos e mosquitos	Arredondamento, encolhimento e lise
Cowpox virus	*Poxviridae*	*Orthopoxvirus*	dsDNA	~222 kbp	350-250 nm	Roedores selvagens, gatos, seres humanos	Lise celular (placa de lise)
Mayaro virus	*Togaviridae*	*Alphavirus*	ssRNA	11.8 kb	70 nm	Primatas não humanos e mosquitos	Cytoplasmic blebbing, rounding, shrinkage and lysis
Niemeyer virus	*Mimiviridae*	*Mimivirus*	dsDNA	~1.3 Mb	~750 nm	Amebas de vida livre	Arredondamento celular e perda de núcleo
Orpheovirus	*Unassigned*	*Unassigned*	dsDNA	~1.4 Mb	1.3 µm	Amebas de vida livre	Afilamento e aumento da motilidade
Pandoravirus	*Unassigned*	*Unassigned*	dsDNA	~2.5 Mb	1 µm	Amebas de vida livre	Arredondamento celular e perda de núcleo
Tupanvirus	*Mimiviridae*	*Unassigned*	dsDNA	1.44 - 1.51 Mb	> 2 µm	Amebas de vida livre	Formação de cachos
Vaccinia virus	*Poxviridae*	*Orthopoxvirus*	dsDNA	~190 kbp	350-250 nm	Bovinos, búfalos, humanos	Lise celular (placa de lise)
Yellow fever virus	*Flaviviridae*	*Flavivirus*	(+)sense RNA	~11 kb	40-50 nm	Primatas não humanos, humanos, mosquitos (*Aedes* e *Haemagogus*)	Lise celular (placa de lise)

Diversidade

Os vírus são muito diversos, vêm em várias formas, tamanhos, diferentes materiais genéticos. Recursos selecionados ao longo da evolução para que consigam infectar seus hospedeiros.



6. CONSIDERAÇÕES FINAIS

As ferramentas *in sílico* são grandes aliadas nas previsões de fenômenos biológicos, e sem esse tipo de recurso, dificilmente seria possível analisar todo o grande volume de dados que são gerados atualmente, por exemplo, pela genômica, transcriptômica e proteômica (KUSHWAHA *et al.*, 2017). Esses dados requerem análises computacionais sofisticadas para tirar conclusões enquanto as previsões permitem tomadas de recurso em menor tempo, poupando recursos financeiros e até mesmo biológicos, no caso de testes *in vivo*, bem como o desgaste da própria equipe de pesquisa em extensos trabalhos *in vitro*.

Ainda assim, é comum que estudos realizados *in sílico* não sejam validados em etapas posteriores dos ensaios biológicos. Por mais avançadas que sejam as análises de bioinformática, a complexidade de um organismo é dificilmente reproduzível pelos algoritmos. Portanto, mesmo que possamos aplicar as análises de transcriptoma, integrá-las e identificar genes associados a patogênese viral e/ou alvos terapêuticos como apresentado nesta tese, haverá ainda uma necessidade de se testar se estas previsões podem ser validadas *in vitro* e *in vivo*.

Nas análises integrativas de transcriptoma aqui apresentadas, há uma limitação clara: a dificuldade de homogeneização das amostras. Visto que temos células de diferentes linhagens e até mesmo de tecidos distintos, encontrar fatores comuns é particularmente desafiador. Há de se compreender que o tempo de infecção é um fator limitante, visto que o vírus manipulará a expressão da célula hospedeira de forma a favorecer a sua replicação e conforme as necessidades do seu ciclo. Integrar os transcriptomas de células que estão em etapas distintas do ciclo de replicação viral pode representar a exclusão da análise genes chave do processo em etapas especificas do ciclo.

Além de se considerar o tempo e o ciclo de infecção, há também uma grande heterogeneidade em relação a multiplicidade de infecção (MOI) das amostras de transcriptômica. É sabido que a MOI tem grande impacto na análise, pois ela pode determinar o tipo de ciclo de infecção, com MOIs altas, temos maior probabilidade de termos todas as células infectadas, e portanto, todas estariam em um ciclo síncrono de infecção, enquanto com MOIs mais baixas, temos a possibilidade de acompanhar

mais ciclos mais com células de uma mesma amostra em diferentes momentos do ciclo, ou seja, um ciclo assíncrono de infecção.

Portanto, mesmo que os protocolos aqui pospostos tenham sido capazes de identificar genes diferencialmente expressos em células neuronais humanas infectadas com ZIKV e trazido pistas para o entendimento da patogênese viral; Ou ainda permitido a identificação de dois alvos terapêuticos e seus respectivos inibidores para o controle da doença causada pelo SARS-CoV-2. Há uma necessidade de se desenvolver algoritmos para a seleção e agrupamento destes conjuntos de dados de transcriptoma de acordo com as condições experimentais em que os dados foram gerados e isso enriqueceria significativamente os trabalhos que utilizam esta abordagem de análise integrativa de transcriptoma.

Apesar dessas limitações, os protocolos parecem eficientes em encontrar genes frequentemente diferencialmente expressos em células infectadas e podem ser uteis para o entendimento de novas doenças virais emergentes, seja na compreensão da biologia e patogênese viral, quanto para a identificação de alvos terapêuticos e reposicionamento de fármacos. Permanece, contudo, a necessidade de se testar as conclusões destes trabalhos em ensaios biológicos, promovendo o silenciamento dos genes frente a uma infecção pelo vírus em questão e o seu impacto na patogênese. Bem como os próprios ensaios antivirais, para se avaliar os impactos da inibição dos alvos pelas moléculas propostas avaliando na produção de SARS-CoV-2.

A pandemia de COVID-19, doença causada pelo SARS-CoV-2, trouxe para a sociedade várias mudanças no seu dia a dia: adotamos o uso de máscaras e intensificamos o trabalho e estudo remotos. Assim como trouxe mudanças, também evidenciou o quanto grande parte da população está afastada da compreensão do que são os vírus. Essa dificuldade em entender estes organismos tem sido detectada mesmo entre pessoas com acesso à educação, pois muitas vezes os vírus são erroneamente associados estritamente a doenças, e se tornam um tanto abstratos no imaginário do aluno, por se restringir a figuras e micrografias e pela dificuldade de se elaborar aulas práticas seguras e de baixo custo.

O “*Virus Goes Viral*” foi desenvolvido como uma ferramenta para superar estas limitações que representam uma falha na transferência de conhecimento entre o que se faz na virologia e o que é passado ao aluno. O Kit de lâminas traz um material a

ser observado pelo aluno, trazendo para a discussão questões como a diversidade dos vírus, sua especificidade e o próprio ciclo de replicação. O material de suporte foi elaborado para guiar o instrutor, mas também permitir que os alunos manipulem o kit sozinhos.

Ainda que o *Virus Goes Viral* seja um material de baixo custo, ele ainda está limitado ao uso em salas de aula com microscópios que permitam com objetivas de aumento de pelo menos 100 vezes. Isso é uma limitação em particular para escolas em regiões carentes e com poucos recursos. Além disso, como revelado pela própria pandemia, com a difusão do ensino remoto emergencial, devemos pensar em soluções alternativas para as aulas de virologia, para momentos em que os alunos não possam acessar os laboratórios de aulas práticas.

Como perspectivas, o *Virus Goes Viral* tem potencial de ser ampliado, com lâminas adicionas que permitam a comparação com outros microrganismos, como bactérias, por exemplo. Há também a possibilidade da criação de roteiros de aulas práticas que utilizem outras abordagens de ensino, menos expositivas e mais investigativas, permitindo que os alunos elaborem hipóteses e cheguem a conclusões ao manipular o material. Ficando a cargo do educador escolher uma abordagem para as suas aulas.

7. CONCLUSÕES

De forma geral, este trabalho apresenta protocolos de pesquisa para o uso de análises integrativas de transcriptômica para dois vírus emergentes. Para o ZIKV, esse protocolo foi utilizado em busca de genes associados ao processo de patogênese viral em células neurais humanas, permitindo que se determinasse genes associados a neurogênese e a apoptose diferencialmente expressos nas células infectadas pelo ZIKV. Para o SARS-CoV-2, este protocolo foi revistado e atualizado, assim, foi possível determinar dois alvos a serem inibidos por fármacos conhecidos e que tem potencial de inibir também a replicação do SARS-CoV-2 em células infectadas. Ao passo que os vírus gigantes de amebas forneceram não só pistas sobre o processo de coevolução dos vírus, mas também uma ferramenta educacional inédita com potencial de expandir tópicos educacionais para superar limitações dos estudos dos vírus em sala de aula, permitindo uma abordagem mais pratica e de uso seguro em aulas de virologia, com foco em tornar o ambiente de aprendizagem mais inspirador.

REFERÊNCIAS

AKASHI, Motohiro *et al.* Visualization of giant virus particles and development of “VIRAMOS” for high school and university biology course. *Biochemistry and Molecular Biology Education*, v. 47, n. 4, p. 426–431, 2019.

ALMEIDA, June D; TYRRELL, D A J. The Morphology of Three Previously Uncharacterized Human Respiratory Viruses that Grow in Organ Culture. *J. Gen. Virol*, v. 1, p. 175–178, 1967.

ALQUTAMI, Fatma; SENOK, Abiola; HACHIM, Mahmood. COVID-19 Transcriptomic Atlas: A Comprehensive Analysis of COVID-19 Related Transcriptomics Datasets. *Frontiers in Genetics*, v. 12, p. 2500, 22 dez. 2021.

APUKE, Oberiri Destiny; OMAR, Bahiyah. Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, v. 56, p. 101475, 1 jan. 2021.

BERKHOUT, Ben; COOMBS, Kevin M. Quantitative omics and its application to study virus-host interactions-a new frontier. *Frontiers in Microbiology*, v. 4, n. MAR, p. 31, 2013.

BONILAURI, Paolo; RUGNA, Gianluca. Animal Coronaviruses and SARS-COV-2 in Animals, What Do We Actually Know? *Life*, v. 11, n. 2, p. 1–17, 1 fev. 2021. Disponível em: </pmc/articles/PMC7914637/>. Acesso em: 18 ago. 2022.

BORATTO, Paulo Victor M. *Et al.* A Brief History of Giant Viruses’ Studies in Brazilian Biomes. *Viruses*, v. 14, n. 2, 2022.

BOS, L. Hundred years of Koch’s Postulates and the history of etiology in plant virus research. *Netherlands Journal of Plant Pathology*, v. 87, n. 3, p. 91–110, 1981. Disponível em: <https://link-springer-com.lama.univ-amu.fr/article/10.1007/BF01976645>. Acesso em: 22 jun. 2022.

BRANT, Ayslan Castro *et al.* SARS-cov-2: from its discovery to genome structure, transcription, and replication. *Cell & Bioscience 2021 11:1*, v. 11, n. 1, p. 1–17, 19 jul. 2021. Disponível em: <https://cellandbioscience.biomedcentral.com/articles/10.1186/s13578-021-00643-z>. Acesso em: 18 ago. 2022.

CARLSON, Colin J *et al.* Climate change increases cross-species viral transmission risk. *Nature*, v. 607, n. 7919, p. 555–562, jul. 2022.

CHAN, Jasper F W *et al.* Zika fever and congenital Zika syndrome: An unexpected emerging arboviral disease. *Journal of Infection*, v. 72, n. 5, p. 507–524, 2016. Disponível em: <http://dx.doi.org/10.1016/j.jinf.2016.02.011>.

CHERRY, James *et al. Feigin and Cherry’s Textbook of Pediatric Infectious Diseases, 8th Edition*. [S.l: s.n.], 2019. Disponível em: <https://books.google.fr/books?Id=z-zidwaaqbaj&pg=PT6615&redir_esc=y#v=onepage&q&f=false>. Acesso em: 18 ago. 2022.

CINEGAGLIA, Naiara C. *Et al.* Integrative transcriptome analysis identifies

deregulated microrna-transcription factor networks in lung adenocarcinoma. *Oncotarget*, v. 7, n. 20, p. 28920–28934, 12 abr. 2016. Disponível em: <https://www.oncotarget.com/article/8713/text/>. Acesso em: 26 jan. 2022.

CLOKIE, Martha R.J. *et al.* Phages in nature. *Bacteriophage*, v. 1, n. 1, p. 31, jan. 2011. Disponível em: </pmc/articles/PMC3109452/>. Acesso em: 22 jun. 2022.

COLSON, Philippe *et al.* Mimivirus: Leading the way in the discovery of giant viruses of amoebae. *Nature Reviews Microbiology*, v. 15, n. 4, p. 243–254, 2017.

COOK, Noble David. *Born to die : disease and New World conquest, 1492-1650 / Noble David Cook*. Cambridge ; New York: Cambridge University Press, 1998. Disponível em: <http://www.loc.gov/catdir/toc/cam026/97025064.html>. (New approaches to the Americas.).

DAAMEN, Andrea R. *Et al.* Comprehensive transcriptomic analysis of COVID-19 blood, lung, and airway. *Scientific Reports 2021 11:1*, v. 11, n. 1, p. 1–19, 29 mar. 2021. Disponível em: <https://www.nature.com/articles/s41598-021-86002-x>. Acesso em: 1 fev. 2023.

ENQUIST, L. W. Virology in the 21st Century. *Journal of Virology*, v. 83, n. 11, p. 5296–5308, jun. 2009.

FAISST, Steffen. PROPAGATION OF VIRUSES | Animal. *Encyclopedia of Virology*. Cambridge, Massachusetts: Academic Press, 1999. P. 1408–1413. Disponível em: </pmc/articles/PMC7173615/>. Acesso em: 22 jun. 2022.

FALLIS, A.G. *A brief guide to emerging infectious diseases and zoonoses*. [S.l: s.n.], 2013. V. 53. Disponível em: <https://apps.who.int/iris/handle/10665/204722>. Acesso em: 31 jan. 2023.

FAUCI, Anthony S. Emerging and reemerging infectious diseases: the perpetual challengeno Title. *Acad Med.*, v. 80, n. 12, p. 1079–85, 2005.

GORBALENYA, Alexander E. *Et al.* The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-ncov and naming it SARS-cov-2. *Nature Microbiology*, v. 5, n. 4, p. 536–544, 2020.

GREENE, Ciara M.; MURPHY, Gillian. Quantifying the Effects of Fake News on Behavior: Evidence From a Study of COVID-19 Misinformation. *Journal of Experimental Psychology: Applied*, v. 27, n. 4, p. 773–784, 2021.

HARDENBROOK, Nathan J.; ZHANG, Peijun. A structural view of the SARS-cov-2 virus and its assembly. *Current Opinion in Virology*, v. 52, p. 123–134, 1 fev. 2022.

HAYS, Jo Nelson. *Epidemics and pandemics: their impacts on human history*. Santa Barbara, CA: ABC-CLIO, Inc, 2005. Disponível em: <https://books.google.is/books?id=gye8qt-ks1kc&lpg=PP1&pg=PR4#v=onepage&q&f=false>.

HEGENBERG, Leonidas. *Doença: um estudo filosófico*. [S.l.]: Editora FIOCRUZ, 1998.

HOSHIDA, Yujin *et al.* Integrative Transcriptome Analysis Reveals Common Molecular Subclasses of Human Hepatocellular Carcinoma. *Cancer Research*, v. 69, n. 18, p. 7385–7392, 15 set. 2009. Disponível em:

<https://cancerres.aacrjournals.org/content/69/18/7385>. Acesso em: 26 jan. 2022.

HUI, David S. *Et al.* The continuing 2019-ncov epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*, v. 91, p. 264–266, 1 fev. 2020. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31953166/>. Acesso em: 24 jan. 2022.

ICTV. *Coronaviridae*. Disponível em: <https://ictv.global/report_9th/rnapos/Nidovirales/Coronaviridae>. Acesso em: 18 ago. 2022.

JAMALI MOGHADAM, Saeed Reza *et al.* Zika virus: A review of literature. *Asian Pacific Journal of Tropical Biomedicine*, v. 6, n. 12, p. 989–994, 2016. Disponível em: <http://dx.doi.org/10.1016/j.apjtb.2016.09.007>.

JEAN BELTRAN, Pierre M *et al.* Proteomics and integrative omic approaches for understanding host-pathogen interactions and infectious diseases. *Molecular systems biology*, v. 13, n. 3, p. 922, mar. 2017. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28348067/>. Acesso em: 26 jan. 2022.

KARAMANOU, Marianna *et al.* From miasmas to germs: a historical approach to theories of infectious disease transmission Dai miasmi ai germi: un approccio storiografico alle teorie della trasmissione delle malattie infettive. *Le infezioni in Medicina*, v. 1, n. 1, p. 52–56, 2012. Disponível em: <https://www.researchgate.net/profile/Antonis_Kousoulis/publication/223957556_From_miasmas_to_germs_A_historical_approach_to_theories_of_infectious_disease_transmission/links/0fcfd505a01e81aa71000000.pdf>.

KEESING, Felicia *et al.* Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature 2010 468:7324*, v. 468, n. 7324, p. 647–652, 1 dez. 2010. Disponível em: <https://www.nature.com/articles/nature09575>. Acesso em: 31 jan. 2023.

KEILIN, D. The Leeuwenhoek Lecture - The problem of anabiosis or latent life: history and current concept. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, v. 150, n. 939, p. 149–191, 1959.

KUSHWAHA, U.K.S. *et al.* Role of Bioinformatics in Crop Improvement. *Global Journal of Science Frontier Research: Agriculture and Veterinary*, v. 17, n. 1, p. 12–23, 2017.

LA SCOLA, Bernard *et al.* A giant virus in amoebae. *Science*, v. 299, n. 5615, p. 2033, 2003.

LAI, Chih Cheng *et al.* Severe acute respiratory syndrome coronavirus 2 (SARS-cov-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, v. 55, n. 3, p. 105924, 1 mar. 2020. Disponível em: </pmc/articles/PMC7127800/>. Acesso em: 18 ago. 2022.

LECOQ, Hervé. Découverte du premier virus, le virus de la mosaïque du tabac : 1892 ou 1898 ? *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie*, v. 324, n. 10, p. 929–933, 2001.

LELAND, Diane S.; GINOCCHIO, Christine C. Role of Cell Culture for Virus Detection in the Age of Technology. *Clinical Microbiology Reviews*, v. 20, n. 1, p. 49, jan. 2007. Disponível em: </pmc/articles/PMC1797634/>. Acesso em: 22 jun. 2022.

LIAO, Qibin *et al.* Identification of mirna-mrna crosstalk in CD4+ T cells during HIV-1 infection by integrating transcriptome analyses. *Journal of Translational Medicine*, v. 15, n. 1, p. 1–11, 21 fev. 2017. Disponível em: <https://translational-medicine.biomedcentral.com/articles/10.1186/s12967-017-1130-y>. Acesso em: 26 jan. 2022.

LOWE, Rohan *et al.* Transcriptomics technologies. *Plos Computational Biology*, v. 13, n. 5, 1 maio 2017. Disponível em: </pmc/articles/PMC5436640/>. Acesso em: 19 jan. 2023.

LU, Roujian *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet (London, England)*, v. 395, n. 10224, p. 565–574, 22 fev. 2020. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/32007145/>. Acesso em: 18 ago. 2022.

LUDWIG, Stephan; ZARBOCK, Alexander. *Coronaviruses and SARS-cov-2: A Brief Overview*. *Anesthesia and Analgesia*. [S.l.]: Wolters Kluwer Health. Disponível em: </pmc/articles/PMC7173023/>. Acesso em: 18 ago. 2022. , 4 abr. 2020

LUNA, Expedito J. A. A emergência das doenças emergentes e as doenças infecciosas emergentes e reemergentes no Brasil. *Revista Brasileira de Epidemiologia*, v. 5, n. 3, p. 229–243, dez. 2002. Disponível em: <http://www.scielo.br/j/rbepid/a/m9mysbmfvb4ztkdj3tbx9sg/?Lang=pt>. Acesso em: 24 jan. 2023.

MAHMOOD, Jemilah; GUINTO, Renzo R. *Impacts of Climate Change on Human Health: Emerging Evidence and Call to Action. The Malaysian journal of medical sciences : MJMS*. Malaysia: [s.n.]. , out. 2022

MASCOLA, John R.; GRAHAM, Barney S.; FAUCI, Anthony S. SARS-cov-2 Viral Variants - Tackling a Moving Target. *JAMA - Journal of the American Medical Association*, v. 325, n. 13, p. 1261–1262, 2021.

MD TAIMUR KHAN, Mohiuddin *et al.* The ``Regulator'' Function of Viruses on Ecosystem Carbon Cycling in the Anthropocene. *Frontiers in Public Health* | *www.frontiersin.org*, v. 10, p. 858615, 2022. Disponível em: <www.frontiersin.org>. Acesso em: 15 fev. 2023.

MOROZOVA, Olena; HIRST, Martin; MARRA, Marco A. Applications of new sequencing technologies for transcriptome analysis. *Annual Review of Genomics and Human Genetics*, v. 10, p. 135–151, 28 ago. 2009. Disponível em: <https://www.annualreviews.org/doi/abs/10.1146/annurev-genom-082908-145957>. Acesso em: 31 jan. 2023.

MOSLEHI, Roxana *et al.* Integrative transcriptome analysis reveals dysregulation of canonical cancer molecular pathways in placenta leading to preeclampsia. *Scientific Reports 2013 3:1*, v. 3, n. 1, p. 1–11, 30 ago. 2013. Disponível em: <https://www-nature-com.lama.univ-amu.fr/articles/srep02407>. Acesso em: 26 jan. 2022.

MURGOLO, Nicholas *et al.* SARS-cov-2 tropism, entry, replication, and propagation:

Considerations for drug discovery and development. *Plos Pathogens*, v. 17, n. 2, p. 1–18, 2021. Disponível em: <http://dx.doi.org/10.1371/journal.ppat.1009225>.

MUSHEGIAN, A. R. Are There 1031 Virus Particles on Earth, or More, or Fewer? *Journal of Bacteriology*, v. 202, n. 9, p. 2192–2197, 9 abr. 2020. Disponível em: </pmc/articles/PMC7148134/>. Acesso em: 31 jan. 2023.

NELSON, N. J. Microarrays have arrived: gene expression tool matures. *Journal of the National Cancer Institute*, v. 93, n. 7, p. 492–493, 4 abr. 2001. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/11287436/>. Acesso em: 31 jan. 2023.

PAEZ-ESPINO, David *et al.* Uncovering Earth's virome. *Nature*, v. 536, n. 7617, p. 425–430, 17 ago. 2016. Disponível em: <https://www.nature.com/articles/nature19094>. Acesso em: 11 fev. 2021.

PETERSEN, Eskild *et al.* Rapid Spread of Zika Virus in The Americas - Implications for Public Health Preparedness for Mass Gatherings at the 2016 Brazil Olympic Games. *International Journal of Infectious Diseases*, v. 44, n. May 2015, p. 11–15, 2016. Disponível em: <http://dx.doi.org/10.1016/j.ijid.2016.02.001>.

ROCHA, Yasmim Mendes *et al.* The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review. *Journal of Public Health (Germany)*, p. 1–10, 9 out. 2021. Disponível em: <https://link.springer.com/article/10.1007/s10389-021-01658-z>. Acesso em: 1 maio 2023.

RODRIGUES, Rodrigo A.L. *et al. An anthropocentric view of the virosphere-host relationship*. *Frontiers in Microbiology*. [S.l: s.n.]. , 2017

ROUX, Simon *et al.* Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature*, v. 537, p. 689–693, 2016.

SCLIAR, Moacyr. História do conceito de saúde. *Physis: Revista de Saúde Coletiva*, v. 17, n. 1, p. 29–41, abr. 2007. Disponível em: <http://www.scielo.br/j/physis/a/wntwlvwqrfbscbzcywv9wgq/?Lang=pt>. Acesso em: 1 maio 2023.

SHEREEN, Muhammad Adnan *et al.* COVID-19 infection: Emergence, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*, v. 24, p. 91–98, 1 jul. 2020.

SHI, Yi; GAO, George F. Structural Biology of the Zika Virus. *Trends in Biochemical Sciences*, v. Xx, p. 1–14, 2017. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0968000417300488>.

SIMON, Charles E. *THE FILTERABLE VIRUSES. The Lancet*. [S.l: s.n.]. Disponível em: <https://journals.physiology.org/doi/10.1152/physrev.1923.3.4.483>. Acesso em: 22 jun. 2022. , 1 out. 1912

SONG, Byung Hak *et al.* Zika virus: History, epidemiology, transmission, and clinical presentation. *Journal of Neuroimmunology*, v. 308, p. 50–64, 15 jul. 2017.

STERN-GINOSSAR, Noam *et al.* Translational control in virus-infected cells. *Cold Spring Harbor Perspectives in Biology*, v. 11, n. 3, p. A033001, 2019. Disponível em: <www.cshperspectives.orghttp://cshperspectives.cshlp.org/Downloadedfrom>.

Acesso em: 31 jan. 2023.

SUN, Guihua *et al.* Comparative transcriptomic analysis of SARS-cov-2 infected cell model systems reveals differential innate immune responses. *Scientific Reports 2021 11:1*, v. 11, n. 1, p. 1–12, 25 ago. 2021. Disponível em: <https://www.nature.com/articles/s41598-021-96462-w>. Acesso em: 1 fev. 2023.

TAYLOR, L. H.; LATHAM, S. M.; WOOLHOUSE, M. E.J. Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 356, n. 1411, p. 983–989, 29 jul. 2001. Disponível em: <https://royalsocietypublishing.org/doi/10.1098/rstb.2001.0888>. Acesso em: 31 jan. 2023.

TSOUCALAS, Gregory *et al.* Epilepsy, Theories and Treatment Inside Corpus Hippocraticum. *Current Pharmaceutical Design*, v. 23, n. 42, p. 6369–6372, 14 fev. 2018.

VAN HELVOORT, Ton. The construction of bacteriophage as bacterial virus: Linking endogenous and exogenous thought styles. *Journal of the History of Biology*, v. 27, n. 1, p. 91–139, 1994. Disponível em: <https://www.jstor.org/stable/4331297>. Acesso em: 22 jun. 2022.

VILLARREAL, Luis P.; WITZANY, Guenther. Viruses are essential agents within the roots and stem of the tree of life. *Journal of Theoretical Biology*, v. 262, n. 4, p. 698–710, 2010. Disponível em: <http://dx.doi.org/10.1016/j.jtbi.2009.10.014>.

WANG, Qihui *et al.* Structural and Functional Basis of SARS-cov-2 Entry by Using Human ACE2. *Cell*, v. 181, n. 4, p. 894- 904.e9, 14 maio 2020. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/32275855/>. Acesso em: 18 ago. 2022.

WANG, Zhong; GERSTEIN, Mark; SNYDER, Michael. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, v. 10, n. 1, p. 57–63, jan. 2009.

WEAVER, Scott C. *Et al.* Zika virus: History, emergence, biology, and prospects for control. *Antiviral Research*, v. 130, p. 69–80, 2016. Disponível em: <http://dx.doi.org/10.1016/j.antiviral.2016.03.010>.

WOOLHOUSE, Mark *et al.* Human viruses: Discovery and emeraence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 367, n. 1604, p. 2864–2871, 2012.

WOOLHOUSE, Mark E.J.; GOWTAGE-SEQUERIA, Sonya. Host range and emerging and reemerging pathogens. *Emerging Infectious Diseases*, v. 11, n. 12, p. 1842–1847, 2005. Disponível em: <https://wwwnc.cdc.gov/eid/article/11/12/05-0997_article>. Acesso em: 31 jan. 2023.

WU, Fan *et al.* A new coronavirus associated with human respiratory disease in China. *Nature 2020 579:7798*, v. 579, n. 7798, p. 265–269, 3 fev. 2020. Disponível em: <https://www-nature-com.lama.univ-amu.fr/articles/s41586-020-2008-3>. Acesso em: 24 jan. 2022.

YANG, Chee Hing *et al.* Influenza A virus upregulates PRPF8 gene expression to increase virus production. *Archives of Virology*, v. 162, n. 5, p. 1223–1235, 1 maio 2017. Disponível em: <https://link.springer.com/article/10.1007/s00705-016-3210-3>. Acesso em: 31 jan. 2023.

YUN, Sang-Im; LEE, Young-Min. Zika virus: An emerging flavivirus. *Journal of Microbiology*, v. 55, n. 3, p. 204–219, 2017. Disponível em: <http://link.springer.com/10.1007/s12275-017-7063-6>.

ZANLUCA, Camila; DOS SANTOS, Claudia Nunes Duarte. Zika virus – an overview. *Microbes and Infection*, v. 18, n. 5, p. 295–301, 2016. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1286457916000496>.

ZHOU, Zhijian; QIU, Ye; GE, Xingyi. The taxonomy, host range and pathogenicity of coronaviruses and other viruses in the Nidovirales order. *Animal Diseases*, v. 1, n. 1, p. 1–28, 2021.

ZHU, Na *et al.* A Novel Coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.*, v. 382, n. 8, p. 727–733, 20 fev. 2020.

OUTRAS ATIVIDADES DESENVOLVIDAS DURANTE O DOUTORADO

Artigos completos publicados em periódicos indexados

GONÇALVES, RICARDO LEMES*, **DE SOUZA, GABRIEL AUGUSTO PIRES***; TERCETI, MATEUS DE SOUZA; DE CASTRO, RENATO FRÓES GOULART; SILVA, BRENO DE MELLO; NOVAES, ROMULO DIAS; MALAQUIAS, LUIZ COSME COTTA; COELHO, LUIZ FELIPE LEOMIL. Integrative transcriptome analysis of SARS-CoV-2 human-infected cells combined with deep learning algorithms identifies two potential cellular targets for the treatment of coronavirus disease. Braz J Microbiol, [s.V], [s.P], 2022.

*Primeira autoria compartilhada entre estes autores

BORATTO, PAULO VICTOR M. ; SERAFIM, MATEUS SÁ M. ; WITT, AMANDA STÉPHANIE A. ; CRISPIM, ANA PAULA C. ; AZEVEDO, BRUNA LUIZA DE ; **SOUZA, GABRIEL AUGUSTO P. DE** ; AQUINO, ISABELLA LUIZA M. DE ; MACHADO, TALITA B. ; QUEIROZ, VICTÓRIA F. ; RODRIGUES, RODRIGO A. L. ; BERGIER, IVAN ; CORTINES, JULIANA REIS ; FARIAS, SAVIO TORRES DE ; SANTOS, RAÍSSA NUNES DOS ; CAMPOS, FABRÍCIO SOUZA ; FRANCO, ANA CLÁUDIA ; ABRAHÃO, JÔNATAS S. . A Brief History of Giant Viruses? Studies in Brazilian Biomes. Viruses-Basel, v. 14, p. 191, 2022.

RODRIGUES, JOÃO FRANCISCO VITÓRIO; **DE SOUZA, GABRIEL AUGUSTO PIRES**; ABRAHÃO, JÔNATAS SANTOS ; AMARAL, RAÍNE PIVA ; DE CASTRO, RENATO FRÓES GOULART ; MALAQUIAS, LUIZ COSME COTTA ; COELHO, LUIZ FELIPE LEOMIL . Integrative transcriptome analysis of human cells treated with silver nanoparticles reveals a distinct cellular response and the importance of inorganic elements detoxification pathways. BIOCHIMICA ET BIOPHYSICA ACTA-GENERAL SUBJECTS, v. 1866, p. 130116, 2022.

SOUZA, GABRIEL AUGUSTO PIRES DE; ROCHA, RAÍSSA PRADO ; GONÇALVES, RICARDO LEMES ; FERREIRA, CYNTIA SILVA ; SILVA, BRENO DE MELLO ; CASTRO, RENATO FRÓES GOULART DE ; RODRIGUES, JOÃO FRANCISCO VITÓRIO ; JÚNIOR, JOÃO CARLOS VILELA VIEIRA ; MALAQUIAS, LUIZ COSME COTTA ; ABRAHÃO, JÔNATAS SANTOS ; COELHO, LUIZ FELIPE LEOMIL . Nanoparticles as Vaccines to Prevent Arbovirus Infection: A Long Road Ahead. PATHOGENS, v. 10, p. 36, 2021.

DE SOUZA, GABRIEL AUGUSTO PIRES; QUEIROZ, VICTÓRIA FULGÊNCIO ; COELHO, LUIZ FELIPE LEOMIL ; ABRAHÃO, JÔNATAS SANTOS . Alohomora! What the entry mechanisms tell us about the evolution and diversification of giant viruses and their hosts. Current Opinion in Virology, v. 47, p. 79-85, 2021.

DE SOUZA, GABRIEL AUGUSTO PIRES; QUEIROZ, VICTÓRIA FULGÊNCIO ; LIMA, MAURÍCIO TEIXEIRA ; DE SOUSA REIS, ERIK VINICIUS ; COELHO, LUIZ FELIPE LEOMIL ; ABRAHÃO, JÔNATAS SANTOS . Virus goes viral: an educational kit for virology classes. Virology Journal, v. 17, p. 13, 2020.

PIRES DE SOUZA, GABRIEL AUGUSTO; SALVADOR, EZEQUIEL APARECIDO ; DE OLIVEIRA, FERNANDA ROZA ; COTTA MALAQUIAS, LUIZ COSME ; ABRAHÃO, JONATAS SANTOS ; LEOMIL COELHO, LUIZ FELIPE . An in silico integrative protocol for identifying key genes and pathways useful to understand emerging virus disease pathogenesis. VIRUS RESEARCH, v. 15, p. 197986, 2020.

Capítulos de livros publicados

PIRES DE SOUZA, GABRIEL AUGUSTO; SACCHETTO, L. ; DRUMOND, B. P. ; ABRAHÃO, JONATAS SANTOS ; LEITE, T. C. R. ; SILVA, BRENO DE MELLO ; PEREIRA, A. C. T. C. ; FERREIRA, G. P. ; MALAQUIAS, L. C. C. ; COELHO, L. F. L. . Nanotechnology applied in the control and diagnosis of Zika virus and its vectors.

Zika Virus Impact, Diagnosis, Control, and Models. 1ed.Massachusetts, EUA: Academic Press, 2021, v. 2, p. 417-426.

Pedidos de Patentes depositados no INPI

Nanopartículas adjuvantes para uso em composições de vacinas
(BR 10 2022 002426 0)

Participação em eventos científicos

VI Simpósio de Microbiologia da UFMG – CONECTA SIM:Microbiologia Interligada: “VIRUSES GO VIRAL”: DESENVOLVIMENTO DE UM KIT PARA EDUCAÇÃO EM VIROLOGIA” (Apresentação Oral);

VII Simpósio de Microbiologia da UFMG – CONECTA SIM 2020: MULTIMERIZAÇÃO DE UMA PROTEÍNA QUIMÉRICA E SEU POTENCIAL PARA A INDUÇÃO DE ANTICORPOS CONTRA O ZIKA VÍRUS (Resumo expandido).

Revisor de periódico

Nanomedicine: Nanotechnology, Biology and Medicine (4)

International Journal of Biological Macromolecules (2)