



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS BIOLÓGICAS – CCB
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E BIOCÊNCIAS

Vilmar Benetti Filho

**ESTUDO DA APLICABILIDADE DE SEQUENCIAMENTO *SHOTGUN* POR
NANOPOROS PARA DETECÇÃO VIRAL EM ESGOTO SANITÁRIO DE
CANASVIEIRAS - FLORIANÓPOLIS-SC**

Florianópolis

2023

Vilmar Benetti Filho

**ESTUDO DA APLICABILIDADE DE SEQUENCIAMENTO SHOTGUN POR
NANOPOROS PARA DETECÇÃO VIRAL EM ESGOTO SANITÁRIO DE
CANASVIEIRAS - FLORIANÓPOLIS-SC**

Dissertação submetida ao Programa de Pós
Graduação em Biotecnologia e Biociências da
Universidade Federal de Santa Catarina para a
obtenção do título de Mestre em Biotecnologia e
Biociências

Orientador: Prof. Dr. Glauber Wagner

Coorientadoras: Profa. Dra. Gislaine Fongaro

Profa. Dra. Jaqueline Góes de Jesus

Florianópolis

2023

Benetti Filho, Vilmar

ESTUDO DA APLICABILIDADE DE SEQUENCIAMENTO SHOTGUN POR NANOPOROS PARA DETECÇÃO VIRAL EM ESGOTO SANITÁRIO DE CANASVIEIRAS - FLORIANÓPOLIS-SC / Vilmar Benetti Filho ; orientadora, Dr. Glauber Wagner, coorientadora, Dra. Gislaiane Fongaro, coorientador, Dra. Jaqueline Góes de Jesus, 2023.
90 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Programa de Pós-Graduação em Biotecnologia e Biociências, Florianópolis, 2023.

Inclui referências.

1. Biotecnologia e Biociências. 2. Epidemiologia Baseada em Águas Residuais (WBE). 3. Vírus Entéricos. 4. SARS-CoV-2. 5. MinION. I. Wagner, Dr. Glauber. II. Fongaro, Dra. Gislaiane. III. Góes de Jesus, Dra. Jaqueline IV. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Biotecnologia e Biociências. V. Título.

Vilmar Benetti Filho

**ESTUDO DA APLICABILIDADE DE SEQUENCIAMENTO *SHOTGUN* POR
NANOPOROS PARA DETECÇÃO VIRAL EM ESGOTO SANITÁRIO DE
CANASVIEIRAS - FLORIANÓPOLIS-SC**

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca
examinadora composta pelos seguintes membros:

Prof.^a. Dr^a Doris Sobral Marques Souza
Universidade Federal de Santa Catarina

Prof. Dr. Renato Simões Moreira
Instituto Federal de Santa Catarina

Certificamos que esta é a versão **original e final** do trabalho de conclusão que foi julgado
adequado para obtenção do título de mestre em Biotecnologia e Biociências.

Coordenação do Programa de Pós-Graduação

Prof. Dr. Glauber Wagner
Orientador

Florianópolis, 2023.

Dedico este trabalho a todos aqueles que perderam alguém na pandemia de COVID-19.

AGRADECIMENTOS

Meus agradecimentos a todos e todas que contribuíram de forma direta ou indireta para a realização deste trabalho. Gostaria de agradecer, primeiramente, aos meus pais, Leoni e Vilmar, pela educação e valores que me ensinaram. Vocês são meus heróis. Aparentemente levei a sério quando falavam *seu trabalho é estudar*. Agradeço às minhas irmãs, Luana e Beatris, pelo apoio e incentivo. Ter uma irmã doutora e outra engenheira me inspiraram nesta carreira, apesar destes títulos serem as menores de suas qualidades. Agradeço ao meu avô João, cuja alegria e carinho aquecem meu coração. Me orgulho muito desta família e amo muito vocês.

Agradeço ao amor da minha vida, Beatriz Figueredo, sem você tudo isso seria muito mais difícil e menos prazeroso. Obrigado por compartilhar a vida comigo. Gostaria de agradecer aos Figueredos, Luiz e Marisete, que me acolherem desde o primeiro dia que me conheceram, sempre gentis, alegres e carinhosos. Minha outra família que aprendi a amar.

Agradeço aos meus orientadores, Glauber, Gislaine e Jaqueline.

Glauber, obrigado pela orientação, parceria, por sempre acreditar na gente. Obrigado por me incluir neste mundo da bioinformática e permitir fazer parte desta equipe excepcional. Realizamos muitas coisas juntos e eu não sei nem expressar o quanto eu aprendi desde minha iniciação científica. Você foi e é fundamental na minha formação. Conte sempre comigo.

Professora Gislaine, você aceitou me coorientar em um período muito importante da sua vida. Apesar de fisicamente ausente, as reuniões que tivemos foram muito produtivas e esclarecedoras. Admiro seu poder de direcionamento, objetividade e gestão de equipe, sempre me senti bem com meus amigos LVAnos.

Professora Jaqueline, gostaria de agradecer por ter aceitado me coorientar, mesmo que a distância. Você é uma inspiração para todos e todas pesquisadoras no Brasil, foi uma honra escutar suas palestras, realizar reuniões e trabalhar junto. Agradeço também a todos do projeto Amazônia +10, que conheci nestas reuniões. Em especial a Valquíria, com quem obtive várias discussões promissoras.

Agradeço aos meus amigos imbatíveis e tristes pelas tardes mais leves, pelos rolês duvidosos, memes internos e BDCs. Obrigado Karin, Taty, Day, Renato, Guilherme, Eric, Carol, Beatriz, Fernando, Felipe, Jaime, Bruna e Maiara.

Agradeço aos meus amigos e colegas da Proto a qual compartilhei angústias e desesperos, mas também brindes e comemorações. Obrigado Ana *craudinha*, Ana Paula, Amabs, Laryssa, Luiza Smit, Luiza Conti e Helena.

Agradeço aos *losers* dos meus amigos, apesar da gente não se encontrar muito depois da graduação, ainda gosto de vocês. Obrigado por serem quem vocês são. Garbo, Gu, Lígia, Ali, Rafa, Rossito e Laura sei que sempre posso contar com vocês.

Agradeço aos demais amigos que a bio me deu, especialmente Gabs, Clóvis, Thiago, Cabral, Klein, Flash, Max.

Um abraço especial para o Guile, Luiz e Eduardo, amigos que levarei para a vida toda.

Agradeço aos meus amigos e colegas LVAnos, sempre dispostos a ajudar e cooperar. Sempre me senti à vontade com vocês, agradeço a receptividade e alegria com que trabalhamos juntos. Obrigado Beatriz, Rafael, Mariana, Doris, Giulia, Helena.

Agradeço a todos que contribuíram para realização do projeto, em especial aos meus amigos Rafael, Dayane e Fernando, que me auxiliaram diretamente na empreitada do *wet lab*. Aprecio muito a ajuda e amizade de vocês, com certeza, para um bioinformata como eu, acostumado ao ar condicionado, as idas ao LVA foram mais tranquilas na companhia de vocês.

Agradeço aos professores do Laboratório de Protozoologia, Professor Mário, Professor Edmundo, Professora Patrícia Stoco, Professora Patrícia Quaresma, que contribuíram com discussões que auxiliaram na minha formação científica. Agradeço aos colaboradores, pelas produções realizadas neste período. Sempre é bom aprender algo novo - de fagos às galinhas.

Agradeço aos membros do projeto VigEAI, as quais gentilmente foram cedidas as amostras para execução deste projeto. Agradeço às instituições participantes do projeto: Serviço Intermunicipal de Água e Esgoto (SIMAE); Companhia Catarinense de Águas e Saneamento (CASAN); Consórcio Intermunicipal de Saneamento Ambiental - CISAM Meio-Oeste; Agência Reguladora; Intermunicipal de Saneamento (ARIS); Diretoria de Vigilância Epidemiológica (DVE - Itajaí); Serviço Municipal de Água, Saneamento (SEMASA -Itajaí); Universidade do Oeste de Santa Catarina (UNOESC); Universidade do Contestado (UnC); Vigilância Sanitária Municipal – Município de Capinzal (VISA); Universidade Federal de Santa Catarina – Campus Joinville e Curitiba; e as prefeituras de Joinville, Concórdia, Itajaí e Florianópolis, as quais auxiliaram com coleta e transporte das amostras cedidas para realização deste projeto. Agradeço a equipe analítica do projeto: Dayane Azevedo Padilha; Dra. Doris Sobral Marques Souza, Henrique Borges da Silva Grisard; Fernando Hartmann Barazzetti; Julia Kinetz Wachter; Dr. Marcos André Schörner; Marcus Vinícius Duarte Rodrigues; Rafael Dorighello Cadamuro; Dra. Aline Viancelli; Dra. Eliandra Mirlei Rossi; Dr. Glauber Wagner; Dra. Maria Luiza Bazzo; Dra. Simone Malutta; Dr. William Michelin; e a professora Dra. Gislaïne Fongaro pela coordenação do projeto.

Queria agradecer a todos que vieram antes. Ciência é colaborativa. Uma pergunta não surge do nada, uma hipótese não pode ser testada sem métodos estabelecidos e não adianta olhar para frente sem entender onde você está. Obrigado a todos os professores e professoras, doutores e doutoras, mestras, mestres e iniciações científicas que vieram antes de mim e edificaram o conhecimento que temos hoje. Em especial quem fez isso de forma aberta. Aos demais, sinto muito, usei o *sci-hub*. Fica aqui o agradecimento especial a Alexandra Elbakyan que criou esta formidável ferramenta.

Agradeço às instituições financiadoras. Primeiramente à CAPES, instituição que concedeu minha bolsa e tornou este trabalho possível. Agradeço também ao grupo Mulheres do Brasil, à FAPESC, à Magalu e à iniciativa *Genomic Day*, que financiaram insumos para realização deste trabalho.

Agradeço à UFSC pelo privilégio de ser estudante em um programa de pós-graduação de excelência nesta cidade maravilhosa.

Voltaremos mais fortes.

"Seek Strength. The rest will follow." – King Vendrick

RESUMO

A vigilância genômica é uma importante ferramenta para detecção e monitoramento da dispersão de patógenos, e pode ser utilizada para o acompanhamento de mutações, emergência de variantes e identificação de mecanismos de escape imunológico. No contexto da saúde única, em que a saúde humana, animal e do ecossistema estão interligadas e interdependentes, a vigilância genômica possui aplicabilidade à detecção de disseminação de agentes microbianos resistentes a tratamento. A detecção de vírus em esgoto por qPCR já foi empregada para monitoramento de surtos de poliomielite, Ebola e, recentemente, SARS-CoV-2. O protocolo de concentração viral baseado no uso de Polietileno Glicol (PEG) como agente concentrador é amplamente utilizado para ensaios que visam concentrar e na sequência detectar tais partículas usando métodos moleculares e de cultura celular. Florianópolis, por ser uma cidade turística, apresenta um aumento populacional no verão, fator que aumenta a descarga de esgoto e possibilidade de transmissão de vírus de rota orofecal, fato que torna o esgoto sanitário uma ferramenta potencial para monitorar circulação e introdução viral no ambiente. Desta forma, este projeto buscou analisar o viroma total do esgoto de Canasvieiras, de dez meses do ano de 2021, como um ensaio acerca da aplicação da metodologia de concentração por PEG, que já é empregada para detecção molecular, para sequenciamento por nanoporos pela metodologia *shotgun*. Dois experimentos de sequenciamento foram realizados utilizando vírus controles, para: (i) detecção de concentração mínima necessária para recuperação de sequências virais; (ii) avaliação de protocolo de PEG seguido de pós protocolos de clarificação e remoção de DNA bacteriano livre. Levando em conta estes resultados, o protocolo PEG + clorofórmio e butanol (CB) foi empregado para concentração viral de dez amostras de dez meses da estação de tratamento de esgoto de Canasvieiras do ano 2021. Estas amostras foram sequenciadas, sem adição de quaisquer controles virais, pela metodologia *shotgun* em equipamento MinION Mk1C. Resultados prévios de detecção dos vírus da Hepatite A, Mastadenovírus humano (Adenovírus) e SARS-CoV-2 por qPCR foram utilizados como controles positivos das amostras, a fim de verificar a recuperação dos mesmos no sequenciamento. Em caráter comparativo, seis amostras de esgoto de outras cidades do estado de Santa Catarina foram utilizadas para obtenção do genoma de SARS-CoV-2, por meio de *amplicons* do genoma viral, sequenciados pela tecnologia Illumina. Diferentes concentrações de vírus controles não refletiram em diferenças na recuperação de fragmentos de leitura virais. A aplicação do protocolo de clarificação associado ao método de concentração por PEG recuperou, proporcionalmente, maior número de fragmentos de leitura viral, com cerca de 26% do total sequenciado. O emprego desta metodologia nas amostras de Canasvieiras/Florianópolis-SC, recuperou apenas os fagos *Shigella phage* SfMu e *Escherichia phage* Mu. Não foram recuperados vírus previamente identificados por metodologia molecular nas mesmas amostras. O protocolo de concentração PEG mesmo que associado ao CB, sem otimização experimental e variação de componentes (pH, temperatura, tempo de centrifugação) não apresentou eficiência na recuperação de vírus previamente detectados por RT-qPCR e qPCR nas amostras de esgoto de Canasvieiras. As sequências de SARS-CoV-2 obtidas a partir de amostras de esgoto não reconstruíram completamente o genoma do vírus, indicando desafios como inespecificidade de primers e degradação de material genético. Há, portanto, a perspectiva para desenvolvimento da otimização do processo baseado em PEG, teste de novas metodologia e estudos de iniciadores para PCR em multiplex visando o sequenciamento dos de genomas virais alvo a serem aplicados na vigilância genômica de vírus entéricos ou de rota ambiental.

Palavras-chave: Epidemiologia Baseada em Águas Residuais (WBE), polietileno glicol, Vírus Entéricos, SARS-CoV-2, MinION Tecnologia Oxford Nanopore (ONT)

ABSTRACT

Genomic surveillance is an important tool for detecting and monitoring the spread of pathogens and can be used to monitor mutations, and the emergence of variants and identify immune escape mechanisms. In the context of one health, in which human, animal and ecosystem health are interconnected and interdependent, genomic surveillance has applicability to detect the spread of treatment-resistant microbial agents. Virus detection in sewage by qPCR has already been used to monitor outbreaks of polio, Ebola and, recently, SARS-CoV-2. The viral concentration protocol based on the use of Polyethylene Glycol (PEG) as a concentrating agent is widely used for assays that aim to concentrate and then detect such particles using molecular and cell culture methods. Florianópolis, as a tourist city, experiences a population increase in the summer, a factor that increases sewage discharge and the possibility of transmission of viruses via the orofecal route, a fact that makes sanitary sewage a potential tool for monitoring circulation and viral introduction into the environment. In this way, this project sought to analyze the total virome of sewage from Canasvieiras, from ten months of the year 2021, as a test on the application of the PEG concentration methodology, which is already used for molecular detection, for nanopore sequencing using the methodology shotgun. Two sequencing experiments were carried out using control viruses, to: (i) detect the minimum concentration necessary to recover viral sequences; and (ii) evaluate of PEG protocol followed by post-clarification protocols and removal of free bacterial DNA. Taking these results into account, the PEG + chloroform and butanol (CB) protocol was used for the viral concentration of ten samples from ten months of the Canasvieiras sewage treatment plant in the year 2021. These samples were sequenced, without adding any viral controls, using the shotgun methodology on MinION Mk1C equipment. Previous results of the detection of Hepatitis A viruses, human Mastadenovirus (Adenovirus) and SARS-CoV-2 by qPCR were used as positive controls for the samples, in order to verify their recovery in sequencing. On a comparative basis, six sewage samples from other cities in the state of Santa Catarina were used to obtain the SARS-CoV-2 genome, through viral genome amplicons, sequenced using Illumina technology. Different concentrations of control viruses did not reflect differences in the recovery of viral reading fragments. The application of the clarification protocol associated with the PEG concentration method recovered, proportionally, a greater number of viral reading fragments, with around 26% of the total sequenced. The use of this methodology in samples from Canasvieiras/Florianópolis-SC recovered only the Shigella phage SfMu and Escherichia phage Mu phages. No viruses previously identified by molecular methodology were recovered from the same samples. The PEG concentration protocol, even when associated with CB, without experimental optimization and variation of components (pH, temperature, centrifugation time) was not efficient in recovering viruses previously detected by RT-qPCR and qPCR in sewage samples from Canasvieiras. SARS-CoV-2 sequences obtained from sewage samples did not completely reconstruct the virus genome, indicating challenges such as primer unspecificity and degradation of genetic material. There is, therefore, the prospect of developing optimization of the PEG-based process, testing new methodology and studying primers for multiplex PCR aimed at sequencing target viral genomes to be applied in the genomic surveillance of enteric or environmental viruses.

Keywords: Wastewater-Based Epidemiology (WBE), Polyethylene Glycol (PEG), Enteric Viruses, SARS-CoV-2, MinION - Oxford Nanopore (ONT)

LISTA DE FIGURAS

Figura 1. Diagrama de um Sistema de vigilância epidemiológica	17
Figura 2. Diferenças entre as abordagens para sequenciamento de vírus.	23
Figura 3. Classificação viral de acordo com Baltimore (1970).....	25
Figura 4. Localização da estação de tratamento de esgoto de Canasvieiras, Florianópolis. ...	33
Figura 5. Esquema geral dos processos para sequenciamento de viroma direcionado e viroma total	35
Figura 6. Desenho experimental utilizado para obtenção do protocolo de sequenciamento shotgun de viroma total em plataforma de sequenciamento nanoporos	37
Figura 7. Comparativo de recuperação de controle viral nos experimentos de definição do protocolo de sequenciamento shotgun por nanoporos.....	47
Figura 8. Cobertura média dos genomas obtidos a partir do viroma direcionado para SARS-CoV-2.	60

LISTA DE TABELAS

Tabela 1. Comparativo de custo benefício das metodologias de sequenciamento	20
Tabela 2. Classificação de Baltimore	24
Tabela 3. Amostras utilizadas para sequenciamento do viroma total.....	32
Tabela 4. Amostras utilizadas para sequenciamento por amplicons de SARS-CoV-2	33
Tabela 5. Quantificação de carga viral em cópias gênicas/L-1 a partir de amostras de esgoto de Canasvieiras, Florianópolis, no ano de 2021.	42
Tabela 6. Número de fragmentos de leitura mapeados em cada genoma viral - dados do primeiro experimento.	48
Tabela 7. Número de fragmentos de leitura virais classificados - segundo experimento.....	51
Tabela 8. Número de fragmentos de leitura mapeados em cada genoma viral - dados das amostras de Canasvieiras no ano de 2021.	54
Tabela 9. Resultados de viroma direcionado para SARS-CoV-2 a partir de amostras do projeto VigEAI de 2021.	58
Tabela 10. Projeção da população flutuante nos bairros de Florianópolis, Santa Catarina.....	83
Tabela 11. Quantificação de DNA e RNA e relações de pureza após extração de ácidos nucleicos.	84
Tabela 12. Quantificação de DNA complementar realizada por meio do equipamento Qubit.	85
Tabela 13. Proporção de sequenciamento e classificação taxonômica nas duplicatas em abordagem com três concentrações de controles virais.....	86
Tabela 14. Proporção de sequenciamento e classificação taxonômica na abordagem de descontaminação associada ao PEG.	86
Tabela 15. Proporções de sequenciamento e classificação taxonômica nas amostras de oriundas de Canasvieiras no período de 2021.	87
Tabela 16. Dados de saída dos sequenciamentos por shotgun em plataforma MinION.	87

LISTA DE ABREVIATURAS E SIGLAS

CASAN	Companhia Catarinense de Águas e Saneamento
CDC	<i>Centers for Disease Control and Prevention</i>
COVID-19	<i>Coronavirus Disease 2019</i>
CT	<i>Cycle Threshold</i>
DNA	Ácido desoxirribonucleico
dsDNA	<i>Double-stranded DNA</i>
dsRNA	<i>Double-stranded RNA</i>
dsRNA-RT	<i>Double-stranded RNA reverse-transcribing</i>
EDTA	Ácido etilenodiaminotetracético
ESBL	Beta-Lactamase de Espectro Estendido
ETE	Estação de Tratamento de Esgoto
FDA	<i>Federal Drug Administration</i>
GC	<i>Genomic copies</i>
gDNA	<i>Genomic DNA</i>
HAdV	<i>Human Adenovirus</i>
HCl	Ácido Clorídrico
HIV	<i>Human Immunodeficiency Virus</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
ICTV	<i>International Committee on Taxonomy of Viruses</i>
ITS	<i>Internal Transcribed Spacer</i>
JCR	<i>Jelly Roll capsid protein</i>
LVA	Laboratório de Virologia Aplicada
M	Molar
MAG	<i>Metagenome-Assembled Genome</i>
Mb	Megabase
MERS-CoV	<i>Middle East Respiratory Syndrome</i>
MHV-3	<i>Murine Hepatitis Virus 3</i>
mL	Mililitro
MNV	<i>Murine Norovirus</i>
NCBI	<i>National Center for Biotechnology Information</i>
ONT	<i>Oxford Nanopore Technologies</i>
PBS	<i>Phosphate-buffered saline</i>
PCR	<i>Polymerase Chain Reaction</i>

PEG	Polietileno Glicol
PFU	Unidade formadora de placa
RefSeq	<i>Reference Sequences</i>
RNA	Ácido ribonucleico
RT-qPCR	<i>Real Time quantitative PCR</i>
RVA	Rotavírus A
SARS-CoV-2	<i>Severe Acute Respiratory Syndrome Coronavirus 2</i>
ssDNA	<i>Single-stranded DNA</i>
ssRNA (+)	<i>Positive-sense single-stranded RNA</i>
ssRNA (-)	<i>Negative-sense single-stranded RNA</i>
ssRNA-RT	<i>Single-stranded RNA reverse-transcribing</i>
SUS	Sistema Único de Saúde
UFSC	Universidade Federal de Santa Catarina
VigEAI	Vigilância Epidemiológica Ambiental Integrativa
vMAG	<i>Viral Metagenome-Assembled Genome</i>
WIMP	<i>What Is in My Pot</i>
°C	Graus Celsius
µL	Microlitro

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Vigilância epidemiológica.....	14
1.1.1	Conceito e contexto histórico	14
1.1.2	Sistema de vigilância Epidemiológica	16
1.2	Vigilância Genômica.....	18
1.2.1	Sequenciamento genético	19
1.3	Metagenômica e Viroma	20
1.3.1	Características e classificação dos vírus.....	23
1.3.2	Vírus entéricos e respiratórios.....	27
1.3.3	Vigilância viral em esgoto.....	27
2	OBJETIVOS	31
2.1	Objetivo Geral.....	31
2.2	Objetivos Específicos.....	31
3	MATERIAL E MÉTODOS	32
3.1	Delineamento experimental.....	32
3.2	Processamento das amostras de esgoto	38
3.3	Ensaio moleculares	38
3.3.1	Extração de ácidos nucleicos e síntese de DNA complementar (cDNA).....	38
3.3.2	Quantificação dos controles virais	39
3.3.3	Quantificação de vírus selvagens em esgoto de Canasvieiras por qPCR e RT-qPCR	39
3.4	Sequenciamento genético.....	40
3.4.1	Análises em Bioinformática	40
4	RESULTADOS E DISCUSSÃO	42
4.1	Monitoramento da contaminação viral presente no esgoto coletado em Canasvieiras	42
4.2	Recuperação viral a partir do sequenciamento <i>shotgun</i> por nanoporos	46
4.3	Sequenciamento das amostras de esgoto de Canasvieiras do ano de 2021	52
4.4	Sequenciamento direcionado para SARS-CoV-2.....	57
5	CONCLUSÕES.....	61
6	REFERÊNCIAS.....	62
7	APÊNDICE A – MATERIAIS COMPLEMENTARES.....	73
7.1	Publicações E Participações Em Eventos.....	73
7.1.1	Artigos Completos Publicados Em Periódicos.....	73
7.1.2	Resumos publicados em anais de eventos	80
7.1.3	Participação Em Eventos.....	82
7.2	Curso De Curta Duração Ministrado.....	82
8	APÊNDICE B – MATERIAIS COMPLEMENTARES	83

9	ANEXOS.....	88
	ANEXO 1 - Protocolo de Sequenciamento Nanopore <i>Rapid sequencing gDNA</i> - barcoding (SQK-RBK004)	88

1 INTRODUÇÃO

1.1 Vigilância epidemiológica

1.1.1 Conceito e contexto histórico

O conceito de vigilância epidemiológica remete ao monitoramento contínuo da prevalência de patógenos em uma população, com o objetivo de conhecer e controlar a disseminação de doenças, bem como identificar a emergência de novas infecções (Frerichs, 1991). A capacidade de disseminação de uma doença infecciosa na população distingue os conceitos de surto, epidemia e pandemia. Surto é definido como dois ou mais casos relacionados a um mesmo agente infeccioso, sugerindo uma fonte comum de infecção em um período (Straif-Bourgeois, Ratard e Kretzschmar, 2014). A epidemia é definida por um número de casos diferente do esperado, de acordo com uma tendência histórica, em um local e em um determinado período de tempo. Por vezes, os limites entre os conceitos de surto e epidemia podem se sobrepor (Straif-Bourgeois, Ratard e Kretzschmar, 2014). No entanto, em termos de tempo, os surtos frequentemente encerram os ciclos de infecção em um curto período (Straif-Bourgeois, Ratard e Kretzschmar, 2014). Pandemia é o termo utilizado para designar o extremo da disseminação de uma doença, que ultrapassa fronteiras internacionais e sinaliza para falta de controle da emergência de novos casos, bem como a urgência de ações de contenção de dispersão (Straif-Bourgeois, Ratard e Kretzschmar, 2014).

Como estes conceitos se relacionam com uma taxa de casos ou infecções esperadas de acordo com um padrão histórico, a partir deste mesmo padrão surgem os conceitos de áreas endêmicas e endemia. Podemos definir uma área endêmica como uma região em que, historicamente, há relatos da ocorrência de uma certa doença. Uma endemia, portanto, reflete um período de ocorrência de casos de uma moléstia, em área endêmica, mas acima do esperado – de acordo com um histórico de observações (Straif-Bourgeois, Ratard e Kretzschmar, 2014).

A origem dos estudos epidemiológicos foi estabelecida com o trabalho de John Graunt, no contexto da peste bubônica de 1662 (Graunt, 1662, apud Rothman, 1981). A partir de certidões de nascimento e atestados de óbito dos cidadãos de Londres em um período de 50 anos, sazonalidades de casos de peste bubônica foram evidenciados. Desta maneira, Graunt relatou períodos de susceptibilidade à doença relacionado a alterações de condições ambientais e, frente a estas evidências, descartou crenças populares que consideravam, dentre outras coisas, eventos astrológicos como agentes causadores da peste. O trabalho de Grant é considerado o passo inicial para a epidemiologia como conhecemos hoje, justamente por organizar dados

populacionais de uma doença, extrair padrões históricos e identificar fatores causais da emergência de novos casos (Almeida Filho, 1986; Morabia, 2013). A catalogação dos dados é o primeiro passo para que um estudo epidemiológico seja realizado, sendo de grande importância a disponibilidade dos dados para os estudos de Grant, assim como nas observações de Jhon Snow, considerado o pai da epidemiologia moderna, que em 1854 durante um surto de cólera (Cameron e Jones, 1983), demonstrou que a origem da água estava diretamente relacionada com maiores casos de cólera na população, estabelecendo um canal entre o estudo epidemiológico e a tomada de decisão dos agentes públicos (Rothman, 1981; Tulchinsky, 2018). Outro cientista importante para o estudo epidemiológico foi William Farr, que, também no Reino Unido, no século XIX, estabeleceu uma rotina de monitoramento de casos e óbitos em períodos de surtos, possibilitando a identificação de novas epidemias (Barata, 2022; Eyles, 1973).

No Brasil, um dos primeiros relatos históricos de surtos relevantes para o desenvolvimento da epidemiologia no país foi o da febre amarela em 1685 em Recife, Pernambuco (Costa *et al.*, 2011; Franco, 1969). O surto foi controlado por medidas que para a época pareciam controversas e incluíam restrições de circulação de doentes, limpeza do ar e de áreas públicas. Apesar de desconhecidos, o agente etiológico, o vetor e o mecanismo de dispersão da doença, o levantamento de infecções e estabelecimento de medidas de controles surtiram efeito e podem ser descritos como uma prática de controle, ainda que rudimentar (Costa *et al.*, 2011). Em um segundo episódio de surto de casos de febre amarela, em 1850, as medidas de coleta de dados e controle das infecções fundaram bases para o monitoramento epidemiológico em termos institucionais e governamentais. Naquela década, foram criados órgãos de acompanhamento e inspeção das condições dos indivíduos infectados e, concomitantemente, foi fundado o Serviço de Estatística Demográfica do Brasil, atuante na documentação das causas de mortes, visando a estimativa dos danos causados pela febre amarela (Costa *et al.*, 2011).

No século XX, Oswaldo Cruz foi personagem central no combate às doenças infecciosas. Ele atuou no estabelecimento de políticas de controle da febre amarela, peste bubônica e varíola - através das chamadas campanhas sanitárias, no Rio de Janeiro a partir de 1903 (Costa *et al.*, 2011), que obtiveram sucesso, apesar de aplicações controversas. Uma delas foi a vacinação em massa de forma compulsória, que teve profundos impactos sociais, de modo que culminou no fato histórico conhecido como a Revolta da Vacina. Apesar disso, o programa para erradicação da febre amarela obteve sucesso em 1907 (Fundação Oswaldo Cruz, s.d. A

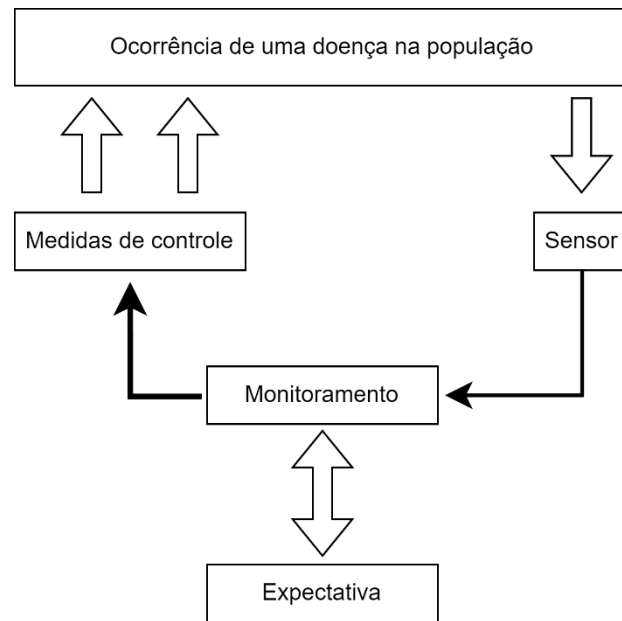
trajetória do médico dedicado à ciência). Novos surtos de febre amarela ocorreram em 1928-29 (Costa *et al.*, 2011), mas sem registros de casos urbanos desde 1942 (Cardoso e Navarro, 2007).

No século XX, destacam-se ainda as criações do Instituto Pasteur em São Paulo, em 1903, no âmbito do estudo e desenvolvimento de soros antirrábicos e para vigilância e ação contra zoonoses (Barata, 2022); assim como, a criação do Instituto Soroterápico Federal no Rio de Janeiro (que posteriormente receberia o nome de Fundação Oswaldo Cruz), atuante na pesquisa e vigilância acompanhamento de doenças infecciosas e negligenciadas (Fundação Oswaldo Cruz, s.d., História). Ainda neste contexto, é importante ressaltar a realização da campanha nacional para controle e erradicação da varíola, iniciada na década de 1960, que fez parte do esforço global, e contou com ações de vacinação em massa da população e vigilância dos casos suspeitos, que culminaram na erradicação nacional da doença em 1973 e de forma global em 1977 (Chagas, 2008; Fernandes, Chagas e Souza, 2011).

1.1.2 Sistema de vigilância Epidemiológica

Um sistema de vigilância epidemiológica (Figura 1) consiste na coleta, registro e análise sistemática de dados bem como a disseminação desta informação de forma eficiente para os órgãos de saúde. A partir da documentação de casos de doenças na população, informações de relevância epidemiológicas são extraídas como: grupos de risco, taxas de infecção, índices de mortalidade, sintomas característicos e sazonalidade de casos, que caracterizam os perfis das doenças (Frerichs, 1991; Reintjes e Krickeberg, 2009). A coleta sistemática destes dados age como um sensor, constantemente mensurando a suscetibilidade da população a determinados patógenos e estabelecendo se estas medidas obedecem a uma tendência histórica ou indicadores, podendo ser desde a nível local ou global. O retorno deste sistema contínuo melhora o sistema de vigilância e auxilia na tomada de decisão dos órgãos de saúde em casos de aumento do número de infecções (Frerichs, 1991). Estas decisões podem ser a nível individual, como o estabelecimento de quarentena para infectados, ou a nível de grupo, como medidas de vacinação em massa e restrições de circulação. Cada doença acompanhada por meio de monitoramento epidemiológico tem suas particularidades, que alteram os métodos de coleta dos dados, análises e exploração da informação (Reintjes e Krickeberg, 2009).

Figura 1. Diagrama de um Sistema de vigilância epidemiológica



Legenda: Fluxo de análise e monitoramento dos dados epidemiológicos em uma população. As setas escuras representam o fluxo, as setas claras representam o efeito de um fator/processo (retângulos) em outro processo/fator.

Fonte: Adaptado de Frerichs, 1991.

A eficiência de um sistema de vigilância epidemiológica depende da quantidade e frequência da coleta de dados e disponibilidade dos mesmos (Reintjes e Krickeberg, 2009). Neste sentido, doenças sazonais ou com baixos números de casos requerem a obtenção de dados relacionados de maneira indireta, que se dá por meio da vigilância vetorial e de animais reservatórios dos agentes etiológicos de doenças zoonóticas, assim como a vigilância de grupos sentinela para inferência da circulação dos patógenos na população (Reintjes e Krickeberg, 2009). O monitoramento dos ambientes e animais que interagem com a população relacionam a vigilância em um contexto de saúde única, ou seja: a saúde da população está diretamente ligada com a saúde ambiental e animal que interagimos. Assim, contaminações da água; emergência de zoonoses; disseminação de resistências antimicrobianas; e ampliação de *habitat* de vetores, são alguns dos fatores que contribuem para ocorrência de doenças infecciosas (Cunningham, Daszak e Wood, 2017; McEwen e Collignon, 2018).

A notificação obrigatória de doenças infecciosas é uma forma de expandir a captação dos dados e evitar subnotificações (Coelho *et al.*, 2016). O Sistema Único de Saúde brasileiro (SUS) possui uma lista de doenças, agravos e eventos de saúde pública cuja notificação, assim que atestado o diagnóstico positivo, é compulsória (Brasil, 2022). Dentre estas, estão incluídas infecções virais causadoras de hepatites, infecções por Zika vírus, Chikungunya vírus, Dengue vírus (DENV), SARS-CoV-2, MERS-CoV, HIV, poliovírus, hantavírus, influenza,

orthopoxvirus (agente causador da varíola), Ebola, Marburg virus, Lassae virus, vírus da febre amarela e vírus da varicela. Através da notificação compulsória os dados epidemiológicos são atualizados constantemente. Para realização da notificação, o diagnóstico correto é requerido e, dependendo do grau de risco biológico relacionado ao vírus, poucos laboratórios podem manipulá-los (Brasil, 2023). Neste sentido, o desenvolvimento científico e tecnológico faz parte do sistema de vigilância epidemiológica, uma vez que o melhor entendimento acerca da biologia dos vírus otimiza e propõe novas metodologias de diagnóstico, que inclusive podem facilitar a sua utilização em locais com pouca infraestrutura e recursos.

1.2 Vigilância Genômica

Em uma evolução natural das metodologias, identificar a presença ou ausência de um agente etiológico circulante não é mais suficiente para vigilância e controle da disseminação das doenças. Desta forma, a vigilância genômica emerge com o objetivo de identificar variações genéticas relevantes para dispersão e severidade com que o agente infeccioso acomete a população (CDC, 2022; Hill *et al.*, 2021). As variantes podem ser caracterizadas por um conjunto de mutações chave no genoma viral, observadas a nível populacional (CDC, 2022). Um genoma é definido como a sequência de nucleotídeos que fazem parte do material genético de qualquer organismo ou vírus. Na infecção viral, durante os ciclos replicativos, erros inerentes de replicação do material genético podem se acumular no genoma dos vírus, gerando pequenas alterações na expressão de proteínas (Knipe e Howley, 2013). Alterações que conferem alguma vantagem na interação com o hospedeiro, ou na evasão do sistema imune, são selecionadas e passam a prevalecer em uma população viral (Knipe e Howley, 2013). Acompanhar o surgimento de novas mutações e o estabelecimento de novas variantes é importante para identificar escapes vacinais (perda de eficácia da vacina frente a uma nova variante); investigar aumento de casos e mortes; assim como rastrear a origem de casos alóctones (Li *et al.*, 2021; Peteranderl, Herold e Schmoldt, 2016).

Vírus com material genético composto por RNA estão mais susceptíveis a mutações em seu genoma em comparação às partículas virais com genoma composto por DNA. Grande parte das mutações ocorre durante a replicação do material genético, e é estimado que vírus de RNA possuam uma taxa de mutação, ou seja a frequência com que um nucleotídeo será replicado de forma errônea, de 10^{-6} a 10^{-4} , enquanto que a taxa para vírus de DNA é de cerca de 10^{-8} a 10^{-6} (Sanjuán; Domingo-Calap, 2016). Fatores como o curto período de gerações, que pode ser resumido a um único ciclo de replicação em uma célula, contribuem para que fatores de seleção possuam uma pressão significativamente maior em vírus do que em organismos celulares. De

modo que, mesmo infecções virais que possuem tratamento vacinal precisam ser atualizadas constantemente, bem como o monitoramento das variantes virais circulantes (Sanjuán; Domingo-Calap, 2016).

A vigilância genômica é empregada para o acompanhamento de cepas de Influenza A desde 1997 (Lam e Pybus, 2018); foi utilizada no programa de erradicação da poliomielite (Dowdle *et al.*, 1999); vigilância de ebola (Gire *et al.*, 2014) e Zika (Faria *et al.*, 2017); solucionar surtos de infecções hospitalares (Price *et al.*, 2014); identificação de origens do surto de MERS-CoV (Sabir *et al.*, 2016); e recentemente para acompanhamento de mudanças de variantes do SARS-CoV-2 na pandemia de COVID-19 (do inglês, *coronavirus disease 2019*) (Li *et al.*, 2021).

A obtenção do genoma do patógeno e identificação de variante é realizado a partir do sequenciamento genético da amostra. (Li *et al.*, 2021), e análise de bioinformática que identificam variantes na(s) sequência(s) genômicas do agente etiológico, relevantes para sua transmissibilidade ou infectividade.

1.2.1 Sequenciamento genético

As técnicas de sequenciamento de DNA, iniciadas com o método de primeira geração de Sanger, na década de 70 (Sanger, Nicklen e Coulson, 1977), e aprimoradas até os dias atuais, são um marco para as ciências ômicas. As metodologias de sequenciamento de segunda geração, cujos expoentes são as plataformas Illumina (MiSeq, HiSeq e NovaSeq), permitiram que grande volume de dados genômicos fossem gerados com menor custo, comparadas às metodologias baseadas em Sanger (Conrads e Abdelbary, 2019; Liu *et al.*, 2012). A terceira geração de sequenciadores (TGS, *Third Generation Sequencing*) trouxe como possibilidade o sequenciamento sem a necessidade de amplificar o material genético de uma amostra. Isto possibilitou que este processo fosse realizado em campo, através do sequenciador portátil - tecnologia estabelecida pela *Oxford Nanopore* através da plataforma MinION (Conrads e Abdelbary, 2019) e PacBio (Xiao; Zhou, 2020).

Um dos principais diferenciais entre as metodologias de sequenciamento é o tamanho dos fragmentos sequenciados, de forma que este caracteriza um divisor entre tecnologias que geram *short reads* (pequenos fragmentos de leitura) e *long reads* (longos fragmentos de leitura). As metodologias de sequenciamento baseadas em *short reads*, como as plataformas da Illumina, possuem a característica de gerar enorme volume de dados de fragmentos considerados pequenos, comparativamente aos *long reads*, com tamanho de fragmento gerado de até 600 bases. As tecnologias com base em *long reads*, como o MinION, possuem

capacidade para gerar dados de fragmentos maiores, 10 mil bases em média, mas com redução na quantidade de sequências (Amarasinghe *et al.*, 2020). A avaliação da qualidade dos fragmentos de leitura gerados no sequenciamento é realizada por meio do cálculo de incerteza na anotação de cada base. Este cálculo expressa a qualidade *phred*, cujo valor é inversamente proporcional à probabilidade de erro, de modo que expressa a acurácia de cada base sequenciada (Ewing e Green, 1998).

Apesar da diferença entre tecnologias de leituras pequenas e longas, ambas podem ser empregadas para um objetivo comum, como análise de viroma total ou direcionado. O custo é a principal limitação na realização do sequenciamento em larga escala, de modo que a metodologia baseada em nanoporos menor custo, quando comparada a plataforma Illumina (Ogden, Vasiljevic e Prost, 2021; Pervez *et al.*, 2022). A diferença de custo entre as metodologias se dá em grande parte pelo valor atrelado ao equipamento, além das células de leitura (chamadas de *flowcell*) que são descartáveis (Tabela 1). Em um sistema de vigilância genômica, o custo é um fator crucial, de modo que a otimização destas metodologias viabiliza a realização de maior número de amostras, proporcionando melhor amostragem das variantes circulantes na população. Os custos da vigilância genômica tem sido gerenciados por meio da identificação de grupos sentinelas que sejam representativos de grandes populações (Padilha *et al.*, 2023), além do uso de amostragem ambiental, que por meio da metagenômica possa auxiliar na identificação de patógenos infectocontagiosos circulantes.

Tabela 1. Comparativo de custo benefício das metodologias de sequenciamento

Plataforma	Custo da plataforma	Custo da rodada (por Mb)	Tamanho do fragmento de leitura	Acurácia
Illumina MiSeq	\$125.000,00	\$0.5	2x 300pb	99.99%
ONT MinION	\$1.000,00	\$1.9 a \$9	200pb a 1Mb	>99%

Legenda: Custo por rodada considerando insumos utilizados na preparação da biblioteca. As unidades monetárias estão expressas em dólar (\$). As unidades de tamanho de leitura estão expressas em pares de base (pb) e milhões de bases (Mb). ONT – *Oxford Nanopore*. Fonte: Adaptado de Ogden, Vasiljevic e Prost, 2021.

1.3 Metagenômica e Viroma

A metagenômica é a área das ciências ômicas que se dedica ao estudo da abundância e diversidade das comunidades microbianas. O viroma, constitui o estudo de um subgrupo destas comunidades: os vírus (Ghosh, Mehta e Khan, 2018). Por meio do sequenciamento genético e pelo uso da técnica do *shotgun*, obtém-se o material genético total presente em determinada

amostra (Quince *et al.*, 2017). A técnica de *shotgun* não envolve amplificação de material genético e, portanto, preserva as abundâncias dos vírus presentes na amostra.

A metodologia *shotgun* é utilizada para sequenciamento de viomas totais, que também pode ser obtido, alternativamente, por protocolos de amplificação aleatória de material genético com objetivo de aumentar o número e a qualidade das sequências obtidas (Quince *et al.*, 2017). Em ambas as metodologias, para análise de viroma total, é realizada a classificação taxonômica das leituras, denominada análise baseada em fragmentos de leitura (do inglês, *read-based analysis*) (Smith *et al.*, 2022) utilizando pipelines automatizados, como Kraken2 (Wood, Lu e Langmead, 2019) e Kaiju (Menzel, Ng e Krogh, 2016), que verificam a similaridade entre as sequências obtidas e genomas disponíveis em bancos de dados. A partir da análise de similaridade é atribuída a classificação taxonômica. O sequenciamento de um *pool* de material genético ambiental possibilita a identificação e quantificação de táxons, sem a necessidade de reconstruir, computacionalmente, os genomas encontrados, apesar desta metodologia também ser utilizada para classificação taxonômica (Escobar-Zepeda, León, De e Sanchez-Flores, 2015).

O viroma total inclui as diversas formas de vírus presentes em determinado ambiente. Vírus patogênicos, de interesse médico, veterinário e agrícola, são o principal foco (Ji *et al.*, 2020), haja vista que sequências de genomas potencialmente novos esbarram na fronteira do conhecimento científico. Análises exploratórias de viomas totais já foram realizadas com amostras de ambientes extremos (Hesse *et al.*, 2017; López-Bueno *et al.*, 2009) águas residuais (Nieuwenhuijse *et al.*, 2020), oceanos (Corte, De *et al.*, 2019; Hurwitz e Sullivan, 2013), além de órgãos e fluídos humanos (Edridge *et al.*, 2019; AlAttas *et al.*, 2020; Li *et al.*, 2020), apresentando a composição viral de cada ambiente e suas respectivas abundâncias.

Genomas montados a partir de amostras ambientais são denominados MAGs, (do inglês, *metagenome assembled genomes*) (Setubal, 2021), de tal forma que genomas virais obtidos com a mesma abordagem são denominados vMAGs (do inglês, *viral metagenome assembled genomes*), e constituem uma sequência representativa de vírions de mesmo táxon presentes no ambiente (Kieft *et al.*, 2022). A montagem dos genomas virais também é empregada na busca de sequências virais não caracterizadas e variantes, através de montagens *de novo* (*do zero*, em latim), ou seja, sem a utilização de um genoma de referência. As montagens resultam em *contigs*, sequências contíguas resultantes da sobreposição e concatenação de diversas leituras de sequenciamento, que podem representar um fragmento ou segmento de um genoma viral (Khan Mirzaei *et al.*, 2021). Duas métricas essenciais para avaliar a qualidade de um genoma são cobertura e profundidade. A cobertura diz respeito à porcentagem do genoma ou gene de

referência que foi obtido na montagem da amostra; e a profundidade representa, em média, quantas vezes cada base do genoma foi sequenciada (Dominguez Del Angel *et al.*, 2018). As sequências contíguas são classificadas, posteriormente, com base em sequências conhecidas depositadas em bancos de dados como as bases de dados de referência do Centro Nacional de informações Biotecnológicas estadunidense - NCBI (RefSeq) (O’Leary *et al.*, 2016); a base de dados do Uniprot (Bateman *et al.*, 2023) que inclui proteínas curadas manualmente – SwissProt e anotadas automaticamente por métodos computacionais - TrEMBL.

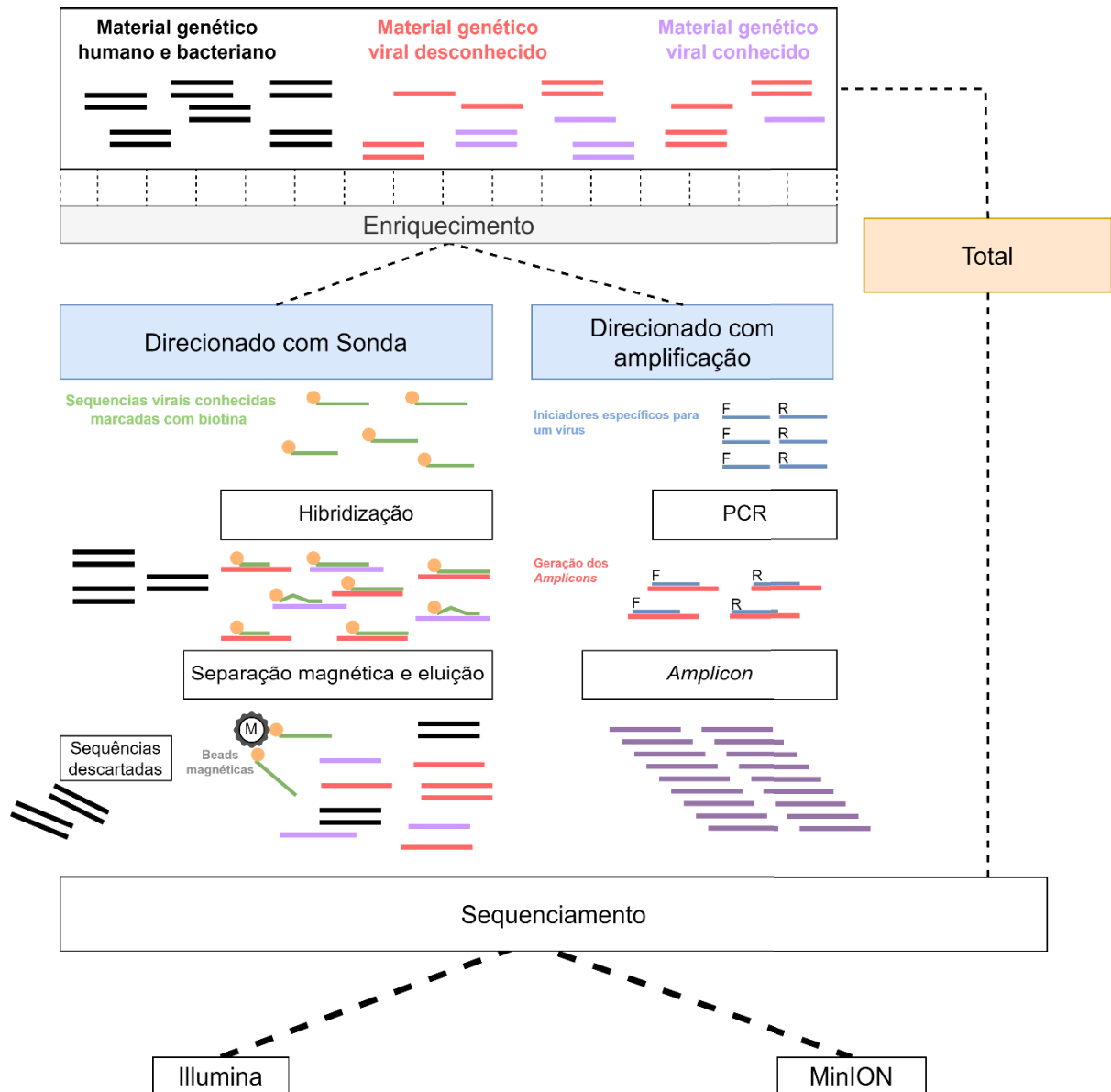
Independentemente da abordagem utilizada, baseada em leitura ou montagem, a classificação taxonômica assume vieses do banco de dados utilizado (Polson, Wilhelm e Wommack, 2011; Roux *et al.*, 2018). Genomas pouco caracterizados podem ser confundidos com falsos negativos, uma vez que há pouco material na literatura para identificação destas sequências raras.

Comunidades microbianas também são caracterizadas através da técnica de *amplicon*, em que se utiliza amplificação por PCR (do inglês, *Polymerase Chain Reaction*) de um gene marcador. A subunidade 16S do RNA ribossomal (rRNA) é o gene marcador de bactérias, a subunidade 18s rRNA de organismos eucariotos e o espaçador interno transcrito (do inglês, *Internal Transcribed Spacer - ITS*) é utilizado para classificar fungos (Ghosh, Mehta e Khan, 2018).

A ausência de um gene marcador universal para vírus, faz com que a análise a partir de *amplicons* seja impossibilitada para análise de diversidade viral total de um ambiente. Porém, *amplicons* de vírus de interesse são empregados em uma metodologia denominada viroma direcionado (Pogka *et al.*, 2022). Esta metodologia consiste na realização do sequenciamento direcionado a um vírus de interesse, a qual, a partir da amostra ambiental ou clínica, realiza-se a amplificação por PCR do genoma viral, sequenciamento e montagem do mesmo (Pogka *et al.*, 2022; Zhan *et al.*, 2021). Esta técnica foi empregada durante a pandemia de COVID-19 para o sequenciamento do vírus SARS-CoV-2 utilizando sequenciamento por nanoporos (Kim *et al.*, 2022). De fato, os primeiros genomas de SARS-CoV-2 obtidos no Brasil em 2020 utilizaram esta abordagem (Jesus *et al.*, 2020). Ainda, a fim de aumentar a recuperação do viroma, é possível isolar material genético viral a partir da utilização de sondas que hibridizam nas sequências dos genomas virais. Ou seja, por complementariedade, alinham a sequência da sonda com a sequência do genoma viral de forma que o material genético é selecionado. Esta tecnologia é recente e utiliza sequências-consenso depositadas em bases de dados públicos para seleção e construção das sondas, adicionando um nível de seletividade ao processo (Briese *et*

al., 2015). O comparativo entre as metodologias de viroma total e as abordagens de enriquecimento para obtenção de viroma direcionado está demonstrada na Figura 2.

Figura 2. Diferenças entre as abordagens para sequenciamento de vírus.



Legenda: Diferenças entre as metodologias de enriquecimento em comparação com o viroma total.
Fonte: Autor, adaptado de Pogka *et al.*, 2022

1.3.1 Características e classificação dos vírus

Os vírus compreendem o agrupamento mais abundante de entidades biológicas no planeta (Simmonds; Aiewsakun, 2018), podendo ser encontrados em todos os ecossistemas e apresentando uma grande diversidade (Hendrix, 2002). Vírus são agentes infecciosos, parasitas intracelulares obrigatórios, que infectam organismos dos três reinos da árvore da vida (Hendrix, 2002). A estrutura básica de um vírion, a partícula viral, consiste em um material genético (de

RNA ou DNA), envolto por uma estrutura proteica denominada capsídeo. Alguns vírus apresentam outra camada lipídica protetora externa ao capsídeo, o envelope (Knipe e Howley, 2013; Louten, 2016).

Vírus envelopados adquirem uma membrana externa ao capsídeo que advém das membranas das células hospedeiras de origens diversas, como organelas, membranas nucleares ou vesículas – a depender do ciclo replicativo do vírus. Este ciclo libera vírions envelopados a partir de brotamento da membrana e difere dos não envelopados em que as partículas virais são liberadas com a lise da célula hospedeira. Entretanto, uma grande quantidade de vírus envelopados no ambiente intracelular pode desencadear a lise celular (Knipe e Howley, 2013; Santos, 2015).

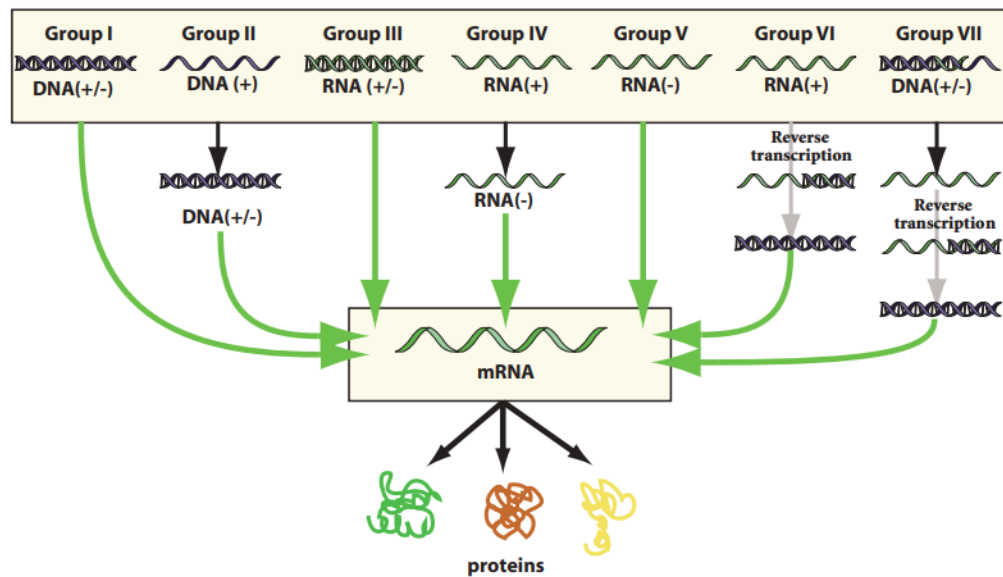
Os vírus podem ser classificados de acordo com o tipo de seu material genético, sendo a classificação realizada por David Baltimore pioneira na década de 1970 (Baltimore, 1971; Louten, 2016). A classificação de Baltimore distingue os vírus em sete agrupamentos (Tabela 2, Figura 3).

Tabela 2. Classificação de Baltimore

Grupo	Designação	Membros	Característica
I	dsDNA	HSV, HAdV	Material genético composto por DNA dupla fita
II	ssDNA	Circovírus	Material genético composto por DNA simples fita
III	dsRNA	RV	Material genético composto por RNA dupla fita
IV	ssRNA (+)	Pólio vírus, HAV, SARS-CoV-2	Material genético composto por RNA simples fita de sentido positivo
V	ssRNA (-)	Influenza, Ebola vírus	Material genético composto por RNA simples fita de sentido negativo
VI	ssRNA-RT	HIV	Material genético composto por RNA simples fita que realizam retrotranscrição
VII	dsRNA-RT	HBV	Material genético composto por DNA dupla fita que realizam retrotranscrição

Legenda: HSV – Herpes vírus; HAdV – Adenovírus Humano; HAV – Vírus da Hepatite A; HBV – Vírus da Hepatite B; HIV – Vírus da Imunodeficiência Humana; RV – Rotavírus; SARS-CoV-2 – *Severe Acute Respiratory Coronavirus 2*. Fonte: Adaptado de Brito e Pinney, 2017 e Fields, 2013.

Figura 3. Classificação viral de acordo com Baltimore (1970).



Fonte: SIB Swiss Institute of Bioinformatics (2020). Disponível em: <https://viralzone.expasy.org/254>. Acesso em: 28 ago. 2023.

Outra característica genética importante dos vírus de RNA é a presença e número de segmentos do genoma. Esta característica está presente em 11 famílias de vírus, incluindo famílias relevantes para a saúde humana como Orthomyxoviridae e Reoviridae, táxons dos vírus Influenza e vírus da família retroviridae Retrovírus, respectivamente (McDonald *et al.*, 2016). Os segmentos são moléculas de RNA e podem gerar rearranjos quando, em uma coinfeção de diferentes variantes em uma mesma célula, ocorre a mistura de segmentos de variantes distintas durante o encapsulamento do genoma viral (Horimoto e Kawaoka, 2005; McDonald *et al.*, 2016). A reorganização destes segmentos, denominada mutações em *drift*, é um mecanismo que gera variabilidade genética e é um fator que contribui para que saltos de hospedeiro ocorram, como nos surtos de gripe aviária, causados por variantes H5N1 e H9N2 do vírus Influenza A (Righetto; Filippini, 2018).

O genoma compacto dos vírus carrega um sistema otimizado em que cópias de uma mesma proteína são transcritas e traduzidas para formar o capsídeo, em um sistema de subunidades, arranjadas e montadas posteriormente na estrutura protetora do genoma. Isto implica em uma simetria natural deste processo, descrito pela primeira vez por Watson e Crick (1956).

A forma do capsídeo de vírus pode ser enquadrada na classificação de isométrico, filamentososo ou complexo, considerando vírus não envelopados. Os isométricos podem ser

definidos como estruturas esféricas, mas com faces triangulares compostas por oligômeros de uma ou mais subunidades proteicas. As formas filamentosas possuem simetria helicoidal, enquanto as formas complexas apresentam combinações isométricas e filamentosas. Vírus envelopados adicionam uma variedade a esta classificação, uma vez que proteínas presentes no envelope podem alterar sua morfologia (Knipe e Howley, 2013; Parvez, 2020). A simetria do capsídeo pode ser visualizada por microscopia eletrônica e é utilizada para diferenciação morfológica, a qual é uma importante ferramenta na caracterização e classificação viral (Knipe e Howley, 2013; Santos, 2015).

Diferentemente do que ocorre com organismos procariotos e eucariotos, os vírus não compartilham um ancestral comum, que ocasiona na falta de um gene universal, marcador do processo evolutivo (Koonin, Senkevich e Dolja, 2006). Vírus possuem genes com papéis importantes em seu ciclo, que aproximam alguns grupos não relacionados. Estes genes são o mais próximo de um marcador e são denominados genes característicos (do inglês, *hallmark genes*) (Koonin, Senkevich e Dolja, 2006). Um exemplo é a proteína JRC (do inglês, *Jelly-Roll capsid protein*), componente do capsídeo de vírus com formato icosaedro presente desde fagos à vírus que infectam arqueobactérias e mimivírus (vírus com tamanho de genoma gigante) (Scola, La *et al.*, 2008), dentre outros.

Desta forma, a taxonomia viral não é linear e combina o uso de características genéticas, de formato do capsídeo, presença de envelope, e tipo de hospedeiro para estabelecimento dos táxons (Adams *et al.*, 2017; Lefkowitz *et al.*, 2018). Estes critérios são utilizados pela Comissão internacional de Taxonomia Viral (do inglês, *International Committee on Taxonomy of Viruses*) – ICTV (<https://talk.ictvonline.org/>) para classificação de vírus. Classificações abaixo do nível de espécie (que incluem cepas, linhagens e variantes) não são englobados pelo ICTV. Assim, a classificação taxonômica viral difere da monofilética clássica - baseada na presença de um ancestral comum a qual todos os indivíduos de um táxon compartilham uma ou mais características. A classificação politética estabelece que um conjunto de propriedades caracteriza um táxon viral, e os membros pertencentes a este compartilham uma série destas características, sem nenhuma característica principal e determinante (Regenmortel, Van, 1990). Na atualização publicada em 10 de junho de 2023 (Zerbini *et al.*, 2023) o ICTV reconhecia 11.273 espécies de vírus, incluindo 2.818 gêneros, 264 famílias, 72 ordens, 40 classes, 17 filos e 10 reinos.

1.3.2 Vírus entéricos e respiratórios

Os vírus entéricos são caracterizados por realizar seu ciclo replicativo em células do trato gastrointestinal, de forma que podem ser liberados nos excrementos e transmitidos pela via orofecal. Estes vírus são causadores de conjuntivites, meningites e gastroenterites - comorbidades com baixa mortalidade em adultos, mas de grande relevância para saúde infantil e de imunocomprometidos (Cuevas-Ferrando et al., 2022; Oude Munnink & Van der Hoek, 2016). De forma mais severa, vírus entéricos podem causar hepatites e paralisia. Neste contexto, destacam-se os adenovírus (HAdV), como os vírus da hepatite A (HAV) e E (HEV); rotavírus (RV); norovírus; e astrovírus (Hellmér *et al.*, 2014).

Os vírus respiratórios, apesar de não utilizarem a rota entérica como principal via de replicação, podem ser excretados pelas fezes e estar presentes no esgoto das cidades. Apesar disto, a via orofecal não é a principal fonte de transmissão (Wigginton, Ye e Ellenberg, 2015). Historicamente, vírus respiratórios possuem potencial pandêmico e são detectáveis em águas residuais e esgoto, como em surtos causados pela Síndrome respiratória do Oriente Médio (do inglês, *Middle East Respiratory Syndrome* - MERS), SARS Coronavírus (do inglês, *Severe Acute Respiratory Syndrome*), influenza humana e aviária, dentre outros (Wigginton, Ye e Ellenberg, 2015).

1.3.3 Vigilância viral em esgoto

O esgoto das cidades pode ser uma ferramenta de monitoramento e vigilância dos vírus circulantes na população, não apenas os que envolvem a via entérica, mas também vírus respiratórios que são excretados nas fezes. Segundo dados de 2010 do Instituto Brasileiro de Geografia e Estatística (2017), no país, 2.211 municípios (39,6%) não possuíam rede de esgotamento sanitário, sendo 164 desses municípios em Santa Catarina (55,6% dos municípios do estado). Em Florianópolis, 87,8% dos domicílios possuem cobertura de esgotamento sanitário. (IBGE portal cidades <https://cidades.ibge.gov.br/>). No censo de 2022, Florianópolis consta com 537 mil habitantes. Por ser uma cidade turística, no período do verão, a população da capital do estado de Santa Catarina salta para 800 mil pessoas na estimativa de 2007 e um milhão de pessoas nas estimativas para 2023 (Dados plano municipal integrado de saneamento básico do município de Florianópolis, Apêndice B – Tabela 10). Em Canasvieiras, seguindo o mesmo documento, a população triplica no período de temporada, nos meses de dezembro a março. O sistema de coleta e tratamento de esgoto é o mesmo, mas com volume de entrada muito maior, fato que sobrecarrega o sistema, acarretando na descarga de matéria orgânica

diretamente nos corpos d'água. A mistura do volume de esgoto com chuvas de verão causa episódios de cheias e extravasamento dos rios, favorecendo contaminações (Cesa e Duarte, 2010). Este fato, aliado ao excedente populacional, más práticas de higiene e o calor, facilitam a disseminação de vírus, especialmente entéricos, acarretando em surtos frequentes no verão de Florianópolis (G1 SC, 2023; Governo de Santa Catarina - Secretaria de Saúde, 2016; NSC Total, 2021; Victoria *et al.*, 2010).

A detecção de vírus no esgoto por métodos moleculares – como PCR e RT-qPCR - requer a concentração das partículas virais diluídas na água. Pela diferença de carga, os vírions tendem a aderir a partículas em suspensão presentes no esgoto (Knipe e Howley, 2013). Para obtenção do material genético, a concentração destas partículas e a liberação dos vírions aderidos a elas é essencial. Metodologias baseadas em filtração e centrifugação, tais como ultracentrifugação, filtração por membrana e concentração por Polietileno Glicol (PEG) são empregadas. Destes, o método de concentração por PEG destaca-se como uma das abordagens com menor custo de execução (Gholipour *et al.*, 2022; Salvo *et al.*, 2021).

A concentração por PEG, foi proposta como método para a purificação de partículas virais (Hebert, 1963), baseado em coluna de separação, com base nas diferentes densidades das partículas virais. Desta forma, um gradiente de concentração de PEG separa fases com diferentes vírus presentes, de acordo com seu tamanho. A associação do PEG com centrifugação otimizou o processo para obtenção de concentrados virais (Clark, 1970). A eficiência de separação está diretamente relacionada com a solubilidade das partículas virais durante o processo de centrifugação, estabilizada pelo uso de diferentes concentrações de PEG e NaCl. Sua aplicação foi relatada para concentração de vírus de plantas (Hebert, 1963), vírus Epstein-Barr (Adams, 1973), vírus entéricos como HAV e Rotavírus (Lewis e Metcalf 1988) e recentemente utilizado para concentração de SARS-CoV-2 em amostras de esgoto (Fongaro, Stoco, et al., 2021). O protocolo de Lewis e Metcalf, utilizado neste projeto, descreve eficiência de recuperação viral de 20%, considerando os vírus HAV.

A detecção de vírus em esgoto por métodos moleculares já foi empregada na detecção de Adenovírus (Bibby e Peccia, 2013; Schlindwein, Rigotto, Simões, C. M. O., *et al.*, 2010; Wong, Onan e Xagorarakis, 2010), Rotavírus (Bibby e Peccia, 2013; Schlindwein, Rigotto, Simões, C. M. O., *et al.*, 2010), Norovírus (Wong, Onan e Xagorarakis, 2010), HIV (Bibby e Peccia, 2013), Hepatite A (Bibby e Peccia, 2013; Jebri *et al.*, 2012; Wong, Onan e Xagorarakis, 2010), Hepatite C (Bibby e Peccia, 2013), Papillomavírus (Bibby e Peccia, 2013), Herpesvírus (Schlindwein, Rigotto, Simões, C. M.O., *et al.*, 2010), Poliomavírus (Schlindwein, Rigotto, Simões, C. M. O., *et al.*, 2010), Poliovírus (Schlindwein, Rigotto, Simões, C. M. O., *et al.*,

2010) e recentemente no monitoramento de SARS-CoV-2 (Fongaro, Stoco, *et al.*, 2021; Gholipour *et al.*, 2022; Serra-Compte *et al.*, 2021). Em Florianópolis, diversos trabalhos vêm documentando a ocorrência de vírus, principalmente entéricos, em amostras de águas e esgoto (Garcia *et al.*, 2012; Gholipour *et al.*, 2022; Moresco *et al.*, 2012; Mortari *et al.*, 2023).

Neste contexto, em meio a uma pandemia, o projeto de vigilância epidemiológica ambiental integrativa VigEAI, surgiu no laboratório de virologia aplicada da UFSC (LVA). Este projeto visou detectar a presença de RNA do vírus SARS-CoV-2 em amostras de esgoto de diversas cidades de Santa Catarina, como estratégia para monitoramento ambiental no período pandêmico. O projeto contou com foi realizado em parceria com as prefeituras e companhia de águas do estado, a CASAN. Através da detecção molecular por RT-qPCR, detecções e quantificações serviram de indicativo para menor circulação do vírus pandêmico. Boletins mensais foram gerados no período, contribuindo para o monitoramento de SARS-CoV-2 no estado de Santa Catarina (Wachter *et al.*, 2023).

No contexto da vigilância genômica, durante a década de 1990, o monitoramento de vírus no esgoto por sequenciamento foi utilizado durante o esforço global para erradicação da poliomielite em países como Índia, Afeganistão e Paquistão, por meio de isolados (Asghar *et al.*, 2014). Ainda, o material genético de outros vírus de interesse médico, como o zika, ebola e norovírus, já foram detectados em esgotos de cidades (Bibby *et al.*, 2015; Muirhead *et al.*, 2020; Victoria *et al.*, 2010), apesar de fatores limitantes como: pH extremo, salinidade, alta incidência de radiação ultravioleta; e efeitos de diluição - causados pelas chuvas e pela contribuição de matéria orgânica de organismos não infectados (Hernroth *et al.*, 2002). Apesar destas barreiras, a amostragem de esgoto não requer aprovação por comitês de ética e permite analisar a composição viral de aglomerados populacionais, como bairros e cidades inteiras (Nieuwenhuijse *et al.*, 2020). Vírus presentes no esgoto das cidades podem ser utilizados como indicadores da presença e dispersão de doenças causadas por estes patógenos. Durante a pandemia de COVID-19, causado pelo vírus SARS-CoV-2, estudos com amostras de esgoto detectaram o vírus presente em algumas cidades antes de casos positivos serem detectados (Lahrich *et al.*, 2021). Em Florianópolis, um estudo realizado em 2020, detectou a presença do RNA viral em amostras de esgoto cerca de três meses antes do primeiro caso de COVID-19 registrado oficialmente, evidenciando a presença do vírus circulante (Fongaro, Stoco, *et al.*, 2021).

Tendo em vista protocolos de concentração viral por PEG já estabelecidos na vigilância molecular de vírus em esgoto, hipotetizamos que a adaptação desta metodologia em conjunto

com um protocolo de sequenciamento de baixo custo permitirá aplicar esta metodologia na vigilância genômica de vírus de interesse em saúde pública.

2 OBJETIVOS

2.1 Objetivo Geral

Avaliar a aplicabilidade do método de concentração baseado em polietilenoglicol (PEG) associado ao sequenciamento *shotgun* por nanoporos para identificação de vírus.

2.2 Objetivos Específicos

- Avaliar a viabilidade do protocolo de concentração por PEG para sequenciamento *shotgun* de vírus em plataforma de sequenciamento por nanoporos;
- Aprimorar o protocolo de concentração por PEG para melhor recuperação de fragmentos de leitura virais;
- Identificar genomas virais selvagens a partir de amostras de esgoto coletadas em Canasvieiras – Florianópolis em dez meses do ano de 2021.

3 MATERIAL E MÉTODOS

3.1 Delineamento experimental

Dez amostras de esgoto (Tabela 3) coletadas na estação de tratamento de esgoto (ETE) de Canasvieiras – Florianópolis/Santa Catarina (Figura 4), foram selecionadas, cedidas pela Companhia Catarinense de Águas e Saneamento (CASAN) ao Laboratório de Virologia Aplicada (LVA/UFSC), como parte do projeto de Vigilância Epidemiológica Ambiental Integrativa - VigEAI (<https://lvapli.paginas.ufsc.br/vigeai/>). Estas amostras foram utilizadas para sequenciamento *shotgun* visando obtenção do viroma total. Outras seis amostras dos municípios de Concórdia, Itajaí e Joinville e Florianópolis (Lagoa da Conceição) foram utilizadas para sequenciamento do viroma direcionado para SARS-CoV-2 (Tabela 4). Para estes, foram utilizados como critério de seleção, a quantificação da carga viral na amostra. Foram coletadas 2L de esgoto bruto das amostras, ou seja, antes de entrar na ETE.

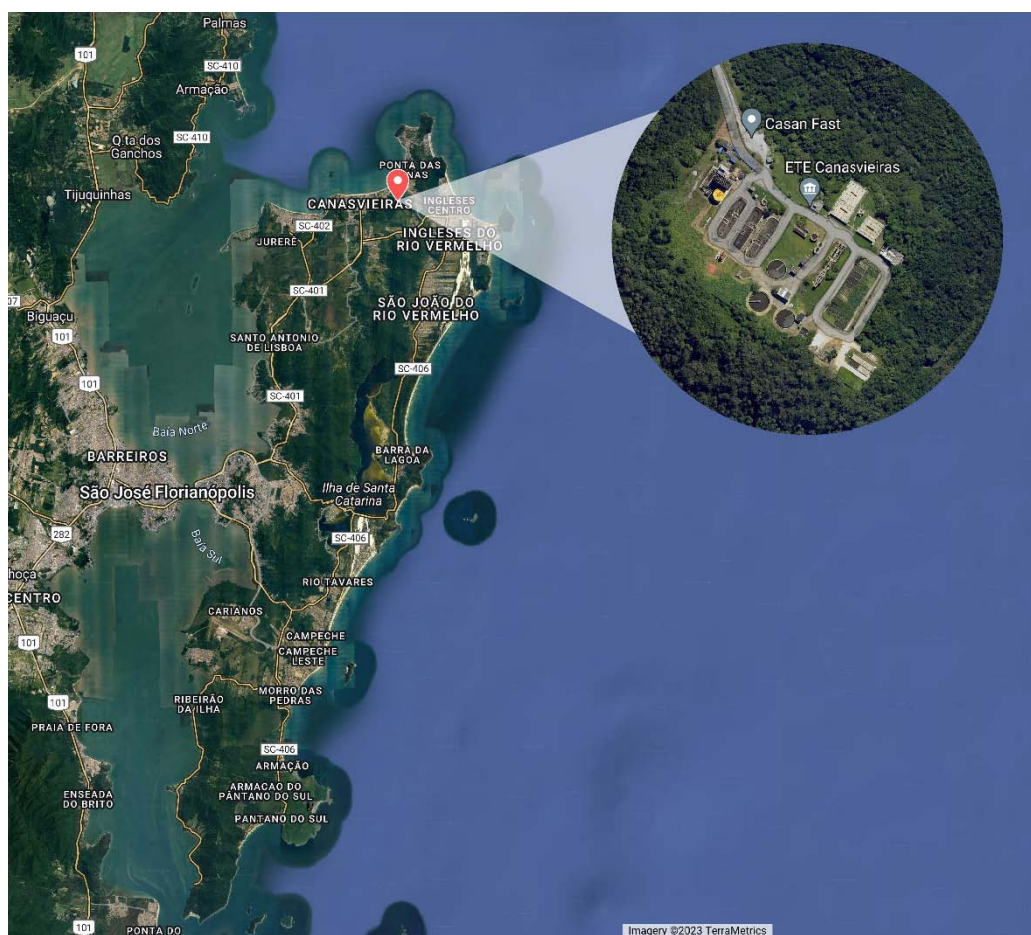
Tabela 3. Amostras utilizadas para sequenciamento do viroma total

Amostra	Data de coleta
Janeiro	01/06/2021
Fevereiro	02/02/2021
Março	15/03/2021
Abril	19/04/2021
Maiο	03/05/2021
Junho	07/06/2021
Agosto	02/08/2021
Outubro	18/10/2021
Novembro	22/11/2021
Dezembro	06/12/2021

Tabela 4. Amostras utilizadas para sequenciamento por *amplicons* de SARS-CoV-2

Amostra	Município	Data de coleta	Carga Viral (GC/L ⁻¹)
CON_09_1	Concórdia	08/09/2021	1.000
ITA_09_2	Itajaí	09/09/2021	10.000
JOI_10_1	Joinville	28/10/2021	10.000
ITA_11_4	Itajaí	09/11/2021	10.000
JOI_12_2	Joinville	06/12/2021	10.000
FLO_12_5	Florianópolis	13/12/2021	1.000

Legenda: Amostras que passaram pelo processo de amplificação para genoma do SARS-CoV-2 e foram sequenciadas por metodologia Illumina. A carga viral é expressa em Cópias Gênicas/Litro.

Figura 4. Localização da estação de tratamento de esgoto de Canasvieiras, Florianópolis.

Legenda: Localização da estação de tratamento de esgoto de Canasvieiras, Florianópolis. indicada pelo ponto vermelho (Ponto de georreferenciamento: -27.43138248031413, -48.43416827205362). Em recorte uma aproximação da estação. Fonte: Autor, 2023.

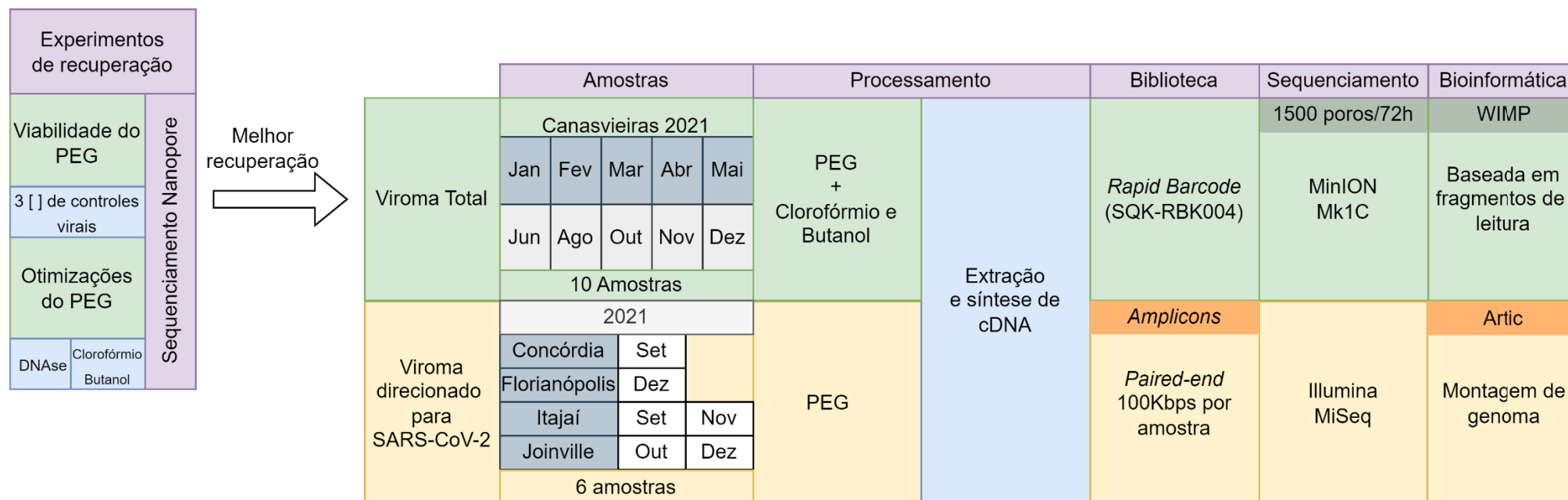
Três experimentos foram realizados para avaliar a recuperação de fragmentos de leitura virais, utilizando a metodologia de concentração polietilenoglicol (PEG) em conjunto com o

sequenciamento *shotgun* por nanoporos. Esta metodologia de concentração foi escolhida devido à exequibilidade e por fazer parte da rotina de monitoramento molecular do Laboratório de Virologia Aplicada da UFSC. Uma visão geral dos processos está demonstrada na Figura 5. Desta forma, sequenciamentos foram realizados para o emprego desta metodologia para: (1) detecção de um limite de concentração detectável no sequenciamento *shotgun* e; (2) segundo aferir o efeito da redução de inibidores em relação a recuperação de fragmentos de leitura virais. A partir destes experimentos, foi estabelecido o método utilizado para obtenção de viroma total para as amostras de Canasvieiras.

O primeiro experimento (Figura 6 – item 1) objetivou obter um limite de concentração mínima de cópias gênicas/mL, para recuperação de fragmentos de leitura virais no sequenciamento *shotgun*. O segundo experimento (Figura 6 – item 2) visou aprimorar o protocolo de concentração PEG para sequenciamento utilizando a abordagem de redução de inibidores orgânicos, pelo processo de clarificação por clorofórmio e butanol, e redução de DNA bacteriano utilizando tratamento com DNase.

Com base nestes dois experimentos, utilizou a concentração por PEG em conjunto com a clarificação por clorofórmio butanol na concentração das dez amostras de esgoto da ETE Canasvieiras do período de 2021 valendo-se apenas da amostra de esgoto, sem adição de controle viral para sequenciamento.

Figura 5. Esquema geral dos processos para sequenciamento de viroma direcionado e viroma total



Legenda: Esquema geral dos processos realizados neste trabalho. No primeiro bloco, foram realizados experimentos de sequenciamento para análise da recuperação viral em sequenciamento *shotgun* em plataforma nanoporos: No quadro a esquerda, Primeiro foram realizados experimentos para análise da viabilidade do método de concentração PEG utilizando três concentrações de controles diferentes e; segundo, foi realizado otimizações do protocolo PEG, utilizando protocolos com clorofórmio e butanol e DNase. Os resultados do primeiro bloco foram aplicados para sequenciamento das amostras de Canasvieiras na abordagem de viroma total, dispostos no bloco a direita, parte superior. Alternativamente, foi realizado viroma direcionado para SARS-CoV-2 utilizando seis amostras de quatro cidades diferentes, cujo esquema está disposto no segundo bloco, porção inferior. PEG – Polietileno Glicol; [] – concentração; Kbps – Mil pares de base; WIMP – *What Is in My Pot*. Fonte: Autor (2023).

Cinco vírus foram utilizados como controles de contaminação viral nos experimentos de sequenciamento apresentados na Figura 6. As alíquotas virais foram produzidas no LVA/UFSC. A partir de estoques do laboratório, alíquotas virais foram inoculadas em uma monocamada de celular, utilizando as seguintes linhagens celulares: A549 para HAdV-2 e L929 para MHV-3 seguindo protocolo descrito por Cadamuro (2022); RAW 246.7 para produção de MNV-1 seguindo protocolo descrito por de Abreu Corrêa e colaboradores (2012); ATCC CCL 81 para produção das alíquotas virais de HSV-2 seguindo protocolo descritos por Feltrin (2014); FRHK-4 para RVA seguindo protocolo descrito por Schlindwein e colaboradores (2010).

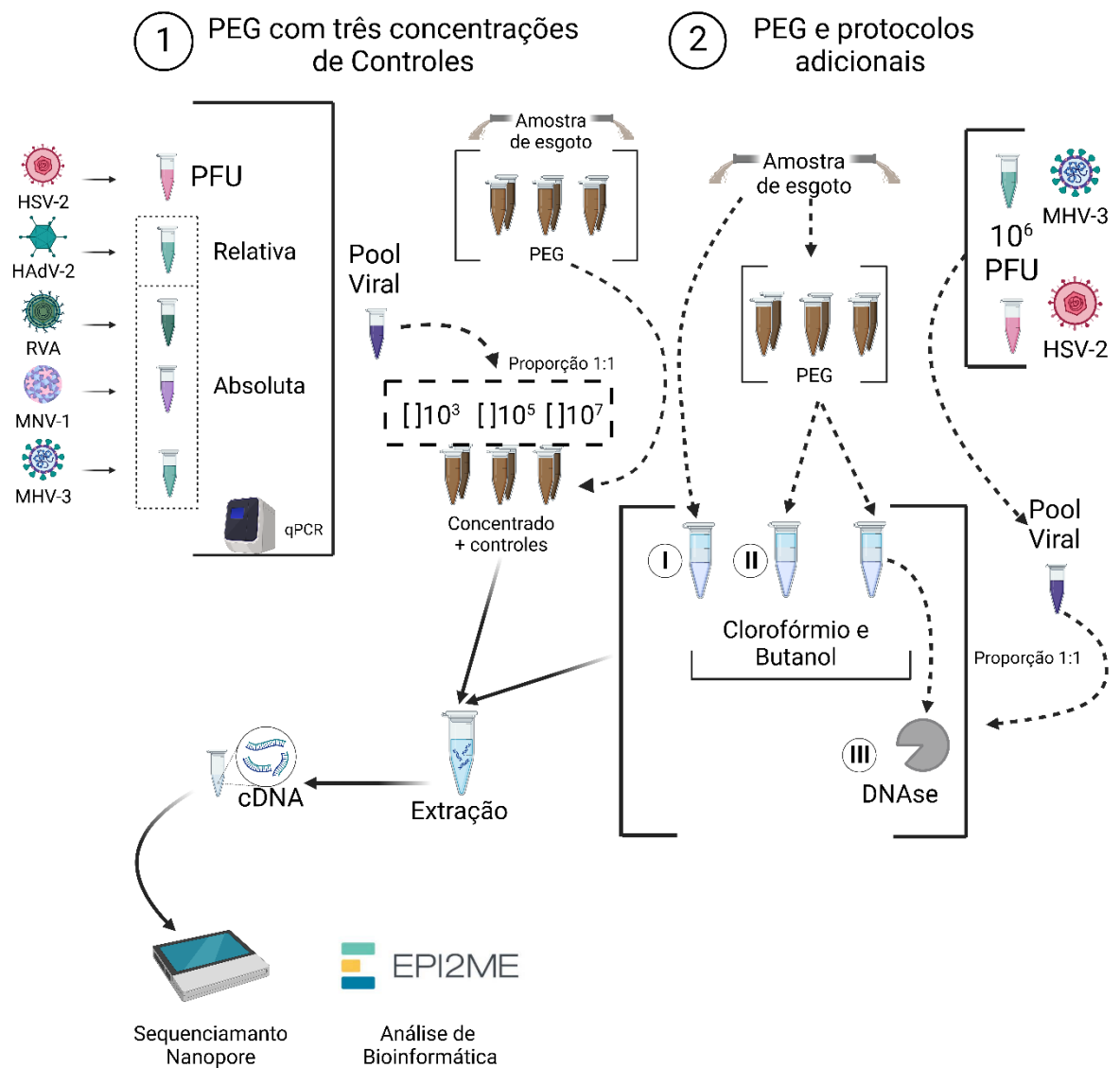
A determinação dos títulos virais foi realizado por meio dos ensaios de placa e pela contagem de unidades formadoras de placa expressos em PFU/mL seguindo protocolo descrito por Burlison Chambers e Wiedbrauk (1992).

As alíquotas virais foram quantificadas de forma absoluta por RT-qPCR, em cópias genômicas (CG)/mL em: 5×10^{11} para Vírus da Hepatite Murina-3 (MHV3); $2,4 \times 10^8$ para Rotavírus A (RVA); $1,8 \times 10^7$ para Norovírus Murino (MNV-1). A alíquota viral de HAdV-2 foi quantificada de forma relativa, pelo valor de CT da qPCR, de forma que o CT 36 foi estimado em cerca de 10^2 CG/mL. Mesmo em baixa concentração, o controle foi mantido. Uma concentração previamente conhecida (10^{10} CG/mL) da alíquota viral produzida de HSV-2, foi utilizada como valor de quantificação, baseado na titulação.

Para o primeiro experimento de sequenciamento, (Figura 6 – item 1) os vírus foram inoculados em proporção 1:1 (vírus: esgoto), antes da extração de ácidos nucleicos, com volume final para extração de 200 μ L. Cinco vírus controles foram utilizados neste experimento, em concentração 10^7 GC/mL para os vírus MHV-3, MNV-1, RV, adicionado a HSV-2 em 10^{10} GC/mL com volume completado por alíquota de HAdV-2, quantificado em 10^2 GC/mL. Deste *pool* foram geradas diluições, utilizando amostras de esgoto como diluente para atingir as concentrações de 10^5 e 10^3 GC/mL dos vírus MHV-3, MNV-1 e RVA, e respectivas diluições dos vírus HSV-2 e HAdV-2. Amostras controles foram adicionadas contendo *pool* viral e água destilada nas mesmas concentrações.

No segundo experimento (Figura 6 – item 2), foram inoculados apenas MHV-3 e HSV-2 proporção 1:1 (vírus: esgoto), antes da extração de ácidos nucleicos, com volume final para extração de 200 μ L. Utilizou-se como base a concentração conhecida de 10^{10} CG/mL obtido a partir da contagem de unidades formadoras de placa, como valor de concentração do pool viral composto por HSV-2 e MHV-3.

Figura 6. Desenho experimental utilizado para obtenção do protocolo de sequenciamento *shotgun* de viroma total em plataforma de sequenciamento nanoporos



Legenda: Diagrama dos sequenciamentos realizados para estabelecimento do protocolo a ser realizado com as amostras de esgoto de Canasvieiras-SC. (1) Sequenciamento com diferentes concentrações de controles virais. (2) Sequenciamento para avaliação de métodos complementares ao PEG: (I) Apenas tratamento com clorofórmio e butanol (CB); (II) PEG + CB; (III) PEG + CB + tratamento com DNase. Controles utilizados: MHV-3 – Hepatite murino 3; MNV-1 – Norovírus murino; RVA – Rotavírus A; HAdV-2 – Adenovírus tipo 2. HSV-2 – Herpes vírus humano tipo 2. PEG - Polietileno Glicol. Criado com biorender.com. Fonte: Autor (2023).

3.2 Processamento das amostras de esgoto

As partículas virais presentes no esgoto foram concentradas por PEG, utilizando-se o protocolo de Lewis e Metcalf, (1988), com adaptações feitas por Wu et al., 2020. Foram utilizados 25 mL de cada amostra de esgoto, o pH foi ajustado para 4,5, utilizando HCl 1M, permanecendo em agitação por 30 minutos. Posteriormente, foi realizada centrifugação a 4°C/15 min a 6.500 rpm. O sobrenadante foi descartado e o precipitado foi ressuscitado em 10 mL de tampão glicina (0,25M pH 9,5). Um novo ciclo de agitação foi realizado por mais 30 minutos, seguido de nova centrifugação, a 4°C 6500 rpm/30 minutos. O sobrenadante foi transferido para um novo tubo em que foi adicionado 1mL de PEG 6.000 24% preparado em 1,2M de NaCl (24g de PEG 6000 em 1L de NaCl). Essa solução (sobrenadante + PEG) foi incubada, sob agitação, por 2 h/4°C e centrifugada a 10.000 rpm a 4°C/90 min. Após a centrifugação o sobrenadante foi descartado e o material sedimentado foi ressuscitado em 5 mL de PBS (do inglês, *Phosphate-Buffered Saline*) pH 7,0 armazenado a -20°C para posterior extração dos ácidos nucleicos.

Variações do protocolo:

- A clarificação com clorofórmio e butanol utilizou a proporção 1:1, por meio da adição de 200 µL de clorofórmio e de butanol ao mesmo volume de amostra, previamente processada pelo método de concentração PEG. Posteriormente, realizou-se uma centrifugação a 5.000 rpm a 4°C/30 minutos. Este processo gerou três fases visíveis, em ordem ascendente de densidade: proteica, solvente e fase lipídica. Obteve-se a fase proteica e esta foi utilizada para extração de material genético.
- O tratamento com DNase (Invitrogen, MA, EUA), foi realizado utilizando 2,5 µL de DNase para cada 1ml de amostra com 100 µL de tampão 10X. As amostras foram incubadas a 37°C/1 h e posteriormente foi adicionado EDTA (Ácido etilenodiaminotetracético) em volume de 2,5 µL para inativação da enzima. Após incubação a 75°C/10min, as amostras foram armazenadas a -20°C até a extração dos ácidos nucleicos.

3.3 Ensaios moleculares

3.3.1 Extração de ácidos nucleicos e síntese de DNA complementar (cDNA)

As extrações foram realizadas utilizando *kit* comercial QIAamp® Viral RNA Mini (QIAGEN, CA, EUA), seguindo indicações do fabricante e eluído em 50 µL de água RNase

free. A concentração de DNA e RNA foi determinada através do equipamento NanoVue (General Electric) a partir de 1 µL (Apêndice B – Tabela 11).

A síntese de cDNA foi realizada em protocolo adaptado de Green e Sambrook (2019). De forma breve, uma solução foi criada utilizando, para uma amostra, 1 µL de iniciadores aleatórios 50 µM (*random hexamers* - Invitrogen, MA, EUA), 0,33 µL do iniciador oligo (dT) 50 µM (Invitrogen, MA, EUA), 10 µM de dNTP (Invitrogen, MA, EUA) e 3,66 µL de água RNase *free*. Esta solução teve 7 µl de amostra adicionada, em seguida foi incubada a 65°C/5 min. Após este processo, outra solução contendo: 4 µL tampão 5x SSIV Buffer (Invitrogen, MA, EUA); 1 µL DTT 100mM (Invitrogen, MA, EUA); 1 µL de inibidor de ribonuclease RNaseOUT™ 40 U/µL (Invitrogen, MA, EUA); 1 µL de enzima transcriptase reversa SuperScript IV 200 U/µL (ThermoFisher, MA, EUA); foi adicionada a amostra e realizado os termo-ciclos de 23°C/10 min, seguido de 50°C/50 minutos e finalizando com 80°C/10 minutos.

3.3.2 Quantificação dos controles virais

Os ácidos nucleicos das alíquotas virais controles foram quantificados através de reações de RT-qPCR (Vírus com genoma composto por RNA) e qPCR (Vírus com genoma composto por DNA) valendo-se do *kit AgPath-ID™*, equipamento *One Step RT PCR (Applied Biosystems)* e *QuantiNova Probe RT-PCR kit (QIAGEN)*. Os iniciadores e sondas utilizados seguiram os descritos nos seguintes protocolos: para Adenovírus tipo 2 (HAdV-2) (Hernroth *et al.*, 2002); MHV-3 (Besselsen, Wagner e Loganbill, 2002); MNV-1 (Baert *et al.*, 2008); Rotavírus (Zeng *et al.*, 2008).

3.3.3 Quantificação de vírus selvagens em esgoto de Canasvieiras por qPCR e RT-qPCR

No âmbito do projeto VigEAI, resultados prévios de detecção de vírus em amostras de esgoto foram incorporados a este trabalho. No projeto, as amostras de esgoto para vigilância molecular foram concentradas pelo método de PEG. Os ácidos nucleicos foram extraídos e quantificados para iniciadores específicos dos vírus da Hepatite A (HAD), Adenovírus (HAdV) e coronavírus (SARS-CoV-2) por meio de reações de RT-qPCR utilizando o kit OneStep qPCR Quantinova (QIAGEN, Alemanha). Para reação RT-qPCR para SARS-CoV-2 utilizou-se as regiões N1 e N2 do genoma viral como alvos de amplificação com o método TaqMan® (Lu *et al.*, 2020). O controle positivo utilizado foi o vírus da hepatite murina 3 (MHV-3).

3.4 Sequenciamento genético

Os sequenciamentos seguiram o protocolo de montagem de biblioteca de DNA genômico para sequenciamento *shotgun* descrito no protocolo de sequenciamento - *Rapid sequencing gDNA - barcoding* (SQK-RBK004) da Oxford Nanopore (Anexo 1). Inicialmente, as amostras foram quantificadas pelo Qubit (Invitrogen, MA, EUA) utilizando 1 μ L de cDNA das amostras (Anexo B – Tabela 12). Com base na quantificação, padronizou-se para obtenção de 400 ng em volume total de 7,5 μ L. Neste, 2,5 μ L de *barcodes* do kit SQK-RBK004 (Oxford Nanopore) foram adicionados - de forma individual em cada amostra. As amostras foram incubadas no termociclador em um ciclo de 30°C/1 minuto, seguido de 80°C/1 minuto para ligação dos *barcodes* nos ácidos nucleicos. Em seguida, em um único tubo, as amostras foram ressuspendidas com *beads* AMPure XP (Agcourt) utilizando volume igual ao total da biblioteca, que corresponde ao volume somado de todas as amostras. Este volume difere em cada experimento, uma vez que foram realizados três sequenciamentos, com nove, quatro e dez amostras, respectivamente.

Para o sequenciamento direcionado para SARS-CoV-2, as amostras concentradas pelo PEG passaram por amplificação do material genético do vírus de SARS-CoV-2 utilizando iniciadores do protocolo Illumina (Eden e Sim, 2020) (Illumina Inc, San Diego, CA, USA), para produção de quatorze *amplicons*, processados com Illumina DNA Prep (Illumina, San Diego, CA, USA). Esta metodologia é a mesma utilizada na vigilância genômica do vírus em Santa Catarina a partir de amostras clínicas (Padilha *et al.*, 2022). O sequenciamento foi realizado por meio da plataforma Illumina com biblioteca *paired-end* com output de 100 mil fragmentos de leitura pela a empresa BiomeHub.

3.4.1 Análises em Bioinformática

Os fragmentos de leitura obtidos do sequenciamento de leituras longas foram processados pela pipeline *fast-Basecalling* do próprio equipamento (MinION Mk1c – Oxford Nanopore), que utiliza o software Guppy 5.0.7 (Wick, Judd e Holt, 2019). Neste *workflow*, utilizou-se como valores mínimos: qualidade *phred* média de 8, e tamanho mínimo de fragmento de 200 bases. Fragmentos de leitura em desacordo com estes parâmetros foram descartados. Os fragmentos de cada amostra foram separados considerando o respectivo *barcode* adicionado na preparação da biblioteca por meio do software Guppy. A análise de composição taxonômica foi realizada por meio da pipeline automatizada da EPI2ME, utilizando o *workflow* WIMP versão 2023.06.13-1865548 (do inglês, *What is in my pot*). A pipeline usa a

base de dados RefSeq do NCBI (O’Leary *et al.*, 2016) em conjunto com o *software* centrifuge (Kim *et al.*, 2016) para estabelecer a classificação de cada fragmento de leitura. Atribuições taxonômicas como *root* e *cellular organism* foram desconsideradas e contabilizadas em conjunto com sequências não classificadas.

Os genomas do viroma direcionado para SARS-CoV-2 foram montados pela metodologia descrita por Padilha e colaboradores (2023). Brevemente, os fragmentos de leitura passaram por um controle de qualidade para remoção de sequências de baixa qualidade, fragmentos de leitura de origem humana (NCBI/RefSeq id: GCF_000001405.13) e adaptadores de sequenciamento. Após este processo, os fragmentos de leitura foram alinhados no genoma de referência de SARS-CoV-2 (NCBI/RefSeq id: NC_045512.2), para gerar a sequência consenso, utilizando o BWA-MEM v. 0.7.17 (Li e Durbin, 2009) iVar v 1.3 (Grubaugh *et al.*, 2019) e a suíte de programas SAMtools v 1.7 (Danecek *et al.*, 2021). O programa *web* IGV foi utilizado para visualização dos alinhamentos (Robinson *et al.*, 2011).

4 RESULTADOS E DISCUSSÃO

4.1 Monitoramento da contaminação viral presente no esgoto coletado em Canasvieiras

O monitoramento das contaminações por HAdV, HAV e SARS-CoV-2, por (RT)-qPCR, nas amostras de esgoto coletadas na ETE Canasvieiras em 2021 ao longo de 12 meses, realizado no escopo do projeto VigEAI, revelou as presenças de HAdV e SARS-CoV-2 (Tabela 5). HAdV de outubro a dezembro/2021, ao passo que SARS-CoV-2 foi detectado em janeiro, fevereiro e abril/2021.

Tabela 5. Quantificação de carga viral em cópias gênicas/L-1 a partir de amostras de esgoto de Canasvieiras, Florianópolis, no ano de 2021.

Mês de coleta	HAV	HAdV	SARS-CoV-2	SARS-CoV-2 Quantificação
Janeiro	-	-	Detectado	10 ²
Janeiro	-	-	Detectado	10 ³
Fevereiro	-	-	Detectado	10 ³
Março	-	-	-	-
Abril	-	-	Detectado	-
Maiο	-	-	-	-
Junho	-	-	-	-
Agosto	-	-	-	-
Outubro	-	4,0 E10 ⁵	-	-
Novembro	-	8,0 E10 ⁴	-	-
Dezembro	-	4,0 E10 ⁵	-	-

Legenda: HAV – Vírus da Hepatite A; HAdV – Adenovírus. SARS-CoV-2 – *Severe Acute Respiratory Syndrome coronavirus 2*. As quantificações estão expressas em (Carga Viral em Cópias Gênicas/L⁻¹)

Fonte: Projeto VigEAI – Laboratório de Virologia aplicada (dados não publicados)

Com base nos resultados de qPCR e RT-qPCR foi possível quantificar a carga viral em cópias gênicas por L. O vírus da Hepatite A (HAV) também foi testado, mas sem resultados positivos para o período. O sul do Brasil não compreende área endêmica de HAV, com baixos índices de infecção (Grandi, Lopez e Burattini, 2022). A cobertura de esgotamento sanitário, é considerada elevada em Florianópolis, comparada a realidade brasileira, sendo um fator que reduz a rota de transmissão do vírus da hepatite A (IBGE, 2020). A liberação de partículas virais

por paciente infectado por HAV nas fezes é de cerca de 10^{10} partículas virais por grama de fezes. A detecção do vírus já foi relatada em estudos anteriores com águas de Florianópolis (Rigotto *et al.*, 2010), indicando possível supressão das transmissibilidades do vírus ao longo do tempo.

Entretanto, é necessário ressaltar que o HAV é um vírus envelopado de RNA, que possui baixa resistência em águas de esgotamento sanitário, uma vez que seu capsídeo é neutralizado na presença de detergentes (Casanova; Weaver, 2015). Este processo faz com que o material genético seja exposto e degradado. Desta forma, apesar da alta taxa de partículas virais liberadas, o tempo entre a liberação dos vírions e a coleta é um fator que influencia diretamente na recuperação deste vírus.

A incorporação da vacina contra o HAV ao cronograma brasileiro de vacinação em 2014, em esquema de dose única, contribuiu para redução das infecções, especialmente em crianças (Grandi, Lopez e Burattini, 2022; Ministério da Saúde, 2013). Importante ressaltar que a vacina contra HAV é baseada no vírus atenuado replicante, de forma que sua presença no esgoto pode contrastar origens vacinais com variantes patogênicas (Shouval, 2019).

Diversos trabalhos têm buscado a detecção de HAV por metodologia molecular em amostras de esgoto, mundialmente (Gholipour *et al.*, 2022). Destaca-se o trabalho realizado no Líbano, que indicou fontes de água utilizadas para consumo como fontes de transmissão do vírus após mais de 600 casos positivos serem detectados em 2022 (Uwishema *et al.*, 2022). Na Argentina, o HAV foi detectado em conjunto com o vírus da hepatite E (HEV) (Lo Castro *et al.*, 2023). A detecção do HAV em esgoto indica para a disseminação do vírus demonstrando a relevância do monitoramento ambiental para a epidemiologia da doença.

Com relação aos adenovírus, estes permanecem estáveis no ambiente e são descritos na literatura como um dos vírus frequentemente encontrados em esgoto sanitário (Thompson *et al.*, 2003). A persistência dos adenovírus no ambiente, é um fator importante para que ele seja elencado como na lista de candidatos a contaminação (CCL) da agência americana de proteção ambiental (EPA), que elenca microorganismos conhecidos por indicar contaminação d'água e sinalizar para problemas de saúde na população (EPA, 2022). Desta forma, é possível utilizar a detecção de HAdV no ambiente como um indicador de poluição. Em esgoto bruto não tratado, esta informação não possui relevância, mas para o monitoramento ambiental de fontes de água potável é um indicador importante.

Em 2012, Garcia e colaboradores quantificaram em Concórdia - oeste do estado de Santa Catarina, HAdV de agosto de 2010 a janeiro de 2011, com resultados positivos em todos os meses amostrados e carga viral média de 10^7 cópias genômicas por L^{-1} (Garcia *et al.*, 2012).

Sch lindwein e colaboradores, em estudo de 2010 detectaram a presença de HAdV em diferentes fontes de água da Ilha de Santa Catarina, com resultados positivos para a viabilidade do vírus em cultura celular. A quantificação realizada pelo grupo apresentou valores entre 10^4 , 10^5 e 10^2 cópias genômicas/ml em águas tratada de esgoto nos meses de outubro, novembro e dezembro de 2010, respectivamente. Os dados relatados por Sch lindwein (2010) se relacionam com os valores obtidos na quantificação atual, apesar da diferença temporal, demonstrando a persistência de infecções por adenovírus em Florianópolis.

As detecções nos meses de outubro a dezembro, escapam do período de maior circulação de pessoas na Ilha. Considerando o período pandêmico de COVID-19, a circulação de turistas foi reduzida no mundo inteiro em decorrência da emergência de saúde pública e dos danos econômicos e à população, ainda em recuperação em 2021 (Ministério do Turismo, 2020). Em Santa Catarina, ainda em 2021 medidas de restrições eram impostas, como o decreto estadual 1172 de 26 de fevereiro de 2021 que instaurou isolamento social e restrição de circulação por três dias em decorrência do aumento de casos de COVID-19 (Governo de Santa Catarina, 2021).

Em setembro de 2021, a Diretoria de Vigilância Epidemiológica (DIVE) emitiu uma nota alertando para aumento de casos de doenças diarreicas agudas no estado nos meses anteriores. Estas doenças têm como principal causa a ingestão de água contaminada, e pode ter como agentes causadores virais, norovírus, adenovírus e rotavírus (Diretoria de Vigilância Epidemiológica, 2021). Apesar das detecções de HAdV nos meses finais de 2021, nenhum registro de surto foi registrado no município.

Nos primeiros meses do ano de 2021, a maior carga viral de SARS-CoV-2 no esgoto pode estar relacionada a uma maior liberação de partículas virais acarretado pelo maior número de pessoas infectadas pelo vírus. É relatado que, apesar de um vírus respiratório, o SARS-CoV-2 possui capacidade de replicação nos enterócitos acarretando em liberação de partículas virais nas fezes de pessoas infectadas durante meses após os sintomas respiratórios cessarem (Lamers et al., 2020). Diferentemente do que aconteceu com HAdV, a menor circulação de pessoas não refletiu em menor detecção, justamente devido a excepcionalidade do período pandêmico e alto número de pessoas infectadas.

Em 2021 Florianópolis apresentou 55.507 casos confirmados de COVID-19, com 740 óbitos. Nos meses em que foi constatada a presença de SARS-CoV-2 em esgoto: janeiro, fevereiro e abril, foram notificados 9.442, 10.944 e 3.423 casos, respectivamente. Especificamente, no centro de saúde de Canasvieiras foram relatados apenas sete casos positivos de janeiro a abril (Prefeitura de Florianópolis, 2023a). No início de 2021, prevalecia

a variante Gama no estado, que, em conjunto com a variante Delta que a sucedeu obtiveram maiores taxas de mortalidade em Santa Catarina, em um cenário inicial de implementação de vacinas contra SARS-CoV-2 (Padilha *et al.*, 2022). A vacinação em massa no Brasil começou em 17 de janeiro de 2021, fator que contribuiu para menor disseminação do vírus, redução de internações e mortalidade, e, conseqüentemente diminuição da carga viral no esgoto, diretamente relacionada ao número de pessoas infectadas (Santos *et al.*, 2023). Florianópolis iniciou a vacinação em 20 de janeiro e ao final de 2021 possuía mais de 482 mil pessoas vacinadas com a primeira dose, 442 mil com a segunda dose e 112 mil vacinadas com a dose de reforço (Prefeitura de Florianópolis, 2023b).

A detecção da presença de SARS-CoV-2 em esgoto foi relatada pela primeira vez na Austrália, como conceito para monitoramento do vírus em escala populacional (Ahmed *et al.*, 2020) uma vez que a detecção individual saturou pela alta demanda de testes. Diversos grupos no Brasil utilizaram da metodologia de detecção por RT-qPCR para atestar a presença do vírus e inferir prevalência de infectados de uma determinada rede abastecida pela coleta de esgoto (Fongaro, Rogovski, *et al.*, 2021; Freitas Bueno, *et al.*, 2022; Prado *et al.*, 2021). Em Florianópolis, Fongaro e colaboradores (2021) identificaram a presença do vírus no esgoto três meses antes do primeiro caso oficial no município, datado de março de 2020, a partir de uma amostra do centro da cidade, pela mesma metodologia de concentração viral empregada neste trabalho. Estes estudos, em consonância com os resultados obtidos expressam a eficácia da vigilância molecular em esgoto e reforçam sua importância no acompanhamento de surtos e identificação da presença de patógenos como fator para tomada de decisão por parte do poder público.

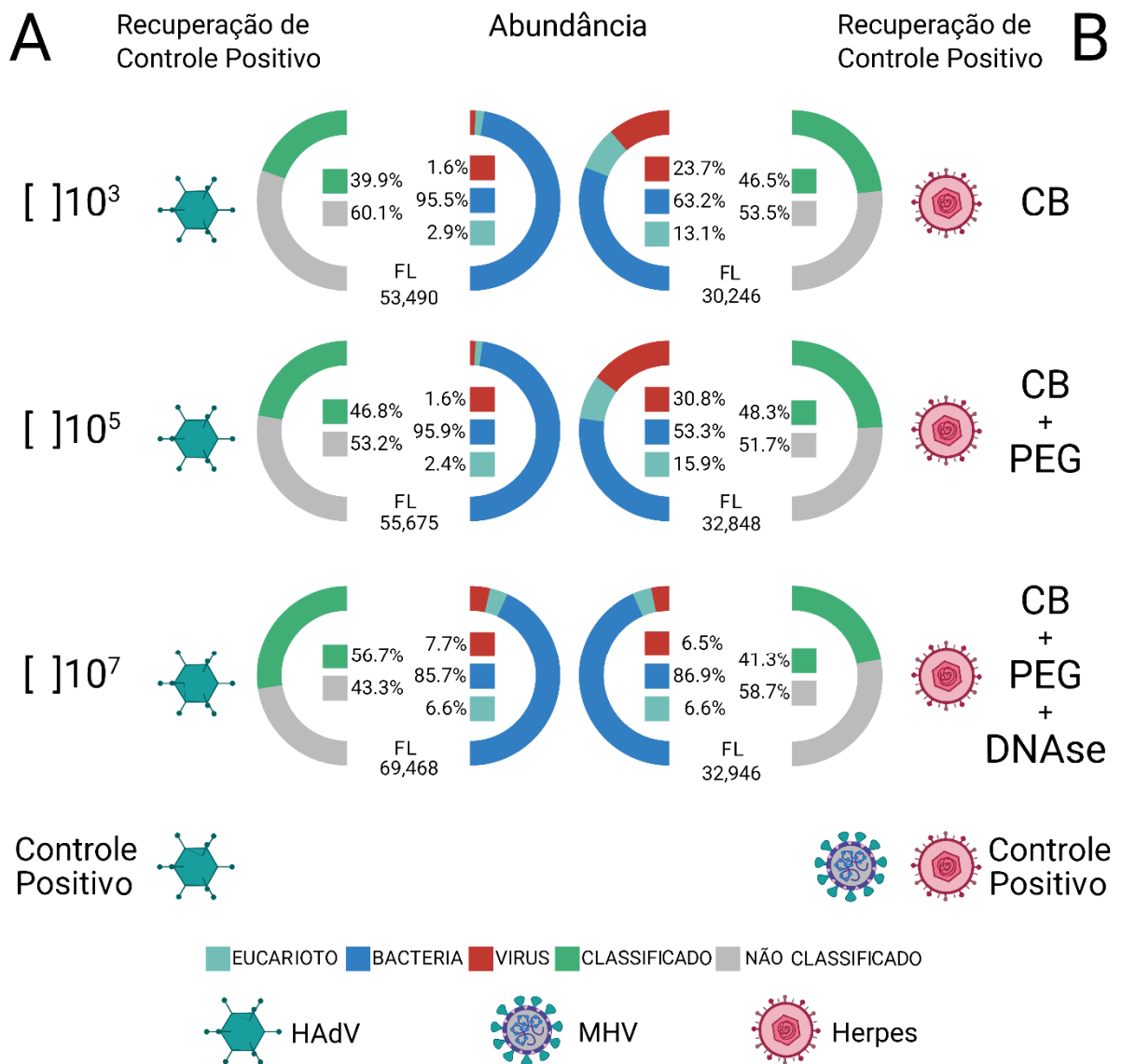
A natureza da amostra, o esgoto bruto não tratado, não permite afirmar a eficiência da ETE Canasvieiras na inativação viral, uma vez que a análise antes e depois do tratamento do efluente não foi realizado e foge do escopo do estudo.

4.2 Recuperação viral a partir do sequenciamento *shotgun* por nanoporos

Os resultados promissores obtidos na pandemia de COVID-19 com a vigilância de SARS-CoV-2 lançaram luz acerca da importância do conhecimento e monitoramento dos subtipos virais de agentes patogênicos (Jesus *et al.*, 2020). Neste sentido, a adaptação da metodologia empregada no diagnóstico RT-qPCR de HAV, HAdV e SARS-CoV-2 no esgoto de Canasvieiras foi testada para sequenciamento. O primeiro experimento de sequenciamento testou diferentes concentrações de vírus controle objetivando identificar a concentração viral mínima para sequenciamento de vírus em esgoto por meio de *shotgun* em plataforma MinION Mk1C. O sequenciamento realizado com três concentrações de cinco vírus controles diferentes, em amostra recente de esgoto (2023) não mostrou diferença significativa nas proporções de fragmentos de leitura obtidas entre as concentrações. A Figura 7 – item A expressa os dados apresentados no Anexo B – Tabela 13. Dos cinco controles utilizados - MHV-3, RVA, HAdV-2, MNV-1 e HSV-2 – Apenas Adenovírus foi recuperado nas duplicatas amostrais. Considerando o histórico de detecção previamente relatado, a identificação de HAdV-2 pode mesclar vírions presentes naturalmente na amostra com o controle viral adicionado. Os controles positivos apresentaram contaminação por material genético bacteriano, de forma que um comparativo entre eles e as amostras de esgoto não pode ser realizado.

Foram identificados fragmentos de leitura de adenovírus 1 e 5 nas amostras controle 10^7 , esgoto 1 10^7 e esgoto 2 10^5 . *Human mastadenovirus C* foi identificado como mais abundante em duas amostras - *esgoto 1 10^7* e *esgoto 2 10^5* . O controle viral utilizado, HAdV-2 está inserido no táxon *mastadenovirus C* – Tabela 6. A indeterminação do sorotipo não garante a recuperação do controle viral adicionado, apesar de abundante em dois tratamentos. Dois fagos *Mu* com hospedeiros de *Shigella* e *Escherichia coli*, também foram identificados em todas as amostras. As demais identificações se reduzem a poucos fragmentos de leitura. Apesar de duplicatas, as amostras não apresentam o mesmo padrão em quantidade de fragmentos de leitura virais sequenciados. Isto fica evidente nas amostras esgoto 1 e 2 de concentração 10^7 em que foram identificados 3.558 e 324 fragmentos de leitura virais, respectivamente.

Figura 7. Comparativo de recuperação de controle viral nos experimentos de definição do protocolo de sequenciamento shotgun por nanoporos



Legenda: Resultados de classificação e abundância relativa nos táxons bacterianos, virais e eucariotos. FL – Corresponde ao total de fragmentos de leitura. As quantificações estão expressas em proporção aos fragmentos de leitura classificados. Classificações como “*root*” e “*cellular organism*” não foram consideradas. Em (A) os dados do sequenciamento com diferentes concentrações virais e em (B) o experimento associando PEG a protocolos de clarificação por Clorofórmio Butanol (CB) e tratamento com DNase. A recuperação viral dos controles não focou nos subtipos virais, apenas a presença do táxon.

Tabela 6. Número de fragmentos de leitura mapeados em cada genoma viral - dados do primeiro experimento.

Vírus	Controle [10⁷]	Controle [10⁵]	Controle [10³]	Esgoto 1 [10⁷]	Esgoto 2 [10⁷]	Esgoto 1 [10⁵]	Esgoto 2 [10⁵]	Esgoto 1 [10³]	Esgoto 2 [10³]
Human mastadenovirus C	2126	13	2	2986	58	2	2183	37	1
<i>Shigella</i> phage SfMu	355	39	66	189	112	351	196	420	166
<i>Escherichia</i> phage Mu	212	75	60	296	146	69	34	56	46
Human adenovirus 1	40	-	-	45	1	-	44	2	-
Human adenovirus 5	35	-	-	31	1	-	27	-	-
Muvirus	3	-	-	-	-	-	-	-	-
Mastadenovirus	1	-	-	-	-	-	-	-	-
Moloney murine leukemia virus	1	-	-	-	-	-	-	-	-
<i>Escherichia</i> phage D108	-	-	-	2	1	1	-	2	1
uncultured crAssphage	-	-	-	2	3	-	-	1	-
Human mastadenovirus D	-	-	-	1	-	-	-	-	-
Titi monkey adenovirus ECC-2011	-	-	-	1	-	-	-	-	-
<i>Mus musculus</i> mobilized endogenous polytropic provirus	-	-	-	1	-	-	1	-	-
<i>Klebsiella</i> virus KP32	-	-	-	-	1	-	-	-	-
Kagunavirus	-	-	-	-	1	1	-	-	-
<i>Mycoplasma</i> phage phiMFV1	-	-	-	-	-	1	-	-	-
Pandoravirus salinus	-	-	-	-	-	-	-	1	-
Yersinia phage vB_YenP_AP10	-	-	-	-	-	-	-	4	-
Teetrevirus	-	-	-	-	-	-	-	1	-
<i>Aeromonas</i> phage 4_L372XY	-	-	-	-	-	-	-	-	1
<i>Vírus indeterminado</i>	2	-	-	4	-	-	1	-	-
Total	2775	127	128	3558	324	425	2486	524	215

Legenda: [x] – Log de concentração em cópias genômicas

Observamos que apesar de quatro *logs* de diferença entre a concentração mais baixa (10^3) e mais alta (10^7) de vírus controles, isso não se refletiu em mais fragmentos de leitura nas duplicatas mais concentradas. Uma das possíveis explicações para este fenômeno é a degradação do material genético, bem como a supressão dos mesmos pela quantidade de DNA bacteriano nas amostras, que se sobressai em relação ao material genético viral, como observado nos experimentos de recuperação. Neste experimento, obtivemos problemas com quatro dos cinco controles virais, os quais nenhum fragmento de leitura associado foi sequenciado, possivelmente relacionado a degradação do material genético viral quando adicionado a amostra de esgoto.

A alíquota de Adenovírus, que possuía o menor log de cópias gênicas por mL foi a única recuperada, demonstrando que mesmo em baixas concentrações, este vírus pode resistir à ação de agentes degradantes.

O adenovírus é detectável no esgoto em caso de excreção por parte de pacientes infectados - como apresentado na Tabela 5. Desta forma, as abundâncias relatadas podem mesclar a presença do vírus selvagem com o controle adicionado. Os resultados não demonstraram um padrão de recuperação relacionado a concentração de modo que a transposição da metodologia e rotina laboratorial executada na identificação de vírus por metodologias moleculares foi descartada.

Diante disso e levando em conta que o esgoto é um ambiente repleto de matéria orgânica que reflete em inibidores de reação e degradação do material genético, a realização de um procedimento para clarificação da amostra, por meio do clorofórmio butanol associado ao PEG foi empregado. Neste segundo experimento, também foi avaliada a aplicação da enzima DNase como agente ativo na redução da proporção de material genético bacteriano, executado em associação ao protocolo de concentração e clarificação.

Como resultados deste segundo experimento (B), obtivemos proporções de fragmentos de leitura virais superiores, comparativamente ao primeiro experimento (A) – Figura 7 e Anexo B Tabela 14. O controle positivo novamente apresentou contaminação de material genético bacteriano, mas com a recuperação dos dois controles. Os tratamentos não apresentaram fragmentos de leitura do controle viral com material genético de RNA – MHV-3.

O controle viral HSV-2 foi recuperado e classificado em *Alphaherpesvirus 1* e 2, de forma que o 1 apresentou a maior quantidade de fragmentos sequenciados, seguido de fagos *Mu* (Tabela 7). Uma vez que o vírus da HSV-2 depende de contato via mucosa para transmissão além de não possuir ciclo em enterócitos, de forma que a detecção de vírus selvagem em esgoto

é descartada (Knipe e Howley, 2013). Assim, justifica-se que a classificação deste vírus nos subtipos 1 e 2 caracteriza um erro de anotação devido à similaridade dos mesmos.

Proporcionalmente, o tratamento com CB e, PEG + CB recuperaram mais fragmentos de leitura associados a vírus, com 23,7% e 30,8% do total de leituras, respectivamente. O tratamento combinado de PEG + CB + DNase recuperou cerca de 6% de fragmentos de leitura virais, inferior a abundância viral encontrada nos outros tratamentos. Interessantemente, o tratamento utilizando apenas CB recuperou mais controles virais do que o PEG + CB + DNase, indicando que possivelmente o processo de concentração pode expor o material genético viral, pela destruição do capsídeo, de forma que a ação da DNase reduziu a proporção de fragmentos não apenas bacterianos, mas também virais. Este resultado ressalta a importância da manutenção da partícula viral para utilização de protocolos de redução de DNA bacteriano por DNase.

É esperada a diminuição de material genético, uma vez que uma enzima de degradação foi utilizada, no entanto, genomas bacterianos maiores e células lisadas sofreriam maior ação da enzima em um cenário em que as partículas virais apresentassem capsídeo íntegro. Apesar da grande maioria dos genomas bacterianos serem expressos em milhares de bases, sendo poucos os que atingem milhões de bases (Ochman; Caro-Quintero, 2016), dentre estes estão bactérias abundantes no esgoto, como *E. coli* (~5Mb) e *Shigella* spp (~5Mb) e *Klebsiella* spp. (~6Mb) – NCBI Genome. Desta forma, esperava-se que a ação da enzima favorecesse cópias gênicas em menor número no ambiente, reduzindo a abundância destes organismos com tamanho de genoma maior. Neste experimento isso não foi observado, embora este princípio de redução do contaminante de forma ativa, visando enriquecer a amostra com o organismo alvo possa ser uma estratégia em futuros estudos. Em sequenciamentos de amostras humanas, a contaminação do hospedeiro é reduzida pelo emprego da DNase e o material genético do organismo alvo é preservado uma vez que este não esteja lisado (Diao *et al.*, 2022). No caso dos vírus em esgoto, o ambiente extremo faz com que a quantidade de vírions viáveis seja reduzida, com exceção a bacteriófagos que realizam seu ciclo replicativo naturalmente neste ambiente, sendo indicadores de contaminação fecal (McMinn, Ashbolt e Korajkic, 2017).

Os resultados de sequenciamento envolvendo a adaptação da descontaminação associada ao PEG mostraram maior proporção de sequenciamento de fragmentos de leitura virais. A associação do método de concentração viral com a remoção da fração lipídica da amostra pode explicar a melhor recuperação viral, tendo em vista a redução de inibidores que influenciam nas reações de PCR para síntese de cDNA e posterior sequenciamento genético (Oikarinen *et al.*, 2009).

Considerando o experimento B (Figura 7) como um todo, a proporção de fragmentos de leitura virais foi maior em todos os tratamentos, quando comparado ao experimento A (Figura 7). É importante ponderar que este resultado foi obtido com concentrações muito altas de vírus controles, uma vez que foram utilizados 10^{10} GC/mL. Apesar da alta concentração, o sequenciamento de fragmentos de leitura do controle de RNA foi limitado, mostrando, portanto, o desafio do sequenciamento de vírus de RNA por meio de *shotgun*. O controle positivo foi o único a apresentar fragmentos de leitura classificados como MHV-3 com 24 leituras. Este resultado pode ser multifatorial, desde a degradação destes vírus quando inoculados na amostra, uma vez que o RNA é mais instável que o material genético de DNA, até o protocolo de síntese de cDNA ineficiente. Em altas concentrações os protocolos de descontaminação recuperaram o vírus controle de DNA para classificação taxonômica, indicando melhores resultados para sequenciamento de amostra ambiental.

Tabela 7. Número de fragmentos de leitura virais classificados - segundo experimento.

Vírus	Controle positivo	CB	PEG + CB	PEG + CB + DNase
Human alphaherpesvirus 1	11427	3180	4619	823
<i>Escherichia</i> phage Mu	133	61	123	32
<i>Shigella</i> phage SfMu	112	40	99	12
Simplexvirus	52	39	17	4
Human alphaherpesvirus 2	44	15	20	5
Murine hepatitis virus	24	-	-	-
<i>Mycoplasma</i> phage phiMFV1	13	3	8	2
Rat coronavirus Parker	10	-	-	-
Chimpanzee herpesvirus strain 105640	7	3	3	1
Gammaretrovirus	2	-	-	-
<i>Mus musculus</i> mobilized endogenous polytropic provirus	2	-	2	-
Leporid alphaherpesvirus 4	1	-	-	-
<i>Escherichia</i> phage D108	1	-	-	-
Propionibacterium phage Wizzo	1	-	-	-
Macaca mulatta polyomavirus 1	1	-	-	-
Shamonda orthobunyavirus	1	-	-	-
Mouse mammary tumor virus	1	-	-	-
unclassified Gammaretrovirus	1	-	-	-
Spumaretrovirinae	1	-	-	-
Rat coronavirus Parker	-	1	-	-
Human coronavirus OC43	-	-	-	1
Sextaecvirus	-	-	1	-
Total	11.834	3.342	4.892	880

Legenda: PEG – Polietileno Glicol; CB – Clorofórmio Butanol.

4.3 Sequenciamento das amostras de esgoto de Canasvieiras do ano de 2021

A partir dos resultados obtidos, o método de PEG + CB foi escolhido para aplicação nas amostras de esgoto de Canasvieiras do período de 2021. O protocolo com melhor recuperação viral nos experimentos realizados utilizando altas concentrações de vírus não refletiu este resultado nas amostras contendo apenas vírus selvagens. A proporção de fragmentos de leitura virais obteve o maior resultado na amostra do mês de junho (3,19%) e menor na amostra de março (0,53% com 26 fragmentos de leitura virais identificados (Apêndice B - Tabela 15).

Os vírus identificados por metodologia molecular (Tabela 5), não foram recuperados no sequenciamento das amostras (Tabela 8). O armazenamento e o tempo entre a detecção por metodologia molecular (realizado no ano da coleta) e o sequenciamento, neste estudo em 2023, podem influenciar na recuperação viral, uma vez que mesmo armazenado em -20 a degradação do material genético pode ocorrer. Uma vez que não foram realizados processos de amplificação, salvo a síntese de cDNA, este fator se intensifica. Além desta problemática, a não recuperação dos vírus previamente detectados sinaliza para a falta de sensibilidade do protocolo aplicado para sequenciamento.

Foram identificados os fagos *Shigella phage* SfMu e *Escherichia phage* Mu. Tendo em vista a abundância de fragmentos de leitura dos demais vírus classificados, não é prudente afirmar a sua correta anotação taxonômica, bem como a presença destes vírus nas amostras.

O sequenciamento de viroma total é um desafio, poucos estudos obtiveram um perfil da comunidade de vírus de um ambiente. O primeiro relato de sequenciamento de viroma total a partir de amostras fecais, recuperou bacteriófagos e elementos repetitivos e móveis, com cerca de 56% de sequências desconhecidas (Breitbart *et al.*, 2003). O estudo maior realizado com amostras de esgoto para obtenção de viroma total, analisou 62 cidades, incluindo as cidades brasileiras de Belém e Belo Horizonte, e obteve em média uma recuperação de fragmentos de leitura virais de 6% com sequenciamento Illumina (Nieuwenhuijse *et al.*, 2020). Neste trabalho, os vírus mais comuns foram bacteriófagos e vírus relacionados a plantas e insetos. A diferença da metodologia de sequenciamento, com ciclos de amplificação randômica que permite gerar maior volume de dados e aumentar as chances de detecção de vírus pouco abundantes. Nieuwenhuijse e colaboradores (2020) relataram presença de vírus de interesse em saúde pública como norovírus, astrovírus e rotavírus, que permitiram identificar sazonalidades de infecções. Cantalupo e colaboradores detectaram fragmentos de leitura de 17 vírus

patogênicas a humanos em amostras de esgoto bruto de Barcelona, Pittsburgh e Addis Ababa, incluindo papilomavírus, poliomavírus, adenovírus 41 e astrovírus (Cantalupo *et al.*, 2011).

O sequenciamento por nanoporos gera menor volume de dados, pelas características de construção de biblioteca de sequenciamento em detrimento de um fragmento de leitura maior. A utilização da plataforma de fragmentos de leitura longos não surtiu efeito no tamanho dos mesmos. A degradação do material genético no ambiente de esgoto impactou no tamanho de fragmento médio que foi de 229 a 331 pares de base considerando os três experimentos (Apêndice B – Tabela 16).

Em um desenho experimental semelhante ao realizado neste projeto, Aw e colaboradores (2014) recuperaram genomas virais a partir de águas residuais de Michigan – Estados Unidos. A metodologia de concentração por PEG, descontaminação por clorofórmio, inclusão de uma etapa de filtração por membrana e uso de tratamento com DNase, foram empregadas em conjunto do sequenciamento Illumina no trabalho de Aw. Os genomas recuperados por este grupo de pesquisa incluem subtipos B, C e F de adenovírus, poliomavírus JC, além de bacteriófagos de mais de 70 hospedeiros (Aw, Howe e Rose, 2014). Desta forma, apesar dos resultados obtidos não recuperarem os resultados da metodologia molecular, há indícios que a metodologia de concentração, que é eficiente para detecção molecular, pode ser empregada para este objetivo, de forma que mais estudos metodológicos são necessários.

O grande número de fragmentos de leituras não classificados é comum na caracterização de comunidades virais e esbarra no conhecimento que está depositado nas bases de dados utilizadas na classificação viral. O *core* de genomas virais depositados nas bases de dados de referência são essencialmente vírus de interesse em saúde, que acarreta em um forte viés de classificação que não reflete a diversidade de vírus nos ambientes (Krishnamurthy e Wang, 2017). Cerca de 65%, em média, dos fragmentos de leitura obtidos do esgoto de Canasvieiras em 2021 não foram classificados, e possivelmente incluindo vírus que estão na nossa fronteira de conhecimento. A fragmentação é um fator que dificulta a classificação (Portik, Brown e Pierce-Ward, 2022) ao mesmo tempo que a literatura descreve porcentagens de fragmentos indeterminados em análises de viroma total de 40 a 90% de acordo com o tipo de amostra (Minot *et al.*, 2013; Ng *et al.*, 2012). Krishnamurthy e Wang chamaram este fenômeno de matéria escura viral, um artefato da maneira como os vírus são identificados, que faz com que vírus desconhecidos passem despercebidos e caiam no limbo de fragmentos indeterminados.

Tabela 8. Número de fragmentos de leitura mapeados em cada genoma viral - dados das amostras de Canasvieiras no ano de 2021.

Vírus	Janeiro	Fevereiro	Março	Abril	Mai	Junho	Agosto	Outubro	Novembro	Dezembro
<i>Shigella</i> phage SfMu	368	60	16	151	351	1020	265	246	112	25
<i>Escherichia</i> phage Mu	319	31	10	127	114	183	49	66	50	7
Pepper chlorotic spot virus	-	-	-	1	-	1	-	-	-	-
<i>Escherichia</i> phage D108	1	-	-	-	-	-	-	-	1	-
Muvirus	-	-	-	-	-	-	-	-	1	-
<i>Achromobacter</i> phage Motura	-	-	-	-	-	-	1	-	-	-
<i>Propionibacterium</i> phage PAC1	-	-	-	-	-	-	-	1	-	-
Human alphaherpesvirus 1	-	-	-	1	-	-	-	-	-	-
<i>Stenotrophomonas</i> phage SMA7	1	-	-	-	-	-	-	-	-	-
<i>Acanthamoeba</i> polyphaga mimivirus	-	-	-	-	-	1	-	-	-	-
Total	689	91	26	280	465	1205	315	313	164	32

Legenda: As contagens representam valores absolutos de fragmentos de leitura associados a cada táxon, em relação a cada mês.

Diante da limitação na obtenção de fragmentos de leitura virais no sequenciamento *shotgun* por nanoporos, alternativas de metodologias de concentração e descontaminação, bem como redução de material genético bacteriano devem ser consideradas para obter o viroma total. Como relação ao protocolo de concentração PEG, fica evidente pelos resultados do projeto VigEAI que a metodologia de concentração por PEG é eficiente para detecção por qPCR e RT-qPCR e permite monitoramento molecular para diagnóstico de presença e ausência de infecções virais.

O uso do PEG como metodologia de concentração para direcionamento de sequenciamento de vírus é descrito na literatura. Como alternativas, concentração por floculação orgânica com leite desnatado, filtração de adsorção monolítica (Hjelmsø *et al.*, 2017; Salvo *et al.*, 2021). Destes, o PEG apresentou maior recuperação de fragmentos de leitura associados as famílias *Adenoviridae*, e *Reoviridae* (famílias dos vírus HAdV e Rotavírus) em estudo de Hjelmsø e colaboradores (2017) por sequenciamento Illumina. O protocolo de filtração por membrana (Katayama) também é utilizado, e é passível de teste e implementação em pesquisas futuras objetivando um protocolo de baixo custo. Apesar disso, é importante ressaltar que a filtração por membrana possui suas limitações, uma delas é a retenção de partículas em suspensão, as quais os vírus se aderem por atração elétrica. Outro fator limitante é quantidade de partículas em suspensão, que pode acarretar na obstrução do filtro. Deste modo, este protocolo é melhor empregado em amostras de águas residuais, com menor quantidade de partículas em suspensão.

O volume de partículas em suspensão, em esgoto bruto, acarreta em retenção das partículas e entupimento do filtro, dificultando o uso do protocolo de filtração por membrana para este tipo de amostra. A utilização das metodologias combinadas pode aumentar a recuperação viral para sequenciamento. Uma perspectiva é a análise de um protocolo que combine a descontaminação por CB, a concentração por PEG e a filtração por membrana. Neste sentido, considerando que os vírus ficam adsorvidos nas partículas sólidas, a inclusão de variações de pH no protocolo PEG permite a liberação de maior número de vírions aderidos as partículas, uma vez que seus pontos isoelétricos diferem (Michen e Graule, 2010). A realização do PEG de forma paralela com variações de pH aumenta o volume de processamento de uma amostra, que pode ser imputada na membrana.

Neste cenário, a membrana retém bactérias, restos celulares e partículas sólidas menores, permitindo que os vírions sejam isolados. Importante mencionar que a recuperação de vírus selvagens, detectados por metodologia molecular deve ser mensurada como forma de avaliação da metodologia proposta. Isto, para evitar o viés dos controles, observado nos

experimentos realizados neste trabalho, cujos controles forma recuperados, mesmo que parcialmente, diferente dos vírus selvagens detectados que não apresentaram fragmentos de leitura associados no sequenciamento.

O protocolo de Ultracentrifugação e ultrafiltração não podem ser considerados dentro do escopo de um protocolo acessível, apesar de eficácia descrita na em ensaios para recuperação de fragmentos de leitura de SARS-CoV-2 em esgoto por sequenciamento Illumina (Jafferli *et al.*, 2021; Salvo *et al.*, 2021).

Analisando a metodologia de sequenciamento, o emprego de *shotgun* depende da carga viral e sua relação de abundância com materiais genéticos de outra natureza presentes no esgoto. Desta forma, uma perspectiva para aprimoramento na detecção é o emprego de técnicas de enriquecimento, com amplificação e direcionamento para genomas virais de interesse. Esta metodologia foi amplamente empregada no monitoramento de SARS-CoV-2, de forma que a prova de conceito está validade. O enriquecimento por captura de material genético viral por sondas, tal como ViroCap-seq (Briese *et al.*, 2015), é uma tecnologia recente, com poucos trabalhos relatados, com resultados de enriquecimento promissores. De fato, Herath e colaboradores relatam uma recuperação viral de mais de 55,12% com a utilização da sonda ViroCap-seq contra 1,17% sem a mesma para sequenciamento *shotgun*, ambos em plataforma de sequenciamento por nanoporos para amostra clínica (Herath *et al.*, 2017).

Considerando a realidade Brasileira e, reforçando a busca por um protocolo baixo custo, o enriquecimento por sondas não é uma alternativa economicamente viável para rotina de monitoramento em esgoto. A abordagem de *amplicons*, apesar de adicionar uma etapa e custos, já foi empregada com sucesso no monitoramento de SARS-CoV-2, e pode ser extrapolada para vírus entéricos. Para isso, é necessário desenvolver e padronizar protocolo de amplificação de genoma completo de vírus entéricos, mapear regiões conservadas e testar a viabilidade de amplificação frente a diferentes variantes virais. Como exemplo de sucesso, o monitoramento de norovírus para e identificação de genótipos circulantes é realizado na China por meio de *amplicons* específicos de cada genótipo e sequenciamento direcionado a partir de amostras de esgoto (Lu, Jing *et al.*, 2021).

4.4 Sequenciamento direcionado para SARS-CoV-2

Uma vez que a detecção de vírus selvagens foi restrita a apenas fagos, a adaptação do protocolo executado para detecção por metodologia molecular requer estudo e desenvolvimento para sequenciamento de viroma total. A alternativa, o viroma direcionado foi realizado para seis amostras de quatro cidades positivas para SARS-CoV-2, no âmbito do projeto VigEAI, utilizando tecnologia de sequenciamento Illumina. Os dados de profundidade e cobertura média estão apresentados na Tabela 9. A quantificação de carga viral para L^{-1} não está diretamente relacionada com a recuperação do genoma em análise de bioinformática. A variação de fragmentos de leitura mapeados no genoma do patógeno dentro do mesmo escopo de carga viral possui uma amplitude grande, indicando que apesar de melhor recuperação esta abordagem ainda apresenta desafios. Um deles é a natureza da amostra, o sequenciamento de SARS-CoV-2 a partir de esgoto foi realizado por diversos grupos de pesquisa no mundo inteiro durante a pandemia de COVID-19. Além da baixa cobertura, exposta na Figura 8, as diferenças de profundidade nos genomas obtidos a partir desta metodologia foram observadas.

Não foram investigadas condições ambientais do período de coleta das amostras, que certamente possuem influência nesta recuperação, como o índice de pluviosidade do período. Outro fator importante é a diversidade de variantes em uma amostra populacional, de forma que os vMAGs obtidos representam toda a diversidade viral circulante do período. Por consequente, é natural que os ácidos nucleicos de variantes apresentem graus de afinidade diferenciais em relação ao conjunto de iniciadores utilizado, reduzindo a eficácia da técnica de amplificação e em parte explica os resultados obtidos.

Ressalta-se que devido a emergência de saúde pública, os iniciadores para uma reação PCR *multiplex* para produção de *amplicons* para montagem do genoma completo de SARS-CoV-2 foram desenvolvidos rapidamente e com constantes atualizações dos iniciadores (Eden *et al.*, 2020). Da perspectiva dos vírus entéricos, muito menos estudados, a realização do viroma direcionado para sequenciamento carecem da disponibilidade destes iniciadores, sendo, portanto, um desafio. Ademais, genomas maiores podem impossibilitar a realização desta técnica para montagem de genomas completos a partir da realização de amplificação por PCR em multiplex. Uma alternativa é a amplificação da região da proteína de superfície, que interage com o hospedeiro e é comumente utilizada na atribuição taxonômica de linhagem, sorotipo e variante, como utilizado, inclusive para a detecção de variantes de SARS-CoV-2 RT-qPCR em uma técnica denominada

Painéis virais são empregados para diagnóstico de presença/ausência dos vírus com alvos específicos. O desenvolvimento de um painel de iniciadores para genotipagem de vírus entéricos é uma perspectiva para desenvolvimento de viroma direcionado. No contexto das infecções por SARS-CoV-2 metodologias de inferência molecular foi empregada para vigilância molecular com iniciadores específicos capazes de diferenciar Variantes de preocupação (Tsui *et al.*, 2022). Esta metodologia é uma alternativa de baixo custo ao sequenciamento, tendo a capacidade limitada de e a emergência de determinação de variantes durante os picos de casos.

Tabela 9. Resultados de viroma direcionado para SARS-CoV-2 a partir de amostras do projeto VigEAI de 2021.

Amostra	Cidade	Mês	Total de FL Brutos	Total de FL Após QC	Total de FL mapeados no genoma de SARS-CoV-2	Cobertura do genoma montado (%)	Carga Viral em L ⁻¹ de esgoto
CON_09_1	Concórdia	Setembro	87.492	43.990	21	1,3	1.000
ITA_09_2	Itajaí	Setembro	119.810	65.114	1.157	6,28	10.000
JOI_10_1	Joinville	Outubro	108.236	53.672	12	1,2	10.000
ITA_11_4	Itajaí	Novembro	108.868	60.127	30	0,1	10.000
JOI_12_2	Joinville	Dezembro	106.772	53.477	37.058	57,1	10.000
FLO_12_5	Florianópolis	Dezembro	128.641	69.843	0	0	1.000

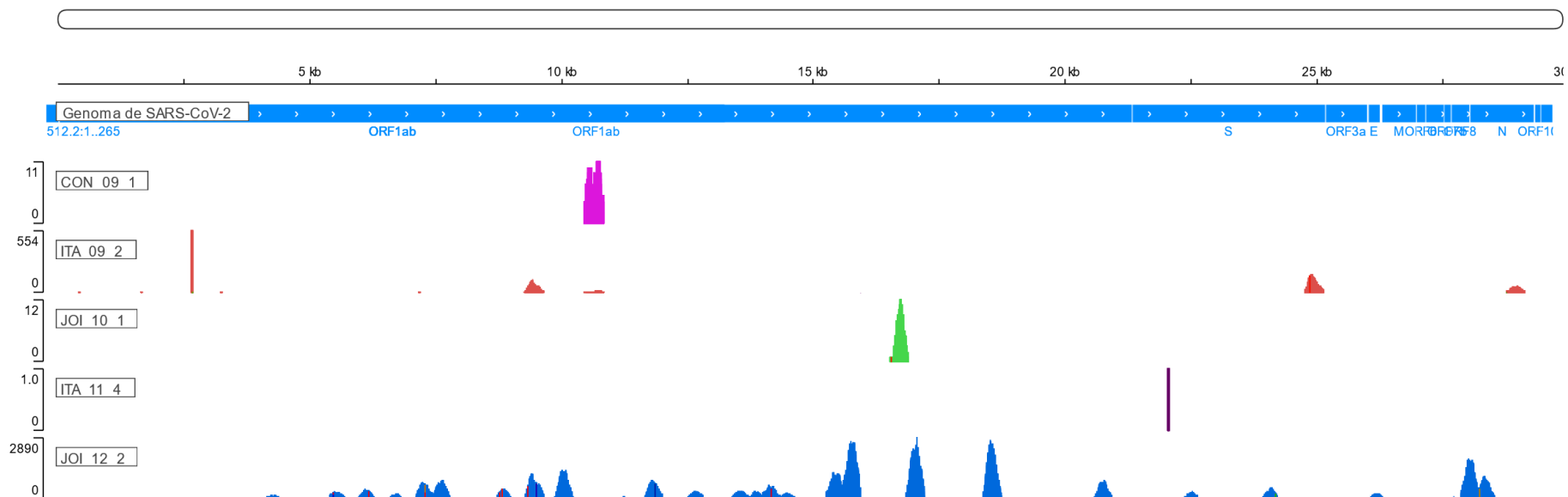
Legenda: Amostras de esgoto obtidas a partir do projeto VigEAI. A quantificação de cópias genômicas foi realizado a partir de quantificação por RT-qPCR. FL – Fragmentos de Leitura; QC – Controle de qualidade.

Fongaro e colaboradores (2021) na primeira detecção de SARS-CoV-2 a partir de amostras de esgoto de novembro de 2019, originárias de Florianópolis apresentaram o sequenciamento de mesma natureza com cobertura de 25% e 95 fragmentos de leitura alinhados no genoma de referência de SARS-CoV-2. Dados do GISAID (do inglês, *Global Initiative on Sharing All Influenza Data*), o repositório mundialmente utilizado para depósito de sequências do vírus causador da pandemia de COVID-19, totaliza 15.892.038 genomas depositados. Destes, 12.741 foram obtidos a partir de amostras ambientais, cerca de 0,08 % dos genomas totais, 22 destes de origem no Brasil (dados <https://gisaid.org/>, Acesso em 22/08/2023). Isto mostra a importância da busca por otimizações e alternativas metodológicas. Apesar de promissora, a vigilância genômica em esgoto ainda caminha para implementação em larga escala, mesmo no vírus de com maior volume de genomas que possuímos (Bonanno Ferraro *et al.*, 2022). O monitoramento realizado até então é focado em estações de tratamento, com cerca

de 57 estudos com genomas de SARS-CoV-2 sequenciados a partir de amostras de esgoto; 8 trabalhos de origem hospitalar e outras origens incluem dormitórios de profissionais de saúde e centro de isolamento para doentes acometidos por COVID-19. Dentre os objetivos destes monitoramentos, ressalta-se a vigilância como sistema precoce de detecção de casos, investigação da diversidade genética do vírus e estimativa da prevalência de casos de COVID-19 na população a partir da carga viral em esgoto (Bonanno Ferraro *et al.*, 2022).

Com a redução de casos graves e o sucesso de vacinação, a vigilância genômica em esgoto é uma alternativa, juntamente com o monitoramento de profissionais de saúde (Padilha *et al.*, 2023) para detecção de linhagens emergentes e estabelecimento de um sistema de alarme precoce. A identificação de novas linhagens circulantes é importante para prever escape imunológico e aumento de número de casos (Harvey *et al.*, 2021; Souho *et al.*, 2022) e seu desenvolvimento para detecção de vírus de interesse em saúde pública podem ajudar na identificação e monitoramento de patógenos.

Figura 8. Cobertura média dos genomas obtidos a partir do viroma direcionado para SARS-CoV-2.



Legenda: Estão representadas as quatro amostras com fragmentos de leitura mapeados no genoma de referência de SARS-CoV-2. No topo o genoma de referência do vírus, com anotação das regiões codificadoras. Abaixo, cada pico representa uma profundidade de sequenciamento obtido na sequência consenso. As escalas de profundidade (a esquerda) variam de acordo com a amostra. Figura gerada a partir do IGV (<https://igv.org/>).

5 CONCLUSÕES

A combinação da metodologia de concentração PEG com pós protocolo de clarificação por CB demonstrou maior recuperação de fragmentos de leitura virais no sequenciamento de vírus controles por *shotgun* em plataforma nanoporos a partir de amostras de esgoto.

A metodologia de concentração PEG associada ao CB não foi eficiente para detecção de vírus selvagens por sequenciamento *shotgun* em plataforma de nanoporos uma vez que não reproduziu resultados positivos obtidos pela detecção por meio de métodos moleculares. Ressalta-se que, para a metodologia molecular, a concentração por PEG é eficiente e financeiramente viável para aplicação em países com condições socioeconômicas desfavoráveis e sua aplicabilidade para obtenção de sequências virais precisa de otimização e estudo.

O viroma total de Canasvieiras, obtido por meio do sequenciamento *shotgun* em plataforma nanoporos, recuperou apenas fagos, o que demonstra a necessidade do desenvolvimento de metodologias de concentração/redução de contaminante bacteriano para aumento de sensibilidade da técnica. Ainda, o viroma direcionado para SARS-CoV-2, a partir de amostras de esgoto obtidas por sequenciamento Illumina, de não obteve cobertura e profundidade suficientes para recuperação dos genomas completos. Este resultado demonstra o desafio de recuperação de fragmentos de leitura de alvos virais a partir de amostras de esgoto de modo que seu desenvolvimento deve ser fomentado, para utilização da amostragem em esgoto para monitoramento epidemiológico a nível genômico.

6 REFERÊNCIAS

- ADAMS A. Concentration of Epstein-Barr virus from cell culture fluids with polyethylene glycol. *J Gen Virol*. p391–394, 1973.
- ADAMS, M. J. *et al.* 50 years of the International Committee on Taxonomy of Viruses: progress and prospects. *Archives of virology*, v. 162, n. 5, p. 1441–1446, 4 maio 2017.
- AHMED, W. *et al.* First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *The Science of the total environment*, v. 728, p. 138764, 1 ago. 2020.
- ALATTAS, S. G. *et al.* Human Virome. *La Prensa Medica Argentina*, v. 106, n. S2, 2020.
- ALMEIDA FILHO, N. DE. Bases históricas da Epidemiologia. *Cadernos de Saúde Pública*, v. 2, n. 3, p. 304–311, set. 1986.
- AMARASINGHE, S. L. *et al.* Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, v. 21, n. 1, 7 fev. 2020.
- ASGHAR, H. *et al.* Environmental Surveillance for Polioviruses in the Global Polio Eradication Initiative. *The Journal of Infectious Diseases*, v. 210, n. suppl_1, p. S294–S303, 1 nov. 2014.
- AW, T. G.; HOWE, A.; ROSE, J. B. Metagenomic approaches for direct and cell culture evaluation of the virological quality of wastewater. *Journal of Virological Methods*, v. 210, p. 15–21, dez. 2014.
- BAERT, L. *et al.* Detection of Murine Norovirus 1 by Using Plaque Assay, Transfection Assay, and Real-Time Reverse Transcription-PCR before and after Heat Exposure. *Applied and Environmental Microbiology*, v. 74, n. 2, p. 543–546, 15 jan. 2008.
- BALTIMORE, D. Viral genetic systems. *Transactions of the New York Academy of Sciences*, v. 33, n. 3, p. 327–332, 1971.
- BARATA, R. B. Epidemiological surveillance: a brief history and the experiences of the United States and the state of São Paulo. *Epidemiologia e serviços de saúde: revista do Sistema Único de Saúde do Brasil*, v. 31, n. 2, p. e2021115, 2022.
- BATEMAN, Alex *et al.* UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, v. 51, n. D1, p. D523–D531, 2023. Disponível em: <<https://academic.oup.com/nar/article/51/D1/D523/6835362>>. Acesso em: 10 abr. 2023.
- BESSELSSEN, D. G.; WAGNER, A. M.; LOGANBILL, J. K. Detection of rodent coronaviruses by use of fluorogenic reverse transcriptase-polymerase chain reaction analysis. *Comparative medicine*, v. 52, n. 2, p. 111–6, abr. 2002.
- BIBBY, K. *et al.* Persistence of Ebola Virus in Sterilized Wastewater. *Environmental Science and Technology Letters*, v. 2, n. 9, p. 245–249, 8 set. 2015.
- BIBBY, K.; PECCIA, J. Identification of Viral Pathogen Diversity in Sewage Sludge by Metagenome Analysis. *Environmental Science & Technology*, v. 47, n. 4, p. 1945–1951, 19 fev. 2013.
- BONANNO FERRARO, G. *et al.* A State-of-the-Art Scoping Review on SARS-CoV-2 in Sewage Focusing on the Potential of Wastewater Surveillance for the Monitoring of the COVID-19 Pandemic. *Food and Environmental Virology*, v. 14, n. 4, p. 315–354, 2 dez. 2022.
- BRASIL. PORTARIA GM/MS Nº 420. Inclusão da síndrome congênita associada à infecção pelo vírus Zika na Lista Nacional de Notificação Compulsória de doenças, agravos e eventos de saúde pública nos

serviços de saúde públicos e privados em todo o território nacional **Diário Oficial da União**. 2022, p. 56–58. Disponível em: <<https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?data=04/03/2022&jornal=515&pagina=56&totalArquivos=85>>. Acesso em: 15 ago 2023.

BRASIL, M. D. S. Guia de Vigilância Genômica do SARS-CoV-2 - Uma abordagem epidemiológica e laboratorial. **Ministério da Saúde**. Brasília, 2023.

BREITBART, M. *et al.* Metagenomic Analyses of an Uncultured Viral Community from Human Feces. **Journal of Bacteriology**, v. 185, n. 20, p. 6220–6223, 15 out. 2003.

BRIESE, T. *et al.* Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. **mBio**, v. 6, n. 5, 30 out. 2015.

BRITO, A. F.; PINNEY, J. W. Protein-protein interactions in virus-host systems. **Frontiers in Microbiology**, v. 8, n. AUG, 17 ago. 2017.

CADAMURO, Rafael Dorighello. Prospecção da atividade enzimática e biológica do secretoma e de metabólitos secundários do fungo *Neofusicoccum parvum* isolado da planta *Avicennia schaueriana* do Mangue do Itacorubi – Ilha de Santa Catarina. (Mestrado em Biotecnologia e Biociências) - Universidade Federal de Santa Catarina (UFSC), 2022.

CAMERON, D.; JONES, I. G. John Snow, the Broad Street Pump and Modern Epidemiology. **International Journal of Epidemiology**, v. 12, n. 4, p. 393–396, 1983.

CANTALUPO, P. G. *et al.* Raw Sewage Harbors Diverse Viral Populations. **mBio**, v. 2, n. 5, nov. 2011.

CARDOSO, T. A. DE O.; NAVARRO, M. B. M. DE A. Emerging and reemerging diseases in Brazil: data of a recent history of risks and uncertainties. **Brazilian Journal of Infectious Diseases**, v. 11, n. 4, p. 430–434, ago. 2007.

CASANOVA, L., WEAVER, S. Inactivation of an Enveloped Surrogate Virus in Human Sewage. **Environmental Science and Technology Letters**, 2, 76-78, 2015.

CDC, C. FOR D. C. AND P. **What is Genomic Surveillance?** 2022. Disponível em: <<https://www.cdc.gov/coronavirus/2019-ncov/variants/genomic-surveillance.html>>. Acesso em: 15 ago 2023.

CESA, M. D. V.; DUARTE, G. M. A qualidade do ambiente e as doenças de veiculação hídrica. **Geosul**, p. 63–78, 2010.

CHAGAS, D. C. Erradicando doenças: De projeto internacional ao Sistema de Vigilância Epidemiológica - a erradicação da varíola no Brasil (1900-1970). **Casa de Oswaldo Cruz-FIOCRUZ**, Rio de Janeiro, 2008.

CLARK, M. F. Polyethylene glycol solubility gradients, a new and rapid technique for investigations of plant viruses. **Virology** p246–247, 1970.

COELHO, F. C. *et al.* Epidemiological data accessibility in Brazil. **The Lancet Infectious Diseases**, v. 16, n. 5, p. 524–525, maio 2016.

CONRADS, G.; ABDELBAR, M. M. H. Challenges of next-generation sequencing targeting anaerobes. **Anaerobe**, v. 58, p. 47–52, 1 ago. 2019.

CORTE, D. DE *et al.* Viral communities in the global deep ocean conveyor belt assessed by targeted viromics. **Frontiers in Microbiology**, v. 10, n. AUG, 2019.

- COSTA, Z. G. A. *et al.* Evolução histórica da vigilância epidemiológica e do controle da febre amarela no Brasil. **Revista Pan-Amazônica de Saúde**, v. 2, n. 1, p. 11–26, mar. 2011.
- CRICK, F. H. C.; WATSON, J. D. Structure of Small Viruses. **Nature**, v. 177, n. 4506, p. 473–475, mar. 1956.
- CUEVAS-FERRANDO, Enric *et al.* Monitoring Human Viral Pathogens Reveals Potential Hazard for Treated Wastewater Discharge or Reuse. **Frontiers in Microbiology**, v. 13, 2022.
- CUNNINGHAM, A. A.; DASZAK, P.; WOOD, J. L. N. One Health, emerging infectious diseases and wildlife: two decades of progress? **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 372, n. 1725, p. 20160167, 19 jul. 2017.
- DANECEK, P. *et al.* Twelve years of SAMtools and BCFtools. **GigaScience**, v. 10, n. 2, 16 fev. 2021.
- DE ABREU CORRÊA, Adriana *et al.* Comparative Inactivation of Murine Norovirus, Human Adenovirus, and Human JC Polyomavirus by Chlorine in Seawater. **Applied and Environmental Microbiology**, v. 78, n. 18, p. 6450–6457, 2012.
- DIAO, Z. *et al.* Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections. **Journal of advanced research**, v. 38, p. 201–212, maio 2022.
- DIRETORIA DE VIGILÂNCIA EPIDEMIOLÓGICA. Nota de Alerta Nº 0015/2021 GEDIM/DIVE/SUV/SES. **Governo de Santa Catarina**, Florianópolis. Disponível em: <<https://dive.sc.gov.br/phocadownload/notas-alerta/notas-alerta-2021/NA015.pdf>>. Acesso: 12 ago 2023.
- DOMINGUEZ DEL ANGEL, V. *et al.* Ten steps to get started in Genome Assembly and Annotation. **F1000Research**, v. 7, p. 148, 5 fev. 2018.
- DOWDLE, W. R. *et al.* Poliomyelitis eradication. **Virus Research**, v. 62, n. 2, p. 185–192, ago. 1999.
- EDEN, J.-S. *et al.* An emergent clade of SARS-CoV-2 linked to returned travellers from Iran. **Virus Evolution**, v. 6, n. 1, 1 jan. 2020.
- EDEN, J.-S.; SIM, E. SARS-CoV-2 Genome Sequencing Using Long Pooled Amplicons on Illumina Platforms. 2020. Available online: <<https://www.protocols.io/view/sars-cov-2-genome-sequencing-using-long-pooled-amp-befyjbpw?step=81>>. Acesso em 10 aug. 2023.
- EDRIDGE *et al.* Viral Metagenomics on Cerebrospinal Fluid. **Genes**, v. 10, n. 5, p. 332, 2019.
- EPA, United States Environment Protection Agency. **Contaminant Candidate List (CCL) and Regulatory Determination**. 2022. Disponível em: <<https://www.epa.gov/ccl/ccl-5-microbial-contaminants>>. Acesso em: 04 out. 2023.
- ESCOBAR-ZEPEDA, A.; LEÓN, A. V. P. DE; SANCHEZ-FLORES, A. The road to metagenomics: From microbiology to DNA sequencing technologies and bioinformatics. **Frontiers in Genetics**, v. 6, n. DEC, p. 348, 2015.
- EWING, B.; GREEN, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. **Genome research**, v. 8, n. 3, p. 186–94, mar. 1998.
- EYLER, J. M. William Farr on the Cholera: The Sanitarian’s Disease Theory and the Statistician’s Method. **Journal of the History of Medicine and Allied Sciences**, v. XXVIII, n. 2, p. 79–100, 1973.
- FARIA, N. R. *et al.* Establishment and cryptic transmission of Zika virus in Brazil and the Americas. **Nature**, v. 546, n. 7658, p. 406–410, 15 jun. 2017.

FELTRIN, Clarissa. **Desenvolvimento e aplicação de um sistema celular repórter para herpes simplex vírus e padronização de uma PCR quantitativa para poliomavírus BK**. 2014. 171 f. Dissertação - Universidade Federal de Santa Catarina (UFSC), 2014.

FERNANDES, T. M. D.; CHAGAS, D. C.; SOUZA, É. M. DE. Variola e vacina no Brasil no século XX: institucionalização da educação sanitária. **Ciência & Saúde Coletiva**, v. 16, n. 2, p. 479–789, fev. 2011.

FONGARO, G.; ROGOVSKI, P.; *et al.* SARS-CoV-2 in Human Sewage and River Water from a Remote and Vulnerable Area as a Surveillance Tool in Brazil. **Food and environmental virology**, 2021.

FONGARO, G.; STOCO, P. H.; *et al.* The presence of SARS-CoV-2 RNA in human sewage in Santa Catarina, Brazil, November 2019. **Science of The Total Environment**, v. 778, p. 146198, jul. 2021.

FRANCO, O. O. **A história da febre amarela no Brasil**. Rio de Janeiro: Ministério da Saúde, 1969.

FREITAS BUENO, R. DE *et al.* Wastewater-based epidemiology: A Brazilian SARS-COV-2 surveillance experience. **Journal of environmental chemical engineering**, v. 10, n. 5, p. 108298, out. 2022.

FRERICHS, R. R. Epidemiologic Surveillance in Developing Countries. **Annual. Rev. Public. Health**, 1991.

FUNDAÇÃO OSWALDO CRUZ. **A trajetória do médico dedicado à ciência**. Disponível em: <<https://portal.fiocruz.br/trajetoria-do-medico-dedicado-ciencia>>. Acesso em: 17 ago 2023.

FUNDAÇÃO OSWALDO CRUZ. **História**. Disponível em: <<https://portal.fiocruz.br/historia>>. Acesso em: 20 ago 2023.

G1 SC, J. C. Surto de diarreia em Florianópolis tem vírus como causa, aponta análise preliminar. 2023. Disponível em: <<https://g1.globo.com/sc/santa-catarina/noticia/2023/01/21/surto-de-diarreia-em-florianopolis-tem-virus-como-causa-aponta-analise-preliminar.ghtml>>. Acesso em: 21 ago 2023.

GARCIA, L. A. T. *et al.* Surveillance of human and swine adenovirus, human norovirus and swine circovirus in water samples in Santa Catarina, Brazil. **Journal of Water and Health**, v. 10, n. 3, p. 445–452, 1 set. 2012.

GHOLIPOUR, S. *et al.* Occurrence of viruses in sewage sludge: A systematic review. **Science of The Total Environment**, v. 824, p. 153886, jun. 2022.

GHOSH, A.; MEHTA, A.; KHAN, A. M. Metagenomic analysis and its applications. **Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics**, v. 1–3, p. 184–193, 1 jan. 2018.

GIRE, S. K. *et al.* Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. **Science**, v. 345, n. 6202, p. 1369–1372, 12 set. 2014.

GOVERNO DE SANTA CATARINA. Decreto nº 1172 de 26 de fevereiro de 2021. **Diário Oficial do estado de Santa Catarina**, 2021. Disponível em: <<https://dados.sc.gov.br/dataset/149a36ac-19c6-47b3-b873-9c0512f7a4db/resource/82ba215c-fd07-4647-8f94-8e1226fdcf25/download/decreto-n-1172-de-25.02.2021.pdf>>. Acesso em ago 2023.

GOVERNO DE SANTA CATARINA. **Secretaria De Saúde**. Vigilância Epidemiológica Monitora Aumento Dos Casos De Diarreia Em Santa Catarina. 2016. Disponível em: <<https://saude.sc.gov.br/index.php/noticias-geral/todas-as-noticias/1627-noticias-2016/4990-vigilancia-epidemiologica-monitora-aumento-dos-casos-de-diarreia-em-santa-catarina>>. Acesso em: 18 ago 2023.

- GRANDI, G.; LOPEZ, L. F.; BURATTINI, M. N. Temporal Trends of Acute Hepatitis A in Brazil and Its Regions. **Viruses**, v. 14, n. 12, 8 dez. 2022.
- GREEN, M. R.; SAMBROOK, J. Synthesis of cDNA Probes from mRNA Using Random Oligonucleotide Primers. **Cold Spring Harbor Protocols**, v. 2019, n. 7, jul. 2019.
- GRUBAUGH, N. D. *et al.* An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. **Genome Biology**, v. 20, n. 1, p. 8, 8 dez. 2019.
- HARVEY, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. **Nature Reviews Microbiology**, v. 19, n. 7, p. 409–424, 1 jul. 2021.
- HELLMÉR, M. *et al.* Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. **Applied and Environmental Microbiology**, v. 80, n. 21, p. 6771–6781, 2014.
- HENDRIX, R. W. Bacteriophages: Evolution of the Majority. **Theoretical Population Biology**, v. 61, n. 4, p. 471–480, 1 jun. 2002.
- HERATH, D. *et al.* Assessing Species Diversity Using Metavirome Data: Methods and Challenges. **Computational and structural biotechnology journal**, v. 15, p. 447–455, 2017.
- HEBERT, T. T. Precipitation of plant viruses by polyethylene glycol. **Phytopathology**. 1963.
- HERNROTH, B. E. *et al.* Environmental Factors Influencing Human Viral Pathogens and Their Potential Indicator Organisms in the Blue Mussel, *Mytilus edulis*: the First Scandinavian Report. **Applied and Environmental Microbiology**, v. 68, n. 9, p. 4523–4533, set. 2002.
- HESSE, U. *et al.* Virome assembly and annotation: A surprise in the namib desert. **Frontiers in Microbiology**, v. 8, n. JAN, 23 jan. 2017.
- HILL, V. *et al.* Progress and challenges in virus genomic epidemiology. **Trends in Parasitology**, v. 37, n. 12, p. 1038–1049, dez. 2021.
- HJELMSØ, M. H. *et al.* Evaluation of Methods for the Concentration and Extraction of Viruses from Sewage in the Context of Metagenomic Sequencing. **PLOS ONE**, v. 12, n. 1, p. e0170199, 18 jan. 2017.
- HORIMOTO, T.; KAWAOKA, Y. Influenza: lessons from past pandemics, warnings from current incidents. **Nature Reviews Microbiology**, v. 3, n. 8, p. 591–600, ago. 2005.
- HURWITZ, B. L.; SULLIVAN, M. B. The Pacific Ocean Virome (POV): A Marine Viral Metagenomic Dataset and Associated Protein Clusters for Quantitative Viral Ecology. **PLoS ONE**, v. 8, n. 2, 28 fev. 2013.
- JAFFERALI, M. H. *et al.* Benchmarking virus concentration methods for quantification of SARS-CoV-2 in raw wastewater. **The Science of the total environment**, v. 755, 10 fev. 2021.
- JEBRI, S. *et al.* Presence and fate of coliphages and enteric viruses in three wastewater treatment plants effluents and activated sludge from Tunisia. **Environmental Science and Pollution Research**, v. 19, n. 6, p. 2195–2201, 20 jul. 2012.
- JESUS, J. G. DE *et al.* Importation and early local transmission of COVID-19 in Brazil, 2020. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 62, 2020.
- JI, P. *et al.* Evaluation of a portable nanopore-based sequencer for detection of viruses in water. **Journal of virological methods**, v. 278, 1 abr. 2020.

- KHAN MIRZAEI, M. *et al.* Challenges of Studying the Human Virome – Relevant Emerging Technologies. **Trends in Microbiology**, v. 29, n. 2, p. 171–181, 1 fev. 2021.
- KIEFT, K. *et al.* vRhyme enables binning of viral genomes from metagenomes. **Nucleic Acids Research**, v. 50, n. 14, p. e83–e83, 12 ago. 2022.
- KIM, D. *et al.* Centrifuge: rapid and sensitive classification of metagenomic sequences. **Genome Research**, v. 26, n. 12, p. 1721–1729, dez. 2016.
- KIM, Hyaekang *et al.* Investigation of SARS-CoV-2 lineages and mutations circulating in a university-affiliated hospital in South Korea analyzed using Oxford Nanopore MinION sequencing. **Osong Public Health and Research Perspectives**, v. 13, n. 5, p. 360–369, 2022.
- KNIPE, D. M.; HOWLEY, PETER. *Fields Virology*. **Wolters Kluwer Health**, 2013.
- KOONIN, E. V.; SENKEVICH, T. G.; DOLJA, V. V. The ancient Virus World and evolution of cells. **Biology direct**, v. 1, p. 29, 19 set. 2006.
- KRISHNAMURTHY, S. R.; WANG, D. Origins and challenges of viral dark matter. **Virus Research**, v. 239, p. 136–142, jul. 2017.
- LAHRICH, S. *et al.* Review on the contamination of wastewater by COVID-19 virus: Impact and treatment. **The Science of the total environment**, v. 751, 10 jan. 2021.
- LAM, T. T.; PYBUS, O. G. Genomic surveillance of avian-origin influenza A viruses causing human disease. **Genome Medicine**, v. 10, n. 1, p. 50, 27 dez. 2018.
- LAMERS, M. M. *et al.* SARS-CoV-2 productively infects human gut enterocytes. **Science (New York, N.Y.)**, v. 369, n. 6499, p. 50–54, 3 jul. 2020.
- LEFKOWITZ, E. J. *et al.* Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). **Nucleic Acids Research**, v. 46, n. D1, p. D708–D717, 4 jan. 2018.
- LEWIS, G. D.; METCALF, T. G. Polyethylene glycol precipitation for recovery of pathogenic viruses, including hepatitis A virus and human rotavirus, from oyster, water, and sediment samples. **Applied and Environmental Microbiology**, v. 54, n. 8, p. 1983–1988, ago. 1988.
- LI, H.; DURBIN, R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics**, v. 25, n. 14, p. 1754–1760, jul. 2009.
- LI, Yanpeng *et al.* Semen virome of men with HIV on or off antiretroviral treatment. **AIDS**, v. 34, n. 6, p. 827–832, 2020.
- LI, J. *et al.* The emergence, genomic diversity and global spread of SARS-CoV-2. **Nature**, v. 600, n. 7889, p. 408–418, 16 dez. 2021.
- LIU, L. *et al.* Comparison of next-generation sequencing systems. **Journal of Biomedicine and Biotechnology**, 2012.
- LO CASTRO, Ivana *et al.* High prevalence of hepatitis A and E viruses in environmental and clinical samples from West Argentina. **The Brazilian Journal of Infectious Diseases**, v. 27, n. 2, p. 102738, 2023.
- LÓPEZ-BUENO, A. *et al.* High diversity of the viral community from an Antarctic Lake. **Science**, v. 326, n. 5954, p. 858–861, 6 nov. 2009.
- LOUTEN, J. *Virus Structure and Classification*. **Essential Human Virology**, p. 19–29, 2016.

- LU, X. *et al.* US CDC Real-Time Reverse Transcription PCR Panel for Detection of Severe Acute Respiratory Syndrome Coronavirus 2. **Emerging Infectious Diseases**, v. 26, n. 8, p. 1654–1665, ago. 2020.
- MCDONALD, S. M. *et al.* Reassortment in segmented RNA viruses: mechanisms and outcomes. **Nature Reviews Microbiology**, v. 14, n. 7, p. 448–460, 23 jul. 2016.
- MCEWEN, S. A.; COLLIGNON, P. J. Antimicrobial Resistance: a One Health Perspective. **Microbiology Spectrum**, v. 6, n. 2, 6 abr. 2018.
- MCMINN, B. R.; ASHBOLT, N. J.; KORAJKIC, A. Bacteriophages as indicators of faecal pollution and enteric virus removal. **Letters in Applied Microbiology**, v. 65, n. 1, p. 11–26, 1 jul. 2017.
- MENZEL, P.; NG, K. L.; KROGH, A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. **Nature Communications**, v. 7, 13 abr. 2016.
- MICHEN, B.; GRAULE, T. Isoelectric points of viruses. **Journal of applied microbiology**, 388–397, 2010.
- MINISTÉRIO DA SAÚDE. **Vacina de Hepatite A**. 2013. Disponível em: <<https://www.gov.br/conitec/pt-br/midias/incorporados/vacinahepatite-final.pdf>>. Acesso em: 22 ago. 2023.
- MINISTÉRIO DO TURISMO. **RELATÓRIO DE IMPACTO da Pandemia de COVID-19 nos setores de turismo e cultura no Brasil**. Brasília: 2020. Disponível em: <https://www.gov.br/turismo/pt-br/aceso-a-informacao/acoes-e-programas/observatorio/estudos-de-impacto/Relatorio_impacto_COVID19_TUR_e_CULT_MTur.pdf>. Acesso em: 25 ago. 2023.
- MINOT, S. *et al.* Rapid evolution of the human gut virome. **Proceedings of the National Academy of Sciences**, v. 110, n. 30, p. 12450–12455, 23 jul. 2013.
- MORABIA, A. Epidemiology's 350th Anniversary. **Epidemiology**, v. 24, n. 2, p. 179–183, mar. 2013.
- MORESCO, V. *et al.* Microbiological and physicochemical analysis of the coastal waters of southern Brazil. **Marine Pollution Bulletin**, v. 64, n. 1, p. 40–48, jan. 2012.
- MORTARI, A. *et al.* Norovirus and rotavirus in surface, malacoculture, and human consumption water in Santa Catarina State, Brazil. **Journal of Water and Health**, v. 21, n. 1, p. 35–46, 1 jan. 2023.
- MUIRHEAD, A. *et al.* Zika Virus RNA Persistence in Sewage. **Environmental Science and Technology Letters**, v. 7, n. 9, p. 659–664, 8 set. 2020.
- NG, T. F. F. *et al.* High Variety of Known and New RNA and DNA Viruses of Diverse Origins in Untreated Sewage. **Journal of Virology**, v. 86, n. 22, p. 12161–12175, 15 nov. 2012.
- NIEUWENHUIJSE, D. F. *et al.* Setting a baseline for global urban virome surveillance in sewage. **Scientific Reports 2020 10:1**, v. 10, n. 1, p. 1–13, 13 ago. 2020.
- NSC TOTAL. Florianópolis registra 79 surtos de diarreia em 2021. 2021. Disponível em: <<https://www.nscotal.com.br/noticias/florianopolis-registra-79-surtos-de-diarreia-em-2021>>. Acesso em: 23 ago. 2023.
- OCHMAN, H.; CARO-QUINTERO, A. Genome Size and Structure, Bacterial. *In: Encyclopedia Of Evolutionary Biology*. Elsevier, 2016. p. 179–185.
- OGDEN, R.; VASILJEVIC, N.; PROST, S. Nanopore sequencing in non-human forensic genetics. **Emerging Topics in Life Sciences**, v. 5, n. 3, p. 465–473, 24 set. 2021.

OIKARINEN, S. *et al.* PCR inhibition in stool samples in relation to age of infants. **Journal of clinical virology**, vol. 44,3, 2009.

O'LEARY, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. **Nucleic Acids Research**, v. 44, n. D1, p. D733–D745, 4 jan. 2016.

OUDE MUNNINK, Bas; VAN DER HOEK, Lia. Viruses Causing Gastroenteritis: The Known, The New and Those Beyond. **Viruses**, v. 8, n. 2, p. 42, 2016.

PADILHA, D. A. *et al.* Emergence of Two Distinct SARS-CoV-2 Gamma Variants and the Rapid Spread of P.1-like-II SARS-CoV-2 during the Second Wave of COVID-19 in Santa Catarina, Southern Brazil. **Viruses**, v. 14, n. 4, p. 695, 27 mar. 2022.

PADILHA, D. A. *et al.* Genomic Surveillance of SARS-CoV-2 in Healthcare Workers: A Critical Sentinel Group for Monitoring the SARS-CoV-2 Variant Shift. **Viruses**, v. 15, n. 4, p. 984, 17 abr. 2023.

PERVEZ, M. T. *et al.* A Comprehensive Review of Performance of Next-Generation Sequencing Platforms. **BioMed Research International**, v. 2022, p. 1–12, set. 2022.

PETERANDERL, C.; HEROLD, S.; SCHMOLDT, C. Human Influenza Virus Infections. **Seminars in Respiratory and Critical Care Medicine**, v. 37, n. 04, p. 487–500, 3 ago. 2016.

POGKA, V. *et al.* Targeted Virome Sequencing Enhances Unbiased Detection and Genome Assembly of Known and Emerging Viruses—The Example of SARS-CoV-2. **Viruses**, v. 14, n. 6, 1 jun. 2022.

POLSON, S. W.; WILHELM, S. W.; WOMMACK, K. E. Unraveling the viral tapestry (from inside the capsid out). **ISME Journal**, v. 5, n. 2, p. 165–168, 2011.

PORTIK, D. M.; BROWN, C. T.; PIERCE-WARD, N. T. Evaluation of taxonomic classification and profiling methods for long-read shotgun metagenomic sequencing datasets. **BMC bioinformatics**, v. 23, n. 1, p. 541, 13 dez. 2022.

PRADO, T. *et al.* Wastewater-based epidemiology as a useful tool to track SARS-CoV-2 and support public health policies at municipal level in Brazil. **Water Research**, v. 191, p. 116810, mar. 2021.

PREFEITURA DE FLORIANÓPOLIS. **Covidômetro**, 2023a. Disponível em: <<https://www.covidometrofloripa.com.br/>>. Acesso em: 20 ago. 2023.

PREFEITURA DE FLORIANÓPOLIS. **Vacinômetro**, 2023b. Disponível em: <https://lookerstudio.google.com/u/0/reporting/0cc4dda4-9bae-4452-9836-efd2ae73d6d3/page/p_py93q3uhsc>. Acesso em: 15 ago. 2023.

PRICE, J. R. *et al.* Whole-Genome Sequencing Shows That Patient-to-Patient Transmission Rarely Accounts for Acquisition of *Staphylococcus aureus* in an Intensive Care Unit. **Clinical Infectious Diseases**, v. 58, n. 5, p. 609–618, 1 mar. 2014.

QUINCE, C. *et al.* Shotgun metagenomics, from sampling to analysis. **Nature Biotechnology**, v. 35, n. 9, p. 833–844, 12 set. 2017.

REGENMORTEL, M. H. V. VAN. Virus Species, a Much Overlooked but Essential Concept in Virus Classification. **Intervirology**, v. 31, n. 5, p. 241–254, 1990.

REINTJES, R.; KRICKEBERG, K. Epidemiologic Surveillance. **Modern Infectious Disease Epidemiology**. Statistics for Biology and Health. Springer, New York, N p. 143–157. 2009

- RIGOTTO, C. *et al.* Assessment of adenovirus, hepatitis A virus and rotavirus presence in environmental samples in Florianopolis, South Brazil. **Journal of Applied Microbiology**, v. 109, n. 6, p. 1979–1987, dez. 2010.
- RIGHETTO, Irene; FILIPPINI, Francesco. Pandemic Avian Influenza and Intra/Interhaemagglutinin Subtype Electrostatic Variation among Viruses Isolated from Avian, Mammalian, and Human Hosts. **BioMed Research International**, v. 2018, p. 1–10, 2018.
- ROBINSON, J. T. *et al.* Integrative genomics viewer. **Nature Biotechnology**, v. 29, n. 1, p. 24–26, 1 jan. 2011.
- ROTHMAN, K. J. The Rise and Fall of Epidemiology, 1950–2000 A.D. **New England Journal of Medicine**, v. 304, n. 10, p. 600–602, 5 mar. 1981.
- ROUX, S. *et al.* Minimum Information about an Uncultivated Virus Genome (MIUViG). **Nature Biotechnology** 2018 37:1, v. 37, n. 1, p. 29–37, 17 dez. 2018.
- SABIR, J. S. M. *et al.* Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. **Science**, v. 351, n. 6268, p. 81–84, jan. 2016.
- SALVO, M. *et al.* Evaluation of low-cost viral concentration methods in wastewaters: Implications for SARS-CoV-2 pandemic surveillances. **Journal of Virological Methods**, v. 297, p. 114249, nov. 2021.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 74, n. 12, p. 5463–5467, 1977.
- SANJUÁN, Rafael; DOMINGO-CALAP, Pilar. Mechanisms of viral mutation. **Cellular and Molecular Life Sciences**, v. 73, n. 23, p. 4433–4448, 2016.
- SANTOS, C. V. B. DOS *et al.* Estimated COVID-19 severe cases and deaths averted in the first year of the vaccination campaign in Brazil: A retrospective observational study. **Lancet regional health. Americas**, v. 17, p. 100418, jan. 2023.
- SANTOS, N. S. DE O. *Virologia Humana. Guanabara Koogan* 3º Ed. 2015.
- SCHLINDWEIN, A. D.; RIGOTTO, C.; SIMÕES, C. M.O.; *et al.* Detection of enteric viruses in sewage sludge and treated wastewater effluent. **Water Science and Technology**, v. 61, n. 2, p. 537–544, 2010.
- SCOLA, B. LA *et al.* The virophage as a unique parasite of the giant mimivirus. **Nature**, v. 455, n. 7209, p. 100–104, 4 set. 2008.
- SERRA-COMPTE, A. *et al.* Elimination of SARS-CoV-2 along wastewater and sludge treatment processes. **Water Research**, v. 202, p. 117435, set. 2021.
- SETUBAL, J. C. Metagenome-assembled genomes: concepts, analogies, and challenges. **Biophysical Reviews**, v. 13, n. 6, p. 905–909, 1 dez. 2021.
- SHOUVAL, Daniel. Immunization against Hepatitis A. **Cold Spring Harbor Perspectives in Medicine**, v. 9, n. 2, p. a031682, 2019.
- SIB, Swiss Institute of Bioinformatics. **The Baltimore classification clusters viruses into families depending on their type of genome**, 2020. Disponível em <<https://viralzone.expasy.org/254>>. Acesso em: 28 ago. 2023.
- SIMMONDS, Peter; AIEWSAKUN, Pakorn. Virus classification – where do you draw the line?. **Archives of Virology**, v. 163, n. 8, p. 2037–2046, 2018.

- SMITH, S. E. *et al.* Emerging technologies in the study of the virome. **Current Opinion in Virology**, v. 54, p. 101231, jun. 2022.
- SOUHO, T. *et al.* Study of the SARS-CoV-2 genomic data generation to evaluate the introduction of genomics in epidemiological surveillance and public health decision making. **The Pan African medical journal**, v. 41, p. 55, 2022.
- STRAIF-BOURGEOIS, S.; RATARD, R.; KRETZSCHMAR, M. Infectious Disease Epidemiology. **Handbook of Epidemiology**, p. 2041, 1 jan. 2014.
- THOMPSON, S. S. *et al.* Detection of Infectious Human Adenoviruses in Tertiary-Treated and Ultraviolet-Disinfected Wastewater. **Water Environment Research**, v. 75, n. 2, p. 163–170, mar. 2003.
- TSUI, W. N. T. *et al.* Molecular detection of SARS-CoV-2 and differentiation of Omicron and Delta variant strains. **Transboundary and emerging diseases**, v. 69, n. 5, p. 1618–1631, set. 2022.
- TULCHINSKY, T. H. John Snow, Cholera, the Broad Street Pump; Waterborne Diseases Then and Now. **Case Studies in Public Health**. Elsevier, 2018. p. 77–99.
- UWISHEMA, Olivier *et al.* Hepatitis A virus outbreak in Lebanon: Is it a matter of concern?. **Annals of Medicine & Surgery**, v. 82, 2022.
- VICTORIA, M. *et al.* Assessment of norovirus contamination in environmental samples from Florianópolis City, Southern Brazil. **Journal of Applied Microbiology**, v. 109, n. 1, p. 231–238, 10 jul. 2010.
- WACHTER, Julia K. *et al.* VIGELAI - Vigilância Epidemiológica Ambiental Integrativa - Laboratório De Virologia Aplicada da UFSC – Brasil: Bases da ferramenta epidemiológica para o monitoramento viral a partir do esgoto sanitário. *In: Qualidade De Vida: Da Saúde Humana À Saúde Ambiental.:* **Instituto de Inteligência em Pesquisa e Consultoria Científica Ltda**, 2023. p. 60–84.
- WICK, R. R.; JUDD, L. M.; HOLT, K. E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. **Genome Biology**, v. 20, n. 1, p. 129, 24 dez. 2019.
- WIGGINTON, K. R.; YE, Y.; ELLENBERG, R. M. Emerging investigators series: The source and fate of pandemic viruses in the urban water cycle. **Environmental Science: Water Research and Technology**, v. 1, n. 6, p. 735–746, 2015.
- WONG, K.; ONAN, B. M.; XAGORARAKI, I. Quantification of Enteric Viruses, Pathogen Indicators, and Salmonella Bacteria in Class B Anaerobically Digested Biosolids by Culture and Molecular Methods. **Applied and Environmental Microbiology**, v. 76, n. 19, p. 6441–6448, out. 2010.
- WOOD, D. E.; LU, J.; LANGMEAD, B. Improved metagenomic analysis with Kraken 2. **Genome Biology**, v. 20, n. 1, 28 nov. 2019.
- WU, F. *et al.* SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases. **mSystems**, v. 5, n. 4, 25 ago. 2020.
- XIAO, Tiantian; ZHOU, Wenhao. The third generation sequencing: the advanced approach to genetic diseases. **Translational Pediatrics**, v. 9, n. 2, p. 163–173, 2020.
- ZENG, S.-Q. *et al.* One-step quantitative RT-PCR for the detection of rotavirus in acute gastroenteritis. **Journal of Virological Methods**, v. 153, n. 2, p. 238–240, nov. 2008.
- ZERBINI, F. M. *et al.* Changes to virus taxonomy and the ICTV Statutes ratified by the International Committee on Taxonomy of Viruses (2023). **Archives of Virology**, v. 168, n. 7, p. 175, 10 jul. 2023.

ZHAN, S. H. *et al.* Target capture sequencing of SARS-CoV-2 genomes using the ONETest Coronavirus Plus. **Diagnostic Microbiology and Infectious Disease**, v. 101, n. 3, p. 115508, nov. 2021.

7 APÊNDICE A – MATERIAIS COMPLEMENTARES

7.1 Publicações E Participações Em Eventos

7.1.1 Artigos Completos Publicados Em Periódicos

- **Emergence of Two Distinct SARS-CoV-2 Gamma Variants and the Rapid Spread of P.1-like-II SARS-CoV-2 during the Second Wave of COVID-19 in Santa Catarina, Southern Brazil**

D.A. PADILHA; **V. B. FILHO**; R. S. MOREIRA; T. A. T. SORATTO; G. A. MAIA ; A. P. CHRISTOFF; F. H. BARAZZETTI; M. A. SCHÖRNER ; F. L. FERRARI; C. L. MARTINS; E. K. KAWAGOE; J. K. WACHTER; P. SACHET ; A. R. BAPTISTELLA ; A. D. SCHLINDWEIN; B. K. COELHO; S. B. FERNANDES; D. B. ROVARIS; M. P. D. DOS ANJOS; F. R. MELO; B. BITTENCOURT; S. CUNHA; K. L. MENEGHETTI; N. WENDT; T. Z. MADALUZ; M. V. D. RODRIGUES; D. S. M. SOUZA; M. H. DE MORAES; R. DE P. BAPTISTA; G. T. SILVA; G. RAZZERA; E. C. GRISARD; P. H. STOCO; L. F. V. DE OLIVEIRA; M. L. BAZZO; G. FONGARO; G. WAGNER.

The western mesoregion of the state of Santa Catarina (SC), Southern Brazil, was heavily affected as a whole by the COVID-19 pandemic in early 2021. This study aimed to evaluate the dynamics of the SARS-CoV-2 virus spreading patterns in the SC state from March 2020 to April 2021 using genomic surveillance. During this period, there were 23 distinct variants, including Beta and Gamma, among which the Gamma and related lineages were predominant in the second pandemic wave within SC. A regionalization of P.1-like-II in the Western SC region was observed, concomitant to the increase in cases, mortality, and the case fatality rate (CFR) index. This is the first evidence of the regionalization of the SARS-CoV-2 transmission in SC and it highlights the importance of tracking the variants, dispersion, and impact of SARS-CoV-2 on the public health systems. **Viruses**, 2022, 14, 695. <https://doi.org/10.3390/v14040695>

- **EpiBuilder: A Tool for Assembling, Searching, and Classifying B-Cell Epitopes**

R. S. MOREIRA; **V. B. FILHO**; N. A. CALOMENO; G. WAGNER; L. C. MILETTI.

Epitopes are portions of a protein that are recognized by antibodies. These small amino acid sequences represent a significant breakthrough in a branch of bioinformatics called immunoinformatics. Various software are available for linear B-cell epitope (BCE) prediction

such as ABCPred, SVMTrip, EpiDope, and EpitopeVec; a well-known BCE predictor is BepiPred-2.0. However, despite the prediction, there are several essential steps, such as epitope assembly, evaluation, and searching for epitopes in other proteomes. Here, we present EpiBuilder (<https://epibuilder.sourceforge.io>), a user friendly software that assists in epitope assembly, classifying and searching using input results of BepiPred-2.0. EpiBuilder generates several output results from these data and supports a proteome-wide processing approach. In addition, this software provides the following features: Chou & Fasman beta-turn prediction, Emini surface accessibility prediction, Karplus and Schulz flexibility prediction, Kolaskar and Tongaonkar antigenicity, Parker hydrophilicity prediction, N-glycosylation domains, and hydropathy. These information generate a unique topology for each epitope, visually demonstrating its characteristics. The software can search the entire epitope sequence in various FASTA files, and it allows to use BLASTP to identify epitopes that eventually have sequence variations. As an EpiBuilder application, we developed a epitope dataset from the protozoan *Trypanosoma brucei gambiense*, the gram-positive bacterium *Clostridioides difficile*, and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). **Bioinformatics and Biology Insights**, 2022. <https://doi.org/10.1177/11779322221095221>

- **Different Swine Production Systems Can Shape Slurry Resistome at Mechanism and Class Levels Based on Swine Manure Evaluation**

L. C. BELTRAME; C. P. ZAMPARETTE; C. FELTRIN; C. R. CUNHA; E. P. COLTRO; G. S. S. ATHAYDE; **V. B. FILHO**; D. C. TÁPPARO; J. MONTEIRO; J. D. KICH; J. K. PALMEIRO; G. WAGNER; G. FONGARO; C. R. ZÁRATE-BLADÉS; T. C. M. SINCERO.

Antimicrobial resistance is a major threat to public health. Antimicrobial use in animal husbandry is a major concern since it can favor an increase in antimicrobial resistance among farms. Herein, we aim to better understand and characterize the main resistome profiles in microbial communities found in pig farms. Sampling of swine manure was performed in two different timepoints (October 2019 and January 2020) in each of the 14 different swine farms, located in the mesoregion of Western Santa Catarina state in Brazil, a pole of swine product production of worldwide importance. Samples were divided into three groups: farms with the opened regimen and no usage of antimicrobials (F1; n = 10), farms with the closed regimen and usage of antimicrobials (F2; n = 16), and farms with the closed regimen and no usage of antimicrobials (F3; n = 2). The metagenomic evaluation was performed to obtain and identify genetic elements related to antimicrobial resistance using nanopore sequencing. We used

ResistoXplorer software to perform composition, alpha and beta diversity, and clustering analysis. In addition, PCR reactions were performed to confirm the presence or absence of seven different beta-lactamase family genes and five phosphoethanolamine transferase gene variants clinically relevant. Our findings based on the identification of resistance genes at the mechanism level showed a prevalence of alteration of the drug target (72.3%) profile, followed by drug inactivation (17.5%) and drug efflux (10.1%). We identified predominantly aminoglycosides (45.3%), tetracyclines (15.9%), and multiclass (11,2%) resistance genes. PCoA analysis indicates differences between F1 and F2 profiles. F2 samples showed increased diversity when compared to the F1 group. In addition, herein we first report the identification of *mcr-4* in a slurry sample (C1F1.1) in Santa Catarina State. In general, our findings reinforce that many factors on the practices of animal husbandry are involved in the resistome profile at the mechanism and class levels. Further studies to better understand microbiome and mobilome aspects of these elements are necessary to elucidate transmission pathways between different bacteria and environments. **Frontiers in Cellular and Infection Microbiology**, Volume 12, 2022. <https://doi.org/10.3389/fcimb.2022.879656>

- **Comparison of functional-oil blend and anticoccidial antibiotics effects on performance and microbiota of broiler chickens challenged by coccidiosis**

P. G. S. PIRES; P. TORRES; T. A. T. SORATTO; V. B. FILHO; L. HAUPTLI; G. WAGNER; C. D. POZZATTI; P. O. MORAES.

This study aimed to compare the effects of different levels of cashew nutshell liquid (CNSL) and castor oil (CNSL–castor oil) with growth-promoting antibiotics associated with anticoccidials in broiler chickens challenged with coccidiosis. In this work, 2520 one-day-old male broiler chicks (Cobb) were randomly assigned to 84 pens, containing 30 birds each. The experimental design was completely randomized, with seven treatments: enramycin (8 ppm), virginiamycin (16.5 ppm), and tylosin (55 ppm); different doses of CNSL–castor oil (0.5, 0.75, and 1.00 kg/t); and a control diet (without additives). All treatments received semduramicin + nicarbazin (500 g/t; Aviax® Plus) from 0 to 28 d and monensin sodium (100 ppm; Elanco) from 29 to 35 days of age, when the feed was without antibiotics. The challenge was introduced at 14 days of age by inoculating broiler chickens with sporulated *Eimeria tenella*, *Eimeria acervulina*, and *Eimeria maxima* oocysts via oral gavage. In addition to performance parameters, intestinal contents were collected at 28 and 42 days of age for microbiota analysis

by sequencing the 16s rRNA in V3 and V4 regions using the Illumina MiSeq platform. Taxonomy was assigned using the SILVA database (v. 138) with QIIME2 software (v. 2020.11). After one week of challenge, the broilers that received tylosin had a higher body weight gain (BWG) than those in the control group ($p < 0.05$), while the other treatments presented intermediate values. At 28 d, the BWG was lower for the control, CNSL–Castor oil 0.5 kg/t, enramycin, and virginiamycin treatments than that in the tylosin treatment. The inclusion of CNSL–Castor oil at concentrations of 0.75 and 1 kg/t acted as an intermediate treatment ($p < 0.05$). For alpha diversity, using the Shannon index, it was possible to observe the effect of age, with substantial diversity at 42 d. The Firmicutes phylum had the highest abundance, with values between 84.33% and 95.16% at 42 d. Tylosin showed better performance indices than other treatments. CNSL–castor oil treatments with concentrations of 0.75 and 1 kg/t showed similar results to those of enramycin and virginiamycin. Furthermore, CNSL–castor oil acted as a modulator of intestinal microbiota, reducing the abundance of pathogenic bacteria. **PLOS ONE**, 2022. <https://doi.org/10.1371/journal.pone.0270350>

- **AnnotaPipeline: An integrated tool to annotate eukaryotic proteins using multi-omics data**

G. A. MAIA; V. B. FILHO; E. K. KAWAGOE; T. A. T. SORATTO; R. S. MOREIRA; E. C. GRISARD; G. WAGNER.

Assignment of gene function has been a crucial, laborious, and time-consuming step in genomics. Due to a variety of sequencing platforms that generates increasing amounts of data, manual annotation is no longer feasible. Thus, the need for an integrated, automated pipeline allowing the use of experimental data towards validation of in silico prediction of gene function is of utmost relevance. Here, we present a computational workflow named AnnotaPipeline that integrates distinct software and data types on a proteogenomic approach to annotate and validate predicted features in genomic sequences. Based on FASTA (i) nucleotide or (ii) protein sequences or (iii) structural annotation files (GFF3), users can input FASTQ RNA-seq data, MS/MS data from mzXML or similar formats, as the pipeline uses both transcriptomic and proteomic information to corroborate annotations and validate gene prediction, providing transcription and expression evidence for functional annotation. Reannotation of the available *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Candida albicans*, *Trypanosoma cruzi*, and *Trypanosoma rangeli* genomes was performed using the AnnotaPipeline, resulting in a higher proportion of annotated proteins and a reduced proportion of hypothetical proteins when

compared to the annotations publicly available for these organisms. AnnotaPipeline is a Unix-based pipeline developed using Python and is available at: <https://github.com/bioinformatics-ufsc/AnnotaPipeline>. **Frontiers in Genetics**, Volume 13, 2022. <https://doi.org/10.3389/fgene.2022.1020100>

- ***Trypanosoma evansi* secretome carries potential biomarkers for Surra diagnosis**

R. S. MOREIRA; N. A. CALOMENO; G. B. NEVES; L. F. N. NASCIMENTO; **V. B. FILHO**; G. WAGNER; L. C. MILETTI.

Trypanosoma evansi is a parasite that is phylogenetically close to *Trypanosoma brucei* and is the causative agent of a disease known as surra. Surra is responsible for a high mortality rate in livestock and large economic losses in the Americas, Africa, and Asia. This work aimed to analyze in vitro secreted proteins from *T. evansi* and identify potential treatment and diagnostic biomarkers for surra diagnosis. Two groups were used. In one group the parasites were purified using a DEAE-Cellulose column and maintained in a secretion medium while in the other group the parasites were not purified. Each group was further divided to be maintained at either 37 °C or 27 °C. We identified 246 proteins through mass spectrometry and found that the temperature appears to modulate protein secretion. We found minimal variations in the protein pools from pure and non-purified sets. We observed an emphasis on proteins associated to vesicles, glycolysis, and cellular homeostasis through the enrichment of GO. Also, we found that most secretome proteins share homologous proteins with *T. b. brucei*, *T. b. gambiense*, *T. vivax*, *T. equiperdum*, and *T. b. rhodesiense* secretome but unique *T. evansi* epitopes with potential biomarkers for surra diagnosis were detected. **Journal of Proteomics**. Volume 272, 2023. <https://doi.org/10.1016/j.jprot.2022.104789>

- **Phenotypic and genotypic characterization of antibiotic resistance of *Salmonella* Heidelberg in the south of Brazil**

L. S. G. SOARES; T. CASELLA; E. K. KAWAGOE; **V. B. FILHO**; W. P. OMORI; M. C. L. NOGUEIRA; G. WAGNER; R. R. de OLIVEIRA; S. R. STAHLHOFER; F. A. FERREIRA; E. C. TONDO; J. D. LINDER.

Salmonella is the main human pathogen present in the poultry chain. *Salmonella* Heidelberg is one of the most important serovars for public health since it has been frequently isolated in broiler chickens from different countries and may present multidrug resistance (MDR). This

study was carried out with 130 *S. Heidelberg* isolates collected from pre-slaughter broiler farms in 2019 and 2020 in 18 cities from three Brazilian states to study relevant aspects regarding their genotypic and phenotypic resistance. The isolates were tested and identified using somatic and flagellar antiserum (0:4, H:2, and H:r), and an antimicrobial susceptibility test (AST) was performed against 11 antibiotics for veterinary use. The strains were typed by Enterobacterial Repetitive Intergenic Consensus (ERIC)-PCR, and representatives of the main clusters of the identified profiles were sequenced by Whole Genome Sequencing (WGS). AST results showed that all isolates were resistant to sulfonamide, 54 % (70/130) were resistant to amoxicillin, and only one was sensitive to tetracycline. Twelve isolates (15.4 %) were MDR. The dendrogram obtained from the ERIC-PCR showed that the strains were grouped into 27 clusters with similarity above 90 %, with some isolates showing 100 % similarity but with different phenotypic profiles of antimicrobial resistance. Identical strains collected on the same farm on other dates were identified, indicating that they were residents. WGS identified 66 antibiotic-resistance genes. The *sul2* (present in all sequenced samples) and *tet(A)* genes were highlighted and validated in the experimental analysis. The *fosA7* gene was also identified in all sequenced samples, but resistance was not observed in the phenotypic test, possibly due to the heteroresistance of the *S. Heidelberg* strains evaluated. Considering that chicken meat is one of the most consumed meats in the world, the data obtained in the present study can corroborate the mapping of the origin and trends of antimicrobial resistance. **International Journal of Food Microbiology**, Volumes 391–393, 2023, <https://doi.org/10.1016/j.ijfoodmicro.2023.110151>

- **Genomic Surveillance of SARS-CoV-2 in Healthcare Workers: A Critical Sentinel Group for Monitoring the SARS-CoV-2 Variant Shift**

D. A. PADILHA; D. S. M SOUZA; E. K. KAWAGOE; **V. B. FILHO**; A. N. AMORIM; F. H. BARAZZETTI; M. A. SHORNER; S. B. FERNANDES; B. K. COELHO; D. B. ROVARIS; M. P. D. dos ANJOS; J. R. MOSER; F. R. MELO; B. B. de SOUZA; D. C. BESSA; F. H. P. S. MENDES; A. C. BOING; A. F. BOING; J. T. de LACERDA; G. V. MOURA; D. C. de BASTIANI; M. H. de MORAES; L. F. V. de OLIVEIRA; R. S. MOREIRA; P. H. STOCO; M. L. BAZZO; G. FONGARO; G. WAGNER.

SARS-CoV-2 genome surveillance is important for monitoring risk groups and health workers as well as data on new cases and mortality rate due to COVID-19. We characterized the circulation of SARS-CoV-2 variants from May 2021 to April 2022 in the state of Santa

Catarina, southern Brazil, and evaluated the similarity between variants present in the population and healthcare workers (HCW). A total of 5291 sequenced genomes demonstrated the circulation of 55 strains and four variants of concern (Alpha, Delta, Gamma and Omicron-sublineages BA.1 and BA.2). The number of cases was relatively low in May 2021, but the number of deaths was higher with the Gamma variant. There was a significant increase in both numbers between December 2021 and February 2022, peaking in mid-January 2022, when the Omicron variant dominated. After May 2021, two distinct variant groups (Delta and Omicron) were observed, equally distributed among the five Santa Catarina mesoregions. Moreover, from November 2021 to February 2022, similar variant profiles between HCW and the general population were observed, and a quicker shift from Delta to Omicron in HCW than in the general population. This demonstrates the importance of HCW as a sentinel group for monitoring disease trends in the general population. *Viruses* 2023, 15, 984. <https://doi.org/10.3390/v15040984>.

- **Bacteriophage-Associated Antimicrobial Resistance Genes in Avian Pathogenic *Escherichia coli* Isolated from Brazilian Poultry**

G. V. T. PILATI; R. D. CADAMURO; **V. B. FILHO**; M. DAHMER; M. A. ELOIS; B. P. SAVI; G. B. C. SALLES; E. C. MUNIZ; G. FONGARO.

Colibacillosis is a disease caused by *Escherichia coli* and remains a major concern in poultry production, as it leads to significant economic losses due to carcass condemnation and clinical symptoms. The development of antimicrobial resistance is a growing problem of worldwide concern. Lysogenic bacteriophages are effective vectors for acquiring and disseminating antibiotic resistance genes (ARGs). The aim of this study was to investigate the complete genome of *Escherichia coli* isolates from the femurs of Brazilian broiler chickens in order to investigate the presence of antimicrobial resistance genes associated with bacteriophages. Samples were collected between August and November 2021 from broiler batches from six Brazilian states. Through whole genome sequencing (WGS), data obtained were analyzed for the presence of antimicrobial resistance genes. Antimicrobial resistance genes against the aminoglycosides class were detected in 79.36% of the isolates; 74.6% had predicted sulfonamides resistance genes, 63.49% had predicted resistance genes against β -lactams, and 49.2% of the isolates had at least one of the tetracycline resistance genes. Among the detected genes, 27 have been described in previous studies and associated with bacteriophages. The

findings of this study highlight the role of bacteriophages in the dissemination of ARGs in the poultry industry. *Viruses* 2023, 15, 1485. <https://doi.org/10.3390/v15071485>.

- **Intestinal health of broilers challenged with *Eimeria* spp. Using functional oil blends in two physical forms with or without anticoccidials**

MORAES, P. O.; PIRES, P. G. S.; FILHO, V. B.; LIMA, A. L. F.; KINDLEIN, L.; TASCETTO, D.; FAVERO, A.; WAGNER, G.

This study aimed to assess the impact of a commercial blend of functional oils, specifically cashew nutshell liquid and castor oil (FO), in two physical forms (solid: P; liquid: S), in comparison to a combination of virginiamycin and anticoccidials on the gut health of broilers challenged with coccidiosis. A total of 1760 1-day-old male chicks were randomly distributed in a study design with eight treatments. The treatments included: a control group (without additive), OFS_0.75_kg/t (FO spray), OFP_1.0_kg/t (FO powder), OFP_1.5_kg/t (FO liquid spray), Sal (anticoccidials), Sal_Vir (virginiamycin and anticoccidials), Sal_OFS_0.5_kg/t (anticoccidials plus FO spray), and Sal_OFP_1.0_kg/t (anticoccidials plus FO powder). All birds were challenged with *Eimeria* spp. at 14 days. The physical form of FO did not affect performance and intestinal health parameters. At 42 days, broilers from the control and OFS_0.75 treatments were the lightest, while those from the Sal_Vir and Sal_OFP_1.0 treatments were the heaviest ($P < 0.05$). FO reduced the presence of *Clostridium perfringens*. The individual phytogetic additives did not prevent weight loss in birds challenged with *Eimeria*, but they mitigated the effects of the infection by modulating the intestinal microbiota. A synergistic effect was observed between the FO and anticoccidials, yielding satisfactory results in substituting virginiamycin. *Scientific Reports*, 13, 14612 (2023). <https://doi.org/10.1038/s41598-023-41743-9>.

7.1.2 Resumos publicados em anais de eventos

1. *Trypanosoma rangeli* transcriptomics: Novel surface proteins isoforms transcribed during *Rhodnius prolixus* infection. 2023.

FIGUEREDO, B.; LIZ, L. V.; KAWAGOE, E. K.; MAIA, G. A.; FILHO V. B.; GRISARD, E. C.; WAGNER, G.; STOCO, P. H. X-Meeting / Brazilian Symposium on Bioinformatics 2023.

2. Fast Protein: An automated pipeline for proteomic analysis. 2023.

MOREIRA, R. S.; **BENETTI FILHO, V.**; MAIA, G. A.; SORATTO, T. A. T.; KAWAGOE, E. K.; WAGNER, G.; MILETTI, LUIZ CLAUDIO. X-Meeting / Brazilian Symposium on Bioinformatics 2023. X-Meeting / Brazilian Symposium on Bioinformatics 2023.

3. Software Demonstration: EpiBuilder: A Tool for Assembling, Searching, and Classifying B-Cell Epitopes. 2023.

MOREIRA, R. S.; **BENETTI FILHO, V.**; CALOMENO, NATHÁLIA ANDERSON; MILETTI, LUIZ CLAUDIO; WAGNER, G. X-Meeting / Brazilian Symposium on Bioinformatics 2023.

4. Genomic surveillance in Santa Catarina: Healthcare workers as sentinel group for monitoring SARS-CoV-2 variants among the population. 2023.

F. H. BARAZZETTI; D. A. PADILHA; D. S. M SOUZA; E. K. KAWAGOE; **V. B. FILHO**; A. N. AMORIM; M. A. SHORNER; S. B. FERNANDES; B. K. COELHO; D. B. ROVARIS; M. P. D. dos ANJOS; J. R. MOSER; F. R. MELO; B. B. de SOUZA; D. C. BESSA; F. H. P. S. MENDES; A. C. BOING; A. F. BOING; J. T. de LACERDA; G. V. MOURA; D. C. de BASTIANI; M. H. de MORAES; L. F. V. de OLIVEIRA; R. S. MOREIRA; P. H. STOCO; M. L. BAZZO; G. FONGARO; G. WAGNER. X-Meeting / Brazilian Symposium on Bioinformatics 2023.

5. Software Demonstration: AnnotaPipeline: An integrated tool to annotate eukaryotic proteins using multi-omics data. 2023.

MAIA, G. A.; **BENETTI FILHO, V.**; KAWAGOE, E. K.; SORATTO, T. A. T.; MOREIRA, R. S.; GRISARD, E. C.; WAGNER, G. X-Meeting / Brazilian Symposium on Bioinformatics 2023.

6. Sewage metavirome: Viral recovery using Polyethylene glycol concentration methods and nanopore sequencing. 2023.

FILHO V. B.; PADILHA, D. A.; CADAMURO, R. D.; BARAZZETTI, F. H.; JESUS, J. G.; FONGARO, G.; WAGNER, G. X-Meeting / Brazilian Symposium on Bioinformatics 2023.

7. Sugar nucleotide biosynthesis pathways diverge among trypanosomatids.

FIGUEREDO, B.; LIZ, L. V.; MAIA, G. A.; **BENETTI FILHO, V.**; GRISARD, E. C.; STOCO, P. H. XXXVII Annual Meeting of the Brazilian Society of Protozoology // XLVIII Annual Meeting on basic research in Chagas Disease, 2022.

8. Sequence of SARS-CoV-2 infection in Southern Brazil.

PADILHA, D. A.; MOREIRA, R. S.; **BENETTI FILHO, V.**; SORATTO, T. A. T.; WACHTER, J. K.; MARTINS, C. L.; BARAZZETTI, F. H.; SCHORNER, M.; FONGARO, G.; WAGNER, G. XXXII Congresso Brasileiro de Virologia - Virologia em Casa, 2021.

7.1.3 Participação Em Eventos

1. **3º LeishSul - Simpósio de Leishmanioses da Região Sul.** Florianópolis – Santa Catarina, 2023. (Simpósio).

2. **X-Meeting / Brazilian Symposium on Bioinformatics,** Curitiba – Paraná, 2023. (Congresso).

3. **VIII Simpósio de Integração das Pós-Graduações do Centro de Ciências Biológicas.** Florianópolis – Santa Catarina, 2022. (Simpósio).

7.2 Curso De Curta Duração Ministrado

1. **Sequenciamento Genético, A prática clínica na identificação de patógenos infectocontagiosos.**

D. A. PADILHA.; **V. B. FILHO.** I Curso de Inverno em Biotecnologia e Biociências – UFSC (2023).

Folha 05/05

DISTRITOS	POPULAÇÃO (habitantes)/ANO														
	2024			2025			2026			2027			2028		
	RES	FLUT	Soma	RES	FLUT	Soma	RES	FLUT	Soma	RES	FLUT	Soma	RES	FLUT	Soma
Sede Ilha	286.537	28.723	315.260	291.720	29.674	321.394	296.067	30.555	326.622	300.478	31.462	331.940	304.955	32.396	337.351
Sede Continente	118.829	13.860	132.689	120.515	14.264	134.779	122.166	14.670	136.836	123.840	15.088	138.928	125.536	15.517	141.053
Barra da Lagoa	9.281	14.251	23.532	9.402	14.650	24.052	9.515	15.042	24.557	9.629	15.444	25.073	9.745	15.857	25.602
Cachoeira Bom Jesus	31.185	80.290	111.475	31.416	82.076	113.492	31.633	83.846	115.479	31.851	85.654	117.505	32.071	87.501	119.572
Campeche	52.732	18.967	71.699	53.947	19.471	73.418	54.940	20.063	75.003	55.951	20.673	76.624	56.980	21.301	78.281
Canasvieiras	38.841	146.516	185.357	39.322	150.519	189.841	39.766	154.437	194.203	40.216	158.457	198.673	40.670	162.581	203.251
Ingleses	52.152	104.502	156.654	52.933	107.611	160.544	53.568	110.488	164.056	54.211	113.443	167.654	54.862	116.476	171.338
Lagoa da Conceição	23.355	28.549	51.904	23.738	29.433	53.171	24.082	30.295	54.377	24.431	31.182	55.613	24.786	32.095	56.881
Pântano do Sul	11.488	4.028	15.516	11.620	4.135	15.755	11.742	4.231	15.973	11.865	4.329	16.194	11.990	4.430	16.420
Ratones	7.418	1.851	9.269	7.577	1.944	9.521	7.720	2.010	9.730	7.866	2.077	9.943	8.015	2.148	10.163
Ribeirão da Ilha	29.671	10.405	40.076	29.853	10.620	40.473	30.023	10.836	40.859	30.194	11.056	41.250	30.366	11.281	41.647
Rio Vermelho	22.781	11.938	34.719	23.441	12.465	35.906	24.034	12.967	37.001	24.642	13.488	38.130	25.266	14.031	39.297
Santo Antônio Lisboa	9.630	2.767	12.397	9.743	2.853	12.596	9.883	2.936	12.819	10.026	3.022	13.048	10.170	3.110	13.280
Soma	693.900	466.646	1.160.546	705.227	479.715	1.184.942	715.139	492.375	1.207.514	725.200	505.375	1.230.575	735.411	518.724	1.254.135

Documento oficial da Prefeitura Municipal De Florianópolis – SC, Secretaria Municipal De Habitação E Saneamento - S M H S A, disponível em: https://www.pmf.sc.gov.br/arquivos/arquivos/pdf/23_09_2010_18.24.35.20a6811bd7b11ef683f3ad54ee16087b.pdf,>. Acesso em: 15/08/2023.

Tabela 11. Quantificação de DNA e RNA e relações de pureza após extração de ácidos nucleicos.

Experimento	Amostra	Concentração		Relação 260/280		Relação 260/230	
		DNA	RNA	DNA	RNA	DNA	RNA
1	Esgoto 1 - [10 ⁷]	21,5	16,8	2,41	2,89	0,01	0,01
1	Esgoto 1 - [10 ⁵]	15,5	12	3,39	3,90	0,01	0,01
1	Esgoto 1 - [10 ³]	20,5	14,4	2,18	3,05	0,02	0,01
1	Esgoto 2 - [10 ⁷]	17	14,8	2,26	2,55	0,01	0,01
1	Esgoto 2 - [10 ⁵]	61	16,4	1,34	3,72	0,03	0,01
1	Esgoto 2 - [10 ³]	55,5	13,6	1,29	3,61	0,03	0,01
1	Pool controles - [10 ⁷]	60	47,2	1,23	1,38	0,03	0,01
1	Pool controles - [10 ⁵]	55	11,2	1,29	1,01	0,03	0,01
1	Pool controles - [10 ³]	58	11,4	1,30	1,91	0,03	0,01
2	CB	37	30	3,3	4	0,02	0,02
2	PEG + CB	47	40	5	4,1	0,02	0,03
2	PEG + DNASE	43	30	3,6	4	0,02	0,02
2	Controle positivo	47	35	3,9	3,8	0,02	0,02
3	Janeiro	34,5	-	2,4	-	0,02	-
3	Fevereiro	21	-	1,7	-	0,03	-
3	Março	35,5	-	4,2	-	0,02	-
3	Abril	22,5	-	1,8	-	0,01	-
3	Mai	37	-	1,8	-	0,02	-
3	Junho	23	-	2,1	-	0,02	-
3	Agosto	28	-	3,5	-	0,02	-
3	Outubro	14,1	-	1	-	0,01	-
3	Novembro	23,5	-	3,1	-	0,01	-
3	Dezembro	45,5	-	1,7	-	0,03	-

Legenda: A relação 260/280 mostra a relação entre material genético e proteína. A relação 260/230 mostra a relação entre material genético e compostos orgânicos.

Tabela 12. Quantificação de DNA complementar realizada por meio do equipamento Qubit.

Experimento	Amostra	ng/μL
1	Esgoto 1 - [10 ⁷]	712,4
1	Esgoto 1 - [10 ⁵]	234,8
1	Esgoto 1 - [10 ³]	262,6
1	Esgoto 2 - [10 ⁷]	963,2
1	Esgoto 2 - [10 ⁵]	368,1
1	Esgoto 2 - [10 ³]	442,8
1	Pool controles - [10 ⁷]	398,0
1	Pool controles - [10 ⁵]	179,1
1	Pool controles - [10 ³]	712,4
2	CB	232,8
2	PEG + CB	218,9
2	PEG + DNASE	151,1
2	Controle positivo	362,2
3	Janeiro	252,7
3	Fevereiro	86,0
3	Março	93,5
3	Abril	96,0
3	Maiο	181,1
3	Junho	158,4
3	Agosto	115,4
3	Outubro	72,0
3	Novembro	70,8
3	Dezembro	126,6

Tabela 13. Proporção de sequenciamento e classificação taxonômica nas duplicatas em abordagem com três concentrações de controles virais.

Amostra	FL Totais	FL não classificados	FL classificados	Classificado (%)	Não classificado (%)	Vírus	Vírus (%)	Eucariota	Eucariota (%)	Bactéria	Bactéria (%)
Controle [] 10 ⁷	69503	38115	31388	45,16	54,84	2775	8,84	3591	11,44	25022	79,72
Controle [] 10 ⁵	27176	17712	9464	34,82	65,18	127	1,34	543	5,74	8794	92,92
Controle [] 10 ³	83946	46407	37539	44,72	55,28	128	0,34	694	1,85	36717	97,81
Esgoto 1 [] 10 ⁷	65737	27862	37875	57,62	42,38	3528	9,31	2948	7,78	31399	82,90
Esgoto 2 [] 10 ⁷	73199	32283	40916	55,90	44,10	2486	6,08	2215	5,41	36215	88,51
Média [] 10⁷	69468	30072,5	39395,5	56,76	43,24	3007	7,70	2581,5	6,60	33807	85,71
Esgoto 1 [] 10 ⁵	46090	24161	21929	47,58	52,42	324	1,48	500	2,28	21105	96,24
Esgoto 2 [] 10 ⁵	65260	35240	30020	46,00	54,00	524	1,75	773	2,57	28723	95,68
Média [] 10⁵	55675	29700,5	25974,5	46,79	53,21	424	1,61	636,5	2,43	24914	95,96
Esgoto 1 [] 10 ³	48362	31051	17311	35,79	64,21	425	2,46	640	3,70	16246	93,85
Esgoto 2 [] 10 ³	58618	32833	25785	43,99	56,01	215	0,83	565	2,19	25005	96,97
Média [] 10³	53490	31942	21548	39,89	60,11	320	1,64	602,5	2,94	20625,5	95,41
Indeterminada	31010	-	-	-	-	-	-	-	-	-	-
Total	568901	-	-	-	-	-	-	-	-	-	-

FL – Fragmentos de leitura; [] – Concentração de cópias genômicas; indeterminada significa amostras sequenciadas, mas sem identificação do *barcode* de origem. Em negrito as médias das duplicatas de esgoto nas diferentes concentrações.

Tabela 14. Proporção de sequenciamento e classificação taxonômica na abordagem de descontaminação associada ao PEG.

Amostra	FL Totais	FL não classificados	FL classificados	Classificado (%)	Não classificado (%)	Vírus	Vírus (%)	Eucariota	Eucariota (%)	Bactéria	Bactéria (%)
Controle positivo	67275	32545	34730	51,62	48,38	11834	34,07	5026	14,47	17870	51,45
CB	30246	16171	14075	46,54	53,46	3342	23,74	1844	13,10	8889	63,15
CB + PEG + DNase	32946	19338	13608	41,30	58,70	880	6,47	905	6,65	11823	86,88
CB + PEG	32848	16995	15853	48,26	51,74	4892	30,86	2516	15,87	8445	53,27
Indeterminada	9418	-	-	-	-	-	-	-	-	-	-
Total	172733	-	-	-	-	-	-	-	-	-	-

FL – Fragmentos de leitura; CB – Tratamento com Clorofórmio Butanol; PEG – Polietileno Glicol; DNase – Tratamento com enzima DNase; [] – Concentração de cópias genômicas. Indeterminada significa amostras sequenciadas, mas sem identificação do *barcode* de origem.

Tabela 15. Proporções de sequenciamento e classificação taxonômica nas amostras de oriundas de Canasvieiras no período de 2021.

Mês	FL Totais	FL não classificados	FL classificados	Classificado (%)	Não classificado (%)	Virus	Vírus (%)	Eucariota	Eucariota (%)	Bactéria	Bactéria (%)
Janeiro	82345	53076	29269	35,54	64,46	689	2,35	1823	6,23	26757	91,42
Fevereiro	28578	20128	8450	29,57	70,43	91	1,08	1092	12,92	7267	86,00
Março	14118	9245	4873	34,52	65,48	26	0,53	527	10,81	4320	88,65
Abril	24127	16452	7675	31,81	68,19	280	3,65	1418	18,48	5977	77,88
Mai	80114	48062	32052	40,01	59,99	465	1,45	1618	5,05	29969	93,50
Junho	125210	87418	37792	30,18	69,82	1205	3,19	2768	7,32	33819	89,49
Agosto	128744	75195	53549	41,59	58,41	315	0,59	2586	4,83	50648	94,58
Outubro	50813	33734	17079	33,61	66,39	313	1,83	1231	7,21	15535	90,96
Novembro	52293	34253	18040	34,50	65,50	164	0,91	1609	8,92	16267	90,17
Dezembro	13250	8519	4731	35,71	64,29	32	0,68	454	9,60	4245	89,73
Indeterminada	44322	-	-	-	-	-	-	-	-	-	-
Total	643914	-	-	-	-	-	-	-	-	-	-

FL – Fragmentos de leitura. Indeterminada significa amostras sequenciadas, mas sem identificação do *barcode* de origem.

Tabela 16. Dados de saída dos sequenciamentos por shotgun em plataforma MinION.

Experimento	Total de FL	Tamanho médio de FL	Bases totais (Mb)	Qualidade <i>phred</i> média	Número de amostras	Média de FL por amostra
1	568.901	314	178.8	11,09	9	63.211
2	172.733	331	57.2	11,01	4	43.183
3	643.914	229	147.8	10,51	10	64.391

FL – Fragmentos de leitura. Mb – Milhões de bases

9 ANEXOS

**ANEXO 1 - Protocolo de Sequenciamento Nanopore *Rapid sequencing gDNA* -
barcoding (SQK-RBK004)**

Overview of the protocol

Rapid Barcoding Kit features

This kit is recommended for users who:

- wish to multiplex samples to reduce price per sample
- need a PCR-free method of multiplexing to preserve additional information such as base modifications
- require a short preparation time
- have limited access to laboratory equipment

Introduction to the Rapid Barcoding Sequencing Kit

This protocol describes how to carry out rapid barcoding of genomic DNA using the Rapid Barcoding Sequencing Kit (SQK-RBK004).

Steps in the sequencing workflow:

Prepare for your experiment

You will need to:

- Extract your DNA, and check its length, quantity and purity.

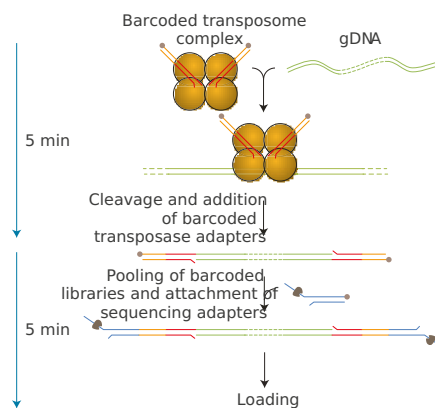
The quality checks performed during the protocol are essential in ensuring experimental success.

- Ensure you have your sequencing kit, the correct equipment and third-party reagents
- Download the software for acquiring and analysing your data
- Check your flow cell to ensure it has enough pores for a good sequencing run

Library preparation

You will need to:

- Tagment your DNA using the Fragmentation Mix RB in the kit; this simultaneously attaches a pair of barcodes to the fragments
- Pool the barcoded samples
- Attach sequencing adapters supplied in the kit to the DNA ends
- Prime the flow cell, and load your DNA library into the flow cell



Sequencing and analysis

You will need to:

- Start a sequencing run using the MinKNOW software, which will collect raw data from the device and convert it into basecalled reads
- Start the EPI2ME software and select a workflow for barcoding

IMPORTANT

Compatibility of this protocol

This protocol should only be used in combination with:

- Rapid Barcoding Sequencing Kit (SQK-RBK004)
- R9.4.1 (FLO-MIN106) flow cells
- Flow Cell Wash Kit (EXP-WSH004)
- RAP Top-Up Kit (EXP-RAP001)
- Sequencing Auxiliary Kit (EXP-AUX001)

Equipment and consumables

Materials

- ~400 ng high molecular weight genomic DNA
- Rapid Barcoding Sequencing Kit (SQK-RBK004)
- Flow Cell Priming Kit (EXP-FLP002)

Consumables

- 1.5 ml Eppendorf DNA LoBind tubes
- 0.2 ml thin-walled PCR tubes
- Nuclease-free water (e.g. ThermoFisher, cat # AM9937)
- Agencourt AMPure XP beads (optional)
- Freshly prepared 70% ethanol in nuclease-free water (optional)
- 10 mM Tris-HCl pH 8.0 with 50 mM NaCl (optional)

Equipment

- Ice bucket with ice
- Microfuge
- Timer
- Thermal cycler or heat block at 30°C and 80°C
- P1000 pipette and tips
- P200 pipette and tips
- P100 pipette and tips
- P20 pipette and tips
- P2 pipette and tips

Optional Equipment

- Standard gel electrophoresis equipment
- Agilent Bioanalyzer (or equivalent)
- Qubit fluorometer (or equivalent for QC check)
- Eppendorf 5424 centrifuge (or equivalent)
- Magnetic rack
- Hula mixer (gentle rotator mixer)

For this protocol, you will need ~400 ng high molecular weight genomic DNA

Input DNA

How to QC your input DNA

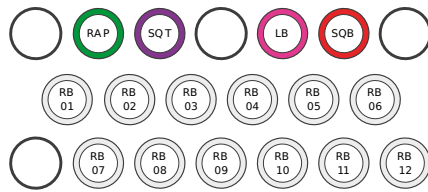
It is important that the input DNA meets the quantity and quality requirements. Using too little or too much DNA, or DNA of poor quality (e.g. highly fragmented or containing RNA or chemical contaminants) can affect your library preparation.

For instructions on how to perform quality control of your DNA sample, please read the [input DNA/RNA QC protocol](#).

Chemical contaminants

Depending on how the DNA is extracted from the raw sample, certain chemical contaminants may remain in the purified DNA, which can affect library preparation efficiency and sequencing quality. Read more about contaminants on the [Contaminants page](#) of the Community.

Rapid Barcoding Kit (SQK-RBK004) contents



RAP : Rapid adapter
 SQT : Sequencing tether
 LB : Loading beads
 SQB : Sequencing buffer
 RB01 : Fragmentation Mix RB 1
 RB02 : Fragmentation Mix RB 2
 RB03 : Fragmentation Mix RB 3
 RB04 : Fragmentation Mix RB 4
 RB05 : Fragmentation Mix RB 5
 RB06 : Fragmentation Mix RB 6
 RB07 : Fragmentation Mix RB 7
 RB08 : Fragmentation Mix RB 8
 RB09 : Fragmentation Mix RB 9
 RB10 : Fragmentation Mix RB 10
 RB11 : Fragmentation Mix RB 11
 RB12 : Fragmentation Mix RB 12

Name	Acronym	Cap colour	No. of vials	Fill volume per vial (µl)
Fragmentation Mix	RB01-RB12	Clear	12, each vial with a unique barcode	20
Rapid Adapter	RAP	Green	1	10
Sequencing Tether	SQT	Purple	1	10
Loading Beads	LB	Pink	1	360
Sequencing Buffer	SQB	Red	1	300

Rapid Barcoding Kit sequences

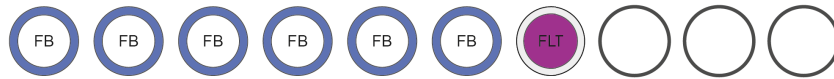
Component	Sequence
RB01	AAGAAAGTTGTCGGTGTCTTTGTG
RB02	TCGATTCCGTTTGTAGTCGTCTGT
RB03	GAGTCTTGTGTCCCAGTTACCAGG

Component	Sequence
RB04	TTCGGATTCTATCGTGTTCCCTA
RB05	CTTGTCAGGGTTTGTGTAACCTT
RB06	TTCTCGCAAAGGCAGAAAGTAGTC
RB07	GTGTTACCGTGGAATGAATCCTT
RB08	TTCAGGGAACAAACCAAGTTACGT
RB09	AACTAGGCACAGCGAGTCTTGTT
RB10	AAGCGTTGAAACCTTTGTCCTCTC
RB11	GTTTCATCTATCGGAGGGAATGGA
RB12	CAGGTAGAAAGAAGCAGAATCGGA

IMPORTANT

Please note that the Sequencing Tether (SQT) tube will NOT be used in this protocol.

Flow Cell Priming Kit contents (EXP-FLP002)

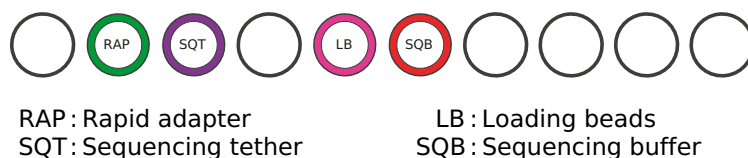


FLB : Flush buffer
 FLT : Flush tether

Name	Acronym	Cap colour	No. of vials	Fill volume per vial (µl)
Flush Buffer	FB	Blue	6	1,170
Flush Tether	FLT	Purple	1	200

The RAP Top-Up Kit (EXP-RAP001) is available to provide enough reagents for another six reactions depending on how the barcodes are used.

This kit contains reagents to be used with any remaining barcodes to load another six sequencing libraries.



Reagent	Acronym	Cap colour	No. of vials	Fill volume per vial (µl)
Rapid Adapter	RAP	Green	1	10
Sequencing Tether	SQT	Purple	1	10
Loading Beads	LB	Pink	1	360
Sequencing Buffer	SQB	Red	1	300

Computer requirements and software

MinION Mk1B IT requirements

Sequencing on a MinION Mk1B requires a high-spec computer or laptop to keep up with the rate of data acquisition. Read more in the [MinION IT Requirements document](#).

MinION Mk1C IT requirements

The MinION Mk1C contains fully-integrated compute and screen, removing the need for any accessories to generate and analyse nanopore data. Read more in the [MinION Mk1C IT requirements document](#).

Software for nanopore sequencing

MinKNOW

The MinKNOW software controls the nanopore sequencing device, collects sequencing data in real time and processes it into basecalls. You will be using MinKNOW for every sequencing experiment. MinKNOW can also demultiplex reads by barcode, and basecall/demultiplex data after a sequencing run has completed.

MinKNOW use

For instructions on how to run the MinKNOW software, please refer to the relevant section in the [MinKNOW protocol](#).

EPI2ME (optional)

The EPI2ME cloud-based platform performs further analysis of basecalled data, for example alignment to the Lambda genome, barcoding, or taxonomic classification. You will use the EPI2ME platform *only* if you would like further analysis of your data post-

basecalling.

EPI2ME installation and use

For instructions on how to create an EPI2ME account and install the EPI2ME Desktop Agent, please refer to the [EPI2ME Platform protocol](#).

Guppy (optional)

The Guppy command-line software can be used for basecalling and demultiplexing reads by barcode instead of MinKNOW. You can use it if you would like to re-analyse old data, or integrate basecalling into your analysis pipeline.

Guppy installation and use

If you would like to use the Guppy software, please refer to the [Guppy protocol](#).

Check your flow cell

We highly recommend that you check the number of pores in your flow cell prior to starting a sequencing experiment. This should be done within three months of purchasing for MinION/GridION/PromethION flow cells, or within four weeks of purchasing for Flongle flow cells. Oxford Nanopore Technologies will replace any flow cell with fewer than the number of pores in the table below, when the result is reported within two days of performing the flow cell check, and when the storage recommendations have been followed. To do the flow cell check, please follow the instructions in the [Flow Cell Check document](#).

Flow cell	Minimum number of active pores covered by warranty
Flongle Flow Cell	50
MinION/GridION Flow Cell	800
PromethION Flow Cell	5000

Library preparation

~10 minutes

Materials

- 400 ng high molecular weight plasmid DNA
- Fragmentation Mix RB01-12
- Rapid Adapter (RAP)

Consumables

- 0.2 ml thin-walled PCR tubes
- Nuclease-free water (e.g. ThermoFisher, cat # AM9937)
- Agencourt AMPure XP beads (optional)
- Freshly prepared 70% ethanol in nuclease-free water (optional)
- 10 mM Tris-HCl pH 8.0 with 50 mM NaCl (optional)

Equipment

- Thermal cycler or heat block at 30°C and 80°C
- Microfuge
- P10 pipette and tips

Optional Equipment

- Magnetic rack
- Hula mixer (gentle rotator mixer)

1 Thaw kit components at room temperature, spin down briefly using a microfuge and mix by pipetting as indicated by the table below:

Reagent	1. Thaw at room temperature	2. Briefly spin down	3. Mix well by pipetting
Fragmentation Mix RB01-12	Not frozen	✓	✓
Rapid Adapter (RAP)	Not frozen	✓	✓
Sequencing Buffer (SQB)	✓	✓	✓*
Loading Beads (LB)	✓	✓	Mix by pipetting or vortexing immediately before use
Flush Buffer (FLB) - 1 tube	✓	✓	✓*
Flush Tether (FLT)	✓	✓	✓

*Vortexing, followed by a brief spin in a microfuge, is recommended for Sequencing Buffer (SQB) and Flush Buffer (FLB).

Please note that the Sequencing Tether (SQT) tube will NOT be used in this protocol. It is provided in the kit for potential future product compatibility.

2 Prepare the DNA in nuclease-free water.

- Transfer ~400 ng genomic DNA into a DNA LoBind tube
- Adjust the volume to 7.5 µl with nuclease-free water
- Mix by flicking the tube to avoid unwanted shearing
- Spin down briefly in a microfuge

3 In a 0.2 ml thin-walled PCR tube, mix the following:

Reagent	Volume
400 ng template DNA	7.5 µl
Fragmentation Mix RB01-12 (one for each sample)	2.5 µl
Total	10 µl

4 Mix gently by flicking the tube, and spin down.

5 Incubate the tube at 30°C for 1 minute and then at 80°C for 1 minute. Briefly put the tube on ice to cool it down.

6 Pool the barcoded samples in a 1.5 ml Eppendorf DNA LoBind tube.

We expect to have about ~10 µl per sample.

IMPORTANT

If barcoding four or more samples, increased throughput can be achieved through cleaning up and concentrating the pooled material using AMPure XP beads as outlined in Steps 6-17. Otherwise, for a more rapid sample preparation, transfer 10 µl of pooled sample from Step 6 into a clean 1.5 ml Eppendorf DNA LoBind tube, and proceed directly to Step 18.

7 Resuspend the AMPure XP beads by vortexing.

8 To the entire pooled barcoded sample from Step 6, add an equal volume of resuspended AMPure XP beads, and mix by flicking the tube.

9 Incubate on a Hula mixer (rotator mixer) for 5 minutes at room temperature.

10 Prepare 500 µl of fresh 70% ethanol in nuclease-free water.

11 Spin down the sample and pellet on a magnet. Keep the tube on the magnet, and pipette off the supernatant.

12 Keep the tube on the magnet and wash the beads with 200 µl of freshly prepared 70% ethanol without disturbing the pellet. Remove the ethanol using a pipette and discard.

13 Repeat the previous step.

14 Spin down and place the tube back on the magnet. Pipette off any residual 70% ethanol. Briefly allow to dry.

15 Remove the tube from the magnetic rack and resuspend pellet in 10 µl of 10 mM Tris-HCl pH 7.5-8.0 with 50 mM NaCl. Incubate for 2 minutes at room temperature.

16 Pellet the beads on a magnet until the eluate is clear and colourless, for at least 1 minute.

17 Remove and retain 10 µl of eluate into a clean 1.5 ml Eppendorf DNA LoBind tube.

- Dispose of the pelleted beads

End of optional steps.

18 Add 1 µl of RAP to 10 µl of barcoded DNA.

19 Mix gently by flicking the tube, and spin down.

20 Incubate the reaction for 5 minutes at room temperature.

END OF STEP

The prepared library is used for loading into the MinION flow cell. Store the library on ice until ready to load.

Priming and loading the SpotON Flow Cell

~10 minutes

Materials

- Flow Cell Priming Kit (EXP-FLP002)
- Sequencing Buffer (SQB)
- Loading Beads (LB)

Consumables

- 1.5 ml Eppendorf DNA LoBind tubes
- Nuclease-free water (e.g. ThermoFisher, cat # AM9937)

Equipment

- MinION Mk1B
- SpotON Flow Cell
- P1000 pipette and tips
- P100 pipette and tips
- P20 pipette and tips
- P10 pipette and tips

IMPORTANT

Please note that the Sequencing Tether (SQT) tube will NOT be used in this protocol.

TIP

Priming and loading a flow cell

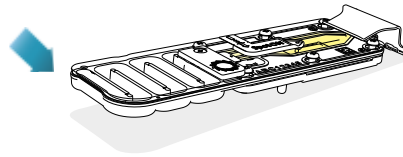
We recommend all new users watch the [Priming and loading your flow cell](#) video before your first run.

- 1 Thaw the Sequencing Buffer (SQB), Loading Beads (LB), Flush Tether (FLT) and one tube of Flush Buffer (FB) at room temperature before mixing the reagents by vortexing, and spin down at room temperature.
- 2 To prepare the flow cell priming mix, add 30 μ l of thawed and mixed Flush Tether (FLT) directly to the tube of thawed and mixed Flush Buffer (FB), and mix by vortexing at room temperature.

3 Open the MinION device lid and slide the flow cell under the clip.

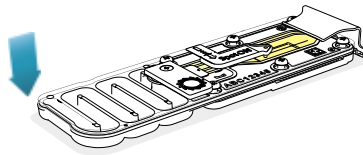
Press down firmly on the flow cell to ensure correct thermal and electrical contact.

1a
Insert the flow cell into the device under the clip and press down firmly.



Key:
● Storage buffer
● Priming mix
● DNA library

1b
Insert the flow cell into the device under the clip and press down firmly.



Optional Action

Complete a flow cell check to assess the number of pores available before loading the library.

This step can be omitted if the flow cell has been checked previously.

Please see the MinKNOW [Flow cell check protocol](#) for more information

4 Slide the priming port cover clockwise to open the priming port.

IMPORTANT

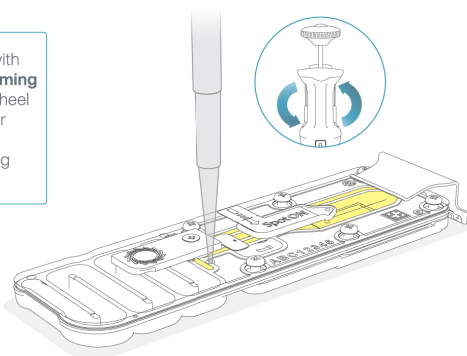
Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 μl , and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores.

5 After opening the priming port, check for a small air bubble under the cover. Draw back a small volume to remove any bubbles (a few μl):

1. Set a P1000 pipette to 200 μl
2. Insert the tip into the priming port
3. Turn the wheel until the dial shows 220-230 μl , to draw back 20-30 μl , or until you can see a small volume of buffer entering the pipette tip

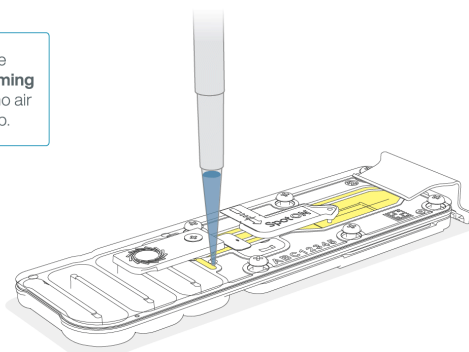
Note: Visually check that there is continuous buffer from the priming port across the sensor array.

3 Insert a P1000 pipette with an empty tip into the **Priming port**. Turn the pipette wheel to draw back 20-30 μl or until you can see a small volume of buffer entering the pipette tip.



- 6 Load 800 μl of the priming mix into the flow cell via the priming port, avoiding the introduction of air bubbles. Wait for 5 minutes. During this time, prepare the library for loading by following the steps below.

4 Slowly load 800 μl of the priming mix into the **Priming port**. Ensure there are no air bubbles in the pipette tip.



Wait 5 minutes before proceeding to the next step.

- 7 Thoroughly mix the contents of the Loading Beads (LB) tubes by vortexing.

- 8 In a new tube, prepare the library for loading as follows:

Reagent	Volume
Sequencing Buffer (SQB)	34 μl
Loading Beads (LB), mixed immediately before use	25.5 μl
Nuclease-free water	4.5 μl
DNA library	11 μl
Total	75 μl

Note: Load the library onto the flow cell immediately after adding the Sequencing Buffer (SQB) and Loading Beads (LB) because the fuel in the buffer will start to be consumed by the adapter.

IMPORTANT

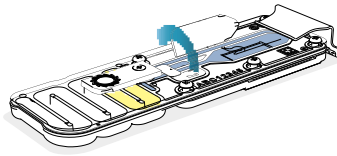
The Loading Beads (LB) tube contains a suspension of beads. These beads settle very quickly. It is vital that they are mixed immediately before use.

9 Complete the flow cell priming:

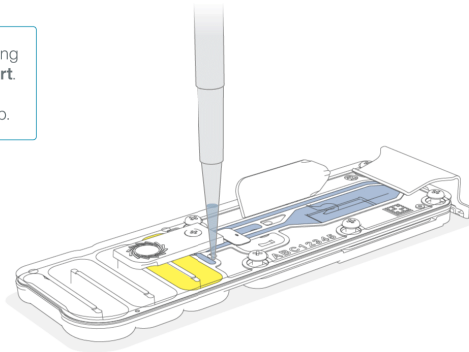
1. Gently lift the SpotON sample port cover to make the SpotON sample port accessible.
2. Load **200 μ l** of the priming mix into the flow cell via the priming port (**not** the SpotON sample port), avoiding the introduction of air bubbles.

Note: Load the library as soon as possible after this step.

5 Gently flip open SpotON sample port cover.



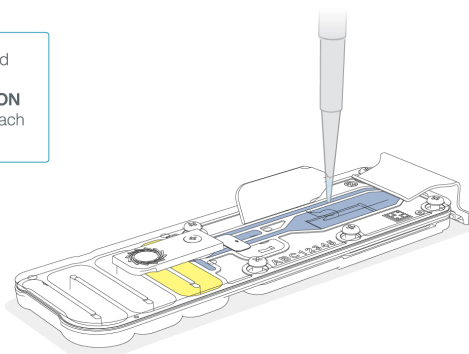
6 Load 200 μ l of the priming mix into the **Priming Port**. Ensure there are no air bubbles in the pipette tip.



10 Mix the prepared library gently by pipetting up and down just prior to loading.

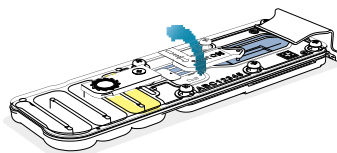
11 Add 75 µl of sample to the flow cell via the SpotON sample port in a dropwise fashion. Ensure each drop flows into the port before adding the next.

7 Pipette mix the prepared library and load 75 µl dropwise into the **SpotON** sample port, ensuring each drop flows into the port.

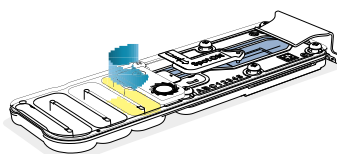


12 Gently replace the SpotON sample port cover, making sure the bung enters the SpotON port, close the priming port and replace the MinION device lid.

8 Gently replace the **SpotON** sample port cover.



9 Gently close the **Priming port**.



Data acquisition and basecalling

Overview of nanopore data analysis

For a full overview of nanopore data analysis, which includes options for basecalling and post-basecalling analysis, please refer to the [Data Analysis](#) document.

How to start sequencing

The sequencing device control, data acquisition and real-time basecalling are carried out by the MinKNOW software. It is assumed you have already installed MinKNOW on your computer, or that you are using the MinIT device for data acquisition and basecalling. There are three options for how to carry out sequencing:

1. Data acquisition and basecalling in real-time using MinKNOW on a computer

Follow the instructions in the [MinKNOW protocol](#) beginning from the "Starting a sequencing run" section until the end of the "Completing a MinKNOW run" section.

2. Data acquisition and basecalling in real-time using the GridION device

Follow the instructions in the [GridION user manual](#).

3. Data acquisition and basecalling in real-time using the MinION Mk1C device

Follow the instructions in the [MinION Mk1C user manual](#).

4. Data acquisition and basecalling in real-time using the MinIT device

Follow the instructions in the [MinIT protocol](#).

5. Data acquisition using MinKNOW on a computer and basecalling at a later time using MinKNOW or Guppy

Follow the instructions in the [MinKNOW protocol](#) beginning from the "Starting a sequencing run" section until the end of the "Completing a MinKNOW run" section. **When setting your experiment parameters, set the *Basecalling* tab to OFF.** After the sequencing experiment has completed, follow the instructions in the [Post-run analysis](#) section of the [MinKNOW protocol](#) or the [Guppy protocol](#) starting from the "Quick Start Guide for Guppy" section.

Downstream analysis

Post-basecalling analysis

There are several options for further analysing your basecalled data:

1. EPI2ME platform

The EPI2ME platform is a cloud-based data analysis service developed by Metrichor Ltd., a subsidiary of Oxford Nanopore Technologies. The EPI2ME platform offers a range of analysis workflows, e.g. for metagenomic identification, barcoding, alignment,

and structural variant calling. The analysis requires no additional equipment or compute power, and provides an easy-to-interpret report with the results. For instructions on how to run an analysis workflow in EPI2ME, please follow the instructions in the [EPI2ME protocol](#), beginning at the "Starting data analysis" step.

2. EPI2ME Labs tutorials and workflows

For more in-depth data analysis, Oxford Nanopore Technologies offers a range of bioinformatics tutorials and workflows available in EPI2ME Labs, which are available in the [EPI2ME Labs](#) section of the Community. The platform provides a vehicle where workflows deposited in GitHub by our Research and Applications teams can be showcased with descriptive texts, functional bioinformatics code and example data.

3. Research analysis tools

Oxford Nanopore Technologies' Research division has created a number of analysis tools, which are available in the Oxford Nanopore [GitHub repository](#). The tools are aimed at advanced users, and contain instructions for how to install and run the software. They are provided as-is, with minimal support.

4. Community-developed analysis tools

If a data analysis method for your research question is not provided in any of the resources above, please refer to the [Bioinformatics section of the Resource centre](#). Numerous members of the Nanopore Community have developed their own tools and pipelines for analysing nanopore sequencing data, most of which are available on GitHub. Please be aware that these tools are not supported by Oxford Nanopore Technologies, and are not guaranteed to be compatible with the latest chemistry/software configuration.

Ending the experiment

Materials

- Flow Cell Wash Kit (EXP-WSH004)

- 1 **After your sequencing experiment is complete, if you would like to reuse the flow cell, please follow the Wash Kit instructions and store the washed flow cell at 2-8°C, OR**

The [Flow Cell Wash Kit protocol](#) is available on the Nanopore Community.

TIP

We recommend you to wash the flow cell as soon as possible after you stop the run. However, if this is not possible, leave the flow cell on the device and wash it the next day.

- 2 **Follow the returns procedure by washing out the flow cell ready to send back to Oxford Nanopore.**

Instructions for returning flow cells can be found [here](#).

IMPORTANT

If you encounter issues or have questions about your sequencing experiment, please refer to the Troubleshooting Guide that can be found in the online version of this protocol.

Issues during DNA/RNA extraction and library preparation

Below is a list of the most commonly encountered issues, with some suggested causes and solutions.

We also have an FAQ section available on the [Nanopore Community Support](#) section.

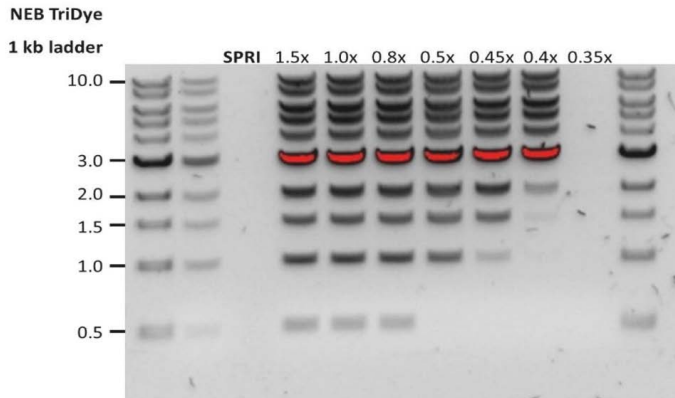
If you have tried our suggested solutions and the issue still persists, please contact Technical Support via email (support@nanoporetech.com) or via [LiveChat](#) in the Nanopore Community.

Low sample quality

Observation	Possible cause	Comments and actions
Low DNA purity (Nanodrop reading for DNA OD 260/280 is <1.8 and OD 260/230 is <2.0-2.2)	The DNA extraction method does not provide the required purity	The effects of contaminants are shown in the Contaminants Know-how piece. Please try an alternative extraction method that does not result in contaminant carryover. Consider performing an additional SPRI clean-up step.
Low RNA integrity (RNA integrity number <9.5 RIN, or the rRNA band is shown as a smear on the gel)	The RNA degraded during extraction	Try a different RNA extraction method). For more info on RIN, please see the RNA Integrity Number Know-how piece .
RNA has a shorter than expected fragment length	The RNA degraded during extraction	Try a different RNA extraction method). For more info on RIN, please see the RNA Integrity Number Know-how piece . We recommend working in an RNase-free environment, and to keep your lab equipment RNase-free when working with RNA.

Low DNA recovery after AMPure bead clean-up

Observation	Possible cause	Comments and actions
Low recovery	DNA loss due to a lower than intended AMPure beads-to-sample ratio	<ol style="list-style-type: none">1. AMPure beads settle quickly, so ensure they are well resuspended before adding them to the sample.2. When the AMPure beads-to-sample ratio is lower than 0.4:1, DNA fragments of any size will be lost during the clean-up.

Observation	Possible cause	Comments and actions
Low recovery	DNA fragments are shorter than expected	The lower the AMPure beads-to-sample ratio, the more stringent the selection against short fragments. Please always determine the input DNA length on an agarose gel (or other gel electrophoresis methods) and then calculate the appropriate amount of AMPure beads to use. 
Low recovery after end-prep	The wash step used ethanol <70%	DNA will be eluted from the beads when using ethanol <70%. Make sure to use the correct percentage.

The VoITRAX run terminated in the middle of the library prep

Observation	Possible cause	Comments and actions
The green light was switched off or An adapter was used to connect the VoITRAX USB-C cable to the computer	Insufficient power supply to the VoITRAX	The green LED signals that 3 A are being supplied to the device. This is the requirement for the full capabilities of the VoITRAX V2 device. Please use computers that meet the requirements listed on the VoITRAX V2 protocol .

The VoITRAX software shows an inaccurate amount of reagents loaded

Observation	Possible cause	Comments and actions
The VoITRAX software shows an inaccurate amount of reagents loaded	Pipette tips do not fit the VoITRAX cartridge ports	TRainin 20 µl or 30 µl and Gilson 10 µl, 20 µl or 30 µl pipette tips are compatible with loading reagents into the VoITRAX cartridge. Rainin 20 µl is the most suitable.
The VoITRAX software shows an inaccurate amount of reagents loaded	The angle at which reagents are pipetted into the cartridge is incorrect	The pipetting angle should be slightly greater than the cartridge inlet angle. Please watch the demo video included in the VoITRAX software before loading.

Issues during the sequencing run

Below is a list of the most commonly encountered issues, with some suggested causes and solutions.

We also have an FAQ section available on the [Nanopore Community Support](#) section.

If you have tried our suggested solutions and the issue still persists, please contact Technical Support via email (support@nanoporetech.com) or via [LiveChat](#) in the Nanopore Community.

Fewer pores at the start of sequencing than after Flow Cell Check

Observation	Possible cause	Comments and actions
MinKNOW reported a lower number of pores at the start of sequencing than the number reported by the Flow Cell Check	An air bubble was introduced into the nanopore array	After the Flow Cell Check it is essential to remove any air bubbles near the priming port before priming the flow cell. If not removed, the air bubble can travel to the nanopore array and irreversibly damage the nanopores that have been exposed to air. The best practice to prevent this from happening is demonstrated in this video .
MinKNOW reported a lower number of pores at the start of sequencing than the number reported by the Flow Cell Check	The flow cell is not correctly inserted into the device	Stop the sequencing run, remove the flow cell from the sequencing device and insert it again, checking that the flow cell is firmly seated in the device and that it has reached the target temperature. If applicable, try a different position on the device (GridION/PromethION).
MinKNOW reported a lower number of pores at the start of sequencing than the number reported by the Flow Cell Check	Contaminations in the library damaged or blocked the pores	The pore count during the Flow Cell Check is performed using the QC DNA molecules present in the flow cell storage buffer. At the start of sequencing, the library itself is used to estimate the number of active pores. Because of this, variability of about 10% in the number of pores is expected. A significantly lower pore count reported at the start of sequencing can be due to contaminants in the library that have damaged the membranes or blocked the pores. Alternative DNA/RNA extraction or purification methods may be needed to improve the purity of the input material. The effects of contaminants are shown in the Contaminants Know-how piece . Please try an alternative extraction method that does not result in contaminant carryover.

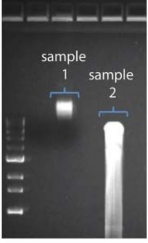
MinKNOW script failed

Observation	Possible cause	Comments and actions
MinKNOW shows "Script failed"		Restart the computer and then restart MinKNOW. If the issue persists, please collect the MinKNOW log files and contact Technical Support.

Pore occupancy below 40%

Observation	Possible cause	Comments and actions
Pore occupancy <40%	Not enough library was loaded on the flow cell	5–50 fmol of good quality library can be loaded on to a MinION Mk1B/GridION flow cell. Please quantify the library before loading and calculate mols using tools like the Promega Biomath Calculator , choosing "dsDNA: µg to pmol"
Pore occupancy close to 0	The Ligation Sequencing Kit was used, and sequencing adapters did not ligate to the DNA	Make sure to use the NEBNext Quick Ligation Module (E6056) and Oxford Nanopore Technologies Ligation Buffer (LNB, provided in the SQK-LSK109 kit) at the sequencing adapter ligation step, and use the correct amount of each reagent. A Lambda control library can be prepared to test the integrity of the third-party reagents.
Pore occupancy close to 0	The Ligation Sequencing Kit was used, and ethanol was used instead of LFB or SFB at the wash step after sequencing adapter ligation	Ethanol can denature the motor protein on the sequencing adapters. Make sure the LFB or SFB buffer was used after ligation of sequencing adapters.
Pore occupancy close to 0	No tether on the flow cell	Tethers are added during flow cell priming (FLT tube). Make sure FLT was added to FB before priming.

Shorter than expected read length

Observation	Possible cause	Comments and actions
Shorter than expected read length	Unwanted fragmentation of DNA sample	<p>Read length reflects input DNA fragment length. Input DNA can be fragmented during extraction and library prep.</p> <ol style="list-style-type: none"> 1. Please review the Extraction Methods in the Nanopore Community for best practice for extraction. 2. Visualise the input DNA fragment length distribution on an agarose gel before proceeding to the library prep.  <p>In the image above, Sample 1 is of high molecular weight, whereas Sample 2 has been fragmented.</p> <ol style="list-style-type: none"> 3. During library prep, avoid pipetting and vortexing when mixing reagents. Flicking or inverting the tube is sufficient.

Large proportion of recovering pores

Observation	Possible cause	Comments and actions
<p>Large proportion of recovering pores (shown as dark blue in the channels panel and duty time plot)</p>	<p>Contaminants are present in the sample</p>	<p>Some contaminants can be cleared from the pores by the unblocking function built into MinKNOW. If this is successful, the pore status will change to "single pores". If the portion of recovering pores (unavailable pores in the extended view) stays large or increases:</p> <ol style="list-style-type: none"> 1. A nuclease flush can be performed, or 2. Run several cycles of PCR to try and dilute any contaminants that may be causing problems. <div style="text-align: center;"> </div> <p>The duty time plot above shows an increasing proportion of "recovering" pores over the course of a sequencing experiment</p>

Large proportion of inactive pores

Observation	Possible cause	Comments and actions
<p>Large proportion of inactive pores (shown as light blue in the channels panel and duty time plot. Pores or membranes are irreversibly damaged)</p>	<p>Air bubbles have been introduced into the flow cell</p>	<p>Air bubbles introduced through flow cell priming and library loading can irreversibly damage the pores. Watch the Priming and loading your flow cell video for best practice</p>
<p>Large proportion of inactive pores</p>	<p>Certain compounds co-purified with DNA</p>	<p>Known compounds, include polysaccharides, typically associate with plant genomic DNA.</p> <ol style="list-style-type: none"> 1. Please refer to the Plant leaf DNA extraction method. 2. Clean-up using the QIAGEN PowerClean Pro kit. 3. Perform a whole genome amplification with the original gDNA sample using the QIAGEN REPLI-g kit.
<p>Large proportion of inactive pores</p>	<p>Contaminants are present in the sample</p>	<p>The effects of contaminants are shown in the Contaminants Know-how piece. Please try an alternative extraction method that does not result in contaminant carryover.</p>

Reduction in sequencing speed and q-score later into the run

Observation	Possible cause	Comments and actions
Reduction in sequencing speed and q-score later into the run	Fast fuel consumption is typically seen when the flow cell is overloaded with library (~5-50 fmol of library is recommended).	Add more fuel to the flow cell by following the instructions in the MinKNOW protocol . In future experiments, load lower amounts of library to the flow cell.

Temperature fluctuation

Observation	Possible cause	Comments and actions
Temperature fluctuation	The flow cell has lost contact with the device	Check that there is a heat pad covering the metal plate on the back of the flow cell. Re-insert the flow cell and press it down to make sure the connector pins are firmly in contact with the device. If the problem persists, please contact Technical Services.

Failed to reach target temperature

Observation	Possible cause	Comments and actions
MinKNOW shows "Failed to reach target temperature" (37°C for Flow Cell Check, 34°C for sequencing on MinION Mk 1B/PromethION flow cells, and 35°C for sequencing on Flongle)	The instrument was placed in a location that is colder than normal room temperature, or a location with poor ventilation (which leads to the flow cells overheating)	MinKNOW has a default timeframe for the flow cell to reach the target temperature. Once the timeframe is exceeded, an error message will appear and the sequencing experiment will continue. However, sequencing at an incorrect temperature may lead to a decrease in throughput and lower q-scores. Please adjust the location of the sequencing device to ensure that it is placed at room temperature with good ventilation, then re-start the process in MinKNOW. Please refer to this FAQ for more information on MinION Mk 1B temperature control.

Guppy - no input .fast5 was found or basecalled

Observation	Possible cause	Comments and actions
No input .fast5 was found or basecalled	<i>input_path</i> did not point to the .fast5 file location	The <i>--input_path</i> has to be followed by the full file path to the .fast5 files to be basecalled, and the location has to be accessible either locally or remotely through SSH.
No input .fast5 was found or basecalled	The .fast5 files were in a subfolder at the <i>input_path</i> location	To allow Guppy to look into subfolders, add the <i>--recursive</i> flag to the command

Guppy - no Pass or Fail folders were generated after basecalling


Observation	Possible cause	Comments and actions
-------------	----------------	----------------------

Observation	Possible cause	Comments and actions
No Pass or Fail folders were generated after basecalling	The <code>--qscore_filtering</code> flag was not included in the command	The <code>--qscore_filtering</code> flag enables filtering of reads into Pass and Fail folders inside the output folder, based on their strand q-score. When performing live basecalling in MinKNOW, a q-score of 7 (corresponding to a basecall accuracy of ~80%) is used to separate reads into Pass and Fail folders.

Guppy - unusually slow processing on a GPU computer

Observation	Possible cause	Comments and actions
Unusually slow processing on a GPU computer	The <code>--device</code> flag wasn't included in the command	The <code>--device</code> flag specifies a GPU device to use for accelerate basecalling. If not included in the command, GPU will not be used. GPUs are counted from zero. An example is <code>--device cuda:0 cuda:1</code> , when 2 GPUs are specified to use by the Guppy command.

MinIT - the MinKNOW interface is not shown in the web browser

Observation	Possible cause	Comments and actions
The MinKNOW interface is not shown in the web browser	Browser compatibility issue	Always use Google Chrome as the browser to view MinKNOW. Alternatively, instead of typing <code>//mt-xxxxxx</code> (x is a number) in the address bar, type in the generic IP address, 10.42.0.1, which identifies the MinIT Wi-Fi router.
The MinKNOW interface is not shown in the web browser	The MinIT Wi-Fi was not used for connecting to the computer or mobile device	<p>Make sure the computer or mobile device is using the MinIT Wi-Fi. It should be shown as MT-xxxxxx (x is a number) on the underside label on the MinIT:</p>  <p>Disable the Ethernet connection from the computer or mobile device as needed. If necessary, contact your IT department to determine if the MinIT Wi-Fi is blocked (MinIT generic IP: 10.42.0.1). Please white-list MinIT as needed.</p>
The MinKNOW interface is not shown in the web browser	The MinIT was not on the same network that the computer was connected to.	Make sure that the wall sockets used by the Ethernet cables from the MinIT and computer belong to the same local network.

MinIT - the MinIT software cannot be updated

Observation	Possible cause	Comments and actions
The MinIT software cannot be updated	The firewall is blocking IPs for update	Please consult your IT department, as the MinIT software requires access to the following AWS IP ranges . Access to the following IP addresses is also needed: 178.79.175.200 96.126.99.215

Observation	Possible cause	Comments and actions
The MinIT software cannot be updated	The device already has the latest version of the software	Occasionally, the MinIT software admin page displays "updates available" even when the software is already up-to-date. Please compare the version listed on the admin page with the one on the Software Downloads page . Alternatively, SSH into the MinIT through a SSH Client (e.g. Bitvise or Putty, as described in the MinIT protocol) on a Windows computer or the terminal window on a Mac, run the command, <code>dpkg -l grep minit</code> , to find out the version of the MinIT software and <code>sudo apt update</code> if an update is needed. If the issue still persists, please contact Technical Services with details of the error.