



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA CELULAR
PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

**Calicivírus humanos: pesquisa em amostra de esgoto,
construção de clone e modelagem estrutural**

Karoline dos Anjos

Brasília, 2017



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA CELULAR
PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

Calicivírus humanos: pesquisa em amostra de esgoto, construção de clone e modelagem estrutural

Karoline dos Anjos

Orientador: Dr. Tatsuya Nagata

Tese de doutorado apresentada ao Programa de Pós-Graduação em Biologia Molecular, do Departamento de Biologia Celular, do Instituto de Ciências Biológicas da Universidade de Brasília como parte dos requisitos para obtenção do título de Doutor em Biologia Molecular.

Brasília, 2017

Tese de autoria de Karoline dos Anjos, intitulada “Calicivírus humanos: pesquisam em amostra de esgoto, construção de clone e modelagem estrutural”, apresentada como parte dos requisitos para obtenção do título de Doutora em Biologia Molecular pela Universidade de Brasília, em 29 de Junho de 2017, aprovada pela banca abaixo assinada:

Dr. Tatsuya Nagata (orientador)

Dr. João Alexandre Ribeiro Gonçalves Barbosa (examinador)

Dra. Rosana Blawid (examinadora)

Dra. Menira Borges de Lima Dias e Souza (examinadora)

Dr. Bergmann Morais Ribeiro (examinador)

*A vida é assim:
Esquenta e esfria,
Aperta daí afrouxa,
Sossega e depois desinquieta,
O que ela quer da gente
É Coragem.
(Guimarães Rosa)*

Lista de Figuras

Figura 1 Distribuição proporcional de causa de morte entre crianças menores de 5 anos.	11
Figura 2 Esquema dos principais vírus causadores de gastroenterites.	12
Figura 3 Relações filogenéticas entre os membros da família <i>Caliciviridae</i> .	13
Figura 4 Esquema representando os genomas dos cinco diferentes gêneros membros da família <i>Caliciviridae</i> .	15
Figure 1 Workflow from collection to data analysis. (Artigo I)	41
Figure 2 Illustration of all the reads from each sample. (Artigo I)	42
Figure 3 Illustration of different hosts viruses detected on sewage. (Artigo I)	42
Figura 5 Representação do capsídeo de <i>Norwalk virus</i> .	52
Figura 6 Imagens dos capsídeos lado a lado comparados.	52
Figura 7 Representação do genoma de norovírus e sua possível forma de tradução e processamento da poliproteína presente na ORF1.	53
Figura 8 Síntese de antígenos do grupo sanguíneo. Proteínas conhecidas como HBGAs podem ser relacionadas a ligação e entrada de norovírus na célula hospedeira.	54
Figura 9 Modificação de pMalc2c.	57
Figura 10 Esquema ilustrando as etapas para cristalização <i>in house</i> com placa de 15 poços e três concentrações diferentes.	59
Figura 11 Análise das identidades filogenéticas dos capsídeos de diferentes genótipos do GII de norovírus.	61
Figura 12 SDS-PAGE de todos os genótipos de norovírus expressos e semi-purificados..	62
Figura 13 Perfil de cromatografia e eletroforese das frações coletadas para GII.21.	63
Figura 14 Perfil de cromatografia e eletroforese das frações coletadas para GII.22.	63
Figura 15 Cristais obtidos a partir da cristalização de GII.21 e GII.22.	64
Figura 16 Estruturas cristalinas de raios X de norovírus GII.21 e GII.22 sem ligante.	67
Figura 17 Interação de fucose com domínio P2 de de norovírus.	69
Figura 18 pcDNA (Invitrogen). Esquema representativo do vetor de expressão em células de mamífero.	78
Figura 19 PCR de Fusão.	79
Figura 20 Células Caco-2 transfectadas com Sapo_CI e pcDNA.	84
Figura 21 Caco-2 controles da transfecção.	85
Figura 22 Imunomarcção de células Caco com anti-P2 e Alexa 488.	86
Figura 23 Controle de imunomarcção de Huh-7 com anti-P2 e Alexa 633.	87
Figura 24 Imunomarcção de células Caco-2 e Huh-7 transfectadas com SapoCI_pcDNA.	88
Figura 25 Citometria de fluxo de células Caco-2 para análise das imunomarcações realizadas com dois diferentes anticorpos primários.	89/90

Resumo

Os calicivírus humanos são os principais causadores de gastroenterites no mundo, atualmente. No entanto, por não possuírem ainda um sistema de replicação viral *in vitro*, diferentes abordagens são desenvolvidas para que se possa obter maiores informações a cerca da biologia desses vírus. A análise ambiental tem demonstrado que assim como os demais vírus entéricos, as partículas dos calicivírus humanos são resistentes e são passíveis de serem detectados a partir de metodologias como RT-PCR em diferentes amostras. No Capítulo I desta tese utilizou-se a tecnologia de *Next Generation Sequencing* para a avaliação da população viral encontrada em um esgoto urbano coletado na Estação de Tratamento de Esgoto Norte, Brasília, DF, Brasil, com intuito de se determinar quais possíveis vírus humanos estariam circulando na população do Distrito Federal. Além de vírus humanos, foram encontrados vírus animais, vegetais e de insetos nas amostras analisadas. No entanto apenas os vírus com hit para entéricos causadores de doença humana tiveram seu genoma completo montado. Dentre os achados pelo menos três genótipos de sapovírus foram encontrados, assim como o novo genótipo epidêmico de norovírus GII.17 e, um vírus novo descrito como astrovirus-tipo, o bastrovirus. No Capítulo II seis diferentes genótipos de norovírus pouco detectados foram utilizados para expressar o domínio protuberante (P) do capsídeo, com intuito de se comparar e compreender possíveis mudanças estruturais e nucleotídicas em relação ao genótipo mais prevalente (GII.4). Como resultados as estruturas cristalizadas de domínio P de norovírus GII.21, GII.22 e GII.22 ligada a fucose, que é um açúcar precursor na diferenciação de HBGA (*histo-blood group antigens*) foram resolvidas. Por fim, no Capítulo III, um clone contendo o cDNA referente ao genoma do sapovírus BR01 foi construído. O uso de microscopia confocal a laser e citometria de fluxo confirmaram a funcionalidade do clone pela positividade de reação entre as células transfectadas e o anticorpo anti-P2 (subdomínio de VP1). Futuros experimentos poderão validar o uso desse clone como ferramenta para genética reversa.

Palavras-chave: Calicivírus humanos, detecção em água de esgoto, expressão de proteínas virais, cristalografia, clone infeccioso, genética reversa

Abstract

Human caliciviruses are the main cause of gastroenteritis in the world today. However, because they do not have a viral multiplication system in cell culture, different approaches are developed so that more information can be obtained about the biology of these viruses. Environmental analysis has shown that, like other enteric viruses, human calicivirus particles are resistant and can be identified from methodologies such as RT-PCR in water or food samples. In Chapter I of this thesis was used the Next Generation Sequencing technology to evaluate the viral population found in an urban sewage collected in the North Waste Water Treatment, Brasilia, DF, Brazil, in order to determine which possible human viruses would be circulating in the population of the Federal District. In addition to human viruses, animal, plant and insect viruses were found in the samples analyzed. However only the hit viruses for enteric-causing human disease had their complete genome assembled. Among the findings, at least three sapovirus genotypes were found, as well as the novel genotype of norovirus GII.17 and, a novel virus described as astrovirus-type, bastrovirus. In Chapter II, six different genotypes of undetected norovirus were used to express the protruding (P) capsid domain, in order to compare and understand possible structural and nucleotide changes in relation to the most prevalent genotype (GII.4). As results the crystallized P domain structures of norovirus GII.21, GII.22 and GII.22 linked to fucose, which is a precursor sugar in the differentiation of HBGA (histo-blood group antigens) were resolved. Finally, in Chapter III, a clone containing the cDNA for the genome of sapovirus BR01 was constructed. The use of laser confocal microscopy and flow cytometry confirmed the functionality of the clone by the reactivity positivity between the transfected cells and the anti-P2 antibody (subdomain of VP1). Future experiments could validate the use of this clone as a tool for reverse genetics.

Key words: Human calicivirus, sewage detection, viral protein expression, crystallography, infectious clone, reverse genetics

Sumário

Lista de Figuras	V
Resumo	VI
Abstract	VI
1. Introdução	10
1.1 Calicivírus	12
1.2 Norovírus	15
1.2.1 Morfologia e Genoma	17
1.2.2 Classificação	18
1.3 Sapovírus	18
1.3.1 Morfologia e Genoma	19
1.3.2 Classificação	20
2. Justificativa	21
3. Objetivos	22
Capítulo I – Detecção de vírus causadores de gastroenterites em água de esgoto não tratada no Distrito Federal	23
1. Nota de apresentação do capítulo	23
2. Introdução	23
Artigo I	26
Artigo II	48
Capítulo II – Biologia Estrutural dos Calicivirus Humanos	51
1. Nota de apresentação do capítulo	51
2. Introdução	51
3. Material e Métodos	56
3.1 Material utilizado nesse estudo	56
3.2 Clonagem	56
3.3 Expressão de proteínas	57
3.4 Purificação	58
3.5 Cristalização	58
3.6 Coleta e Processamento de Dados	60
4. Resultados	61
4.1 Norovírus	61

4.1.1 Alinhamento de aminoácidos referente ao capsídeo de norovírus GII	61
4.1.2 Expressão e Purificação de proteína	62
4.1.3 Cristalização	63
4.1.4 Estruturas	66
5. Discussão	70
Capítulo III - Desenvolvimento de ferramentas básicas aplicadas ao estudo de genética reversa de sapovírus humanos	72
1. Introdução	72
2. Material e Métodos	77
2.1 Clonagem	77
2.2 Construção do clone infeccioso	77
2.4 Transfecção celular	80
2.5 Ensaio de Imunomarcção	81
2.6 Citometria de Fluxo	82
3. Resultados	83
4. Discussão	91
5. CONCLUSÃO GERAL	94
6. Perspectivas Futuras	97
7. Referências Bibliográficas	98

1. Introdução

O estudo da virologia humana vem sendo realizado há pelo menos oito décadas desde a identificação dos vírus no universo microbiano (1). A compreensão da morfologia das partículas virais, suas formas de invasão e evasão celular bem como o estudo de seus ciclos replicativos têm sido o objeto de trabalho de inúmeros pesquisadores ao redor do mundo. Todos esses esforços buscam, sobretudo, gerar conhecimento básico e aplicado para prevenção e cura de possíveis patologias causadas pelos vírus.

Apesar de existirem inúmeros vírus descritos na literatura apenas alguns foram descritos como de infectar e causar danos nos humanos. No entanto, sabe-se que algumas doenças humanas mais devastadoras são ocasionadas por infecções virais, como por exemplo, influenza, sarampo, febre amarela, síndrome da imunodeficiência adquirida (AIDS) e mais recentemente a epidemia de microcefalia em neonatos causadas pelo Zika vírus (1, 2). Além disso estima-se que os vírus sejam associados a pelo menos 20% dos casos de câncer e mortes de milhares de crianças acometidas por infecções virais nos tratos respiratório e intestinal (1).

Nesse contexto, encontram-se os vírus causadores de gastroenterites agudas (GEA), que, assim como estirpes do vírus influenza, causam um elevado índice de morbimortalidade, especialmente em indivíduos menores de cinco anos, idosos e imunocomprometidos (3). Estima-se que cada indivíduo apresente pelo menos um episódio de diarreia por ano. Somente nos Estados Unidos da América, por exemplo, há registros anuais de 21 milhões de indivíduos acometidos por essa patologia (Centro de Controle e Prevenção de Doenças, EUA) (<https://www.cdc.gov/norovirus/trends-outbreaks.html>).

A diarreia tem por definição um aumento no número de evacuações, com fezes aquosas ou de pouca consistência, podendo apresentar, em alguns casos, muco ou sangue. É causada por diferentes agentes etiológicos (vírus, bactérias ou protozoários) sendo em sua maioria autolimitada (com duração entre 2 a 14 dias) podendo apresentar grande perda de fluídos associada diretamente a desidratação. A maioria dos casos são considerados brandos, no entanto, em crianças menores de 5 anos, devido ao seu metabolismo acelerado e menor capacidade de armazenamento de água quando comparadas a jovens e adultos, a possibilidade de desidratação por perda de fluídos é mais acentuada (3).

De acordo com dados apresentados pela OMS (Organização Mundial de Saúde) (<http://www.who.int/mediacentre/factsheets/fs330/en/>) e UNICEF (Fundação das Nações Unidas para a Infância) (<https://data.unicef.org/topic/child-health/diarrhoeal-disease/>) cerca de 9 milhões de crianças menores de 5 anos morrem a cada ano, sendo doenças diarreicas a segunda principal causa destes óbitos, abaixo apenas dos ocasionados por pneumonia. Estima-se que 1 a cada 5 crianças menores de 5 anos venham a óbito por doenças diarreicas. De acordo com estes dados, isso significa que o óbito infantil por doenças diarreicas é maior do que o somatório de óbitos por HIV, malária e sarampo (Fig. 1).

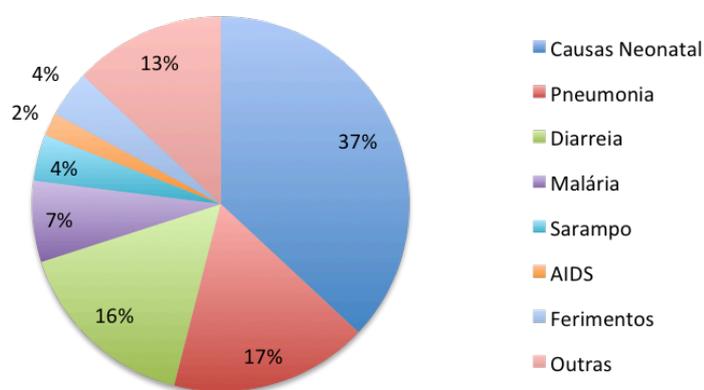


Figura 4 Distribuição proporcional de causa de morte entre crianças menores de 5 anos. Nesse gráfico observa-se que 17% e 16% das mortes entre crianças menores de cinco anos são causadas por pneumonia e diarreia, respectivamente. 13% são denominadas como outras. 7% são causadas por malária. 4% por sarampo e 4% por ferimentos/acidentes. 2% por AIDS. As causas neonatais somam 37% e são variáveis. (Adaptado de Organização Mundial de Saúde, Estimativas do encargo global de doenças, 2004).

Dados estatísticos mostram comprovam que 88% das mortes por doenças diarreicas podem ser relacionadas a água não tratada bem como a condições de saneamento e processos de higienização inadequados. A higienização adequada das mãos apenas com sabão pode reduzir em até 40% a incidência dessa doença (3).

Os principais vírus causadores de gastroenterites agudas são os rotavírus (*Reoviridae*), norovírus e sapovírus (*Caliciviridae*), astrovírus (*Astroviridae*) e adenovírus (*Adenoviridae*) (Fig. 2). No entanto, com o desenvolvimento e aplicação da vacina contra os rotavírus, os calicivírus humanos têm sido os principais causadores de GEA virais no mundo (Centro de Controle e Prevenção de Doenças, EUA) (<https://www.cdc.gov/norovirus/trends-outbreaks.html>). Entretanto, em diversos países, assim como no Brasil, ainda não existe um sistema de vigilância eficaz para detecção e controle desses vírus na população. Um estudo realizado no

ano de 2014 demonstrou a diversidade genotípica dos calicivírus humanos encontrados em todo país (4). Os dados apresentados nesse estudo corroboram as estatísticas mundiais de prevalência e incidência desses vírus.

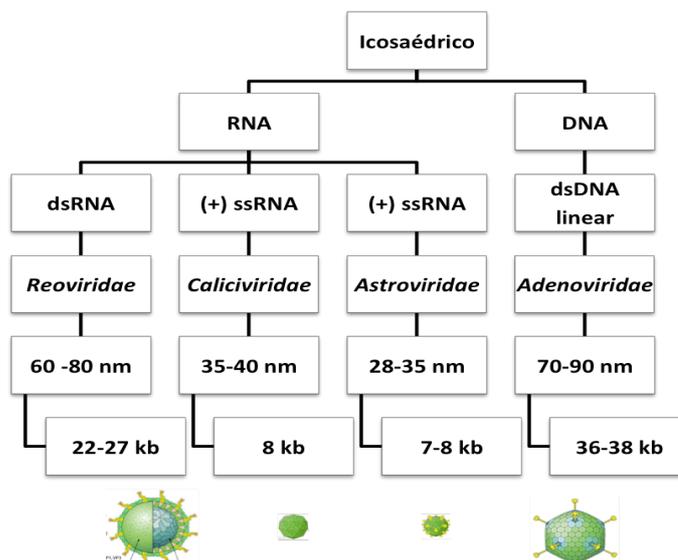


Figura 5 Esquema dos principais vírus causadores de gastroenterites. Esse fluxograma ilustra as principais características de cada família. Além de causarem doenças diarreicas, estes vírus possuem arranjo icosaédrico de seus capsídeos. No entanto apenas os calicivírus e astrovírus conservam algumas similaridades como, um genoma não segmentado com RNA fita simples e partícula viral de até 40 nm.

1.1 Calicivírus

A família *Caliciviridae* é atualmente dividida em cinco gêneros: *Norovirus*, *Sapovirus*, *Lagovirus*, *Vesivirus* e *Nebovirus* (Fig.3), com pelo menos seis outros gêneros em análise para inclusão (<http://www.ictvonline.org/virusTaxonomy.asp>). É descrito na literatura que a história natural desses vírus pode ser datada desde 1932 a partir de infecções descritas em mamíferos aquáticos como, leões marinhos e focas. O primeiro calicivírus a infectar suínos, *Vesicular exanthema of swine virus* (VESV), espécie tipo do gênero *Vesivirus*, surgiu após a alimentação de suínos a partir de carne oriunda de animais marinhos contaminados com calicivírus. Apesar de não ser relatado desde 1959, o VESV permanece como um vírus de importância para saúde pública devido a semelhança de seus sintomas com os ocasionados pelo enterovírus Coxsackie causador da doença das mãos-pés e boca (5, 6).

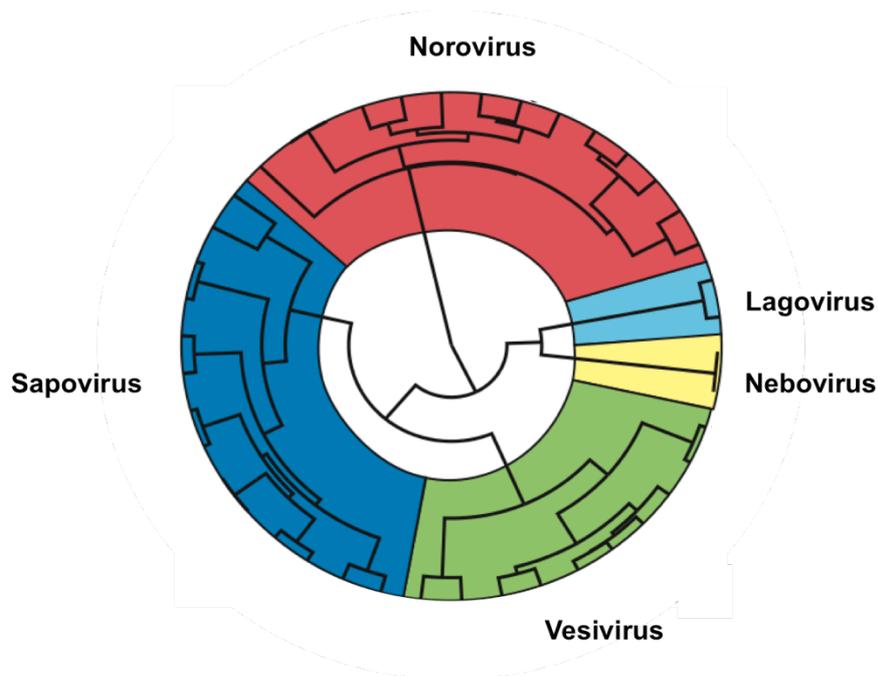


Figura 6 Relações filogenéticas entre os membros da família Caliciviridae. Os norovírus são destacados em vermelho, lagovírus em azul claro, nebovírus em amarelo, vesivírus em verde e sapovírus em azul. (Adaptado de https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/254/caliciviridae-figures).

Os primeiros calicivírus foram inicialmente classificados dentro da família *Picornaviridae* devido à similaridade morfológica entre seus membros e, principalmente, à elevada identidade de suas proteínas 3C-like protease (7). Somente em 1979 o ICTV (Comitê Internacional de Taxonomia de Vírus) aprovou a criação da família *Caliciviridae*. A atual classificação em cinco gêneros foi aprovada pelo ICTV em 2012 e baseia-se em análises estruturais e filogenéticas, considerando principalmente a sequência referente ao gene do capsídeo.

Dentre os gêneros atualmente descritos apenas *Norovirus* e *Sapovirus* infectam humanos. Estes ainda são encontrados em animais suínos, bovinos, caninos, entre outros (8). Os *Vesivirus* são representados por duas espécies, a espécie-tipo *Vesicular exanthema of swine virus* (VESV), e a espécie *Feline calicivirus* (FCV), que são responsáveis por infecções em felinos selvagens e domésticos e são associados a diferentes patologias como doenças respiratórias, estomatite aguda e crônica, artrite aguda e febre hemorrágica, também conhecida como doença sistêmica virulenta associada a FCV (9). Os *Lagovirus* são representados, principalmente, pelos vírus causadores da febre hemorrágica de coelhos, espécie tipo *Rabbit hemorrhagic disease virus* (RHDV). A patologia é descrita como necrose e hemorragia hepáticas, renais e pulmonares. Estima-se que o animal venha a óbito três dias depois do início

da infecção (10). Por fim, o gênero *Nebovirus*, espécie *Newbury-1 virus*, foi o último aprovado pelo ICTV. Embora inicialmente classificado como uma norovirose em bovinos, análises filogenéticas corroboraram para a distinção entre os vírus pertencentes a esse gênero e o genogrupo de norovírus responsável por causar diarreia aguda em bovinos, especialmente em bezerros (11).

A família *Caliciviridae* é formada por vírus esféricos, não envelopados, com um diâmetro entre 27 e 45nm. O capsídeo é formado por uma única proteína, VP1, (para norovírus já se constatou a presença de VP2 como parte do capsídeo (12), e exibe uma simetria icosaédrica composta por 180 unidades organizadas em 90 dímeros oriundos de dois domínios: o domínio S (do inglês *shell*), que envolve o RNA genômico, e, o domínio P (do inglês *protruding*), ainda dividido em dois subdomínios, P1 e P2 (13). Seu RNA genômico é de fita simples, senso positivo, poliadenilado no terminal 3' e de tamanho entre 7300 a 8500 pares de bases. Em geral o genoma apresenta duas ou três fases de leitura aberta (ORF do inglês *open reading frame*). No terminal 5' o RNA apresenta uma proteína viral associada ao genoma (VPg) e sua ORF1 compõe uma poliproteína, que codifica sete proteínas não estruturais e é clivada após tradução da protease viral 3C-like (7). A proteína VP1 é codificada pela ORF2 e traduzida a partir de um RNA mensageiro subgenômico. As diferenças encontradas entre os genomas de cada gênero são apresentadas na Figura 4, com a ressalva de que os norovírus que infectam murinos apresentam uma ORF4 que codifica uma proteína de fator de virulência (VF1) (14) e pelo menos dois genogrupos de sapovírus apresentam uma ORF3 de função ainda desconhecida.

atualmente, esses critérios são meramente ilustrativos e não são utilizados no diagnóstico (16):

1. Ocorrência de vômito em mais da metade dos pacientes acometidos;
2. Período de incubação entre 24 e 48 horas;
3. Período de duração de sintomas entre 12 e 60 horas e;
4. Ausência de bactérias nas amostras fecais.

Em 1993, a primeira sequência da espécie-tipo *Norwalk virus* foi clonada (17), o que proporcionou inúmeros estudos moleculares e estruturais. A expressão da proteína do capsídeo em células de inseto por meio de baculovírus recombinante demonstrou a geração de partículas semelhantes a vírus (do inglês *virus like particles* VLP), apresentando tamanho e morfologia similares, e em alguns casos podendo induzir respostas imunológicas (18, 19). A partir desses estudos foi possível se obter informações como a estrutura cristalográfica do capsídeo (20).

Nos anos 2000 diferentes abordagens foram descritas para uma melhor compreensão dos norovírus humanos e houve um importante avanço quando o norovírus murino foi identificado pela primeira vez (21), logo em seguida proporcionando a primeira descrição de replicação de norovírus em cultura de célula (22). Em 2004 um grupo holandês se propôs a avaliar mais de 20 linhagens celulares entre elas HEK 293T, Caco-2 e Vero. Contudo, eles concluíram que apesar de algumas linhagens permanecerem positivas para norovírus mesmo após a quinta passagem não eram observados efeitos citopáticos ou formação de novas partículas virais nestas culturas (23). Em seguida foi demonstrada a replicação de RNA viral transcrito sob o comando de um promotor T7 em células HEK-293T e BHK (24), assim como o uso de RNA extraído de fezes mostrou-se infeccioso em cultura de células Huh-7 (25).

Os achados mais recentes incluem o desenvolvimento do primeiro sistema de genética reversa para norovírus humanos e dois novos sistemas de cultivo celular. A expressão e formação de progênie viral de um norovírus GII.3 sob o comando de um promotor EF-1 α (fator de alongação) foi obtida sem o auxílio de um vírus ajudante na célula transfectada. Além disso a inserção de proteína verde fluorescente (GFP) na construção permitiu a observação de RNA encapsidado em células transfectadas observadas por microscopia confocal (26).

Sabe-se que norovírus murino é replicado com alta eficiência em células macrofágicas e seguindo esse padrão, em 2014, foi estabelecido um sistema de cultura humana de células linfóides B onde a presença de enterobactérias que expressam antígenos do sistema sanguíneo (do inglês *histo blood group antigens* HBGA) estimulou a infectividade de norovírus GII.4 presentes em amostra fecal não filtrada (27). Esse importante avanço, no entanto, não descartou a possibilidade de que os norovírus possuem tropismo por células que compõe o epitélio intestinal. Dessa forma foi demonstrada, em 2016, a replicação de norovírus GII.4 em um sistema onde células tronco são isoladas das criptas intestinais e estimuladas a se diferenciarem em um complexo multicelular denominado enteróides intestinais humanos (do inglês *human intestinal enteroids* HIEs) (28, 29).

1.2.1 Morfologia e genoma

Assim como os demais calicivírus, os norovírus apresentam morfologia esférica, diâmetro entre 29-32nm, simetria icosaédrica T=3, presença de um domínio mais enterrado, “concha” (S) (do inglês *shell*), que envolve o RNA genômico e, um domínio mais exposto, protuberância (P) (do inglês *protrude*), que se subdivide em P1, porção intermediária, e P2, região mais exposta e com maior variação nucleotídica (20).

O genoma é de fita simples de RNA, senso positivo, de 7,5kb e codifica três ORFs (do inglês *open reading frames*), no caso de norovírus murino é ainda encontrada uma quarta ORF (ORF4) (14). A ORF1 codifica sete proteínas não estruturais, entre elas a RNA polimerase RNA-dependente (NS7), a ORF2 codifica a proteína do capsídeo e a ORF3 uma proteína estrutural menor, essas duas últimas são traduzidas a partir de um RNA subgenômico. A ORF4 foi descrita como responsável pela codificação de uma proteína antagonista de imunidade inata denominada Fator de Virulência 1 (VF1) (14).

1.2.2 Classificação

Os norovírus são classificados em sete genogrupos (GI – GVII) e mais de 40 genótipos. As infecções em humanos em geral são causadas por GI, GII, GIV, com exceção de GII.11, GII.18, GII.19 que infectam suínos e GIV.2 que infectam espécies caninas. GIII infectam vacas e ovelhas, GV infectam ratos e camundongos e, GVI e GVII infectam caninos (21, 30, 31). O norovírus GII.4 é atualmente o genótipo com maior prevalência mundial ainda que GII.17 tenha sido detectado mundialmente no ano de 2015 (8, 32).

A classificação em genogrupos e genótipos é primordialmente baseada em parte da sequência do capsídeo e, ainda, em alguns casos, há a classificação a partir de regiões conservadas da polimerase (8, 33). A ausência de um consenso dificulta a acurácia classificatória, principalmente à luz de diversos estudos que demonstraram a ocorrência de recombinação entre isolados virais, especialmente na região do genoma da junção entre a RNA polimerase RNA-dependente e a sequência do gene do capsídeo viral (34). É importante ressaltar que a sequência de aminoácidos completa do capsídeo pode ser a melhor opção para resolver a diversidade dos norovírus.

1.3 Sapovírus

Sapporo virus é a espécie-tipo do gênero *Sapovirus*. A primeira descrição de sapovírus humanos foi realizada em 1977 em Sapporo, Japão (35). Partículas pequenas e arredondadas foram observadas por microscopia eletrônica após um surto de gastroenterite aguda em um orfanato da região (35). Crianças menores de cinco anos apresentaram diarreia e amostras fecais foram coletadas e analisadas para diagnóstico (35). Em 1980 Saif e colaboradores (36) descreveram por meio de microscopia eletrônica de transmissão vírus que apresentavam estruturas semelhantes a um cálice em sua superfície, oriundos de amostras fecais de leitões com sintomas diarreicos (37). Atualmente, sapovírus foram também identificados em amostras fecais de visom, cães, morcegos e chimpanzés (37-40).

Sintomas diarreicos nem sempre estão presentes em animais e humanos. É possível encontrar na literatura estudos que demonstraram que indivíduos assintomáticos apresentavam elevada carga viral em suas amostras fecais (41). Sabe-

se no entanto, que em inúmeros casos embora os indivíduos apresentem sintomatologia, poucos são aqueles que procuram assistência médica e a rota de infecção entre os animais ainda é pouco compreendida.

Recentemente, sapovírus humanos foram reportados em frutos do mar e em corpos d'água como rios e redes de tratamento de água ao redor do mundo. Além disso, houve um aumento no número de casos de gastroenterites agudas causadas por sapovírus em adultos, embora esses vírus acometam em sua maioria crianças menores de cinco anos (8, 42-44).

Até o presente momento apenas o sapovírus suíno (PSaV Cowden) possui sistema de genética reversa descrito (45), onde foi revelada a necessidade do uso de ácidos biliares para infecção celular e propagação viral, os estudos demonstraram que estes ácidos estariam associados a uma redução no transdutor e ativador de sinal da via de sinalização mediada pela transcrição (STAT-1) (46). Além disso um estudo recente demonstrou que o fator de iniciação eucariótico 4A (eIF4A) é necessário para o afrouxamento da região 5' UTR (não traduzida) do RNA para facilitar a tradução viral (47, 48).

1.3.1 Morfologia e genoma

Sapovírus são descritos como vírus de morfologia esférica com diâmetro entre 30-45 nm. Apresentam as mesmas características icosaédricas descritas para calicivírus (7). O domínio P subdivide-se em P1 (mais interno) e P2 (mais externo – porção globular) (13, 49). Assim como para os demais calicivírus, o domínio S é descrito como a porção mais interna da estrutura do capsídeo envolvendo o RNA genômico. O domínio P é responsável pela composição dos cálices que, observados ao microscópio eletrônico, conferem a partícula uma morfologia comumente descrita como “estrela-de-Davi” (35).

O genoma dos sapovírus é linear, senso positivo, fita simples de RNA com aproximadamente 7500 pares de base, apresenta VPg em seu terminal 5' e cauda de poli-A no terminal 3'. É um genoma não segmentado apresentando de duas a três ORFs (50). A ORF1 codifica uma poliproteína viral que contém regiões de aminoácidos conservadas para sete proteínas não estruturais: NS1; NTPase

⁴⁸¹GOOGIGKT⁴⁸⁸, VPg ⁹⁴²KGKTK⁹⁴⁶ e ⁹⁶⁴DEYDE⁹⁶⁸, protease ¹¹⁶⁹GDCG¹¹⁷², RNA polimerase RNA dependente ¹⁵⁰⁴GLPSG¹⁵⁰⁸ e ¹⁵⁵² YGDD¹⁵⁵⁵, e a proteína do capsídeo ¹⁸⁵⁶PPG¹⁸⁵⁸ (7). O papel da ORF2 e da ORF3 ainda não foi claramente elucidado, mas se comparado aos norovírus, a ORF2 pode codificar uma proteína estrutural menor (VP2) e, a ORF3 pode codificar uma proteína funcional, uma vez que sua sequência apresenta uma região conservada para tradução no terminal 5' GCAAUGG (51, 52).

1.3.2 Classificação

No ano de 2010, o Comitê Científico Internacional de Calicivírus propôs que a classificação dos sapovírus fosse realizada a partir da análise da sequência completa referente ao capsídeo viral (33).

A partir deste critério, sabe-se que os sapovírus podem ser classificados em 5 genogrupos (GI-GV), com a proposta de nove novos genogrupos (GVI-GXIV) (Tabela 1) e que os sapovírus humanos podem ser ainda sub-classificados em pelo menos 16 genótipos (33, 53). Os genogrupos que infectam humanos são GI, GII, GIV e GV, enquanto GIII, GV, GVI, GVII, GVIII, GIX, GX e GXI infectam suínos. Como mencionado anteriormente, sapovírus foram detectados em outros mamíferos que também fazem parte desta classificação. GXII infectam visom, GXIII infectam cães e GIV infectam morcegos. Assim como GV primariamente infecta humanos e só recentemente foi encontrado em suínos, sapovírus pertencentes ao GI foram encontrados em fezes de chimpanzés (40).

Tabela 1 Classificação de *Sapovirus* baseada na sequência de nucleotídeos completa do capsídeo.

Genogrupo	Número de Genótipos	Hospedeiro
I	7	Humano / Chimpanzé
II	7	Humano
III	1	Suíno
IV	1	Humano
V	1	Humano / Suíno
VI	-	Suíno
VII	-	Suíno
VIII	-	Suíno
IX	-	Suíno
X	-	Suíno
XI	-	Suíno
XII	-	Visom
XIII	-	Canino
XIV	-	Morcego

2. Justificativa

Anualmente, milhões de indivíduos são acometidos por doenças diarreicas no mundo inteiro. Dentre estes, estima-se que pelo menos uma centena de milhar venha a óbito. Apesar de uma diversidade de microrganismos estarem associados a essa patologia sabe-se que em países desenvolvidos a principal causa é viral.

Até o presente momento, os rotavírus e os norovírus são os principais agentes virais causadores de gastroenterites agudas no mundo, sendo o norovírus o principal agente etiológico viral em localidades onde a vacina contra rotavírus é distribuída. No entanto, alguns estudos demonstraram, na América Central, em lugares como Nicarágua e Porto Rico, um aumento na incidência da sapovírus em crianças menores de dois anos.

A ausência de um programa nacional de detecção de calicivírus humanos justifica o uso de novas tecnologias e abordagens para geração de dados que descrevam a distribuição desses patógenos no Brasil.

Além disso a ausência de um sistema de cultura celular robusto inviabiliza a aquisição de novas informações sobre a biologia viral levando os cientistas a abordagens mais clássicas como, o uso de ferramentas da biologia estrutural.

Por fim, diferentes abordagens para a construção de um clone infeccioso estão disponíveis para as principais espécies de calicivírus animais, no entanto, para sapovírus humanos nenhuma delas foi descrita até o momento.

3. Objetivos

O objetivo principal desse trabalho é gerar, a partir de diferentes metodologias, informações que resultem em conhecimento básico sobre os calicivírus humanos circulantes no Brasil. Este objetivo divide-se em três principais pontos: 1. detecção de calicivírus humanos em água de esgoto não tratada no Distrito Federal a fim de estimar a circulação desses vírus na região nos anos de 2014 e 2015, 2. elucidação da estrutura tridimensional do domínio P que compõe a proteína do capsídeo dos calicivírus humanos e 3. desenvolver uma ferramenta de genética reversa que possibilite o estudos moleculares mais aprofundados desses vírus a partir de genética reversa.

Os objetivos específicos são:

- Realizar a coleta de efluente sanitário na Estação de Tratamento Norte, Brasília, Distrito Federal.
- Analisar os dados com montagem de genomas para descrição dos possíveis vírus circulantes durante o período de coleta.
- Resolver estruturas domínios P de norovírus pertencentes ao GII de genótipos não epidêmicos.
- Construir um clone infeccioso de sapovírus humano.

CAPITULO I – Detecção de vírus causadores de gastroenterites em água de esgoto não tratada no Distrito Federal

1. Nota de apresentação do capítulo

O capítulo I dessa tese será apresentado com uma breve introdução ao tema, seguida do manuscrito ainda em processo de correção a ser submetido ao *Journal Applied and Environmental Microbiology* será apresentado como parte do critério exigido para realização de defesa de título de Doutora em Biologia Molecular. Além disso será apresentado o manuscrito submetido ao jornal *Genome Announcements*.

2. Introdução

Acredita-se que vírus entéricos possuem um maior tempo de vida quando livres em meio ambiente que os demais vírus devido a sua resistência ao calor, alta pressão e variação de pH (54). Dessa forma o esgoto não tratado torna-se um ambiente enriquecido desses vírus pois concentra uma carga viral representante de uma população de determinada região geográfica.

A análise desse tipo de amostra ambiental pode sugerir a descrição de vírus que não foram relatados ao sistema de saúde e, além disso a detecção desses vírus em amostras não tratadas permite uma escolha para um tratamento adequado dessas águas que são devolvidas ao meio ambiente.

Atualmente sabe-se que o norovírus é o principal agente viral causador de gastroenterites no mundo, e devido ao seu possível potencial zoonótico tornou-se um importante patógeno. Em países como, os Estados Unidos da América (EUA) foi implementado, em 2009, pelo Centro de Controle de Prevenção de Doenças (CDC) o programa CaliciNet que visa não só alimentar e organizar um banco de dados sobre surtos e infecções causadas por esse vírus, mas também funciona como uma ferramenta informativa sobre desenvolvimento e pesquisa em métodos de detecção e tecnologia para compreensão do vírus e possível prevenção. Esse órgão estima que 60% dos surtos de gastroenterites ocorram por infecção por norovírus.

Estudos sobre a presença de norovírus no meio ambiente, incluindo águas como veículo de contaminação, demonstraram a presença de norovírus em água potável, água utilizada para recreação, gelo, entre outros. Propõem-se água de esgoto

como a principal fonte de contaminação não só de águas utilizadas para consumo, mas também água de rios e mares, no entanto, estudos sobre a distribuição de norovírus e demais calicivírus no meio ambiente ainda não foram realizados.

Os sapovírus humanos têm sido relacionados a surtos esporádicos no mundo todo, sendo bastante incidentes em crianças menores de 2 anos (55-60). Além de estarem associados a contaminação de comida os sapovirus foram recentemente evidenciados em amostras clínicas e em água de esgoto coletada no Rio de Janeiro, Brasil (61). Antes disso já haviam sido descritos em amostras ambientais como, águas de rios, efluente tratado e animais bivalves, por exemplo, ostras (56, 62-67).

No Brasil a ausência de sistema de monitoramento para calicivírus humanos contribui para o desenvolvimento de novas abordagens para se obter dados a cerca de circulação desses vírus. Com exceção ao Laboratório de Virologia Comparada e Ambiental vinculado a Fundação Oswald Cruz, que é um instituto nacional de pesquisa brasileiro, as demais informações obtidas a respeito desses vírus são provenientes de estudos realizados em institutos de ensino superior e pesquisa.

A metodologia mais utilizada na detecção desses vírus é a reação em cadeia da polimerase (PCR), quantitativa e em tempo real as quais têm sido extensivamente aplicadas (68-73). Embora seja uma metodologia com alta especificidade a PCR não permite descrever a viabilidade do vírus nessas amostras, e devido a ausência de um cultivo celular esse tipo de questão permanece uma incógnita.

Uma nova abordagem para detecção de vírus em amostras ambientais tem sido o uso da metagenômica, que permite não só detectar vírus conhecidos como também descrever novos vírus (74). Essa abordagem tem sido descrita em trabalhos para esgoto não tratado e também análise de fezes de indivíduos saudáveis. Nesses dois trabalhos particularmente houve uma congruência em alguns resultados como, a presença de vírus de plantas (74, 75). Isso justificado em ambas amostras pelo consumo de vegetais e folhas que podem estar contaminados com diversos microrganismo. Essa comparação apenas reforça o uso de amostras de esgotos urbanos como uma matriz enriquecida de diversidade viral e um bom exemplo de uma amostra representativa de uma população de uma região geográfica específica (74).

No trabalho apresentado a seguir três amostras durante os meses de junho, julho e agosto de 2014 e uma amostra referente ao mês de agosto de 2015 foram coletadas na Estação de Tratamento Norte, Brasília, Brasil. As amostras foram coletadas no tanque antes do tratamento com lodo ativado. Essa estação possui uma

vazão de 450 l/s e trata aproximadamente o esgoto sanitário de uma população de 180 mil habitantes referente a sete regiões administrativas do Distrito Federal.

Essas amostras foram concentradas em baixa rotação e em seguida as possíveis partículas virais foram concentradas por ultracentrifugação e um colchão de 20% de sacarose. O RNA das amostras foi enviado para sequenciamento utilizando método Illumina HiSeq paired-end, Macrogen, Coréia do Sul. Os dados obtidos foram analisados por meio de ferramentas de bioinformática e os resultados encontrados estão descritos no artigo a seguir.

Title:

An overview of the virome in wastewater in Brasilia, Brazil by High-throughput sequencing

Authors:

Dos Anjos, Karoline; Nagata, Tatsuya; Melo; Fernando Lucas

Abstract**Introduction**

The International Committee on Taxonomy of Viruses determines that presently the virosphere is composed of seven orders (*Caudovirales*, *Herpesvirales*, *Ligamenvirales*, *Mononegavirales*, *Nidovirales*, *Picornavirales* and *Tymovirales*) and 111 families represented by more than 9,000 viral sequences (76). However, metagenomic studies suggest that this number is just a small representation of all viruses present in nature (74). Viral metagenomic studies can be performed directly after nucleic acid extraction, and have the potential to identify novel viruses without any previously known sequences and without lab culturing and virus isolation (74, 77-79).

In the past few years environmental samples have become increasingly important sources for virus discovery (74, 78-80). The World Health Organization (WHO) created a guideline regarding the quality of water for consumption that reinforces viruses as important pathogens once waterborne transmission is well known as an important dissemination route for enteric viruses.

It is well known that low temperatures, acidic pH, presence of salts and organic matter in solid suspensions are important factors that can increase the

persistence of enteric viruses in wastewater biomes (81, 82). Some studies demonstrated that water temperature alone modulates the persistence of viruses in groundwater. It has been broadly demonstrated in literature that human enteric viruses, plant viruses and even insect viruses are detected, and even persist in wastewater after treatment (77, 78, 83-85). Traditional guidelines for the assessment of water quality are based on monitoring of indicator fecal bacteria, a method that completely bypasses the virome (67, 86).

In Brazil sanitary surveillance systems monitor rotavirus and adenovirus presence in clinical samples, but viruses like human calicivirus are only studied in research level (87). Recently, in environmental studies performed in Brazil human caliciviruses have been detected and sequenced using a combination of molecular techniques in a water source (61, 88, 89).

In this work a metagenomic approach using the Next generation sequencing (NGS) or High-throughput sequencing method was performed to have a broadly view of viral diversity of the Federal District of Brazil. One untreated wastewater sample was collected per four months and the total RNA was sequenced using paired-end methodology and later libraries were assembled and analyzed to determine what viruses exist in this region and specifically the enteric viruses complete genomes that can represent gastroenteritis outbreaks during those periods. The presence and relevance of the animal, plant and insect viruses in wastewater were also discussed but other viruses as bacteriophages and amoeba or protozoa infecting viruses are not targets in this study.

Methods

Studied Area

The Federal District is the capital of Brazil, located in the Midwest region of the country with a population of 2.8 million. Brasilia city is located in the center of the Federal District with a population of 200.000. The Midwest region is defined by two seasons: rainy season (October to April) and dry-warm season (May to September).

Wastewater samples

Wastewater samples were obtained from the third wastewater treatment plant (wwtp) in the Federal District (ETE Norte) with an average flow of 450 L/s. It collects effluent from seven neighborhoods (Asa Norte, Vila Planalto, Lago Norte, Taquari, Vila Estrutural, Vila Varjão and Torto) that together have approximately 180.000 habitants. The sewage is treated using a conventional system of activated sludge, an aerobic process, and the final effluent is discharged in the Paranoá Lake. The artificial Paranoá Lake has an area of 38.1 km² and an estimated volume of 498.6 x 10⁶ m³.

All sample collections of wastewater were done at the station in the pre-treatment stage, where only a physical barrier is applied to reduce the presence of large debris in the water. The samples were collected on the first week of June (S1), July (S2) and August (S3), 2014 and on August (S4), 2015. For each sample, 1000mL of wastewater were collected and stored in a sterilized glass bottle kept on ice until transportation to the laboratory for immediate processing.

Virus concentration

In order to remove solid matter, the sample of 400 mL was first centrifuged for 20 min, at 5,000 x g at 4°C. Virus enrichment of the supernatant was performed by ultracentrifugation at 142,000 x g for 1h at 4°C (Beckman rotor Type 45Ti), over a 20% sucrose cushion. The pellet was resuspended with 2 mL of Milli-Q water and kept on ice.

Total RNA extraction and Next-Generation Sequencing

RNA extractions were performed using ZR Soil/Fecal RNA MicroPrep (Zymo Research) according to the manufacturer's instructions. Briefly, the concentrated sample (200 µL out of 2 mL) was mixed with lysis buffer (1 mL) and the lysate was passed through three different steps of filtrations using specific columns for large DNA elimination, total RNA elution and polyphenol removal. For each sample, 5µg of total RNA were dried in an RNASTable column (Biomatrix, San Diego, CA, USA) and sent for high-throughput sequencing at Macrogen Inc. (Seoul, Korea), where the samples were resuspended and further treated for rRNA removal using Ribo-Zero rRNA removal kit (Bacteria) (Illumina, San Diego, CA, USA) and cDNA library construction using TruSeq RNA library preparation kit (Illumina). The samples were sequenced at Macrogen Inc. (Seoul, Republic of Korea) using Illumina HiSeq 2000 paired-end method in 5G scale for each sample

Data Analysis

The reads from each sample were quality verified; adapter sequences removed and contigs were *de novo* assembled using CLC Genomics Workbench version 6.3. The commercial program Geneious Biomatters version 9.1.5 was used to perform

Blastx searches using the assembled contigs against the viral genome RefSeq database obtained from GenBank (<http://www.ncbi.nlm.nih.gov/genome/viruses/>).

The results were organized into viral (hit) and non-viral (no hit). The viral hits were further sorted into viruses that infect plants, insects, animals (humans included) and other viruses.

To assemble the largest viral genomic sequences as possible for the gastroenteritis viruses that were found in the samples, contigs associated to each virus were used in the “map to reference” tool of the Geneious program. This way, the de-novo assembly tests all other reads in the dataset for their contiguity to the initial gastroenteritis virus contigs generated. This technique allows for an extension of the original contigs leading to full length or near full-length genomic sequences. All complete genomes assembled were annotated in GenBank under the following numbers (XXXXXX).

Results

Total Data Set

The workflow was shown in Fig 1 from the sample collection to (near)-complete genome assembly. After quality control, the following amounts of trimmed paired reads were obtained: 27,260,854 (June_2014 (S1)); 27,268,978 (July_2014 (S2)); 27,270,556 (August_2014 (S3)) and 26,929,026 (August_2015 (S4)). The great similarity of obtained reads between samples is an indicative of good quality pattern maintained during sample processing.

As shown in Fig 2, virus and virus-related reads composed less than 2% of all contigs. Those numbers were obtained after Blastx search against viral RefSeq with results presented as viral hit and no-hit. It is important to highlight that together the

sum of all contigs of no-hit is 628,025 and this could hide a very gross fraction of yet unknown viruses.

We distributed the viral hits into three different classes: plant, insect and animal-infecting viruses as shown in Fig 3. The dataset from samples S1 was the one with lesser viral hits, however it was the sample that had more contigs assembled to animal viruses including human viruses that represented 3.23% of all viral sequences obtained. Samples S2, S3 and S4 had a similar number of viral hits however the distribution of viral hits between classes, as shown in Fig 3, demonstrated that S2 had similar amounts of contigs for plant and animal-infecting viruses, while S3 and S4 had a similar number of hits between the three classes. Regarding human virus hit the results were inferior to S1 being 1.65 % to S2, 0.68 % to S3 and finally 1.06 % of each respectively total virus contig hit. The results of BlastX are represented in four tables (supplementary table S1, S2, S3, S4).

The results concerning most hit contigs are described below. Where the main focus of these study was to detect and describe the most important acute gastroenteritis viruses, only those had their complete genomes assembled and described.

Human viruses

The most representative human viruses in all samples were human *picobirnavirus*. and husavirus. The human *picobirnavirus* is a small, spherical, non-enveloped virus with two double-stranded RNA genomic segments reported in respiratory and fecal human samples (90). *Husavirus* was found in samples S1 and S3 and *Picobirnavirus*, in S2 and S4. Husavirus named human stool-associated RNA virus was described as a novel genus in the order *Picornavirales* that contains one

single ORF and ssRNA positive sense with approximately 9kb and VPg (viral protein genome associated) at the 5' end followed by type I IRES and a poly(A) tract at the 3' end (91). Enterovirus C of the family *Picornaviridae* was found in samples S1 and S2, with a total of 20 contigs (from 72 to 2209 bps).

The major etiologic agents of human gastroenteritis were detected in all samples and those that had their complete genomes assembled are assigned in Table 1.

Aichi virus (accession number:xxxx) is a picornavirus associated with human acute gastroenteritis that was detected in all sewage samples but S3. All complete genomes were deposited in GenBank under following accession numbers: aichi virus S1 (XXXXXX); aichi virus S2 (XXXXXX); aichi virus S4 (XXXXXX). In all three samples positive for aichi virus the genomes assembled were classified as genotype B and presented nucleotide identity of 98% (S1), 98% (S2) and 96% (S4) to reference genome FJ890523.

Rotavirus A was detected in all samples except S1. All genomic segments of Rotavirus A were found and the longest contig sequence belonged to segment 1 with 1088bp. This virus belongs to the *Reoviridae* family and has a dsRNA genome divided in 11 segments that codify at least 13 proteins. The non-enveloped, non-turreted virion with a triple capsid structure has about 80nm in diameter (92). It is still the major cause of acute gastroenteritis around the world especially in countries where the vaccination is not implemented (93).

Astrovirus, another pathogen that causes gastroenteritis in young children, was detected in all samples. One complete genome of human astrovirus was assembled (accession number: xxxxx) (Table 1). The virus possesses non-enveloped virion with positive ssRNA and three open reading frames (94, 95). The assembled virus genome

of S1, S2, S3 and S4 showed higher identities to astrovirus strain 1 HQ398856.2. Also in S2 two other genomes were assembled, one from strain 4 with 99% identity to KF039912.1 and the other from strain 5 with 97% identity to KF157967.1.

The most representative gastroenteritis viruses are of the *Caliciviridae* family, namely *Norovirus* and *Sapovirus*. They are non-enveloped viruses, ssRNA positive sense, with two or three ORFs and genomes with approximately 8kb with VPg at the 5' end and poly(A) at the 3' end (96). Presently, they are the major viral causative agents of acute gastroenteritis in children under 5 years old and elderly people (17). Two norovirus and three sapovirus genogroups were assembled (Table 1) (accession numbers: xxxxx, xxxxx, xxxx, xxxxx...). The number of contigs and sequence lengths can be found in Supplementary Table S1. These viruses are classified by the capsid gene nucleotide sequence. They are also divided in genogroups and genotypes. The most epidemic norovirus genotype is GII.4 (97) and it was detected in our samples. Since 2014, in China, GII.17 was described as new epidemic genotype because of its persistence in population. Surprisingly this new epidemic genotype was detected in S4 (August_2015) and its entire genome was assembled with 98% nucleotide identity to reference genome KT780404.

Animal viruses

In Figure 3 it is possible to see that a great part of the viral hits of all samples is related to animal viruses. The presence of these in domestic sewage could be related to zoonosis. In all samples, viruses described as pathogens for domestic animals were detected after Blastx such as canine astrovirus and porcine picobirnavirus. Both were detected at least in three out of four samples. Canine astrovirus is a pathogen described as a cause of diarrheic disease in dogs and is

usually detected in mixed infections with canine parvovirus and canine coronavirus, although these two were not identified in any of samples of this study.

The picobirnavirus was first described in Brazil in human fecal samples and from feces of wild black-foot pigmy rice rats (*Oryzomys nigripes*) in 1988. The isolates from human, porcine and those from environmental samples usually present genetic homology.

Plant viruses

Important pathogens in plants were also encountered in our survey such as Bean necrotic virus, Melon yellow virus and Pepper mild mottle virus (PMMoV). The latter is extensively mentioned in literature as a possible wastewater treatment biomarker for quality control (67, 75, 98-101). Viruses like Garlic common latent virus (GCLV) were also found in the samples. GCLV is a *Carlavirus* that belongs to the *Betaflexiviridae* family. It has a non-enveloped, flexuous, filamentous capsid with 470-1000nm length and 13nm in diameter. The linear single-stranded positive RNA has six open reading frames and the genome size ranges between 5.8 to 9kb. Carlaviruses have a 5'end cap and a poly-(A) tract at the 3'end (102).

PMMoV is a tobamovirus that belongs to the *Virgaviridae* family. It is a non-enveloped virion with a rigid helical rod capsid with 18nm in diameter and 300-310nm in length. The genome is linear, single-stranded positive RNA with an average length of 6.3-6.5kb. PMMoV genomes contain four open reading frames and have a cap at the 5'end and a tRNA-like structure at the 3'end (103, 104). PMMoV wasn't detected in sample S1 but in the other samples it was possible to count 25 hits on Blastx with minimum sequence of 70bp and maximum of 809bp.

Insect viruses

In all four samples several invertebrate viruses were detected like *Autographa californica* virus, *Helicoverpa zea* nudivirus, *Solenopsis invicta* and some important members of the family *Dicistroviridae* like Acute bee paralysis virus and Aphid lethal paralysis virus detected in all four samples. Dicistroviruses have ssRNA, positive sense, linear genomes of 8.5-10.2 kb with two open reading frames separated by an internal ribosome entry site (IRES) element and with VPg in their 5' ends and poly A in their 3' ends. They have a non-enveloped capsid of approximately 30nm of diameter (105). Aphid lethal paralysis virus was isolated from bird cherry-oat aphid (*Rhopalosiphum padi*) a well described wheat virus vector (106). This virus was mostly detected in sample S1 with 5 hits on Blastx with minimum sequence of 807bp and maximum of 2,035 bp, all of them had an e-value equal to 0.

Acute bee paralysis virus is described as a virus that can kill pupae and adult honeybees (107). It was found in all samples with a total of 24 hits on Blastx with minimum sequence length of 44 bp and maximum of 1,906 bp with best e-value of 4.49e-160.

Discussion

Viral metagenomic studies of environmental samples have been shown as efficient approaches to detect human and animal-related pathogens but also viruses that are part of the ecosystem in a symbiotic way. The genome sequence information obtained through NGS is valuable especially for uncultured viruses. The main objective of this study was to describe human enteric viruses circulating in Brasilia, Brazil, during dry season months on the year of 2014 and 2015. Unexpectedly, our analysis showed that

the sewage water virome is very useful not only to monitor for human viruses but also for plant viruses.

Using next generation sequence paired-end method the data obtained was first classified between viral match and non-viral match. Here it is important to point out that the majority of viral matches were of phages and giant viruses in all samples as previously observed by Cantalupo et al in 2011 from samples collected in Barcelona (Spain), Pittsburgh (United States of America) and Addis Ababa (Ethiopia)(74).

In this study, at least 49 different human viruses were found by Blastx analysis against a viral RefSeq genome database. Among those, all important human acute gastroenteritis RNA viruses were detected and the most epidemics had examples of complete genome assembled (Table 1). However, many of them were not fully studied due to the lack of contigs or reads or low identities to reference viruses.

Among gastroenteritis viruses, the number of contigs of rotavirus was unexpectedly low, possibly due to the widespread vaccination of children under 2 years old in Brazil. On the other hand, human astrovirus were found in all samples, as the third most common viral cause of acute gastroenteritis in children under 5 years old. This virus is estimated to be involved in 0.5 % to 15 % of all viral enteric outbreaks (2, 108). In detail, astrovirus genome assembled was identified as MLB1 and MLB2 in this study and the first one was first reported in Brazil only in 2015 using clinical samples from South, Southeast and Northeast coastal regions(109). This shows that the wastewater virome analysis is also useful for gastroenteritis virus monitoring.

The *Caliciviridae* family contains two genera that infect humans, *Norovirus* and *Sapovirus*. Both are divided into genogroups and genotypes being genogroup GII for norovirus and GI for sapovirus the most prevalent. Norovirus are considered the

cause of almost 50% of gastroenteritis worldwide (110) and is broadly detected in clinical samples, rivers and wastewater (61, 88, 89, 111). In this study norovirus GI.2, GI.6, GII.4 and GII.17 were identified and their complete genomes were assembled from the sewage samples. The genotype GII.17 is a novel variant that emerged and persisted in Asian countries replacing GII.4 from hospitalized patients (112). Norovirus GII.17 was first detected in Brazil in 2016 in stool samples collected in the North region (113). The complete genome of genotype GII.17 demonstrated 98% identity to the GII.17 reference genome available in GenBank (accession number: KT780404).

Unlike norovirus, the human sapovirus are less described in Brazilian clinical literature (61). Only recently the first environmental study detected the presence of sapoviruses in sewage in the state of Rio de Janeiro by RT-qPCR (61). It has been shown that even though a high number of positive environmental samples for sapoviruses is common, only few positive samples have been found in clinical investigations(61). It is assumed that sapovirus infections of humans are associated to subclinical conditions, as commonly observed for Norovirus GI. In this study, the complete genomes of SaV genotypes GI.1, GI.2, GIV and GV were assembled, corroborating to other studies performed in water and wastewater in Italy, Spain and Japan (114-116).

Among the plant viruses detected it was noticed in half of the samples the presence of Pepper mild mottle virus, a virus that can infect all species in the genus *Capsicum*. Surprisingly this virus was detected in abundance in another RNA virus human feces metagenomic study performed using sample from healthy individuals (75). In this work not only this plant viruses was detected but also contigs were assembled to Tobacco mosaic virus, Tomato mosaic virus, Maize chlorotic mottle

virus, Melon necrotic spot virus and others. Interestingly, all of them have also been previously detected in the gastrointestinal tract of humans (75).

The data input also presented results assembled for important insect viruses from the *Dicistroviridae* family. One of them, Acute bee paralysis virus, was represented in all four samples. This virus is widespread in honeybees (*Apis mellifera*) presenting sub-clinical etiology or a highly virulent pathology that causes death (107). In the southeastern region of Brazil the presence of this viruses were demonstrated by qPCR in correlation to a decrease in the adult bee populations (117).

In conclusion this is the first study performed in Brazil to access viral diversity from raw sewage by only means next generation sequencing, resulting in the detection and description of several human acute gastroenteritis virus pathogens complete genomes. Our results demonstrate the diversity of enteric viruses circulating in the central region of Brazil and offer a simple methodology for human virus surveillance beyond individual clinical samples.

References

1. Fauquet, C.M. and D. Fargette, *International Committee on Taxonomy of Viruses and the 3,142 unassigned species*. Virol J, 2005. **2**: p. 64.
2. Cantalupo, P.G., et al., *Raw sewage harbors diverse viral populations*. MBio, 2011. **2**(5).
3. Bibby, K., E. Viau, and J. Peccia, *Viral metagenome analysis to guide human pathogen monitoring in environmental samples*. Lett Appl Microbiol, 2011. **52**(4): p. 386-92.
4. Edwards, R.A. and F. Rohwer, *Viral metagenomics*. Nat Rev Microbiol, 2005. **3**(6): p. 504-10.
5. Li, Y., et al., *VIP: an integrated pipeline for metagenomics of virus identification and discovery*. Sci Rep, 2016. **6**: p. 23774.
6. Berg, G., *Removal of viruses from sewage, effluents and waters. 2. Present and future trends*. Bull World Health Organ, 1973. **49**(5): p. 461-9.
7. Gerba, C.P., *Virus Occurrence and Survival in the Environmental Waters*, in *Human Viruses in Water*, A. Bosch, Editor. 2007, Elsevier B. V. All rights.

8. Alavandi, S.V. and M. Poornima, *Viral metagenomics: a tool for virus discovery and diversity in aquaculture*. Indian J Virol, 2012. **23**(2): p. 88-98.
9. Bibby, K. and J. Peccia, *Identification of viral pathogen diversity in sewage sludge by metagenome analysis*. Environ Sci Technol, 2013. **47**(4): p. 1945-51.
10. Hrynyszyn, A., M. Skonieczna, and J. Wiszniowski, *Methods for Detection of Viruses in Water and Wastewater*. Advances in Microbiology, 2013(3): p. 7.
11. Iaconelli, M., et al., *One-year Surveillance of Human Enteric Viruses in Raw and Treated Wastewaters, Downstream River Waters, and Drinking Waters*. Food Environ Virol, 2017. **9**(1): p. 79-88.
12. Iturriza-Gomara, M. and S.J. O'Brien, *Foodborne viral infections*. Curr Opin Infect Dis, 2016. **29**(5): p. 495-501.
13. Kitajima, M., et al., *Relative abundance and treatment reduction of viruses during wastewater treatment processes--identification of potential viral indicators*. Sci Total Environ, 2014. **488-489**: p. 290-6.
14. da Silva Polo, T., et al., *Human norovirus infection in Latin America*. J Clin Virol, 2016. **78**: p. 111-9.
15. Fioretti, J.M., et al., *Occurrence of human sapoviruses in wastewater and stool samples in Rio De Janeiro, Brazil*. J Appl Microbiol, 2016. **121**(3): p. 855-62.
16. Teixeira, D.M., et al., *Occurrence of Norovirus GIV in Environmental Water Samples from Belem City, Amazon Region, Brazil*. Food Environ Virol, 2016. **8**(1): p. 101-4.
17. Vieira, C.B., et al., *Viruses Surveillance Under Different Season Scenarios of the Negro River Basin, Amazonia, Brazil*. Food Environ Virol, 2016. **8**(1): p. 57-69.
18. Han, T.H., et al., *Detection of norovirus genogroup IV, klassevirus, and pepper mild mottle virus in sewage samples in South Korea*. Arch Virol, 2014. **159**(3): p. 457-63.
19. Zhang, T., et al., *RNA viral community in human feces: prevalence of plant pathogenic viruses*. PLoS Biol, 2006. **4**(1): p. e3.
20. Hamza, I.A., et al., *Evaluation of pepper mild mottle virus, human picobirnavirus and Torque teno virus as indicators of fecal contamination in river water*. Water Res, 2011. **45**(3): p. 1358-68.
21. Kuroda, K., et al., *Pepper mild mottle virus as an indicator and a tracer of fecal pollution in water environments: comparative evaluation with wastewater-tracer pharmaceuticals in Hanoi, Vietnam*. Sci Total Environ, 2015. **506-507**: p. 287-98.
22. Symonds, E.M., et al., *Faecal pollution along the southeastern coast of Florida and insight into the use of pepper mild mottle virus as an indicator*. J Appl Microbiol, 2016. **121**(5): p. 1469-1481.
23. Song, S.I., et al., *Complete genome sequence of garlic latent virus, a member of the carlavirus family*. Mol Cells, 2002. **14**(2): p. 205-13.
24. Wetter, C., et al., *Pepper mild mottle virus, a tobamovirus infecting pepper cultivars in Sicily*. Phytopathology, 1984(74): p. 5.
25. Rodriguez-Cerezo, E., A. Moya, and F. Garcia-Arenal, *Variability and evolution of the plant RNA virus pepper mild mottle virus*. J Virol, 1989. **63**(5): p. 2198-203.
26. Bonning, B.C., *The Dicistroviridae: An Emerging Family of Invertebrate Viruses*. Virologica Sinica, 2009. **24**(5): p. 12.

27. Liu, S., et al., *Analysis of new aphid lethal paralysis virus (ALPV) isolates suggests evolution of two ALPV species*. J Gen Virol, 2014. **95**(Pt 12): p. 2809-19.
28. de Miranda, J.R., G. Cordoni, and G. Budge, *The Acute bee paralysis virus-Kashmir bee virus-Israeli acute paralysis virus complex*. J Invertebr Pathol, 2010. **103 Suppl 1**: p. S30-47.
29. Oude Munnink, B.B., et al., *A novel genus in the order Picornavirales detected in human stool*. J Gen Virol, 2015. **96**(11): p. 3440-3.
30. Gallimore, C.I., et al., *Detection and characterisation of bisegmented double-stranded RNA viruses (picobirnaviruses) in human faecal specimens*. J Med Virol, 1995. **45**(2): p. 135-40.
31. Sasaki, J., et al., *Construction of an infectious cDNA clone of Aichi virus (a new member of the family Picornaviridae) and mutational analysis of a stem-loop structure at the 5' end of the genome*. J Virol, 2001. **75**(17): p. 8021-30.
32. Yamashita, T., et al., *Complete nucleotide sequence and genetic organization of Aichi virus, a distinct member of the Picornaviridae associated with acute gastroenteritis in humans*. J Virol, 1998. **72**(10): p. 8408-12.
33. Estes, M.K. and A.Z. Kapikian, *Fields Virology*, in *Fields Virology*, R.A. Knipe, et al., Editors. 2007, Lippincott, Williams and Wilkins: Philadelphia. p. 1917-1974.
34. O'Ryan, M., Y. Lucero, and A.C. Linhares, *Rotarix(R): vaccine performance 6 years postlicensure*. Expert Rev Vaccines, 2011. **10**(12): p. 1645-59.
35. Kurtz, J.B., T.W. Lee, and D. Pickering, *Astrovirus associated gastroenteritis in a children's ward*. J Clin Pathol, 1977. **30**(10): p. 948-52.
36. Lee, T.W. and J.B. Kurtz, *Human astrovirus serotypes*. J Hyg (Lond), 1982. **89**(3): p. 539-40.
37. Kapikian, A.Z., et al., *VISUALIZATION BY IMMUNE ELECTRON-MICROSCOPY OF A 27-NM PARTICLE ASSOCIATED WITH ACUTE INFECTIOUS NONBACTERIAL GASTROENTERITIS*. Journal of Virology, 1972. **10**(5): p. 1075-1081.
38. Jiang, X., et al., *Sequence and genomic organization of Norwalk virus*. Virology, 1993. **195**(1): p. 51-61.
39. Siebenga, J.J., et al., *Phylogenetic reconstruction reveals norovirus GII.4 epidemic expansions and their molecular determinants*. PLoS Pathog, 2010. **6**(5): p. e1000884.
40. Teixeira, E.W., et al., *Virus infections in Brazilian honey bees*. J Invertebr Pathol, 2008. **99**(1): p. 117-9.
41. Black, R.E., et al., *Global, regional, and national causes of child mortality in 2008: a systematic analysis*. Lancet, 2010. **375**(9730): p. 1969-87.
42. Bosch, A., R.M. Pinto, and S. Guix, *Human astroviruses*. Clin Microbiol Rev, 2014. **27**(4): p. 1048-74.
43. Xavier Mda, P., et al., *Surveillance of Human Astrovirus Infection in Brazil: The First Report of MLB1 Astrovirus*. PLoS One, 2015. **10**(8): p. e0135687.
44. Patel, M.M., et al., *Noroviruses: a comprehensive review*. J Clin Virol, 2009. **44**(1): p. 1-8.
45. Miagostovich, M.P., et al., *Molecular detection and characterization of gastroenteritis viruses occurring naturally in the stream waters of Manaus, central Amazonia, Brazil*. Appl Environ Microbiol, 2008. **74**(2): p. 375-82.
46. Zhang, X.F., et al., *An outbreak caused by GII.17 norovirus with a wide spectrum of HBGA-associated susceptibility*. Sci Rep, 2015. **5**: p. 17687.

47. Silva, L.D., et al., *Detection and genetic characterization of the emergent GII.17_2014 norovirus genotype among children with gastroenteritis from Northern Brazil*. *Infect Genet Evol*, 2016. **48**: p. 1-3.
48. Di Bartolo, I., et al., *Identification and Genotyping of Human Sapoviruses Collected from Sewage Water in Naples and Palermo, Italy, in 2011*. *Food Environ Virol*, 2013.
49. Kitajima, M., et al., *Detection and genetic analysis of human sapoviruses in river water in Japan*. *Appl Environ Microbiol*, 2010. **76**(8): p. 2461-7.
50. Sano, D., et al., *Quantification and genotyping of human sapoviruses in the Llobregat river catchment, Spain*. *Appl Environ Microbiol*, 2011. **77**(3): p. 1111-4.

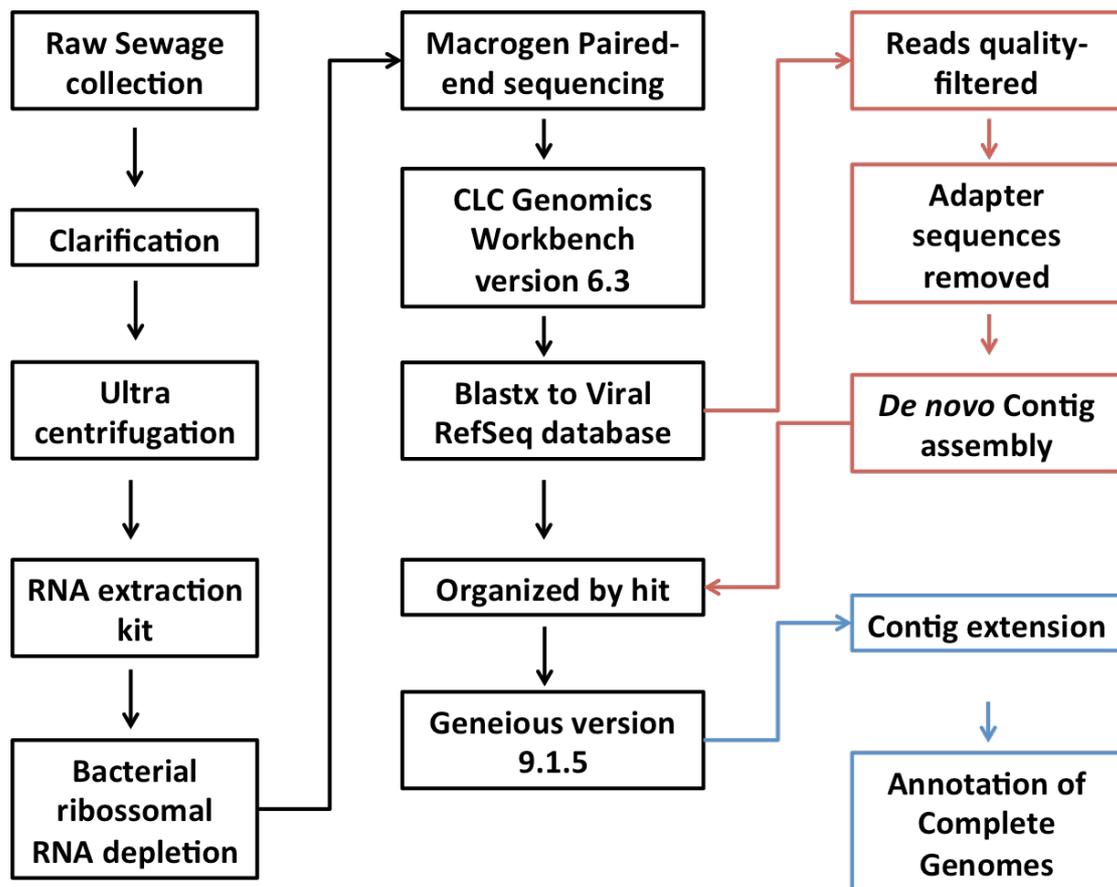


Figure 1 Workflow from collection to data analysis. From left to right, from top to bottom are described all steps performed to achieve the results in this work. The black boxes concern the work performed on the bench from sewage collection until sequence analysis. The red boxes are related to pre data treatment until contigs were assembled. In the blue boxes the two main steps for genome assembling.

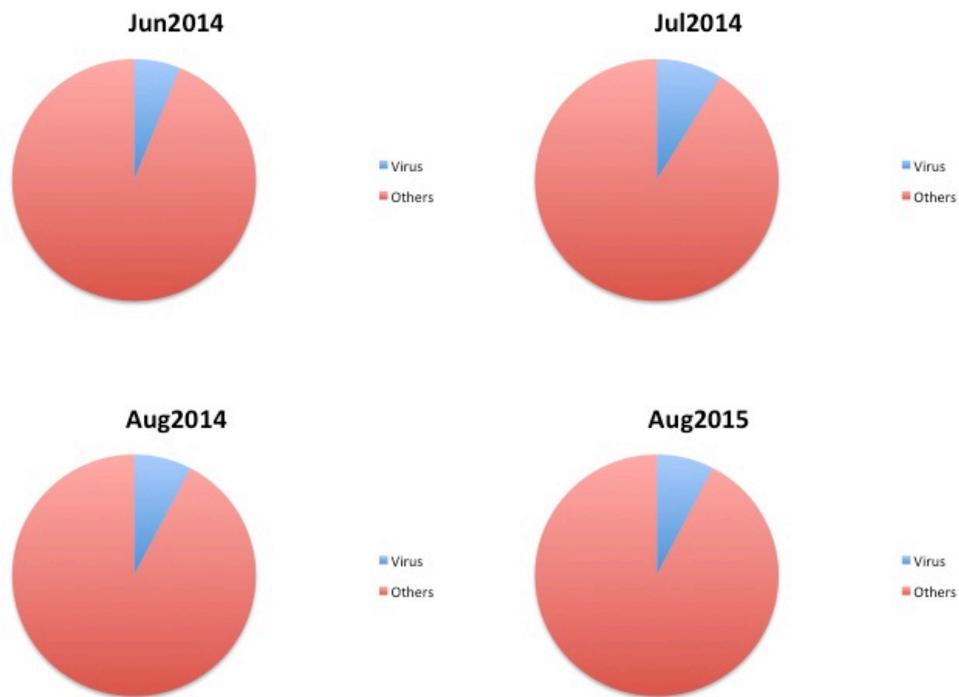


Figure 2 Illustration of all the reads from each sample. Jun 2014 (S1), Jul2014 (S2), Aug2014 (S3) and Aug2015 (S4) had the same profile concerning amount of viruses hit e non viruses.

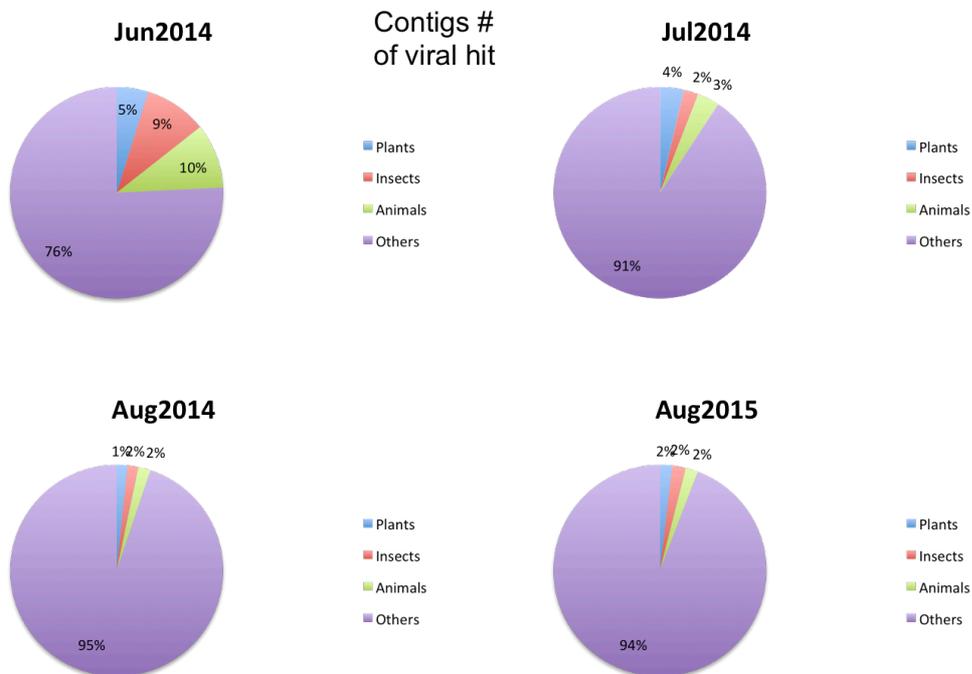


Figure 3 Illustration of different hosts viruses detected on sewage. In blue there is plants, red is insect, green animals and the majority is purple and is composed mostly by phages and giant viruses that were excluded from this work. Animals in this figure are representing the number of human's viruses also.

Table 2 Enteric viruses found in sewage samples that had their complete genome assembled.

Gastroenteritis Viruses Complete Genomes				
	Jun_2014	Jul_2014	Aug_2014	Aug_2015
Aichi virus	✓	✓		✓
Human Astrovirus 1	✓	✓	✓	✓
Human Astrovirus 4		✓		
Human Astrovirus 5		✓		
Norovirus GI.2		✓		
Norovirus GI.6	✓			
Norovirus GII.4	✓	✓		
Norovirus GII.17				✓
Rotavirus A		✓	✓	✓
Sapovirus GI.1	✓	✓	✓	
Sapovirus GI.2		✓		
Sapovirus GIV	✓			
Sapovirus GV	✓			

Table S1 Blast results from S1.

June_2014			
Contig #	Accession #	Sequence size	Virus hit
6	NC_001918	60-1276	Aichi virus 1
5	NC_011400	97-731	Astrovirus MLB1
1	NC_013060	758	Astrovirus VA1
5	NC_012800	2124	Cosavirus A
3	NC_001612	106-574	Enterovirus A
11	NC_001472	67-478	Enterovirus B
18	NC_002058	2209	Enterovirus C
1	NC_001434	660	Hepatitis E virus
1	NC_001489	2225	Hepatovirus A
3	NC_013443	66-348	HMO Astrovirus A
1	NC_001806	481	Human alphaherpesvirus 1
12	NC_001943	54-367	Human astrovirus
4	NC_024694	90-132	Human circovirus VS6600022
4	NC_006577	70-140	Human coronavirus HKU1
1	NC_023984	143	Human cosavirus
1	NC_012802	185	Human cosavirus D
1	NC_012798	71	Human cosavirus E
4	NC_022518	89-104	Human endogenous retrovirus K113
5	NC_030454	69-109	Human enterovirus
5	NC_030454	2187	Human enterovirus
3	NC_009333	94-550	Human gammaherpesvirus 8
1	NC_009334	1015	Human herpesvirus 4 type 2
1	NC_001802	1519	Human immunodeficiency virus 1
11	NC_001897	72-630	Human parechovirus
11	NC_001897	2179	Human parechovirus
7	NC_010810	71-227	Human TMEV-like cardiovirus
30	NC_032480	2893	Husavirus sp.
11	NC_027126	2776	Kilifi Virus
8	NC_030922	98-176	Mamastrovirus 1
19	NC_001959	84-548	Norovirus GI
22	NC_029646	32-598	Norovirus GII
1	NC_029647	242	Norovirus GIV
8	NC_001617	69-184	Rhinovirus A
2	NC_009996	120-184	Rhinovirus C
2	NC_012986	151-642	Salivirus A
2	NC_012957	81-342	Salivirus NG-J1
4	NC_006554	69-447	Sapovirus C12
11	NC_006269	87-712	Sapovirus Hu/Dresden/pJG-Sap01/DE Sapovirus Hu/Nagoya/NGY- 1/2012/JPN
7	NC_027026	120-452	
5	NC_010624	78-978	Sapovirus Mc10
8	NC_008291	68-277	Taterapox virus
3	NC_006998	108-844	Vaccinia virus
2	NC_001611	151-844	Variola virus

Table S2 Blast results from S2.

July_2014			
Contig #	Accession #	Sequence size	Virus hit
9	NC_001918	88-303	Aichi virus 1
1	NC_011400	787	Astrovirus MLB1
1	NC_016155	787	Astrovirus MLB2
14	NC_032426	68-144	Bastrovirus-like_virus/VietNam/Bat/17819_21
1	NC_013443	77	HMO Astrovirus A
3	NC_001943	182-853	Human astrovirus
2	NC_024694	108-174	Human circovirus VS6600022
103	NC_007026	30-556	Human picobirnavirus
25	NC_032480	20-305	Husavirus sp.
18	NC_027126	72-353	Kilifi Virus
4	NC_030922	119-917	Mamastrovirus 1
17	NC_001959	68-351	Norovirus GI
16	NC_029646	80-919	Norovirus GII
1	NC_011501	315	Rotavirus A
2	NC_011502	105-317	Rotavirus A
2	NC_011503	247-325	Rotavirus A
2	NC_011504	172-166	Rotavirus A
2	NC_011505	77-108	Rotavirus A
3	NC_011506	88-894	Rotavirus A
3	NC_011507	393-1088	Rotavirus A
2	NC_011508	101-716	Rotavirus A
2	NC_011509	39-317	Rotavirus A
1	NC_011510	776	Rotavirus A
2	NC_007547	84-89	Rotavirus C
1	NC_007572	69	Rotavirus C
1	NC_006554	117	Sapovirus C12
20	NC_006269	73-642	Sapovirus Hu/Dresden/pJG-Sap01/DE
9	NC_027026	89-604	Sapovirus Hu/Nagoya/NGY-1/2012/JPN
1	NC_010624	105	Sapovirus Mc10
2	NC_001611	67-73	Variola virus

Table S3 Blast results from S3.

August_2014			
Contig #	Accession #	Sequence size	Virus hit
7	NC_011400	92-415	Astrovirus MLB1
2	NC_032426	86-1399	Bastrovirus-like_virus/VietNam/Bat/17819_21
5	NC_001943	72-191	Human astrovirus
1	NC_006273	145	Human betaherpesvirus 5
1	NC_009334	414	Human herpesvirus 4 type 2
1	NC_001897	169	Human parechovirus
14	NC_007026	44-149	Human picobirnavirus
16	NC_032480	76-226	Husavirus sp.
6	NC_027126	74-141	Kilifi Virus
14	NC_030922	52-508	Mamastrovirus 1
2	NC_001959	99-148	Norovirus GI
2	NC_029646	115-121	Norovirus GII
2	NC_011500	132-505	Rotavirus A
1	NC_011501	315	Rotavirus A
1	NC_011502	317	Rotavirus A
2	NC_011503	242-326	Rotavirus A
2	NC_011504	110-175	Rotavirus A
1	NC_011505	198	Rotavirus A
1	NC_011506	893	Rotavirus A
2	NC_011507	444-634	Rotavirus A
2	NC_011508	373-450	Rotavirus A
1	NC_011509	397	Rotavirus A
1	NC_011510	775	Rotavirus A
4	NC_006554	85-194	Sapovirus C12
3	NC_006269	72-170	Sapovirus Hu/Dresden/pJG-Sap01/DE
8	NC_010624	82-139	Sapovirus Mc10
2	NC_008291	56-83	Taterapox virus

Table S4 Blast results from S4.

August_2015			
Contig #	Accession #	Sequence size	Virus hit
5	NC_001918	96-256	Aichi virus 1
6	NC_011400	106-438	Astrovirus MLB1
3	NC_016155	60-121	Astrovirus MLB2
7	NC_013060	67-182	Astrovirus VA1
1	NC_019027	76	Astrovirus VA4
6	NC_032426	69-391	Bastrovirus-like_virus/VietNam/Bat/17819_21
2	NC_002058	72-103	Enterovirus C
4	NC_013443	85-107	HMO Astrovirus A
11	NC_001943	71-725	Human astrovirus
28	NC_007026	62-55	Human picobirnavirus
14	NC_032480	67-318	Husavirus sp.
7	NC_027126	76-381	Kilifi Virus
7	NC_030922	52-1082	Mamastrovirus 1
15	NC_029646	81-530	Norovirus GII
1	NC_011500	413	Rotavirus A
1	NC_011501	315	Rotavirus A
2	NC_011502	139-188	Rotavirus A
2	NC_011503	73-203	Rotavirus A
1	NC_011504	118	Rotavirus A
1	NC_011506	868	Rotavirus A
2	NC_011507	413-621	Rotavirus A
1	NC_011508	835	Rotavirus A
1	NC_011509	147	Rotavirus A
2	NC_011510	242-326	Rotavirus A
2	NC_006269	94-131	Sapovirus Hu/Dresden/pJG-Sap01/DE
1	NC_027026	103	Sapovirus Hu/Nagoya/NGY-1/2012/JPN

Novel bastrovirus genome in raw sewage

Karoline dos Anjos^a, Fernando Lucas Melo^{a#}, Tatsuya Nagata^a

Department of Cell Biology, University of Brasilia, Campus Universitário Darcy

Ribeiro, Brasília, Distrito Federal 70910-900, Brazil^a

Running Title: Novel Bastrovirus in sewage

#Address correspondence to Fernando L. Melo, flmelo@unb.br

The complete genome sequence of a novel bastrovirus was determined. The genome has 5,875 nucleotides and shares 56% nucleotide identity with a bastrovirus-like virus found in bat feces. Astroviruses are one of the most common causes of human acute gastroenteritis. They belong to Astroviridae family and are currently classified into two genus: *Avastrovirus* and *Mamastrovirus*, which infect birds and mammals, respectively (<https://data.ictvonline.org/proposals/ICTV%206th%20Report.pdf>). The viruses are characterized as small rounded particles, non-enveloped, with linear, positive-sense and single-stranded RNA genome (118). Moreover, astrovirus-like viruses have been identified in feces of health humans (119), bats and rats and also in invertebrates (120). Despite most infections caused by astroviruses are usually asymptomatic, the emergence of novel astroviruses and their zoonotic potential should not be ignored. Here we report a draft of a complete genome of a bastrovirus identified in raw sewage sample analyzed in Brasilia, Federal District, Brazil.

The sewage sample was centrifuged in a low rotation to remove the solid matter followed by an ultracentrifugation with 20% sucrose cushion to viral particles concentration. RNA extraction was performed using ZR Soil/Fecal RNA MicroPrep (Zymo Research). The sample was processed for rRNA removal using Ribo-Zero rRNA removal kit (Bacteria) (Illumina, San Diego, CA, USA) and cDNA library construction using TruSeq RNA library preparation kit (Illumina). The samples were sequenced at Macrogen (Seoul, Republic of Korea) using Illumina HiSeq 2000 paired-end method. The raw reads were quality trimmed and *de novo* assembled using CLC Genomics Workbench version 6.3 (<http://www.clcbio.com/products/clc-genomics-workbench/>). The resulting contigs were compared to astrovirus complete genomes available in genbank using BLASTx (121) implemented in Geneious version 9.1.5 (122). One assembled contig were related to bastroviruses, a basal group of astroviruses (2).

The genome consists of 5875 nt and three ORFs were identified: nucleotide position 75 to 4298 (ORF1), 4324 to 5376 (ORF2) and 5406 to 5747 (ORF3). A brief analysis of the ORF1 encoded protein (1407 aa) revealed three distinct domains of v-metil transferase (29 to 428 aa), viral helicase-1 (607 to 849 aa) and viral RdRp-2 (969 to 1334 aa). The analysis ORF2 encoded protein (750 aa) revealed a structural domain of calicivirus coat protein (24 to 137 aa). No domain was found for the ORF3 hypothetical protein (342 aa).

The phylogenetic analysis using all complete bastrovirus genomes (7 from human, 5 from rat, 3 from bat and 1 from porcine) revealed that the Sewage-associated bastrovirus clustered with a highly divergent bastrovirus-like found in bat (KX907135) in Vietnam, which present only two orfs, with 56% nucleotide identity over the entire genome. Despite been probably associated with a mammalian host, it

is impossible from our results to confirm that and further studies need to be performed.

Nucleotide sequence accession number.

The nucleotide sequence has been deposited in GenBank under the accession number MF042208

ACKNOWLEDGMENTS

This work was funded by grants from CNPq (Project #XXXXXX). TN is a CNPq fellow.

REFERENCES

1. **Mendez E, Arias CF.** 2007. Astroviruses, p 981-1000. *In* Howley PM, Knipe DM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (ed), *Fields Virology*, vol I. Lippincott Williams & Wilkins, Philadelphia.
2. **Oude Munnink BB, Cotten M, Canuti M, Deijs M, Jebbink MF, van Hemert FJ, Phan MV, Bakker M, Jazaeri Farsani SM, Kellam P, van der Hoek L.** 2016. A Novel Astrovirus-Like RNA Virus Detected in Human Stool. *Virus Evol* **2**:vew005.
3. **Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ.** 2016. Redefining the invertebrate RNA virosphere. *Nature* doi:10.1038/nature20167.
4. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**:3389-3402.
5. **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A.** 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647-1649.

CAPITULO II – Biologia Estrutural dos Calicivírus Humanos

1. Nota de apresentação do capítulo

O capítulo II dessa tese refere-se ao trabalho realizado no período sanduíche de 2014 a 2015 na Universidade de Heidelberg e Centro de Pesquisa de Câncer da Alemanha (DKFZ). O trabalho foi realizado em colaboração com o Dr. Grant S Hansman.

Número da bolsa fomentada pelo CNPq: SWE 249420/2013-9.

2. Introdução

A obtenção de estruturas tridimensionais de partículas virais constituiu uma importante ferramenta para a maior compreensão das funções biológicas desses microrganismos. Questões a cerca da montagem, desmontagem, evasão da resposta imunológica do hospedeiro, entrada nas células, apresentação do genoma viral para replicação entre outras inúmeras questões têm sido respondidas por meio de estudos estruturais.

As técnicas de microscopia eletrônica e cristalografia com difração de raios-X contribuíram nas últimas décadas para a geração de respostas as questões supracitadas. A microscopia eletrônica com contrastação negativa proveu as primeiras imagens de partículas virais e a classificação a partir da morfologia, enquanto que informações relativas às interações das subunidades do capsídeo em nível atômico foram obtidas a partir de cristalografia com difração de raios-X.

Considerando-se que desde sua primeira identificação em 1970 (15) os calicivírus humanos ainda não possuem uma cultura celular estável que permita estudos a respeito da biologia viral, os dados obtidos por meio da biologia estrutural contribuíram para um maior conhecimento a respeito desses vírus.

Os calicivírus são constituídos morfologicamente de 90 dímeros de VP1 organizados em simetria icosaédrica onde $T=3$ (número triangular) (Fig. 5). A simetria $T=3$ resulta em duas classes de capsômero. Uma composta por monômeros organizados em eixo de simetria de ordem 5 em dois ambientes quasi-equivalentes diferentes denominados A e B, e outro composto por dímeros organizada em torno eixo de simetria de ordem 2 onde as moléculas de VP1 ocupam espaços equivalentes denominado C-C (Fig. 5) (20).

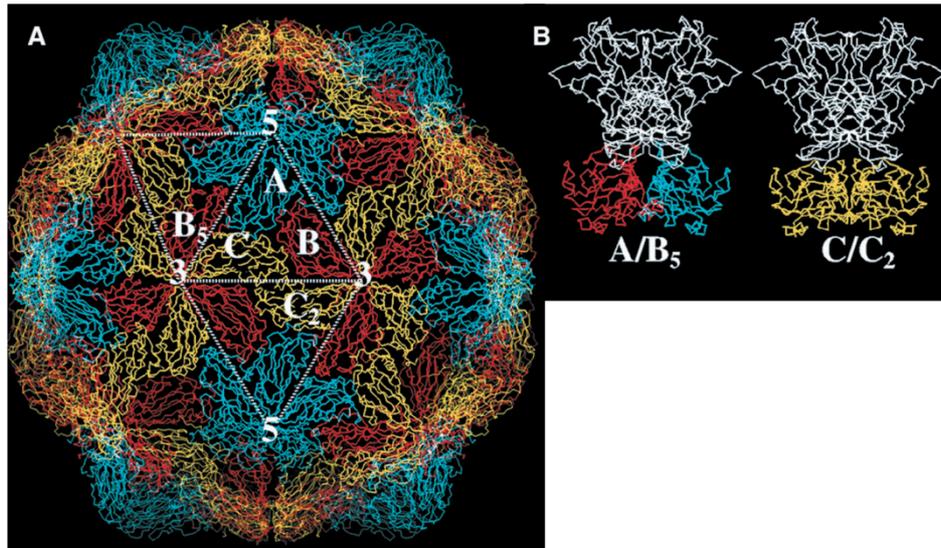


Figura 5 Representação do capsídeo de *Norwalk virus*. (A) Interações entre os domínios S do capsídeo de *Norwalk virus* Ilustração da simetria icosaédrica T=3 com três subunidades quasi-equivalentes, A, B e C, coloridas em azul, vermelho e amarelo respectivamente. Os os vértices dos triângulos indicam 5 e 3 eixos. (B) Dímeros formados entre as subunidades em domínios quasi-equivalentes. A esquerda A/B e a direita C/C. (Modificada de Prasad et al, 1999).

Com o uso de proteínas recombinantes foi possível determinar algumas estruturas de calicivírus humanos e animais. Em 1994, um estudo comparou, utilizando crio-microscopia eletrônica, capsídeos recombinantes expressos em células de inseto, de cada gênero da família *Caliciviridae*. A resolução média entre as quatro estruturas ficou em torno de 19 Å. Foram observadas 180 moléculas de VP1 organizadas em 90 dímeros (Fig. 6). O domínio P, descrito como protuberâncias na superfície das esferas, apresentou uma curvatura mais acentuada para os gêneros *Sapovirus* e *Lagovirus*, enquanto que para os gêneros *Norovirus* e *Vesivirus* o mesmo domínio apresentou extremidades mais arredondadas (13). Um estudo similar comparou um *Vesivirus*, identificado em ovários de camundongos e, outros calicivírus com resolução aproximada de 10 Å. A conclusão no entanto foi de que esse tipo era mais próximo dos sapovírus em morfologia do que dos calicivírus felinos (FCV) pertencentes ao mesmo gênero (123).

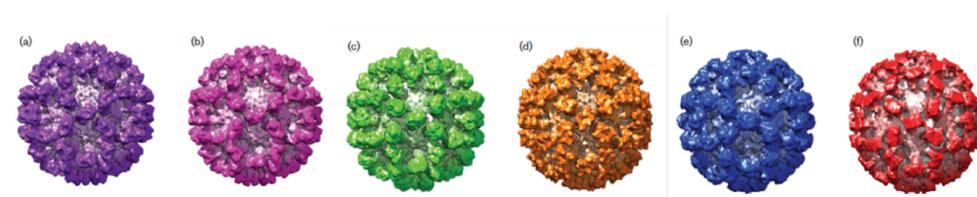


Figura 6 Imagens dos capsídeos lado a lado comparados. (a) vesivírus isolado de camundongo, (b) sapovírus (c) vírus causador da doença hemorrágica em coelhos (RHDV) (d) calicivírus felino (FCV), (e) *Norwalk virus* e (f) norovírus GII.10. (Adaptado de Conley et al 2017).

Em 1999, Prasad e colaboradores determinaram por meio de cristalografia (difração de raios-X) a estrutura do capsídeo recombinante de norovírus numa resolução de 3,5 Å (Fig. 5). A região N-terminal contém 225 aminoácidos que constituem 8 fitas-β antiparalelas e formam o domínio S. Essa estrutura foi observada em outros capsídeos virais, no entanto a região C-terminal contém também 225 resíduos que formam o domínio P que não foram observados em nenhuma outra proteína viral. Foi nesse estudo onde pela primeira vez a estrutura de 6 fitas-β formando um barril foi descrita como o subdomínio P2, região mais variável (em nucleotídeos) do capsídeo de norovírus e demais calicivírus (20).

Nos anos 2000, uma série de estudos foram realizados para determinar proteínas não estruturais e suas atividades, como por exemplo, a protease 3C-like de norovírus humano. Essa proteína foi resolvida numa resolução de 1,5 Å, formando um dímero que contém uma tríade catalítica composta por His30, Glu54 e Cys139, sendo Glu54 fundamental para atividade enzimática (124) (Fig. 7).

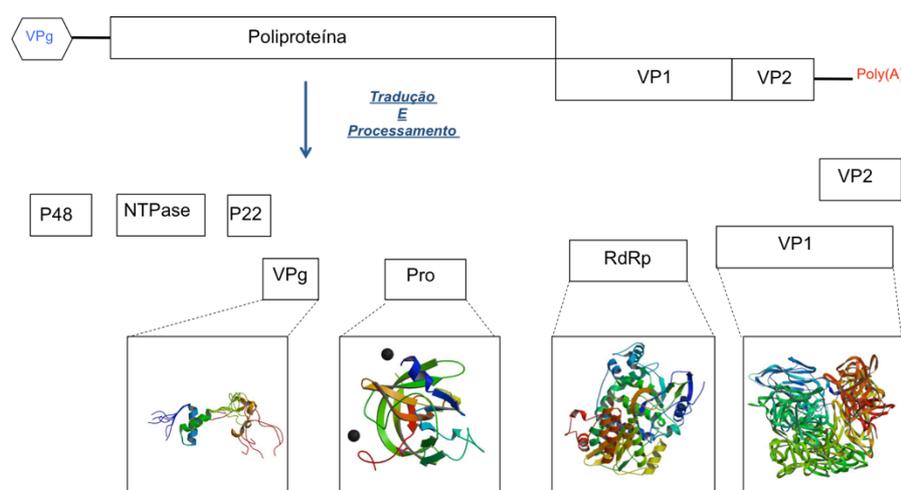


Figura 7 Representação do genoma de norovírus e sua possível forma de tradução e processamento da poliproteína presente na ORF1. O intuito dessa figura é ilustrar as estruturas resolvidas e publicadas no vinculando-as ao seu gene de expressão. As estruturas ilustradas estão disponíveis sob as seguintes identificações: VPg (norovírus murino) PDB: 2M4G; Protease PDB: 2FYQ; RdRp PDB: 1SH0 e VP1 PDB: 1HIM.

A conformação em mão direita de homodímeros da RNA polimerase dependente de RNA (RpRd) foi resolvida numa resolução de 2,3 Å. Nesse trabalho foi demonstrada ainda a síntese de RNA com a formação de dímeros (125). A estrutura da RpRd ligada a um substrato natural foi resolvida numa resolução de 1.8 Å (126) (Fig.7).

Há alguns anos, estudos vêm descrevendo a importância de carboidratos encontrados em superfície de hemácias, epitélios celular da mucosa dos tratos respiratório, geniturinário e digestivo (127). Esses carboidratos são denominados antígenos do sistema sanguíneo (do inglês *histo blood group antigens* HBGA). Sua importância dentre os calicivírus deu-se a partir de sua identificação como possível receptor ou interceptor celular para entrada dos norovírus na células. Esses carboidratos são determinantes do sistema ABO e Lewis, a presença de fucose é comum a todos eles (Fig. 8).

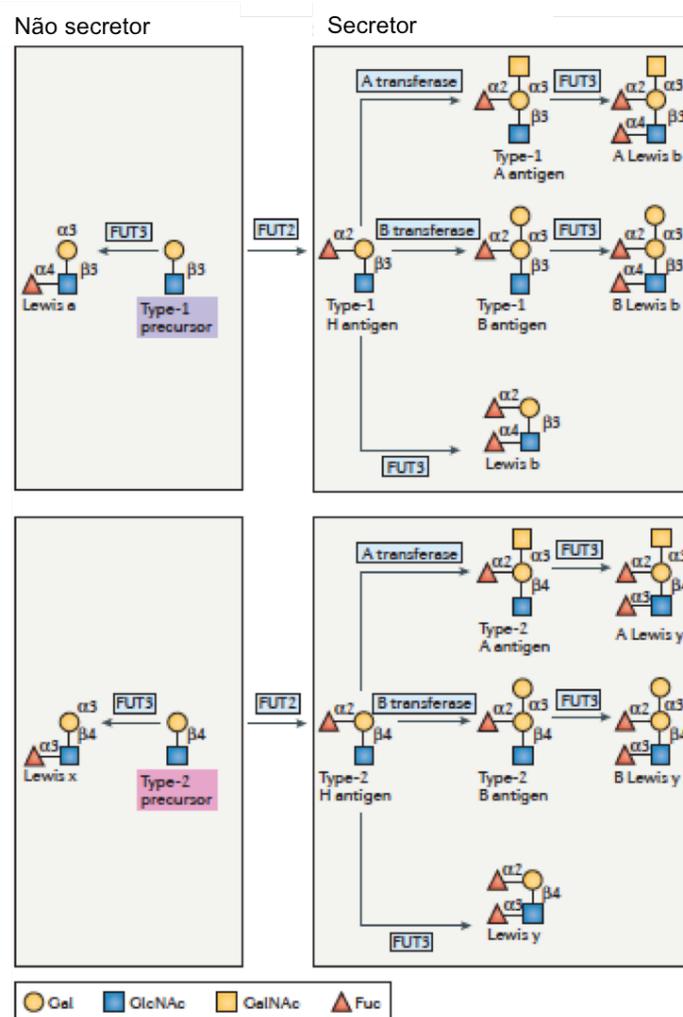


Figura 8 Síntese de antígenos do grupo sanguíneo. Proteínas conhecidas como HBGAs podem ser relacionadas a ligação e entrada de norovírus na célula hospedeira. Os precursores de HBGA de tipo 1 e tipo 2 são modificados por α (1,2) -fucosiltransferase 2 (FUT2) para produzir H HBGAs, que são ainda modificados por transferases A e B para Produzem A e B HBGAs. Lewis HBGAs são determinantes do sistema de grupos sanguíneos de Lewis e são produzidos quando HBGAs são modificados por α (3,4) -fucosiltransferase (FUT3). Indivíduos sem FUT2 funcional são denominados não-secretores, devido à ausência de HBGA A, B e H no corpo, enquanto indivíduos com um FUT2 funcional são denominados secretores. Fuc, Fucose; Gal, galactose; GalNAc, N-acetilgalactosamina; GlcNAc, N-acetilglucosamina. (Modificado de de Graaf et al 2016).

Recentemente, um estudo comparando inúmeros domínios P de norovírus ilustrou a interação destes com carboidratos relacionados a antígenos do sangue (HBGAs). Nele foram definidas as ligações entre norovírus GII.4 encontrados em diferentes epidemias em comparação com norovírus GII.10 e GII.12. As altas resoluções obtidas para essas estruturas determinou interações antes descritas como negativas por imunoenaios (ELISA) (128).

Para sapovírus, apenas a proteína não estrutural RNA polimerase RNA dependente teve sua estrutura resolvida (PDB: 2CKW). Ela apresenta uma porção em formato de mão direita que possui em seu terminal um canal contendo o sítio de atividade enzimática; é visualizada como um semicilindro de diâmetro de 58,4 Å (129).

Existem ainda inúmeras estruturas para calicivírus já resolvidas, no entanto, para sapovírus são necessários maiores estudos e investimentos, enquanto que para norovírus é possível avaliar evolutivamente a persistência epidemiológica de determinados genogrupos a partir de suas estruturas, e ainda assim, esse tipo de análise possibilita confirmação ou sugestão de novas hipóteses de interações proteína viral-proteína hospedeiro, principalmente na ausência de estudos de genética reversa com culturas de células estáveis.

3. Material e Métodos

3.1 Material utilizado nesse estudo

As sequências utilizadas nesse trabalho foram sintetizadas pela Life Technologies GmbH com intuito de otimizá-las para expressão em bactéria (*E. coli*). Os números de acesso no GenBank de cada sequência estão descritos na Tabela 2.

Tabela 3 Número de acesso GenBank das sequências utilizadas nesse estudo.

Genogrupo Genótipo	Número de Acesso (GenBank)
Norovírus	
GII.16	AY502010
GII.18	EU448333
GII.19	AY823306
GII.20	EU373815
GII.21	JN899245
GII.22	AB083780

3.2 Clonagem

Todos os clones foram digeridos pelas enzimas de restrição NotI (NEB) e BamHI (NEB) de acordo com as indicações do fabricante e, em seguida clonados em vetor pMAL-c2X (NEB) modificado (Fig. 9). Após amplificação do plasmídeo por meio de transformação de *E. coli* DH5 α , foi realizada a extração de DNA plasmidial, para posterior sequenciamento na plataforma GATC Biotech AG (Colônia, Alemanha).

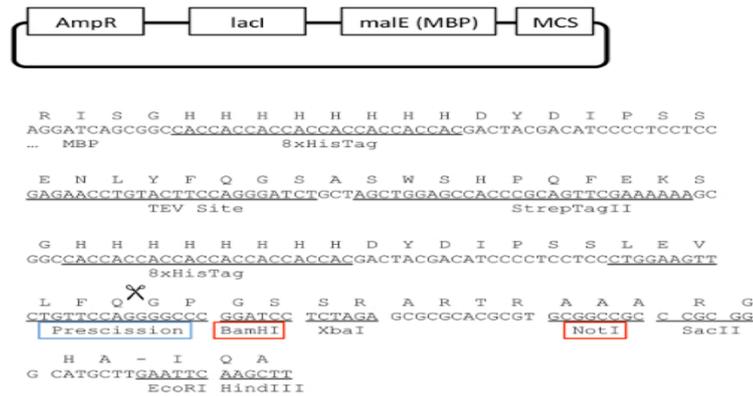


Figura 9 Modificação de pMalc2c. Representação esquemática do pMalc2x com cassete de resistência a ampicilina (AmpR), lac-operon (lacI), proteína de ligação a maltose (MBP) e sítio múltiplo de clonagem (MCS). A sequência do MCS está representada na parte inferior da figura e as enzimas de restrição estão destacadas nas caixas vermelhas e a sequência destacada na caixa azul refere-se ao sítio de reconhecimento da HRV 3C protease. (Modificado de Leuthold et al 2016).

3.3 Expressão de proteínas

Após a confirmação dos plasmídeos por meio de sequenciamento, 100 ng/μL de cada plasmídeo foram utilizados para a transformação da bactéria *E. coli* cepa BL21. Seguindo o protocolo, após o crescimento do transformante em meio nutritivo S.O.C sem antibióticos por 1 h a 37°C e 130 rpm, o volume total foi transferido para um erlenmeyer contendo 100 mL de meio LB suplementado com ampicilina. O mesmo foi disposto em uma incubadora a 37°C e 60 rpm, overnight.

Para expressão em larga escala foram utilizados 9 L de meio LB suplementado com ampicilina nos quais foram distribuídos igualmente a cultura de *E. coli* BL21 transformada com o plasmídeo de interesse citado previamente. Durante aproximadamente 2 h a cultura foi colocada a 37°C (norovírus) e 30°C (sapovírus), a uma rotação de 6000 rpm. A densidade ótica em um comprimento de onda de 600 nm foi então mensurada e ao atingir 0,4 a temperatura da incubadora foi reduzida a 22°C (norovírus) e 15°C (sapovírus). Em seguida, 1 mM de IPTG foi adicionado a cultura para indução da expressão do domínio P. A cultura foi incubada por 18 h.

O volume total foi centrifugado por 15 min a 6500 rpm, 4°C (Beckman coulter rotor JLA 8.1). O pellet foi coletado e armazenado a -20°C.

3.4 Purificação

O *pellet* foi ressuspensionado em 100 mL de PBS 1X e em seguida sonificado por dois minutos (pulsos de 10 segundos em intensidade máxima). Para a primeira etapa de purificação, que consiste em recuperar a proteína de interesse fusionada a MBP (proteína de ligação a maltose), foi utilizada uma resina de Agarose Ni-NTA para proteínas marcadas com histidina (His) (Genaxxon bioscience) e três etapas de lavagem com 10 mM, 20 mM e 50 mM de imidazol em tampão Tris-HCl 25 mM, pH 7,6 e 300 mM NaCl. Para eluição, a coluna foi lavada com o mesmo tampão mas contendo 250 mM de imidazol e o volume foi coletado (o volume utilizado nessa etapa era variável de acordo com a concentração de proteína alcançada, ou seja a cada 10 mL de tampão adicionado a coluna, as gotas coletadas eram mensuradas em um NanoDrop (Thermo Scientific) e quando havia um decréscimo na concentração a adição e coleta de tampões eram interrompidas). A confirmação dessa etapa foi realizada por meio de SDS-PAGE. Em seguida, a proteína contendo MBP e cauda de histidina foi digerida com HRV-3C Protease (Novagen), seguindo as recomendações do fabricante.

Na segunda etapa do procedimento, a proteína clivada foi novamente purificada utilizando coluna de Ni-NTA (Qiagen). Após a lavagem com 10 mM de imidazol em tampão Tris-HCl + NaCl (como descrito acima), a proteína clivada foi coletada e sua presença no eluato confirmada por SDS-PAGE. Para a etapa final, a proteína coletada foi dialisada *overnight* a 4°C no mesmo tampão supracitado.

Como a técnica de cristalização exige um elevado grau de pureza, a etapa final para obtenção dos domínios P foi de purificação utilizando coluna de cromatografia por exclusão de tamanho (Äkta prime) (coluna Superdex 75pg – Hiload 25/600). As frações coletadas foram confirmadas por SDS-PAGE e concentradas até a proteína atingir concentração final de 8 mg/mL.

3.5 Cristalização

O processo de cristalização foi dividido em duas etapas. A primeira foi realizada pela Plataforma de Cristalização de Proteínas (Biochemie-Zentrum der Universität Heidelberg), onde a proteína foi entregue para um rastreamento em 4 placas de 96 poços com diferentes condições para cristalização. O kit utilizado foi

JGSC Core Suite screen (Qiagen). O processo automatizado, realizado equipamento Phoenix RE nano-liter crystallization robot. Foi possível monitorar o processo de cristalização a partir do software Rigaku CrystalTrak. Após uma análise diária era possível determinar quais condições foram compatíveis para formação de cristais¹.

A segunda etapa da cristalização foi realizada manualmente. Placas de 15 poços (EasyXtal 15-Well Tools – QIAGEN) foram montadas utilizando as melhores condições encontradas na plataforma de cristalização (Fig. 10). Como ilustrado na figura 10 cada linha possuía uma concentração de solução mãe variada. A primeira linha possuía a concentração indicada pela plataforma, enquanto que a segunda foi diluída 1:10 e a terceira 1:9. O volume final em cada poço era de 500 μ L. O método de formação do cristal foi o de gota pendurada onde 1 μ L de proteína (4 mg/mL) é misturado com 1 μ L da solução. As placas foram incubadas a 18°C por 2 a 6 dias e todos os dias foram observadas utilizando o estereomicroscópio Nikon SMZ1500.

Para observação de ligação entre domínio P e HBGA foi realizada co-cristalização com 150 mM de cada carboidrato (antígeno H, Lewis Y, A tri, B tri) e também com fucose na mesma concentração.

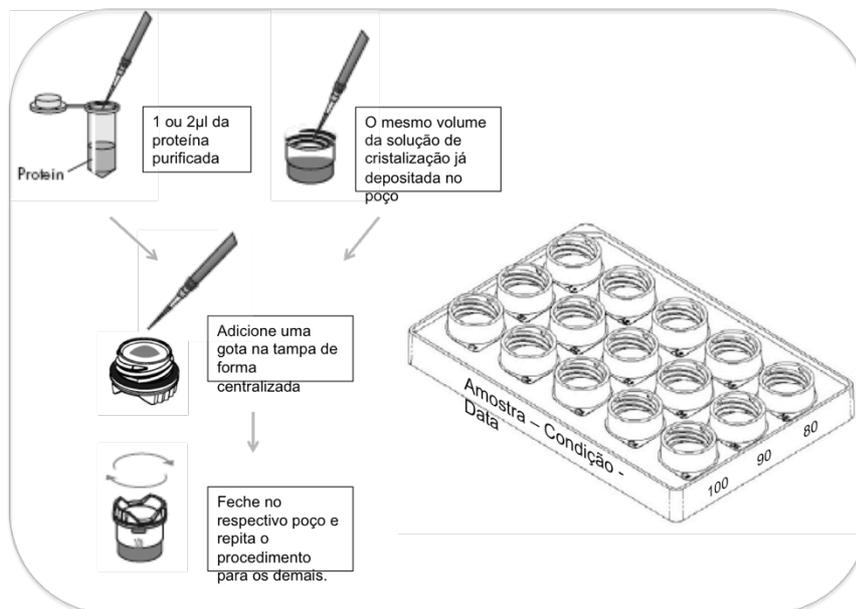


Figura 10 Esquema ilustrando as etapas para cristalização *in house* com placa de 15 poços e três concentrações diferentes.

As condições para cristalização de cada proteína estão descritas na Tabela 3.

Tabela 4 Condições de cristalização para cada proteína expressa.

Condição	Concentração	Nome	pH	Grupo
H12	0,1 M	Ácido Cítrico	5	Sal
	20% m/v	Polietileno Glicol 6000	6,5	Polímero
F5	0,1 M	Acetato de Sódio	5	Sal

3.6 Coleta e processamento de dados

Os dados referentes a difração de raios-X foram coletados no European Synchrotron Radiation Facility (ESRF), Grenoble, França, utilizando os *beamlines* BM30A e ID23-1. Os dados foram processados usando XDS (do inglês *X-ray Detector Software*) (130).

As estruturas foram resolvidas usando substituição molecular (do inglês *molecular replacement*) usando o *software* PHASER (131). A estrutura do domínio P de GII.22 foi resolvida usando o domínio P de GII.10 como modelo. A estrutura de GII.21 foi resolvida usando um monômero de GII.22 como modelo. As estruturas foram refinadas e corrigidas manualmente em várias rodadas utilizando-se o programa COOT (132). Em seguida foram ainda mais refinadas utilizando-se o *software* PHENIX (133). As estruturas foram validadas pelo MolProbity. (134). As imagens foram produzidas com o visualizador tri-dimensional para computador PyMOL (versão 1.12r3pre) (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.).

4. Resultados

4.1 Norovírus

4.1.1 Alinhamento de aminoácidos referente ao capsídeo de norovírus GII

As sequências completas de aminoácidos referentes aos capsídeos de todos os genótipos de norovírus utilizados nesse trabalho foram alinhados usando MUSCLE (Fig. 16A) e em seguida, suas distâncias filogenéticas foram analisadas com *Sequence Demarcation Tool* versão 1.2 (SDTv1.2) (Fig. 16B) (135). Os genótipos GII.10, GII.4 e GII.17 foram inclusos na análise pois, o primeiro foi utilizado para substituição molecular na resolução da estrutura de GII.22 enquanto que GII.4 e GII.17 são referentes aos genótipos mais prevalente e incidente, respectivamente (Fig. 11).

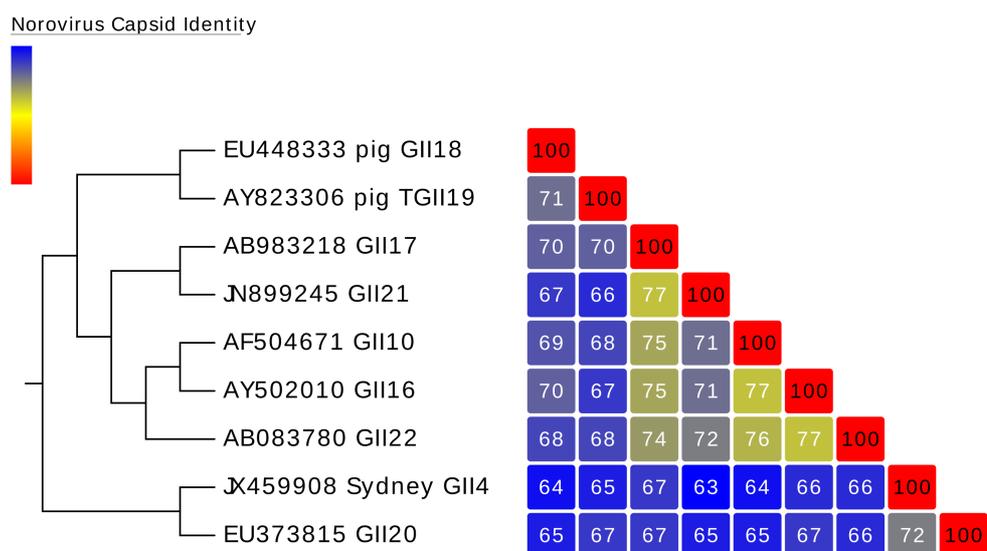


Figura 11 Análise das identidades filogenéticas dos capsídeos de diferentes genótipos do GII de norovírus. Alinhamento realizado com as sequências completas de GII.16, GII.18, GII.19, GII.20, GII.21 e GII.22 com adição de GII.10, GII.4 e GII.17. O genoma é dividido em domínio S formado pelos aminoácidos de 1 a 225 e domínio P de 226 a 568 (em média). A partir do alinhamento foi montada uma árvore filogenética com intuito de ressaltar as identidades na pirâmide colorida a frente onde as distâncias são destacadas de cima para baixo e da esquerda para direita, sendo 100 em vermelho o número referente a relação entre a sequência e ela mesma. Tanto em A quanto em B a nomenclatura contém o número de acesso do GenBank e o genótipo. As sequências referentes a GII.18 e GII.19 são as únicas utilizadas no estudo que representam norovírus que infectam suínos (do inglês *pig*).

4.1.2 Expressão e Purificação de proteína

Como mencionado na Metodologia seis genótipos de norovírus do genogrupo GII pouco detectados foram escolhidos para expressão, purificação e cristalização, GII.16, GII.18, GII.19, GII.20, GII.21 e GII.22.

Em todos os casos foi possível seguir o mesmo desenho experimental até a segunda etapa de purificação, onde a proteína de ligação a maltose é clivada do domínio P (Fig. 12).

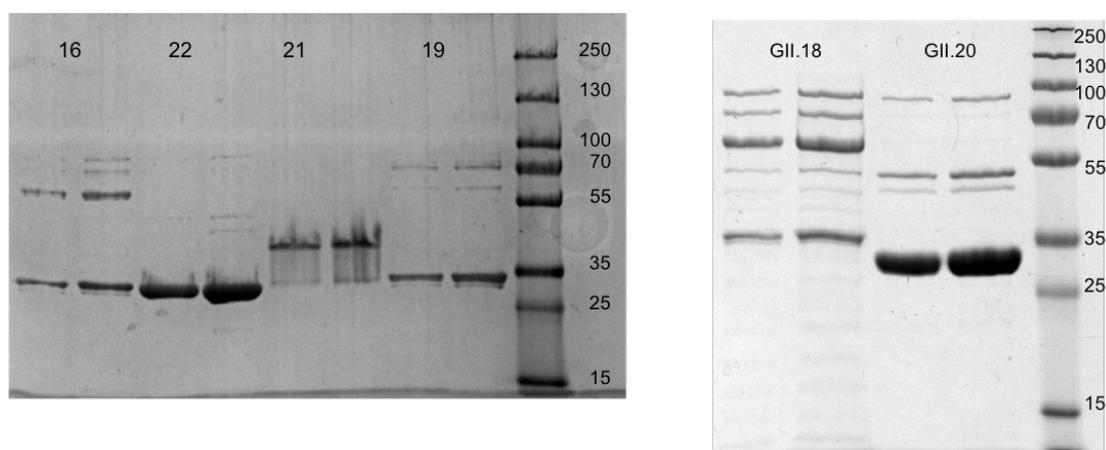


Figura 12 SDS-PAGE de todos os genótipos de norovírus expressos e semi-purificados. Os números acima dos poços são referentes aos genótipos. Esses resultados de eletroforese ilustram a clivagem bem sucedida de todos os domínios P (entre 27 a 40 kDa).

A etapa seguinte de purificação através de cromatografia por exclusão de tamanho foi realizada para todos os domínios P. O domínio P GII.18 não apresentou boas frações após purificação e foi excluído após essa etapa. O GII.19 não apresentou nenhum cristal na plataforma, sendo, também, excluído do estudo.

Os domínios GII.16 e GII.20 foram purificados e concentrados até aproximadamente 4 mg/mL. No entanto após envio a plataforma de cristalização não foram observados bons cristais. Um tentativa de crescimento de cristais *in house* foi aplicada mas após inúmeras tentativas esses dois genótipos também foram descartados do estudo.

Os domínios GII.21 e GII.22 apresentaram bons resultados de purificação por exclusão de tamanho e foram os únicos a serem avaliados durante todo o ensaio. Dessa forma na Figura 13 é possível observar o perfil da cromatografia realizada para

GII.21 enquanto que na Figura 14 um perfil semelhante de cromatografia e eletroforese para GII.22 em SDS-PAGE são observados.

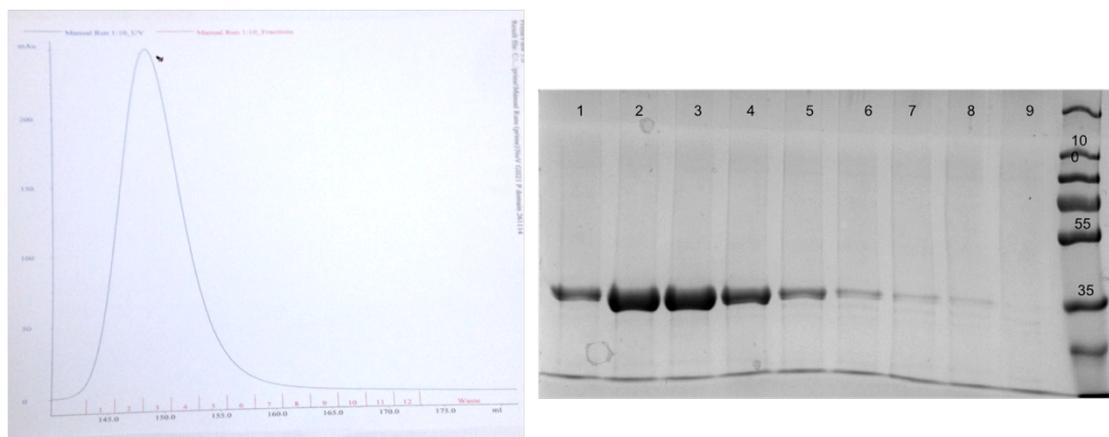


Figura 13 Perfil de cromatografia e eletroforese das frações coletadas para GII.21. A esquerda observam-se 12 frações que compõe o pico utilizado como referência para o domínio P de GII.21. Apenas 9 frações foram avaliadas por eletroforese e são observadas na imagem a direita. Essa análise é levada em conta para avaliação do grau de pureza da amostra.

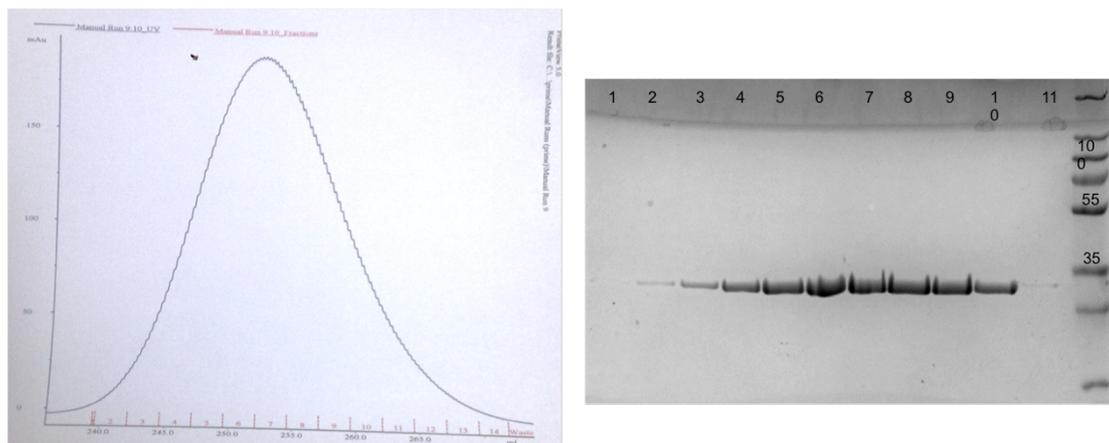


Figura 14 Perfil de cromatografia e eletroforese das frações coletadas para GII.22. A esquerda observam-se 13 frações que compõe o pico utilizado como referência para o domínio P de GII.21. Apenas 11 frações foram avaliadas por eletroforese e são observadas na imagem a direita. Essa análise é levada em conta para avaliação do grau de pureza da amostra

4.1.3 Cristalização

Todos os domínios foram expressos, purificados e enviados a plataforma de cristalização na Universidade de Heidelberg com exceção das proteínas referentes a GII.18 e GII.19. Ambas não foram purificadas de maneira eficiente na última etapa

por exclusão de tamanho. Como mencionado acima GII.16 e GII.20 não apresentaram bons cristais.

Os cristais de GII.21 foram obtidos com mistura 1:1 de proteína e solução mãe nomeada H12 descrita na Tabela 3. Os cristais apresentaram formato de placas alongadas, retangulares (Fig. 15A). Os cristais de GII.22 foram obtidos utilizando-se apenas acetato de sódio (condição F5), adaptada da plataforma para *in house* (Tabela 3). Os cristais apresentaram formato de diamante (Fig. 15B).

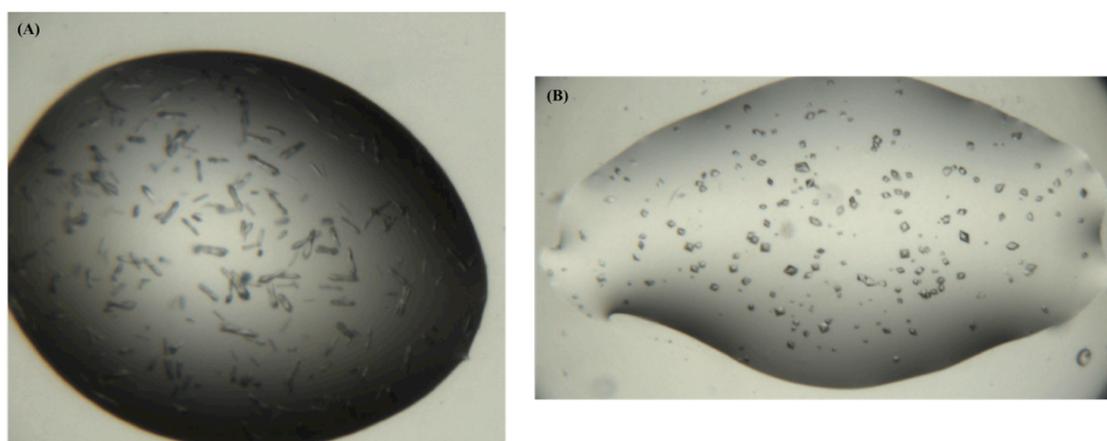


Figura 15 Cristais obtidos a partir da cristalização de GII.21 e GII.22. Imagens dos cristais formados a partir de 4 mg/mL de domínio P de norovírus enviados para a Plataforma de Cristalografia da Universidade de Heidelberg. (A) Cristais alongados em forma de placas (GII.21). (B) Cristais com formato de diamante (GII.22).

Para GII.21 e GII.22 foi realizado o processo de co-cristalização com diferentes HBGAs (antígenos sanguíneos), com antígeno H, Lewis Y, Atri, Btri, entre outros, no entanto não foram obtidos resultados após a difração. Para GII.22 a adição de fucose, o açúcar que está presente na maioria das composições dos diferentes HBGA, gerou cristais em formato diamantes (não mostrados) e um padrão de difração de raios-X com densidade e dados que comprovam sua interação com o domínio GII.22.

O cristal do domínio P de GII.21 mostrou melhor solução no grupos espacial ortorrômbico $P2_12_12_1$ enquanto os cristais tanto da estrutura sem ligante de GII.22 e GII.22 mais fucose apresentaram solução no grupo espacial tetragonal $P4_32_12$ (Tabela 4).

Tabela 4 Coleta de dados e resultados do refinamento das estruturas GII.21, GII.22 e GII.22_Fucose

Parâmetros	Valor ou Descrição		
	GII.21_apo**	GII.22_apo	GII.22_Fuc
Alcance de resolução (Å)	48.56 - 2.17 (2.23 - 2.15)	39.57 - 1.307 (1.35 - 1.31)	49.12 - 1.36 (1.41 - 1.36)
Grupo espacial	P 21 21 21	P 43 21 2	P 43 21 2
Célula Unitária (a, b, c (Å)) (α, β, γ (°))	64.36 164.12 171.11 90 90 90	88.491 88.491 78.21 90 90 90	88.92 88.92 78.69 90 90 90
Reflexões únicas	98683 (9750)	75196 (7416)	68074 (6719)
Completeza (%)	99.36 (97.14)	99.74 (99.85)	99.95 (99.87)
Média I/sigma(I)	17.49 (3.93)	27.58 (2.49)	15.19 (1.96)
Wilson B-factor	31.11	11.75	16.9
Refinamento			
R-work	0.1842 (0.2503)	0.1786 (0.2269)	0.1897 (0.2638)
R-free	0.2179 (0.2828)	0.1954 (0.2356)	0.2106 (0.2888)
Número de átomos não-hidrogênio	14429	2805	2691
Macromoléculas	13867**	2336	2368
Ligantes	40		15
Água	522	469	308
Resíduos de proteína	1810	301	305
RMS(ligações)	0.005	0.006	0.007
RMS(ângulos)	0.97	1.11	1.09
Ramachandran favored (%)	97	97	96
Ramachandran outliers (%)	0	0	0.33
Clashscore	2.21	6.07	1.06
Average B-factor	44.9	17.1	22.3
Macromoléculas	45.2	14.9	21.3
Ligantes	44.5		25.7
Água	38.4	28.3	29.9

* Os valores entre parênteses são referentes a camada de maior resolução. **GII.21 foi resolvido com três dímeros, ou seja, apresenta seis cadeias.

4.1.4 Estruturas

Os domínios P foram observados pela primeira vez por cristalografia e difração de raios X em 1999 (20). Os achados encontrados descreviam a presença de uma estrutura barril- β composta por 6 fitas antiparalelas. Além disso o domínio P se divide em P1 que é composto pelos aminoácidos 224 a 274 e 418 a 530, P2 então torna-se uma inserção entre esses aminoácidos sendo composto por 275 a 417, entretanto, P2 é a região mais externa e mais variável da estrutura, pois, supõe-se sua função como epítipo com função de neutralização.

Além disso, o domínio conserva três hélices α que possuem um loop conservado com os aminoácidos Pro-Pro-Gly para os dois genótipos aqui resolvidos.

O domínios P de GII.21 e GII.22 são aqui apresentados como monômeros a fim de facilitar as análises comparativas, no entanto, sabe-se que naturalmente ocorrem como dímeros (Fig. 16). A resolução de GII.21 foi configurada em três dímeros enquanto que a de GII.22 foi monomérica, como apresentada.

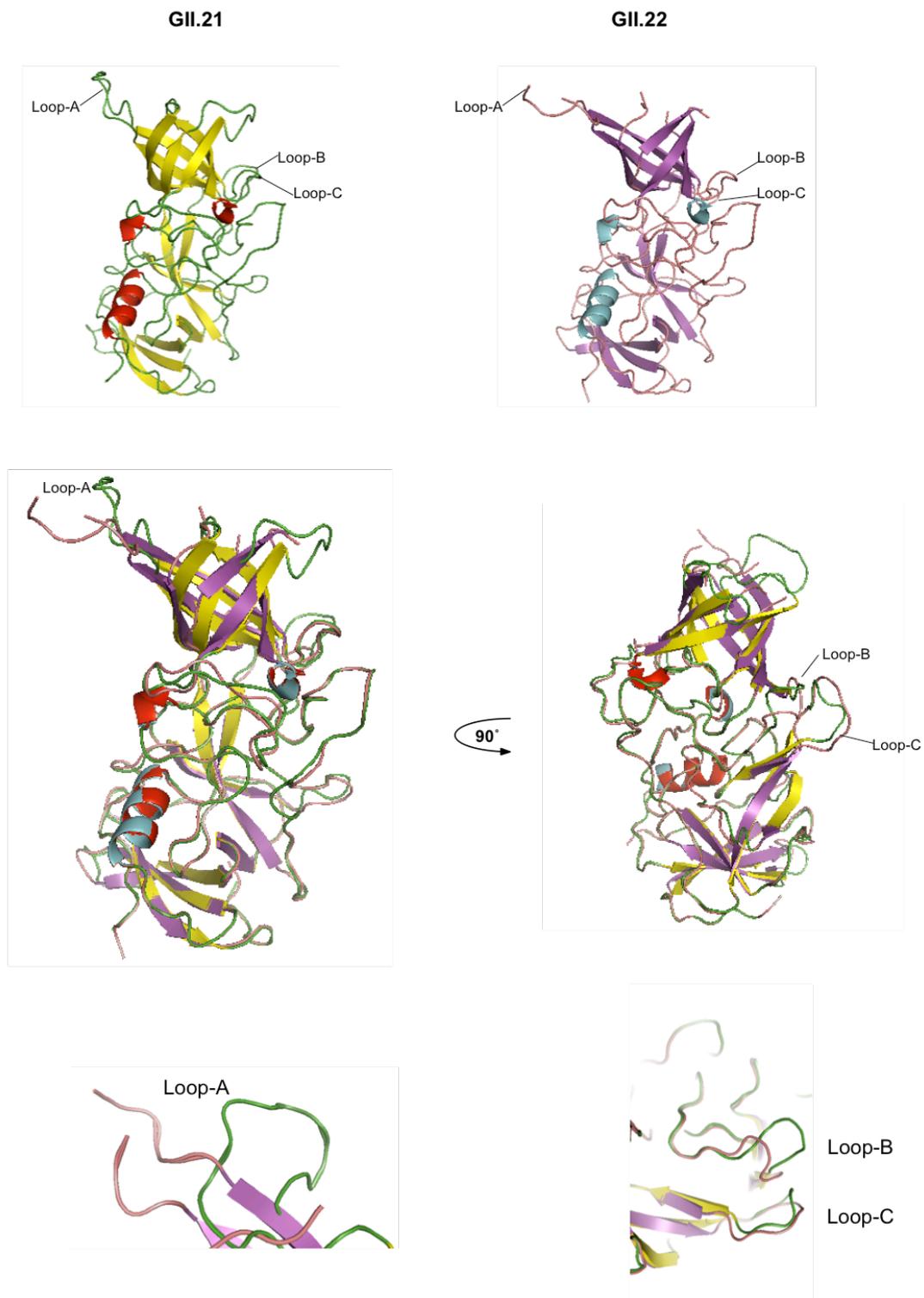
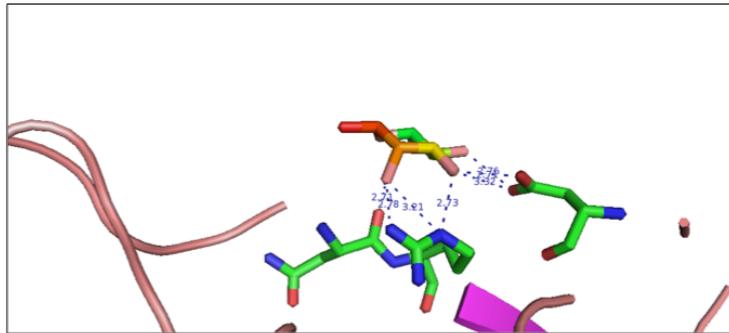


Figura 16 Estruturas cristalinas de raios X de norovírus GII.21 e GII.22 sem ligante. Os monômeros foram sobrepostos para que as diferenças fossem observadas e em seguida a sobreposição foi rotacionada para melhor observação dos loops indicados.

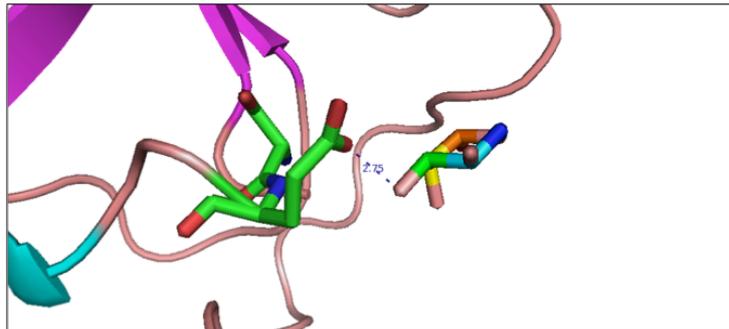
As estrutura do domínio P é composta por três hélices α , sendo duas pequenas posicionadas dentro do domínio P2 com as seguintes sequências TTA e PKL para GII.21 e PAN e PKL para GII.22. Possui ainda 13 fitas β , onde 6 se organizam para compor a estrutura em forma de barril no domínio P2 presente nas duas estruturas apesar da diferença entre aminoácidos.

A região do N-terminal se sobrepõe por 6 aminoácidos idênticos na sequência de P1 para os dois genótipos. O loop observado na porção superior da estrutura de GII.21 que não se sobrepõe a GII.22 é formado por aminoácidos entre 290 a 298, na Figura 21 é possível observar que não há formação desse loop na estrutura de GII.22 (loop posicionado a direita do loop-A). O loop-A é o que mais se destaca forma-se dos aminoácidos entre as posições 337 a 348 para GII.21 e 341 a 351 para GII.22. Na Figura 16 é possível observar que não há sobreposição entre eles e ambos encontram-se em posições diferentes. Além disso foram destacados dois loops na lateral da estrutura (loop-B e loop-C), o primeiro é composto pelos aminoácidos 392 a 403 para GII.21 enquanto que para GII.22 é composto pelos aminoácidos 396 a 401, o segundo é formado pelos aminoácidos 437 a 444 para GII.21 e para GII.22 os aminoácidos 439 a 448. Essa região está destacada na figura 16 pois foi observada como região de ligação para HBGAs na estrutura GII.21 (136).

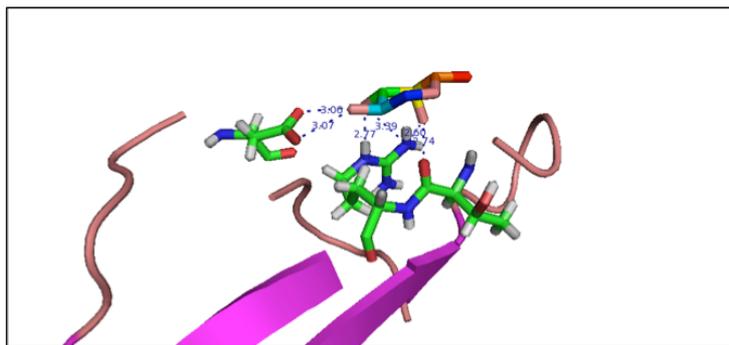
A última análise realizada nesse trabalho foi a resolução da estrutura de domínio P de GII.22 co-cristalizada com fucose (Fig. 17) onde os aminoácidos N351, R352, D380 e I382 foram identificados como as regiões de contato com a fucose, no início da folha- β seguinte ao loop-A (Figura 16) formado por aminoácidos 341 a 451. Pelo menos o aminoácido R352 é conservado quando comparado GII.22 com GII.4 e GII.10, no entanto para GII.21, previamente descrito na literatura a região de contato com a fucose resume-se ao aminoácido 357E (Figura 17).



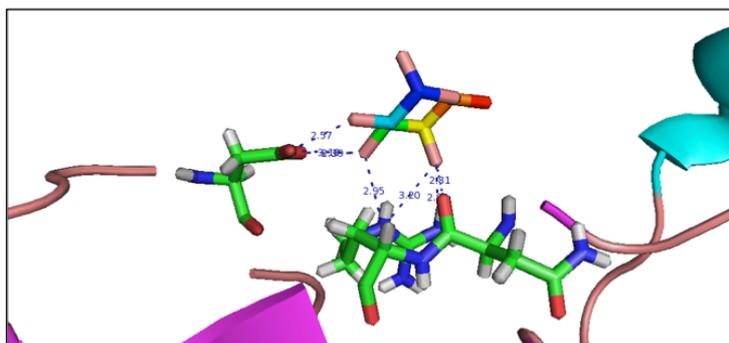
GII.22



GII.21



GII.4



GII.10

Figura 17 Interação de fucose com domínio P2 de de norovírus. Acima a figura representa o domínio P2 destacado da estrutura completa de P com a fucose. O primeiro quadro é um aumento da estrutura GII.22 ligada a fucose resolvida neste trabalho, onde os pontos de contato da fucose são compostos pelos aminoácidos N352, R352 e D381. O terceiro quadro é um aumento da estrutura de GII.4 (PDB:SJ35) onde apenas a fucose foi mantida. Os pontos de contato mostrados pela linha azul são compostos pelos aminoácidos T343, R344 e D373. No quarto quadro a estrutura de GII.10 (PDB:3ONY) é possível observar a fucose interagindo com os aminoácidos N354, R355 e D384. Por fim o quadro referente a estrutura GII.22 (PDB:4RMO) onde uma nova região de interação foi descrita, aqui apenas o aminoácido E357 apresenta contato com a fucose.

5. Discussão

Os calicivírus humanos, com maior destaque, os norovírus, são a principal causa de gastroenterites agudas. Para os norovírus o modelo animal encontrado que produz o maior número de informações é o modelo murino enquanto que para os sapovírus o modelo usual é o modelo suíno (45). No entanto, as informações biológicas adquiridas a partir desses modelos só podem ser transpostas aos modelos de infecção humana com um limite bastante delineado. Sendo assim abrem-se portas para a busca de informações aquém de cultivos celulares e sistemas replicativos. E em uma dessas portas encontra-se a biologia estrutural.

O uso de cristalografia, ressonância magnética nuclear e crio-eleto microscopia têm gerado inúmeros dados a respeito de diversos vírus incluindo os calicivírus. No entanto, até o momento, somente duas estruturas referentes ao capsídeo de sapovírus foram resolvidas, sendo uma delas, uma estrutura quimérica ainda não encontrada em infecções naturais (137, 138). Ambas foram resolvidas a partir de crio-eleto microscopia, sendo a primeira uma estrutura de resolução de 22 Å e a segunda de 8,5 Å.

Além disso, a metodologia aplicada para obtenção das partículas (VLPs do inglês *virus-like particles*) em geral é a expressão de proteínas usando baculovírus recombinante. Essa técnica, bastante utilizada, desde 1997 (139), já foi descrita como pouco eficiente para obtenção de concentrações elevadas de partículas de sapovírus humanos (GI, GII e GV), o que dificulta estudos cristalográficos (140).

Expressar uma proteína de capsídeo pode ser um trabalho dispendioso, no caso dos calicivírus a estrutura em simetria icosaédrica T=3 determina que a sua composição seja de 180 unidades de VP1, dessa forma são 180 cópias de domínios S e 180 cópias de domínio P. O uso de cristalografia com difração de raios X para resolução de proteínas é bastante eficaz quando se têm fontes abundantes de quantidades de proteína e um bom modelo para substituição. Para os sapovírus humanos já foi abordada a dificuldade em ambos os casos, e para os norovírus uma forma mais rápida de se compreender a estrutura do capsídeo e suas possíveis funções foi estudar a unidade de forma individual, separando-a em domínios.

Norovírus ligam-se a HBGA que são carboidratos polimórficos encontrado na saliva e na superfície de diferentes epitélios, incluindo mucosa intestinal (141), sabe-se que pelo menos nove diferentes tipos de HBGA são reconhecidos pelos norovírus

(127) e, são agrupados em tipo ABH ou Lewis de acordo com a posição da fucose (141).

Nesse trabalho seis genótipos pouco detectados do genogrupo GII de norovírus foram escolhidos como objeto de estudo a fim de se avaliar sua estrutura, e possível ligação com HBGAs. No entanto, apenas duas delas resultaram em modelos estruturais resolvidos e apenas GII.22 demonstrou ligação com fucose apesar dos esforços para se avaliar outras possíveis ligações.

Um estudo realizado recentemente demonstrou que o sítio de ligação para GII.21 seria único em relação aos demais, uma vez que não apresenta aminoácidos conservados na região previamente descrita. Enquanto GII.4 apresenta ligação na região dos aminoácidos T344, R345, D374 e G443, GII.21 apresenta ligação entre HBGAs e os aminoácidos W296, S352, S355 e N379, que seria uma região localizada na porção mais superior do domínio P (136). Para GII.22 foi possível demonstrar apenas pontos de contato em N351, R352, D380 e I382, que se aproxima da região de ligação de GII.4 (128).

A ligação com HBGA foi demonstrada para diferentes calicivírus (142-145) evolutivamente isso demonstraria uma manutenção de estratégia entre todos os patógenos que possuem esses carboidratos como receptores. A determinação de ligações com HBGA para diferentes norovírus pode gerar uma importante classificação fenotípica. Além disso estudos avaliando essas ligações podem ser implicados a estudos epidemiológicos. Embora não se saiba a razão específica para GII.4 ser o mais epidêmico considerado a capacidade de ligação de outros genogrupos a HBGAs (146).

Nesse capítulo com intuito de se obter maiores informações sobre a biologia estrutural de calicivírus humanos, foram obtidos dados de duas novas estruturas de domínios P de norovírus, sendo GII.22 ainda não descrita na literatura. Juntas essas informações compõem novos conhecimentos a respeito de calicivírus humanos e criam bases para futuros estudos moleculares e estruturais de ambos os gêneros.

CAPITULO III – DESENVOLVIMENTO DE FERRAMENTAS BÁSICAS APLICADAS AO ESTUDO DE GENÉTICA REVERSA DE SAPOVÍRUS HUMANOS

1. Introdução

O termo genética reversa dentro da genética clássica é utilizado para descrever o processo onde a informação segue um caminho contrário, ou seja, um gene é determinado ou mutado e, o fenótipo resultante é estudado. Para vírus de RNA o termo pode ser utilizado quando há geração de vírus ou VLP (*virus-like particle*) a partir de cDNA, com intuito de se compreender a função de proteínas específicas do genoma viral (147). Estudos de genética reversa permitem uma maior compreensão da biologia viral, gerando dados a respeito das interações entre vírus-hospedeiro e, esses dados são a base para desenvolvimento de vacinas e tratamentos antivirais.

Apesar dos inúmeros avanços nessa área, informações a respeito dos calicivírus são obtidas através dos sistemas desenvolvidos para os vírus animais devido a falta de uma cultura celular robusta que seja permissiva a infecção de calicivírus humanos.

O primeiro sistema de genética reversa descrito para calicivírus foi publicado em 1995 pelos pesquisadores Sosnovtsev e Green. Eles demonstraram a infecção de cultura celular a partir de RNA de calicivírus felino (FCV) transcrito na presença de análogos de CAP-5' *in vitro* (148). Antes, no entanto, nos anos de 1973 e 1988, já havia sido demonstrada infecção de culturas de células com partículas virais purificadas de calicivírus felinos e sapovírus suínos, respectivamente (149, 150).

As estratégias bem sucedidas para esses sistemas são basicamente duas: 1. transcrição e adição de CAP *in vitro* e 2. transfecção de DNA em células expressando T7 RNA polimerase de um vírus ajudante.

Para FCV primeiro foi desenvolvida a estratégia na qual o cDNA viral do genoma completo foi transcrito e adicionado de CAP *in vitro*, esse RNA apresentou-se infeccioso e permitiu a seleção de uma melhor cultura celular (células de rins de felino) isso porque a presença de um receptor específico, JAM-1, aumenta a eficiência da infecção facilitando o aumento no níveis de replicação e amplificação viral (148). Uma forma encontrada para pular a etapa de transcrição *in vitro* foi o desenvolvimento de um sistema onde um vírus ajudante, *Vaccinia virus* modificado,

expressando T7 RNA polimerase foi utilizado para infectar as células de felinos possibilitando a transfecção direta do cDNA do genoma completo do vírus mais o promotor T7 inseridos em um plasmídeo de expressão em células de mamíferos (151). Ambos ensaios permitiram identificar: i) receptores celulares importantes, ii) obrigatoriedade da presença de um análogo a VPg, bem como a iii) integridade do domínio referente a Protease-Polimerase viral (152-154).

Seguindo a cronologia, o segundo calicivírus a ter um sistema de genética reversa desenvolvido foi o sapovírus suíno. Em 1988, já havia sido demonstrada a infecção desses vírus em culturas celulares (150), no entanto, apenas em 2005, utilizando uma abordagem similar a descrita para FCV, foi reportado o primeiro sistema de genética reversa para sapovírus suíno. A permissividade das células LLC-PK (células de rim de suíno) só ocorreu quando foram adicionados ácidos biliares ao meio de cultura. Os vírus formados a partir desse sistema quando injetados em leitões, causaram sintomas atenuados, o que constatou a infectividade das partículas (37, 45). A necessidade de adição dos ácidos biliares para infecção viral já havia sido demonstrada anos antes em estudos que comprovaram a regulação negativa da resposta mediada por interferon na presença de ácidos biliares (46).

Em 2006, o vírus da doença hemorrágica em coelhos (RHDV) teve seu RNA transcrito *in vitro* e transfectado em células de coelho RK13 (células de rim de coelho), além disso o RNA foi demonstrado infeccioso quando injetado diretamente no fígado desses animais (155). Posteriormente, demonstrou-se que, diferentemente de FCV, RHDV não exigia a adição de CAP e nem a presença da sequência do gene referente a VP2 para recuperação de novos vírus (156). Uma importante mudança na construção, como clonar o genoma viral completo sob o comando de um promotor CMV permitiu que houvesse uma produção maior de partículas virais infecciosas (157).

Um importante modelo para norovírus humano é o norovírus murino (MNV), descrito pela primeira vez em 2003, como letal para camundongos imunocomprometidos (21). Sua importância deve-se ao fato de ser o primeiro norovírus com uma cultura celular permissiva amplamente estabelecida (22). Para o desenvolvimento de um sistema de genética reversa para MNV foram aplicadas as mesmas estratégias descritas anteriormente para FCV, no entanto, a recuperação de partículas virais só foi completa quando as células foram infectadas com *Fowlpox*

virus que expressavam T7 RNA polimerase, isso porque o *Vaccinia virus* apresentou um efeito inibitório na replicação de MNV (158).

Atualmente, o sistema de genética reversa de norovírus murino foi modificado para a transfecção do RNA transcrito *in vitro*, pois o uso de um kit para adição de CAP pós-transcrição conhecido como ScriptCap seguido de eletroporação das células RAW264.7 (células macrófagos de camundongos BALB/c) permitiu uma recuperação mais rápida e robusta de partículas de MNV infecciosas (159). Muitos resultados foram obtidos a partir desses sistemas, um deles por exemplo, foi a função das regiões não traduzidas 5'-terminal e 3'-terminal do vírus, que em alguns casos poderia ser deletéria para a replicação viral, quando mutadas (160). Outro importante achado foi a descrição da função da proteína codificada pela ORF4, que atua como um antagonista da resposta imune inata atrasando a expressão de alguns genes que desencadeariam a apoptose celular (14).

Para norovírus humanos inúmeros estudos indicam tentativas de desenvolvimento de uma cultura celular robusta, bem como um sistema efetivo de genética reversa.

O primeiro trabalho bem sucedido utilizando um construção com um gene repórter de norovírus humano consiste em uma linhagem BHK (células de rins de bebe camundongos) capaz de manter a replicação do genoma de *Norwalk virus*, por meio da inserção de um gene que confere resistência a antibiótico na posição de VP1 (161). O RNA extraído dessas células foi capaz de replicar em células de hepatoma celular humano (Huh-7). Sendo então a primeira linhagem a replicar RNA de norovírus humano (161). Esse sistema, no entanto, não permite estudos a partir de modificações no genoma porém, permitiu avaliações quanto aos efeitos de inibição da replicação viral do interferon- α e do antiviral ribavirina (162). O mesmo grupo posteriormente utilizou esse sistema para determinar o papel do colesterol no aumento da infecção viral (163).

De maneira geral tentativas de sistema de genética reversa, similares as descritas para FCV, não foram bem sucedidas para norovírus humanos, ou obtiveram apenas um ciclo de replicação sem novas infecções pós-passagem celular ou, em alguns casos, houve apenas detecção de RNA viral sem partículas ou com formação de VLPs (24, 25, 164).

Recentemente os principais achados consistem de infecção de norovírus humanos em células B, um plasmídeo que contém uma progênie viral com um gene

repórter e, replicação de norovírus humanos em células tronco derivadas de enterócitos humanos (26-28).

Após observarem replicação e presença de proteína estrutural de norovírus murino em células B, foi avaliada a possibilidade de se conseguir replicar norovírus humanos também nesse tipo celular. Para que isso ocorresse amostras fecais positivas para norovírus GII.4 foram utilizadas para a infecção. O grande achado desse trabalho consiste na não filtração da amostra fecal e da constatação de que a presença da bactéria *Enterobacter cloacae*, que expressa HBGA em sua superfície, aumentou os títulos virais na cultura celular. O uso de HBGA sintéticos com vírus oriundos de amostras filtradas também demonstrou um aumento no título viral. Dessa forma concluiu-se não só que norovírus humanos são capazes de infectar células B mas, também HBGAs funcionam como moléculas mediadoras entre o vírus e a membrana celular (27).

Outra forma de cultivo celular descrita recentemente para norovírus humanos envolve o uso de células tronco oriundas de enterócitos humanos coletados durante biopsia de pacientes. Essas células são estimuladas a se diferenciarem em células do epitélio intestinal (enterócitos, células em goblet, células enteroendócrinas e células de Paneth), em uma monocamada, que foi transfectada com norovírus humano presente em amostra fecal. Nesse estudo a exemplo das culturas celulares de sapovírus suínos (45), foi utilizado bile junto ao meio de cultivo. Como resultado observou-se que apenas norovírus que alguns genótipos como GII.3, só eram isolados na presença de bile. Os norovírus GII.4 foram encontrados apenas em enterócitos, o que confirma a permissibilidade dessas células também para esses vírus assim como ocorre para outros vírus entéricos (28).

Similar a outros sistemas de genética reversa, Katayama e colaboradores descreveram um sistema de genética reversa para norovírus humanos GII.3 utilizando-se de GFP como gene repórter inserido na poliproteína, e o genoma completo viral expresso sob o comando do fator de alongação 1- α (EF-1 α) em células COS-7 (células tipo fibroblasto de símio). Esse sistema comporta a replicação do genoma, formação de partícula viral, partículas virais contendo marcação de GFP e, o RNA extraído dessas partículas foi determinado infeccioso (26).

Para sapovírus o ciclo replicativo ainda não é totalmente conhecido e a genética reversa foi estabelecida, até o momento, apenas para sapovírus suínos. Isso ocorre porque esses sistemas não foram elucidados usando genótipos humanos

ainda e, a cultura celular para sapovírus suínos tende a ser laboriosa devido a condições restritivas para sua manutenção (37, 45). O papel da VPg na tradução viral foi recentemente descrito utilizando sapovírus suíno como modelo. Foi demonstrado que existe uma ligação entre VPg e eIF4E e uma associação ao complexo de tradução eIF4F (48).

A ausência de uma cultura celular robusta faz com outras abordagens sejam desenvolvidas para estudo de calicivirus humanos. O uso de genética reversa tem se provado uma importante ferramenta na aquisição de informações a respeito de diferentes vírus de RNA. Conseqüentemente o desenvolvimento de um sistema de genética reversa para sapovírus humano seria a base para a geração de novos conhecimentos a respeito desse gênero viral.

2. Material e Métodos

2.1 Clonagem

Para a construção do clone infeccioso foi utilizado o genoma completo do sapovírus BR01 (número de acesso: AB614356) clonado em pGEM-T easy (Promega) e confirmado previamente (165).

2.2 Construção do clone infeccioso

A sequência do sapovírus BR01 (Sapo_CI) foi digerida por enzima de restrição PacI (NEB) (Fig. 24), de acordo com as indicações do fabricante. A purificação de DNA foi realizada utilizando-se o kit *illustra Gfx PCR DNA and Gel Band Purification* (GE Healthcare Life Sciences). O produto foi armazenado a 4°C.

Para amplificação do vetor pcDNA3.1(+) (Fig. 24) (Thermo Fisher Scientific), gentilmente cedido pelo Dr. Alexandre M. Vieira Machado (Fiocruz – Centro de Pesquisa René Rachou), foram utilizados os primers indicados na Tabela 5, utilizando-se LongAmp Taq DNA polimerase (NEB). O produto de PCR foi purificado, também através do kit *illustra Gfx PCR DNA and Gel Band Purification* (GE Healthcare Life Sciences).

```

          enhancer region (3' end)
689 CATTGACGTC AATGGGAGTT TGTTTGGCA CCAAAATCAA CGGACTTTC CAAAATGCG
          CAAT TATA
749 TAACAACTCC GCCCCATTGA CGCAAATGGG CGGTAGGCGT GTACGGTGGG AGGTCTATAT
          3' end of hCMV putative transcriptional start
809 AAGCAGAGCT CTCTGGCTAA CTAGAGAACC CACTGCTTAC TGGCTTATCG AAATTAATAC
T7 promoter/primer binding site
869 GACTCACTAT AGGGAGACCC AAGCTGGCTA GCGTTAAAC TTAAGCTTGG TACCGAGCTC
          BamHI BclI* EcoR I EcoR V BclI* Not I Xba I
929 GGATCCACTA GTCCAGTGTG GTGGAATTCT GCAGATATCC AGCACAGTGG CGGCCGCTCG
          Xba I Apa I Pme I pcDNA3.1/BGH reverse priming site
989 AGTCTAGAGG GCCCGTTTAA ACCCGTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC
1049 CATCTGTTGT TTGCCCTCC CCGTGCTT CTTGACCCT GGAAGGTGCC ACTCCACTG
          BGH poly(A) site
1109 TCCTTCTTA ATAAAATGAG GAAATTGCAT

```

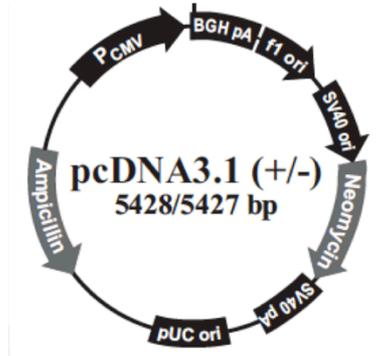


Figura 18 pcDNA (Invitrogen). Esquema representativo do vetor de expressão em células de mamífero.

Tabela 5 Principais primers utilizados para PCR de fusão

DNA alvo	Primer Forward	Sequência
pcDNA 3.1(+)	PacI_BGH_For	TTTAATTAACCGTGCCTTCCTTGACC
Sapo_CI	pcDNA_For_Sapo5	TATATAAGCAGAGCTCTCTGGCTAACGTGATTGGTTAGATGGTTTC CAAGC
	Primer Reverse	
pcDNA 3.1(+)	pcDNA_Rev	GCTTGGAACCATCTAACCAATCACGTTAGCCAGAGAGCTCTGCTT ATATA
Sapo_CI	PacI_M10_oligodT50_Rev	GTTTTCCAGTCACGACTTAATTAATTTTTTTTTTTTTTTTTTTTTT TTTTT

A técnica de PCR de fusão (do inglês *overlap-extension PCR*) foi utilizada nessa estratégia para inserir o cDNA genômico de sapovírus BR01 em pcDNA3.1(+). Para a sobreposição foram utilizados 150 ng/μL de inserto e 50 ng/μL de vetor. Primeiro duas PCRs individuais foram realizadas para gerar cada fragmento (inserto e vetor) utilizando Phusion DNA polimerase (NEB). Para a fusão pelo primeiro PCR, foram utilizados 150 ng/μL de inserto e 75 ng/μL de vetor (proporção molecular 1:1) sem adicionar primers com 12 ciclos na temperatura para anelamento. Nessa etapa ocorre a sobreposição dos fragmentos indicados na Figura 19, que ilustra o processo completo. em seguida, a segunda PCR foi realizada com primers PacI_BGH_For e PacI_M10_oligodT50_Rev que anelam novas extremidades geradas com 20 ciclos de extensão. A última PCR foi realizada com a enzima LongAmpTaq DNA polimerase (NEB).

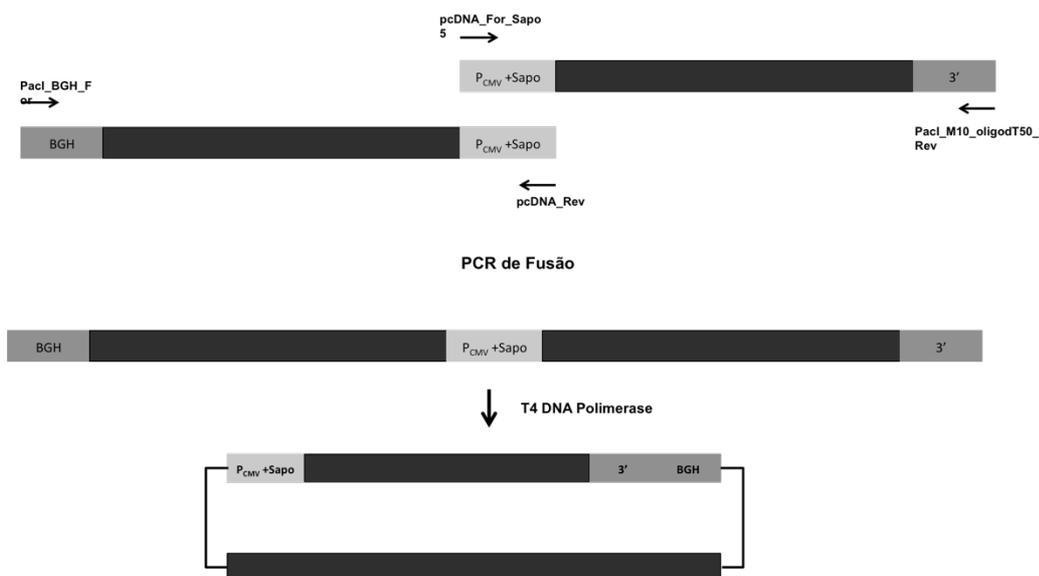


Figura 19 PCR de Fusão. O esquema representa como foi realizada a fusão entre o pcDNA e o genoma de sapovírus. As setas indicam a posição de anelamento dos primers da tabela 5.

O produto de PCR gerado, de aproximadamente 13 kb, foi precipitado com etanol e, em seguida digerido utilizando-se a enzima de restrição PacI (NEB) a 37°C por 16 h. Para purificação completa, um gel de agarose 0,8% foi corrido com todo o volume da reação anterior. Em seguida, as bandas do gel foram excisadas e purificadas utilizando-se QIAEX II Gel Extraction Kit (QIAGEN). Após quantificar o DNA (300 ng/μL), foi adicionada T4 DNA ligase (NEB) para recircularização do produto e posterior transformação. A reação foi realizada seguindo as instruções do fabricante.

Um total de 5μL da reação foi utilizado para transformação de células eletrocompetentes DH10B (Termo Fisher Scientific). Após 90 min de crescimento, 28°C, em meio nutritivo S.O.C sem antibiótico o volume total (1mL) foi distribuído utilizando-se uma alça de Drigalski em uma placa de petri contendo meio LB ágar suplementado com ampicilina (100 mg/mL). Essa placa foi incubada por 18h em estufa 28°C. Todas as colônias foram inoculadas em meio LB líquido suplementado com antibiótico para posterior extração de DNA plasmidial e análise a cerca da presença do plasmídeo referente ao genoma completo de sapovírus por sequenciamento de genoma completo, Macrogen (Coréia do Sul), utilizando-se da técnica de *primer walking* (Tabela 6). O clone positivo foi nomeado Sapo_CI_pcDNA3.1.

Tabela 6 Primers utilizados para confirmação da sequência do genoma completo (primer walking).

Primer	Sequência
Sapo_CI_1	CAAAAGCTAAAGCCAACCG
Sapo_CI_2	TGCATATGGGATTTTG
Sapo_CI_3	TCAACACACAAAATCTGCC
Sapo_CI_4	CAGTAATCGGCAAAGGTG
Sapo_CI_5	CTTCCTTGCTATCACGGG
Sapo_CI_6	CTGGCAATGTATTCAGATAGAGAC
Sapo_CI_7	TTGGAATGACAGAATGCC
Sapo_CI_8	GTGCTCCACGGTAATGG
Sapo_CI_9	GTGGCACAACCTGGAGAAAC

2.3 Transfecção celular

Foram utilizadas para ensaio inicial as células CaCo-2 (adenocarcinoma colorretal) e Huh-7 (carcinoma hepato-celular). A cultura foi mantida em meio DMEM High Glucose (Gibco) com adição de 10% de Soro Fetal Bovino (FBS) (Gibco) e 1% Antibiótico-Antimicótico (anti-anti) (100x) (Gibco), em garrafas de 75 cm² de área (frasco de cultura T75 KASVI) onde, cerca de 10⁶ células eram plaqueadas a cada passagem.

As transfecções foram realizadas em placas de 24 poços (Kasvi), onde, 10⁵ células eram plaqueadas 24 h antes do procedimento, para obtenção de confluência entre 70 a 90% de células aderidas. Para o ensaio 2 µg de DNA extraído a partir de *miniprep.* foram transfectados utilizando-se o reagente Lipofectamina 3000 (Invitrogen), seguindo as indicações do fabricante. Meio sem antibiótico e sem soro foi utilizado quando o reagente era adicionado ao DNA. (Sapo_CI_pcDNA3.1 e pcDNA3.1).

Brevemente, a recomendação do fabricante descreve que para o primeiro dia as células devem ser semeadas nas placas (após crescimento em garrafas de cultivo em passagens abaixo de cinco vezes) em uma concentração de 10⁵ por poço. Após 24 h espera-se uma confluência entre 70 a 90% para realização da transfecção. O DNA em quantidade de 2µg para cada poço foi então diluído em meio DMEM sem antibióticos e sem soro fetal, num volume final de 50 µL por poço, a esse volume foi adicionado 2 µL de reagente P3000 multiplicado ao número de poços a serem transfectados, por exemplo, 3 poços para pcDNA3.1(+) vazio igual a 6µg de DNA mais 150 µL de DMEM mais 6 µL de reagente P3000 em um único tubo. Para preparo da Lipofectamina 3000 são adicionados 0,75 µL de reagente para cada µg de

DNA mais 25 μL de DMEM para cada poço a ser transfectado. Seguindo o exemplo acima foram adicionados 4,5 μL de Lipofectamina 3000 mais 75 μL de DMEM em um único tubo. As duas misturas foram misturadas e mantidas a temperatura ambiente por 15 minutos. Em seguida o complexo foi distribuído em cada poço e DMEM sem antibiótico e sem soro foi adicionado a cada poço até um volume final de 500 μL . As placas foram então levadas a estufa, 37°C com 5% de CO_2 e, analisadas 24 horas e 48 horas pós transfecção (hpt).

As células transfectadas foram observadas diretamente ao microscópio de fluorescência invertida ou confocal antes e depois da imunomarcação.

Além de poços contendo células transfectadas em todos os ensaios células não transfectadas mas, submetidas a crescimento em meio DMEM sem antibióticos e sem soro fetal, foram observadas.

2.4 Ensaio de Imunomarcação

Como anticorpo primário foi utilizado soro de camundongo imunizado com subdomínio P2 de sapovírus, expresso previamente em *E. coli* BL21AI e, anti-dsRNA monoclonal coletado de hibridoma de camundongo, gentilmente cedido pelo Dr. Gorben Pijlman (Universidade de Wageningen, Holanda), e como anticorpo secundário Alexa Fluor 633 Goat Anti-mouse IgG (H+L) (Thermo Fisher Scientific). Após o período de transfecção (24 ou 48 hpt) as células foram lavadas com PBS 1X e fixadas com 3% Paraformaldeído e 0,1% Glutaraldeído em PBS 1X por 10 minutos. Para permeabilização das células foi adicionado PBS 1X, seguido de lavagem com 0,5% Triton por 20 minutos. As células foram então lavadas com PBS 1X, 3 vezes com duração de 15 minutos cada lavagem. O processo de bloqueio foi realizado com adição de PBS 1X mais BSA 3% em cada poço por 30 minutos.

O anticorpo primário (anti-P2 ou anti-dsRNA) foi adicionado (1:50 ou 1:75) em PBS 1X, as placas foram incubadas a temperatura ambiente por 24 h em câmara úmida escura, a 4°C (ou geladeira). O procedimento lavagem foi repetido e então o anticorpo secundário (Alexa 633 anti-mouse IgG) (1:500) adicionado por 3 h. Após processo de lavagem foi adicionado DAPI 10 μM , por 30 minutos, seguido de 3 lavagens com PBS 1X, por 15 minutos. As lamínulas foram então montadas em lamina novas e devidamente identificadas e levadas ao microscópio confocal (Leica TCS – SP5).

2.5 Citometria de Fluxo

O procedimento de transfecção foi realizado como descrito acima, utilizando-se 4 poços de uma placa de 24 poços para cada ensaio, com a obrigatoriedade de um poço de células sem transfecção.

As células transfectadas com pcDNA3-1(+) vazio e Sap0_CI_pcDNA3.1 foram imunomarcadas antes da leitura.

Os poços foram incubados com 100 μ L de 0,05% tripsina-EDTA (Gibco) por 5 minutos a 37°C. A tripsina foi inibida com adição de 500 μ L de DMEM com 10% de soro fetal bovino. As células foram então centrifugadas a 400 x g por 3 minutos.

O equipamento utilizado para leitura foi FACSCalibur (Becton Dickinson). Foram analisados os aspectos de granulidade das populações de célula e a fluorescência de acordo com a sua faixa de excitação e emissão. Utilizando o software FLOWJO (LLC) dois aspectos foram avaliados, a granulidade por tamanho da célula e a intensidade de fluorescência por número de células.

3. Resultados

Após a confirmação dos clone por sequenciamento ensaios referentes a transfecção celular foram realizados em duas diferentes linhagens celular Caco-2 e Huh-7. Uma vez definido o melhor protocolo de transfecção viral a primeira análise realizadas foi a avaliação de possível efeito citopático. Dessa forma utilizando um microscópio de fluorescência invertido as células foram observadas apenas em campo claro, com objetiva 10x, apenas para efeito de comparação.

É possível observar na Figura 27 que as células Caco-2 são células alongadas, com pouco volume citoplasmático, crescem aderidas á superfície de cultivo e estão confluentes após 24 h de semeadura. Observa-se também que as demais imagens adquiridas demonstram uma ruptura nesse padrão morfológico. Nos casos representados as células perderam sua morfologia alongada, em sua maioria, apresentando pouco ou nenhum volume citoplasmático.

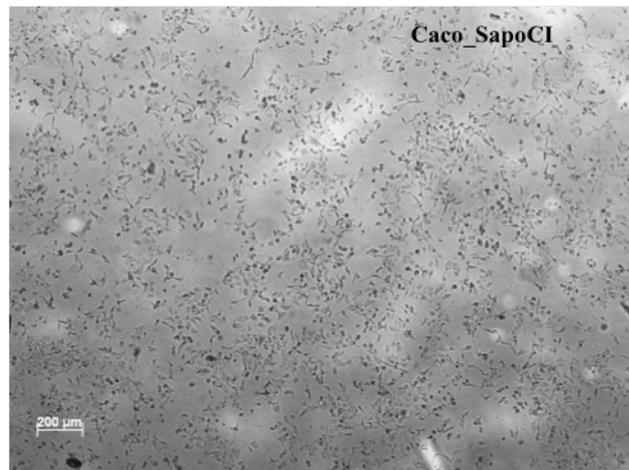
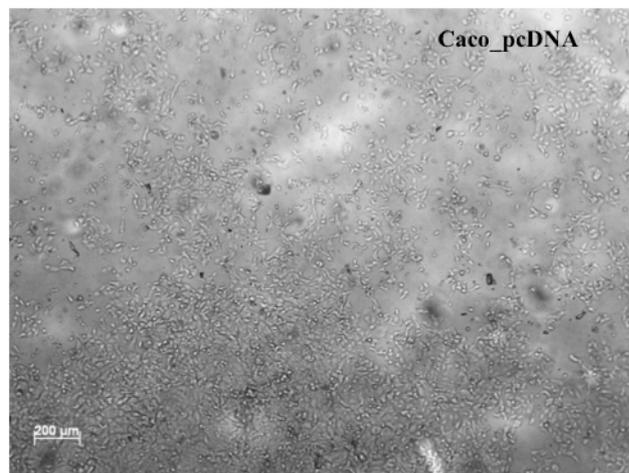
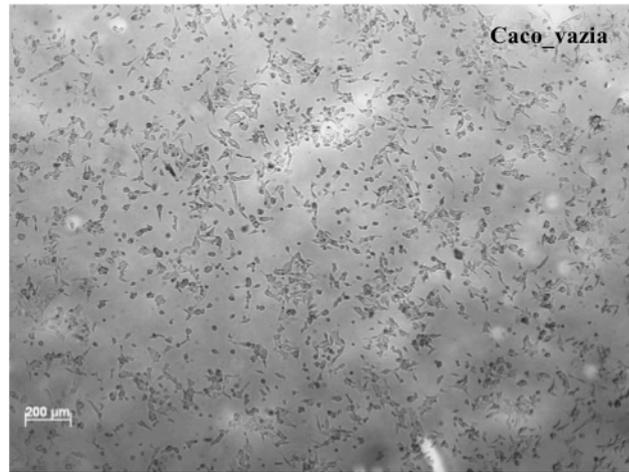


Figura 20 Células Caco-2 transfectadas com Sapo_C1 e pcDNA. Todas as imagens foram capturadas utilizando objetiva 10x. Caco-2 não transfectadas, células transfectadas com plasmídeo pcDNA3.1(+), com Sapo_C1_pcDNA3.1. Todas as imagens foram observadas 24 hpt. As células com Sapo_GFP apresentam perfil de morte celular.

Em todos os ensaios de imunomarcção para células Caco-2 ou Huh-7 pelo menos um plasmídeo controle (sem o genoma viral) foi imunomarcado, bem como Caco-2 vazia foi imunomarcada com anti-P2 e Alexa633. Para maior compreensão os controles foram organizados em figuras separadas. A seguir, a Figura 21 referente aos controles de transfecção de Caco-2 observadas em confocal. A primeira linha representa células que não foram imunomarcadas. As duas seguintes foram imunomarcadas com anti-P2 e Alexa 633 e não apresentaram fluorescência no comprimento de onda analisado, como era esperado.

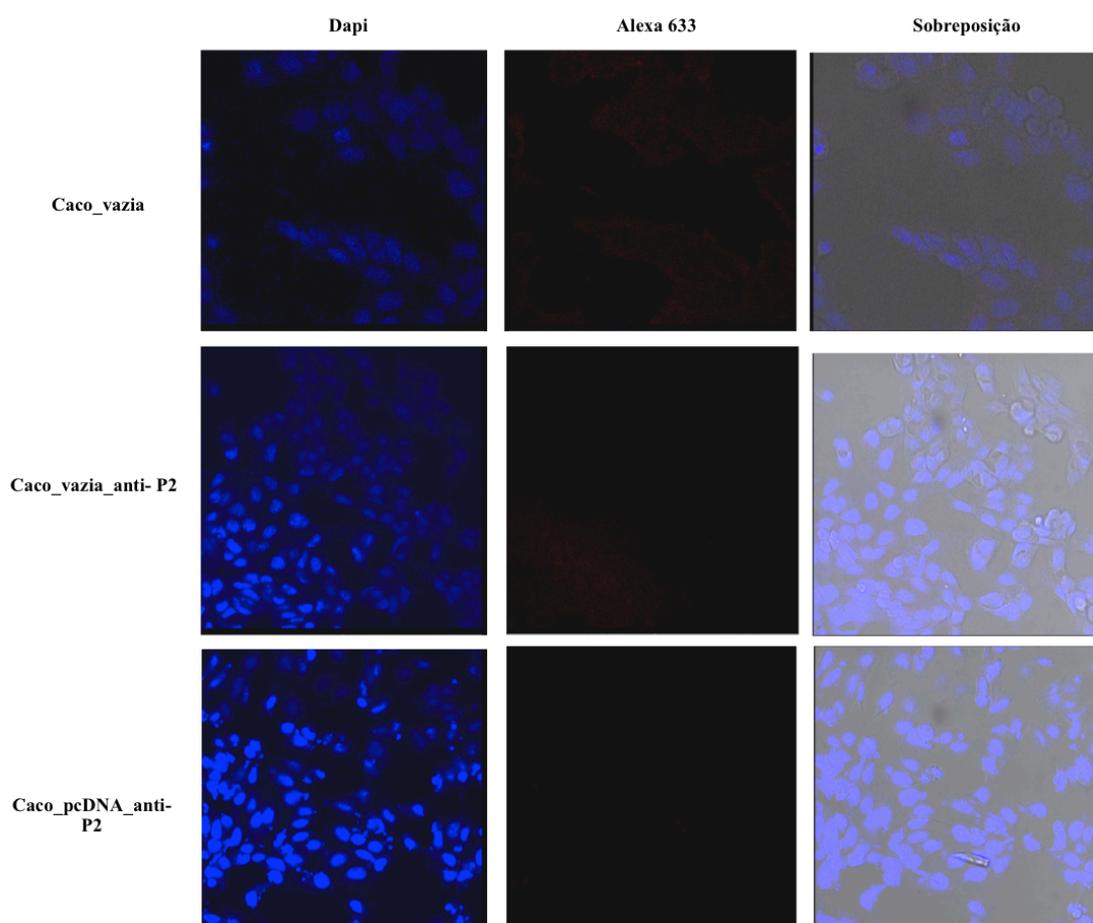


Figura 21 Caco-2 controles da transfecção. Na primeira linha as células não foram marcadas, foram apenas coradas. Na segunda e terceira linhas células Caco não transfectada e células Caco transfectada com plasmídeo pcDNA3.1(+) vazio foram coradas com DAPI e imunomarcadas com anti-P2 e Alexa 633. Como esperado não houve sinal de fluorescência.

As Figuras 22 e 23 representam o mesmo ensaio com apenas anticorpos secundários diferentes, a primeira foi realizada com anti-P2 e Alexa 488 e a segunda com anti-P2 e Alexa 633. Os controles da Figura 22 encontram-se na Figura 21.

Na Figura 22 observa-se a imunomarcção de pcDNA3.1 vazio com anti-P2 e Alexa 488 e, a mesma marcação na segunda linha nas células Caco-2 transfectadas com a construção Sapoci_pcDNA 3.1(+). A ausência de sinal de fluorescência no canal de Alexa 488 para o controle é esperado. A marcação de anti-P2 e Alexa 488 em células transfectada com Sapoci apresenta pouco não-específico.

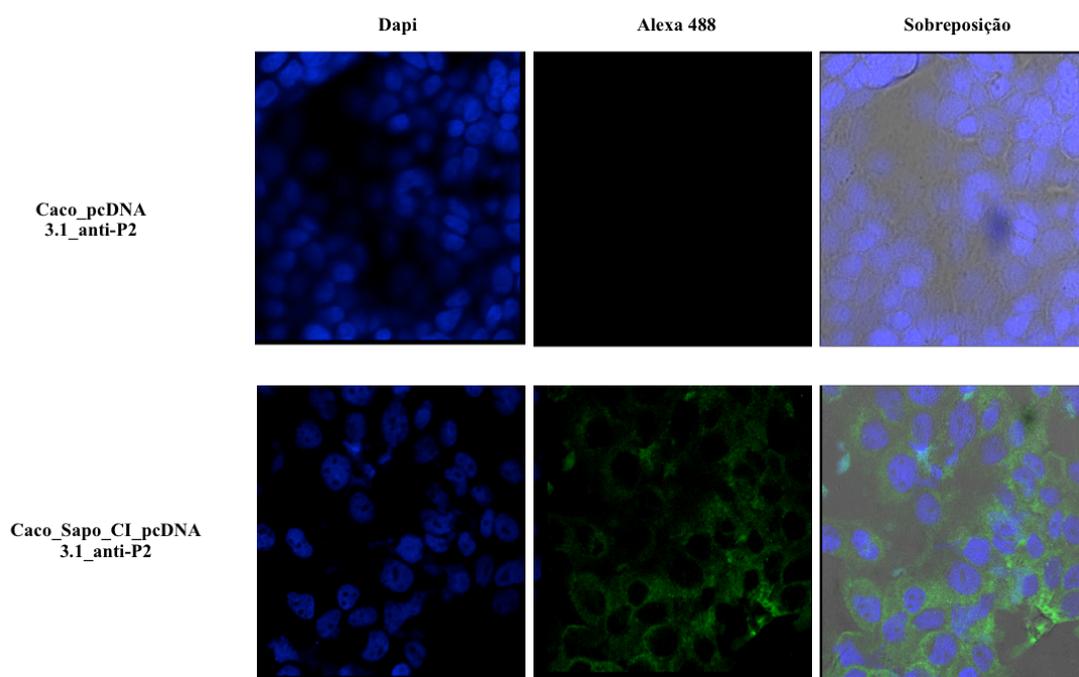


Figura 22 Imunomarcção de células Caco com anti-P2 e Alexa 488. Na primeira linha o controle de pcDNA3.1(+) vazio sem sinal de imunomarcção. Na segunda linha uma imunomarcção do ensaio contendo o genoma viral. Na sobreposição é possível notar um aumento nuclear, no entanto, ainda há volume citoplasmático, algumas marcações com características granuladas mais destacadas e a estrutura da célula parcialmente alongada.

Na Figura 23 são apresentados os controles de transfecção em células Huh-7. Nessa imagem observa-se na primeira linha células não transfectada e não marcadas. Na segunda linha foi realizada a imunomarcção (anti-P2 e Alexa 633) de células transfectadas com pcDNA3.1(+) vazio.

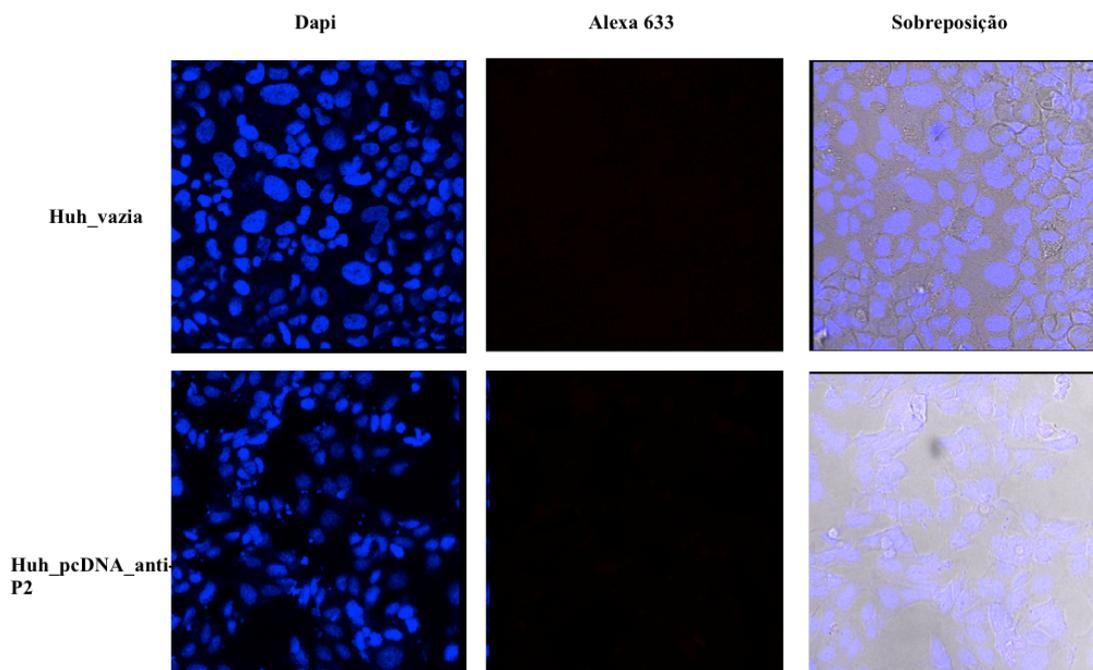


Figura 23 Controle de imunomarcção de Huh-7 com anti-P2 e Alexa 633. Na primeira linha células Huh não marcadas e não transfectada, apenas coradas com DAPI. Na segunda linha imunomarcção de pcDNA3.1(+) vazio sem sinal de fluorescência.

A Figura 24 representa o ensaio de transfecção de células Caco-2 e Huh_7 com as construções que contem o genoma do sapovírus. Na primeira linha tem-se a transfecção de SapoCI_pcDNA3.1 em Caco-2 e na segunda linha SapoCI_pcDNA3.1 em Huh-7. A imunomarcção foi realizada com anti-P2 como anticorpo primário e anti-mouse Alexa633 como secundário. Há um sinal de fluorescência bastante específico na primeira linha e é possível observar a que a marcação apresenta um padrão granuloso. Na segunda linha esse mesmo padrão é observado, porém com um sinal de não-específico.

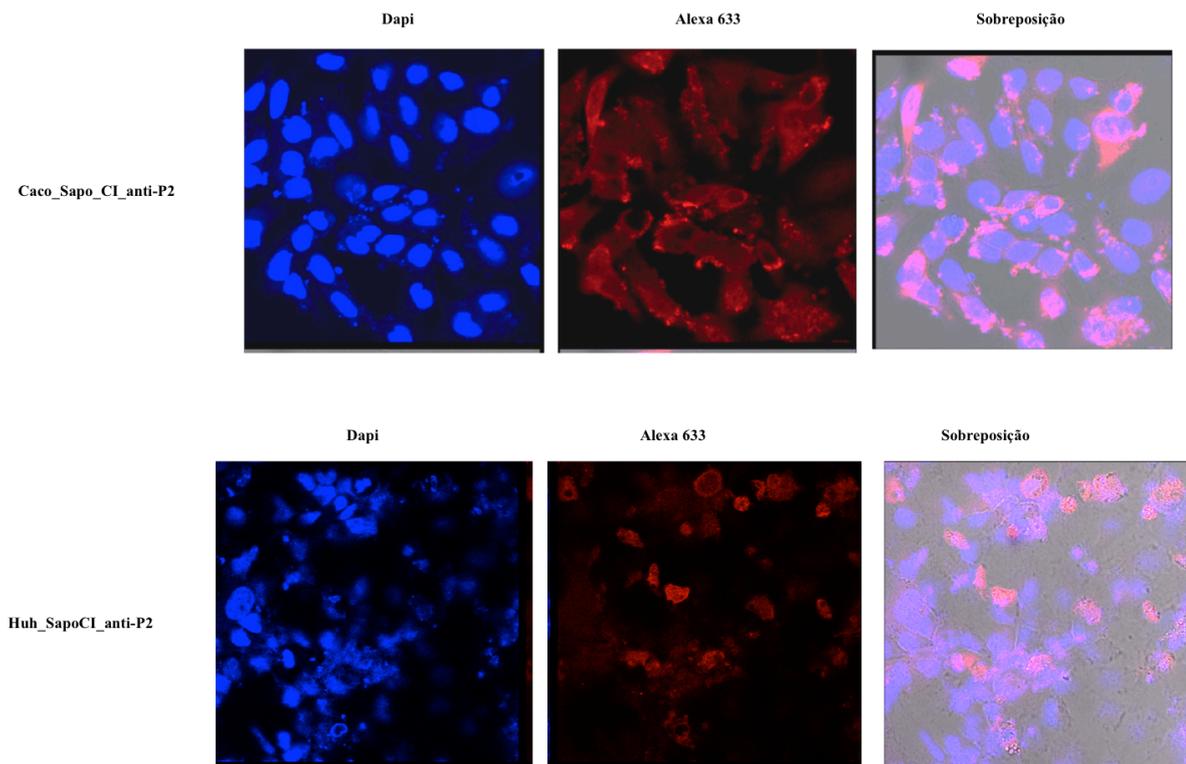


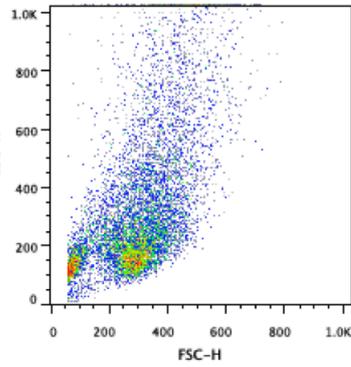
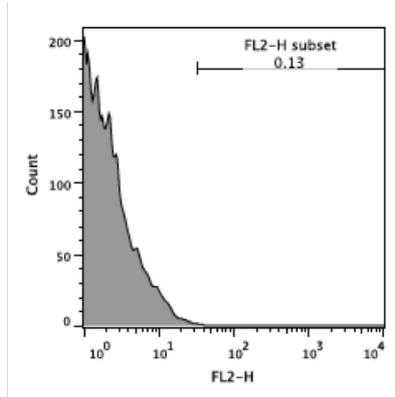
Figura 24 Imunomarcção de células Caco-2 e Huh-7 transfectadas com Sapoci pcDNA. Na primeira linha encontra-se a imunomarcção de Caco-2 e na segunda de Huh-7. O padrão de marcação em ambas é similar apresentando uma granulosidade em volta do núcleo corado com DAPI.

As imagens obtidas a partir de imunomarcção com anti-dsRNA e Alexa 633 foram inconclusivas quando comparadas entre controles e plasmídeos contendo genoma viral. A possibilidade excessivo background ou reação inespecífica justificam a ausência dessas imagens nos resultados. No entanto, as células Caco transfectadas com pcDNA3.1(+) vazio e Sapoci pcDNA3.1 foram imunomarcadas e analisadas por citometria de fluxo. Nesse ensaio foram utilizados os anticorpos primários anti-P2 e anti-dsRNA e, para ambos, o secundário Alexa 633.

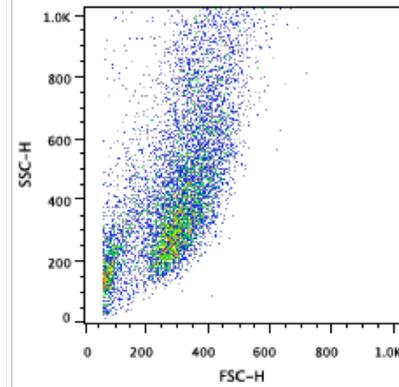
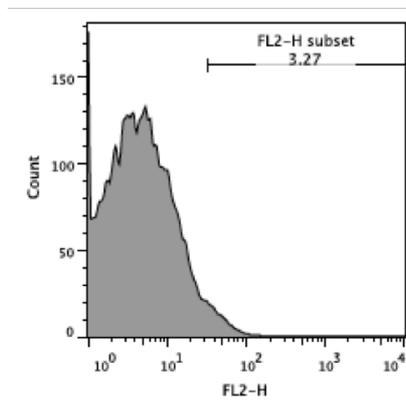
Para as células Caco não transfectada e não imunomarcadas foram detectas 18 células com possível sinal de auto fluorescência. Para o controle pcDNA3.1(+) vazio imunomarcado com anti-P2 num total de 10 mil células 327 estavam marcadas, porém a leitura das células com Sapoci pcDNA3.1 imunomarcadas com anti-P2 apresentou 550 células marcadas. A imunomarcção com anti-dsRNA para pcDNA3.1(+) apresentou 266 células imunomarcadas enquanto que Sapoci pcDNA3.1 apresentou um número de 390 células imunomarcadas.

Na Figura 25 é possível observar o controle e os demais ensaios divididos em análise da intensidade de fluorescência a esquerda, onde a barra indica as células

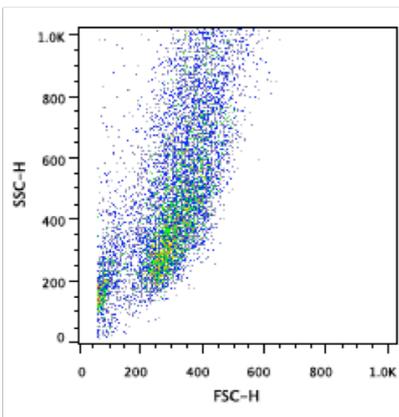
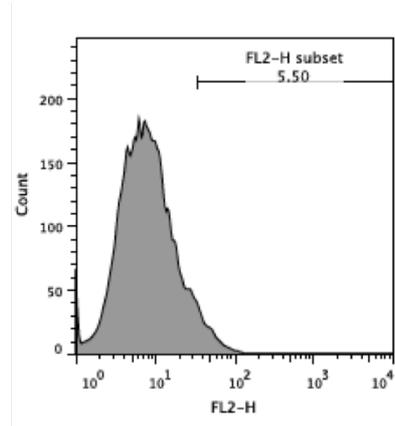
imunomarcadas e o valor a média de células em relação ao total. Enquanto que a direita o gráfico representa a morfologia da célula observada em tamanho por granulidade.



Caco_Control_48h



Caco_pcDNA_48h_anti-P2_Alexa633



Caco_Sapo_48h_anti-P2_Alexa633

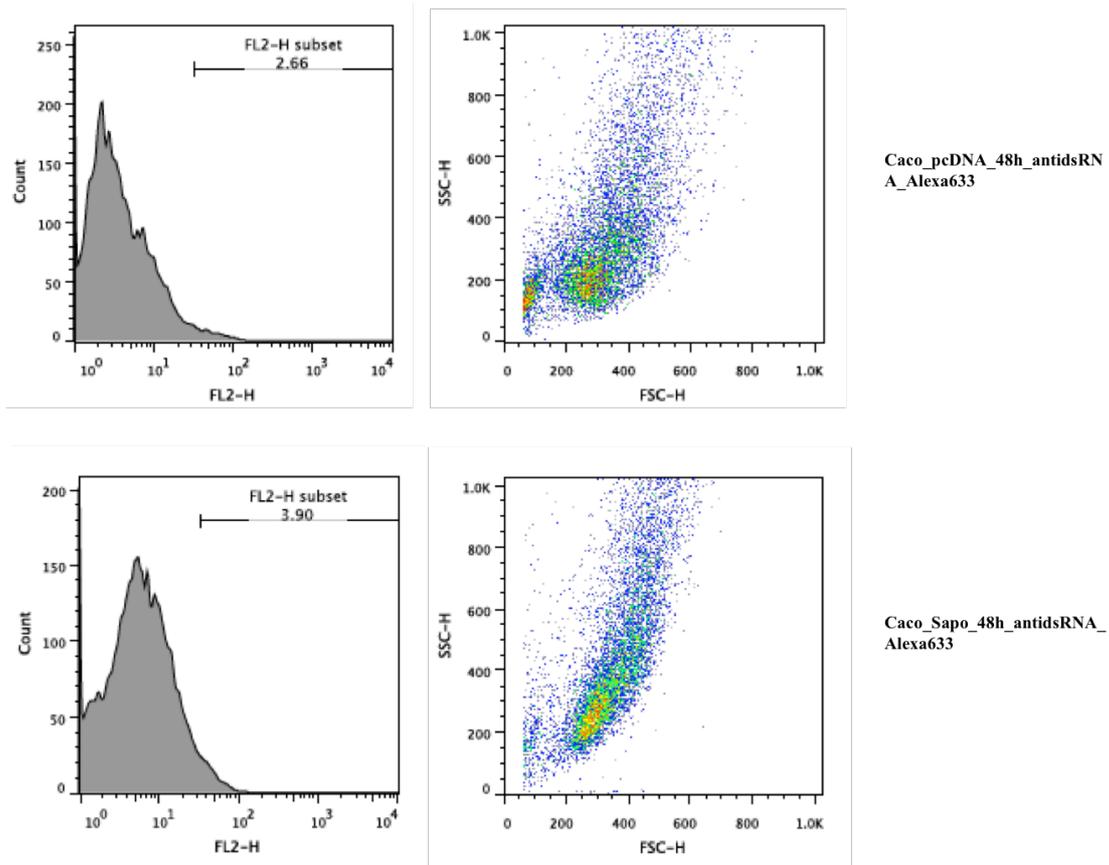


Figura 25 Citometria de fluxo de células Caco-2 para análise das imunomarcações realizadas com dois diferentes anticorpos primários. A primeira linha registra a leitura da célula vazia não imunomarcada. A segunda e a terceira linhas referem-se ao pcDNA vazio e SapoCI, respectivamente, imunomarcados com anti-P2. A quarta e a quinta linhas referem-se ao pcDNA vazio e SapoCI, respectivamente, imunomarcados com anti-dsRNA. A morfologia celular avaliada nos gráficos a direita da figura manteve-se semelhantes. Embora as células transfectadas com SapoCI tenham apresentado um número maior de células marcadas em relação as células transfectadas com pcDNA vazio, a diferença entre esses valores é maior quando avaliada a imunomarcção anti-P2.

4. Discussão

O primeiro sistema de genética reversa foi descrito em 1995 para calicivírus felinos (148) e, apenas dez anos depois, em 2005, o primeiro sistema de genética reversa para o gênero *Sapovirus* foi apresentado (45). No entanto, a exemplo do primeiro o segundo foi baseado no vírus que infecta suínos. Até o presente momento não se obteve um sistema de cultivo celular para calicivírus humanos assim como não há relatos de um sistema de genética reversa para sapovírus humanos.

Recentemente, um sistema de genética reversa sem o uso de vírus “ajudante” foi descrito para norovírus humanos. Essa abordagem foi relatada como mais eficaz do que abordagens que utilizavam-se do sistema que utilizava o vírus vaccínia recombinante expressando T7 polimerase (26). Sabe-se ainda que esse sistema mostrou-se eficaz para diferentes calicivírus mas não para os sapovírus suínos. Que foram replicados em um sistema direto, onde o RNA viral é transcrito e CAP é adicionado *in vitro*, gerando assim partículas virais viáveis capazes de infectar células LLC-PK (45).

O sistema de genética reversa descrito para norovírus GII.3 humanos, mostrou-se eficaz em quatro diferentes linhagens de células (COS-7, Hek293T, Huh-7 e Caco-2), produzindo partículas virais detectadas por microscopia eletrônica, bem como a produção de uma progênie viral marcada com o gene reporte GFP. O uso do fator de alongação 1-alfa substituiu o uso de promotores como T7 polimerase e CMV, que foram previamente descritos como, bem sucedidos, para norovírus murino (MNV) e vírus da febre hemorrágica do coelho (RHDV) (157, 158, 166, 167).

Nesse capítulo é apresentada uma abordagem para o desenvolvimento de um clone infeccioso de sapovírus humano como ferramenta para estudo de genética reversa desses vírus. A abordagem foi a clonagem do cDNA do genoma viral completo sem nenhuma modificação para posterior transfecção celular.

Apenas dois anticorpos primários foram utilizados nesse estudo, um deles, específico, produzido contra o subdomínio P2 da proteína VP1 do capsídeo e, o outro, um monoclonal produzido contra fita dupla de RNA. O primeiro anticorpo indicaria a possível expressão da proteína estrutural viral, enquanto o segundo, indicaria a presença de fita dupla de RNA, que é uma forte evidência de transcrição *de novo*.

Em ambos os casos os anticorpos foram usados para imunomarcação celular e suas reações foram observadas a partir de um critério qualitativo e, a partir de uma

estimativa quantitativa. Ao observar os sinais de fluorescência nas imagens obtidas por meio de microscopia de confocal a laser, pode-se assumir que há uma interação entre ambos anticorpos e as culturas transfectada com cDNA viral. Em um ensaio com células Huh, não apresentados nesse capítulo, observou-se a formação de estruturas similares a corpos apoptóticos num período avaliado entre 3 a 36 hpt. Esse tipo de resultado já foi previamente descrito na literatura, onde a presença de GFP estava relacionada diretamente a progressão do efeito citopático, ou ainda, inúmeras passagens faziam com que o GFP deixasse de ser expresso pelo sistema e, em alguns vírus animais a presença de GFP anulou a infectividade da partícula gerada (26, 168, 169).

O uso do clone denominado SapoCI-pcDNA3.1 apresentou resultados importantes, pois, além de emitir sinais de fluorescência após imunomarcção com anti-P2 e anti-dsRNA, também apresentou menores efeitos citopáticos. Além disso numericamente por análise de citometria de fluxo apresentou um número maior de células marcadas com anti-P2 que é um anticorpo específico para o genogrupo GI.2 utilizado no estudo.

A tentativa de recuperação de partículas virais para visualização por meio de microscopia eletrônica de transmissão não foi bem sucedida, assim como a baixa especificidade observada no Western blot realizado com os sobrenadantes das transfecções e o anticorpo anti-P2 (dados não mostrados). Isso pode ter ocorrido devido a uma ineficiência no processo de concentração e purificação e partículas desse material.

Um ensaio capaz de duplicar a prova de que o cDNA viral está sendo transcrito *de novo* pelo mecanismo celular, seria um ensaio de Northern blot para detecção da fita senso negativo. A positividade desse ensaio concluiria a funcionalidade do sistema, mesmo que não haja formação de partículas.

O uso de anticorpos para diferentes proteínas não estruturais como, VPg, RdRp, possibilitaria a mesma contra prova de uma maneira mais simples, no entanto, esses anticorpos não são disponíveis comercialmente.

É importante que a escolha de se usar um cDNA viral inserido num plasmídeo sob o comando de um promotor CMV é oriunda dos sistemas de genética reversa para sapovírus suíno, norovírus murino e RHDV (45, 156, 166). Apesar de se ter uma construção onde o genoma completo de sapovírus está clonado num promotor que apresenta a sequência do promotor T7 polimerase, a transcrição *in vitro* foi descartada

devido a falta do sistema de adição de CAP *in vitro* não ser acessível nesse momento. No entanto, essa abordagem, se realizada, também trará uma nova ferramenta a ser explorada, bem como, será considerada um ineditismo para os sapovírus humanos, até o presente.

Sendo assim nesse capítulo foram apresentadas evidências de que dois plasmídeos contendo cDNA de sapovírus, um com eGFP inserido como gene repórter, podem funcionar como ferramentas para genética reversa desses vírus. Entretanto, são necessários maiores esforços para confirmação dessa hipótese.

5. Conclusão Geral

Os primeiros calicivírus humanos foram detectados e descritos em 1972 e 1977 (15, 170), nessa mesma década inúmeros outros vírus entéricos foram identificados como, por exemplo, os rotavírus e astrovírus (171, 172). Curiosamente, apenas os calicivírus humanos permanecem sem possuírem uma cultura celular permissiva que possibilite não só ensaios que requeiram isolamento viral, mas também compreensão de infecção-replicação-saída do vírus. Isto posto, outras abordagens foram desenvolvidas para obtenção de informações a respeito desses vírus, que, atualmente, são os maiores causadores de gastroenterites no mundo.

No Brasil os dados a cerca da epidemiologia dos calicivírus humanos ainda não faz parte de um sistema de vigilância nacional, contudo, sabe-se que os calicivírus humanos já foram detectados em rios, lagoas e águas de esgoto e, em muitos casos o objetivo final desses estudos era o monitoramento da contaminação das águas ou análise comparativa entre amostras clínicas e amostras ambientais (61, 111, 173-175). Porém a maior parte dos dados compilados é oriunda das grandes regiões Sudeste, Nordeste, Norte e Sul. Pouco se sabe a respeito da circulação desses vírus na região Centro Oeste. Dessa forma foi proposto uma análise água de esgoto não tratada nos anos de 2014 e 2015, os meses referente ao período de baixas temperaturas e baixo índice pluviométrico foram escolhidos baseados em estudos que os descrevem como meses de maior incidência de calicivirus humanos em países tropicais (87).

Nesse trabalho os resultados obtidos foram além dos calicivírus humanos e demais vírus entéricos. O uso de metagenômica permitiu a identificação de diferentes vírus de plantas e animais. Ainda que não seja um trabalho de vigilância, a partir dos dados coletados pode-se inferir que a população de Brasília é acometida por calicivírus humanos, já bem descritos em trabalhos que avaliaram água de esgoto do Rio de Janeiro, por exemplo (61, 176). Além de sapovírus, que ainda são pouco estudados no país, os norovírus identificados curiosamente apresentaram genótipos descritos no mesmo período em outros países, incluindo a presença do genótipo GII.17 na amostra de 2015 (32, 177, 178).

A análise de amostras de esgoto não tratado resultou em uma descrição dos possíveis vírus circulantes no Distrito Federal e, esse tipo de dado pode proporcionar futuras medidas na escolha do tratamento desse esgoto, para que os vírus sejam eliminados e o risco de contaminação da água onde o efluente é despejado seja

minimizado. Futuros estudos que incluam monitoramento de diferentes meses do ano podem ajudar numa melhor abordagem para conscientização dos cuidados a serem tomados para evitar contaminação, principalmente, de água e alimentos, uma vez que não existe um tratamento com antivirais ou vacinas contra calicivírus humanos.

As informações a respeito das sequências virais são de suma importância principalmente para identificação e classificação viral, dentro dos estudos dos calicivírus humanos utiliza-se a sequência para se obter informações epidemiológicas e muitas vezes tentar compreender o porque de determinado genótipo é mais prevalente que outros. Levando isso em consideração uma abordagem diferente no estudo desses vírus é a análise estrutural.

Nesse trabalho diferentes genótipos de norovírus, pouco incidentes, foram escolhidos com o intuito de se compreender o que poderia causar essa baixa incidência no aspecto estrutural da região mais variável em sequencia desse genótipos.

O que foi demonstrado é que pelo menos um deles, o GII.21 possui uma estrutura diferente do genótipo mais epidêmico o GII.4 e, além disso o genótipo GII.22 apesar de inúmeros esforços e sua estrutura similar a GII.4, não apresentou ligação com HBGAs inteiros mas sim com a fucose, em regiões já determinadas para outro genótipo pouco incidente o GII.10 (179). Esse tipo de informação é importante uma vez que não se sabe ao certo qual a forma de entrada dos norovírus nas células humanas, mas sabe-se que HBGA livres são carboidratos que favorecem essa entrada (27). Sendo assim a resolução de estruturas de norovírus pode permitir a análise evolutiva desses vírus e qual a relação entre as diferentes estruturas do capsídeos e os diferentes carboidratos a quais se ligam.

Por fim uma vez que um vírus é identificado em uma amostra, analisado por sequenciamento, analisado por características morfológicas e estruturais, espera-se poder estudar-se suas interações com as células do hospedeiro. Esse fluxograma, no caso dos calicivírus humanos ainda não é alcançado com facilidade devido a ausência de uma cultura celular robusta. Embora diferentes esforços tenham sido demonstrados recentemente como, o uso de células e enterócitos derivados de células tronco (27, 28), a reprodutibilidade de ambos ainda não é acessível. Desta forma o uso de plasmídeos modificados com o genoma viral são ferramentas simples, de fácil manipulação para geração de informação a respeito do vírus, seu ciclo de replicação e possível interação entre proteínas virais e proteínas do hospedeiro.

Foi descrito nesse trabalho um plasmídeo cuja a similaridade encontra-se na presença do genoma completo do sapovírus BR01, identificado pela primeira vez no Distrito Federal, no ano de 2011 (165). Nos resultados descritos a despeito das abordagens utilizadas para a confirmação da funcionalidade do clone foram observados resultados que caracterizam uma possível transcrição e tradução. Comprovadas pela detecção com anticorpo específico contra a região mais exposta do capsídeo (anti-P2) e ainda a presença de dupla fita de RNA através da imunomarcação analisada através de citometria.

Conclui-se que os calicivírus humanos, norovírus e sapovírus, são vírus causadores de gastroenterites agudas que circulam na população do Distrito Federal, necessitando a análise comparativa com amostras clínicas e estudos que correlacionem possíveis surtos a esses vírus na mesma localidade. Além disso a aplicação de biologia estrutural permite avaliar a diferença morfológica entre diferentes genótipos e criar hipóteses que correlacionem as estruturas e a incidência epidemiológica desses vírus. Durante o desenvolvimento de estudos estruturais é possível melhoras técnicas que permitam a expressão de proteínas em condição nativa não demonstradas ainda e, por fim, a clonagem do cDNA de sapovírus em um vetor de expressão contendo promotor CMV mostrou-se promissora no desenvolvimento de uma ferramenta que permita a genética reversa de sapovírus humanos.

6. Perspectivas Futuras

Nessa tese foram apresentadas três diferentes abordagens para o estudo de virologia básica e, a partir dessas abordagens foram gerados conhecimentos ainda não descritos na literatura até o momento. Espera-se com isso uma contribuição na área de estudos sobre calicivírus humanos e, espera-se a publicação da estrutura de GII.22 e publicação do clone de sapovírus após ensaios que demonstrem a presença da partícula viral.

O uso de metagenoma para detecção de calicivírus poderá ser utilizado como justificativa futura para implementação de uma forma de vigilância clínica e ambiental a respeito dos vírus causadores de gastroenterites circulantes no Distrito Federal.

Espera-se, ainda, a submissão, no banco de proteínas, das estruturas de norovírus resolvidas nesse trabalho.

Por fim, com o uso de microscopia eletrônica de transmissão espera-se provar a presença de partículas virais após a transfecção do clone apresentado, bem como um ensaio de hibridização *in situ* ou um Northern blot para a comprovação de replicação e um possível RNA subgenômico referente ao capsídeos de sapovírus.

7. Referências Bibliográficas

1. **Flint SJ, Enquist LW, Racaniello VR, Skalka AM.** 2008. Principles of Virology, 3 ed, vol 1. AMS Press.
2. **Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, Jha P, Campbell H, Walker CF, Cibulskis R, Eisele T, Liu L, Mathers C, Child Health Epidemiology Reference Group of WHO, Unicef.** 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* **375**:1969-1987.
3. **Unicef, WHO.** 2009. Diarrhoea: Why children are still dying and what can be done, New York.
4. **Fioretti JM, Bello G, Rocha MS, Victoria M, Leite JP, Miagostovich MP.** 2014. Temporal dynamics of norovirus GII.4 variants in Brazil between 2004 and 2012. *PLoS One* **9**:e92988.
5. **Smith AW, Skilling DE, Castello JD, Rogers SO.** 2004. Ice as a reservoir for pathogenic human viruses: specifically, caliciviruses, influenza viruses, and enteroviruses. *Med Hypotheses* **63**:560-566.
6. **Thiel HJ, Konig M.** 1999. Caliciviruses: an overview. *Veterinary Microbiology* **69**:55-62.
7. **Oka T, Katayama K, Ogawa S, Hansman GS, Kageyama T, Ushijima H, Miyamura T, Takeda N.** 2005. Proteolytic processing of sapovirus ORF1 polyprotein. *Journal of Virology* **79**:7283-7290.
8. **Ushijima H, Okitsu S, Khamrin P.** 2011. [Calicivirus]. *Uirusu* **61**:193-203.
9. **Pedersen NC, Elliott JB, Glasgow A, Poland A, Keel K.** 2000. An isolated epizootic of hemorrhagic-like fever in cats caused by a novel and highly virulent strain of feline calicivirus. *Vet Microbiol* **73**:281-300.
10. **Alonso C, Oviedo JM, Martin-Alonso JM, Diaz E, Boga JA, Parra F.** 1998. Programmed cell death in the pathogenesis of rabbit hemorrhagic disease. *Arch Virol* **143**:321-332.
11. **Kaplon J, Guenau E, Asdrubal P, Pothier P, Ambert-Balay K.** 2011. Possible novel nebovirus genotype in cattle, France. *Emerg Infect Dis* **17**:1120-1123.
12. **Oka T, Yokoyama M, Katayama K, Tsunemitsu H, Yamamoto M, Miyashita K, Ogawa S, Motomura K, Mori H, Nakamura H, Wakita T, Takeda N, Sato H.** 2009. Structural and biological constraints on diversity of regions immediately upstream of cleavage sites in calicivirus precursor proteins. *Virology* **394**:119-129.
13. **Prasad BV, Matson DO, Smith AW.** 1994. Three-dimensional structure of calicivirus. *J Mol Biol* **240**:256-264.
14. **McFadden N, Bailey D, Carrara G, Benson A, Chaudhry Y, Shortland A, Heeney J, Yarovinsky F, Simmonds P, Macdonald A, Goodfellow I.** 2011. Norovirus regulation of the innate immune response and apoptosis occurs via the product of the alternative open reading frame 4. *PLoS Pathog* **7**:e1002413.
15. **Kapikian AZ, Wyatt RG, Dolin R, Thornhill TS, Kalica AR, Chanock RM.** 1972. Visualization by immune electron microscopy of a 27-nm particle

- associated with acute infectious nonbacterial gastroenteritis. *J Virol* **10**:1075-1081.
16. **Kaplan JE, Gary GW, Baron RC, Singh N, Schonberger LB, Feldman R, Greenberg HB.** 1982. Epidemiology of Norwalk gastroenteritis and the role of Norwalk virus in outbreaks of acute nonbacterial gastroenteritis. *Ann Intern Med* **96**:756-761.
 17. **Jiang X, Wang M, Wang K, Estes MK.** 1993. Sequence and genomic organization of Norwalk virus. *Virology* **195**:51-61.
 18. **White LJ, Ball JM, Hardy ME, Tanaka TN, Kitamoto N, Estes MK.** 1996. Attachment and entry of recombinant Norwalk virus capsids to cultured human and animal cell lines. *J Virol* **70**:6589-6597.
 19. **Ball JM, Graham DY, Opekun AR, Gilger MA, Guerrero RA, Estes MK.** 1999. Recombinant Norwalk virus-like particles given orally to volunteers: phase I study. *Gastroenterology* **117**:40-48.
 20. **Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK.** 1999. X-ray crystallographic structure of the Norwalk virus capsid. *Science* **286**:287-290.
 21. **Karst SM, Wobus CE, Lay M, Davidson J, Virgin HWt.** 2003. STAT1-dependent innate immunity to a Norwalk-like virus. *Science* **299**:1575-1578.
 22. **Wobus CE, Karst SM, Thackray LB, Chang KO, Sosnovtsev SV, Belliot G, Krug A, Mackenzie JM, Green KY, Virgin HW.** 2004. Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. *PLoS Biol* **2**:e432.
 23. **Duizer E, Schwab KJ, Neill FH, Atmar RL, Koopmans MP, Estes MK.** 2004. Laboratory efforts to cultivate noroviruses. *J Gen Virol* **85**:79-87.
 24. **Asanaka M, Atmar RL, Ruvolo V, Crawford SE, Neill FH, Estes MK.** 2005. Replication and packaging of Norwalk virus RNA in cultured mammalian cells. *Proc Natl Acad Sci U S A* **102**:10327-10332.
 25. **Guix S, Asanaka M, Katayama K, Crawford SE, Neill FH, Atmar RL, Estes MK.** 2007. Norwalk virus RNA is infectious in mammalian cells. *J Virol* **81**:12238-12248.
 26. **Katayama K, Murakami K, Sharp TM, Guix S, Oka T, Takai-Todaka R, Nakanishi A, Crawford SE, Atmar RL, Estes MK.** 2014. Plasmid-based human norovirus reverse genetics system produces reporter-tagged progeny virus containing infectious genomic RNA. *Proc Natl Acad Sci U S A* **111**:E4043-4052.
 27. **Jones MK, Watanabe M, Zhu S, Graves CL, Keyes LR, Grau KR, Gonzalez-Hernandez MB, Iovine NM, Wobus CE, Vinje J, Tibbetts SA, Wallet SM, Karst SM.** 2014. Enteric bacteria promote human and mouse norovirus infection of B cells. *Science* **346**:755-759.
 28. **Ettayebi K, Crawford SE, Murakami K, Broughman JR, Karandikar U, Tenge VR, Neill FH, Blutt SE, Zeng XL, Qu L, Kou B, Opekun AR, Burrin D, Graham DY, Ramani S, Atmar RL, Estes MK.** 2016. Replication of human noroviruses in stem cell-derived human enteroids. *Science* **353**:1387-1393.
 29. **Saxena K, Blutt SE, Ettayebi K, Zeng XL, Broughman JR, Crawford SE, Karandikar UC, Sastri NP, Conner ME, Opekun AR, Graham DY, Qureshi W, Sherman V, Foulke-Abel J, In J, Kovbasnjuk O, Zachos NC,**

- Donowitz M, Estes MK.** 2015. Human Intestinal Enteroids: a New Model To Study Human Rotavirus Infection, Host Restriction, and Pathophysiology. *J Virol* **90**:43-56.
30. **Martella V, Lorusso E, Decaro N, Elia G, Radogna A, D'Abramo M, Desario C, Cavalli A, Corrente M, Camero M, Germinario CA, Banyai K, Di Martino B, Marsilio F, Carmichael LE, Buonavoglia C.** 2008. Detection and molecular characterization of a canine norovirus. *Emerg Infect Dis* **14**:1306-1308.
31. **van Der Poel WH, Vinje J, van Der Heide R, Herrera MI, Vivo A, Koopmans MP.** 2000. Norwalk-like calicivirus genes in farm animals. *Emerg Infect Dis* **6**:36-41.
32. **de Graaf M, van Beek J, Vennema H, Podkolzin AT, Hewitt J, Bucardo F, Templeton K, Mans J, Nordgren J, Reuter G, Lynch M, Rasmussen LD, Iritani N, Chan MC, Martella V, Ambert-Balay K, Vinje J, White PA, Koopmans MP.** 2015. Emergence of a novel GII.17 norovirus - End of the GII.4 era? *Euro Surveill* **20**.
33. **Oka T, Mori K, Iritani N, Harada S, Ueki Y, Iizuka S, Mise K, Murakami K, Wakita T, Katayama K.** 2012. Human sapovirus classification based on complete capsid nucleotide sequences. *Arch Virol* **157**:349-352.
34. **Bok K, Abente EJ, Realpe-Quintero M, Mitra T, Sosnovtsev SV, Kapikian AZ, Green KY.** 2009. Evolutionary dynamics of GII.4 noroviruses over a 34-year period. *J Virol* **83**:11890-11901.
35. **Chiba S, Sakuma Y, Kogasaka R, Akihara M, Horino K, Nakao T, Fukui S.** 1979. OUTBREAK OF GASTROENTERITIS ASSOCIATED WITH CALICIVIRUS IN AN INFANT HOME. *Journal of Medical Virology* **4**:249-254.
36. **Saif LJ, Bohl EH, Theil KW, Cross RF, House JA.** 1980. ROTAVIRUS-LIKE, CALICIVIRUS-LIKE, AND 23-NM VIRUS-LIKE PARTICLES ASSOCIATED WITH DIARRHEA IN YOUNG-PIGS. *Journal of Clinical Microbiology* **12**:105-111.
37. **Guo M, Hayes J, Cho KO, Parwani AV, Lucas LM, Saif LJ.** 2001. Comparative pathogenesis of tissue culture-adapted and wild-type cowden porcine enteric calicivirus (PEC) in gnotobiotic pigs and induction of diarrhea by intravenous inoculation of wild-type PEC. *Journal of Virology* **75**:9239-9251.
38. **Li L, Pesavento PA, Shan T, Leutenegger CM, Wang C, Delwart E.** 2011. Viruses in diarrhoeic dogs include novel kobuviruses and sapoviruses. *Journal of General Virology* **92**:2534-2541.
39. **Tse H, Chan W-M, Li KSM, Lau SKP, Woo PCY, Yuen K-Y.** 2012. Discovery and Genomic Characterization of a Novel Bat Sapovirus with Unusual Genomic Features and Phylogenetic Position. *Plos One* **7**.
40. **Mombo IM, Berthet N, Bouchier C, Fair JN, Schneider BS, Renaud F, Leroy EM, Rougeron V.** 2014. Characterization of a genogroup I sapovirus isolated from chimpanzees in the republic of congo. *Genome Announc* **2**.
41. **Yoshida T, Kasuo S, Azegami Y, Uchiyama Y, Satsumabayashi K, Shiraishi T, Katayama K, Wakita T, Takeda N, Oka T.** 2009. Characterization of sapoviruses detected in gastroenteritis outbreaks and

- identification of asymptomatic adults with high viral load. *J Clin Virol* **45**:67-71.
42. **Hansman GS, Oka T, Katayama K, Takeda N.** 2007. Human sapoviruses: genetic diversity, recombination, and classification. *Rev Med Virol* **17**:133-141.
 43. **Schlenker C, Surawicz CM.** 2009. Emerging infections of the gastrointestinal tract. *Best Practice & Research Clinical Gastroenterology* **23**:89-99.
 44. **Thongprachum A, Khamrin P, Maneeakarn N, Hayakawa S, Ushijima H.** 2016. Epidemiology of gastroenteritis viruses in Japan: Prevalence, seasonality, and outbreak. *J Med Virol* **88**:551-570.
 45. **Chang KO, Sosnovtsev SV, Sosnovtsev SS, Belliot G, Wang Q, Saif LJ, Green KY.** 2005. Reverse genetics system for porcine enteric calicivirus, a prototype sapovirus in the Caliciviridae. *J Virol* **79**:1409-1416.
 46. **Chang KO, Sosnovtsev SV, Belliot G, Kim Y, Saif LJ, Green KY.** 2004. Bile acids are essential for porcine enteric calicivirus replication in association with down-regulation of signal transducer and activator of transcription 1. *Proceedings of the National Academy of Sciences of the United States of America* **101**:8733-8738.
 47. **Hosmillo M, Sweeney TR, Chaudhry Y, Leen E, Curry S, Goodfellow I, Cho KO.** 2016. The RNA Helicase eIF4A Is Required for Sapovirus Translation. *J Virol* **90**:5200-5204.
 48. **Hosmillo M, Chaudhry Y, Kim DS, Goodfellow I, Cho KO.** 2014. Sapovirus Translation Requires an Interaction between VPg and the Cap Binding Protein eIF4E. *Journal of Virology* **88**:12213-12221.
 49. **Chen R, Neill JD, Estes MK, Prasad BVV.** 2006. X-ray structure of a native calicivirus: structural insights into antigenic diversity and host specificity. *Proceedings of the National Academy of Sciences of the United States of America* **103**:8048-8053.
 50. **Clarke IN, Lambden PR.** 2001. The molecular biology of human caliciviruses. *Novartis Foundation Symposium* **238**:180-191; discussion 191-196-180-191; discussion 191-196.
 51. **Schuffenecker I, Ando T, Thouvenot D, Lina B, Aymard M.** 2001. Genetic classification of "Sapporo-like viruses". *Archives of Virology* **146**:2115-2132.
 52. **Farkas T, Zhong WM, Jing Y, Huang PW, Espinosa SM, Martinez N, Morrow AL, Ruiz-Palacios GM, Pickering LK, Jiang X.** 2004. Genetic diversity among sapoviruses. *Archives of Virology* **149**:1309-1323.
 53. **Scheuer KA, Oka T, Hoet AE, Gebreyes WA, Molla BZ, Saif LJ, Wang Q.** 2013. Prevalence of porcine noroviruses, molecular characterization of emerging porcine sapoviruses from finisher swine in the United States, and unified classification scheme for sapoviruses. *J Clin Microbiol* **51**:2344-2353.
 54. **Rzezutka A, Cook N.** 2004. Survival of human enteric viruses in the environment and food. *FEMS Microbiol Rev* **28**:441-453.
 55. **Iizuka S, Takai-Todaka R, Ohshiro H, Kitajima M, Wang Q, Saif LJ, Wakita T, Noda M, Katayama K, Oka T.** 2013. Detection of multiple human sapoviruses from imported frozen individual clams. *Food Environ Virol* **5**:119-125.

56. **Murray TY, Mans J, Taylor MB.** 2013. Human calicivirus diversity in wastewater in South Africa. *J Appl Microbiol* **114**:1843-1853.
57. **Hassan-Ríos E, Torres P, Muñoz E, Matos C, Hall AJ, Gregoricus N, Vinjé J.** 2013. Sapovirus gastroenteritis in preschool center, Puerto Rico, 2011. *Emerg Infect Dis* **19**:174-175.
58. **Borrows CL, Turner PC.** 2014. Seasonal screening for viral gastroenteritis in young children and elderly hospitalized patients: is it worthwhile? *Journal of Hospital Infection* **87**:98-102.
59. **Nidaira M, Taira K, Kato T, Arakaki E, Kyan H, Takara T, Okano S, Kuba Y, Kudaka J, Noda M.** 2014. Phylogenetic Analysis of Sapovirus Detected from an Outbreak of Acute Gastroenteritis on Ishigaki Island (Okinawa Prefecture, Japan) in 2012. *Japanese Journal of Infectious Diseases* **67**:141-143.
60. **Iritani N, Kaida A, Abe N, Kubo H, Sekiguchi JI, Yamamoto SP, Goto K, Tanaka T, Noda M.** 2014. Detection and genetic characterization of human enteric viruses in oyster-associated gastroenteritis outbreaks between 2001 and 2012 in Osaka City, Japan. *J Med Virol* doi:10.1002/jmv.23883.
61. **Fioretti JM, Rocha MS, Fumian TM, Ginuino A, da Silva TP, de Assis MR, Rodrigues Jde S, Carvalho-Costa FA, Miagostovich MP.** 2016. Occurrence of human sapoviruses in wastewater and stool samples in Rio De Janeiro, Brazil. *J Appl Microbiol* **121**:855-862.
62. **Haramoto E, Katayama H, Phanuwat C, Ohgaki S.** 2008. Quantitative detection of sapoviruses in wastewater and river water in Japan. *Lett Appl Microbiol* **46**:408-413.
63. **Kiulia NM, Netshikweta R, Page NA, Van Zyl WB, Kiraithe MM, Nyachio A, Mwenda JM, Taylor MB.** 2010. The detection of enteric viruses in selected urban and rural river water and sewage in Kenya, with special reference to rotaviruses. *J Appl Microbiol* **109**:818-828.
64. **Sano D, Pérez-Sautu U, Guix S, Pintó RM, Miura T, Okabe S, Bosch A.** 2011. Quantification and genotyping of human sapoviruses in the Llobregat river catchment, Spain. *Appl Environ Microbiol* **77**:1111-1114.
65. **Murray TY, Mans J, Taylor MB.** 2013. First detection of human sapoviruses in river water in South Africa. *Water Sci Technol* **67**:2776-2783.
66. **Murray TY, Mans J, van Zyl WB, Taylor MB.** 2013. Application of a competitive internal amplification control for the detection of sapoviruses in wastewater. *Food Environ Virol* **5**:61-68.
67. **Kitajima M, Iker BC, Pepper IL, Gerba CP.** 2014. Relative abundance and treatment reduction of viruses during wastewater treatment processes--identification of potential viral indicators. *Sci Total Environ* **488-489**:290-296.
68. **Zhang C, Niu P, Hong Y, Wang J, Zhang J, Ma X.** 2015. A probe-free four-tube real-time PCR assay for simultaneous detection of twelve enteric viruses and bacteria. *J Microbiol Methods* **118**:93-98.
69. **Yoshida T, Kasuo S, Azegami Y, Uchiyama Y, Satsumabayashi K, Shiraishi T, Katayama K, Wakita T, Takeda N, Oka T.** 2009. Characterization of sapoviruses detected in gastroenteritis outbreaks and identification of asymptomatic adults with high viral load. *Journal of*

- Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology **45**:67-71.
70. **Yan H, Yagyu F, Okitsu S, Nishio O, Ushijima H.** 2003. Detection of norovirus (GI, GII), Sapovirus and astrovirus in fecal samples using reverse transcription single-round multiplex PCR. *Journal of Virological Methods* **114**:37-44.
 71. **Xavier MPTP, Oliveira SA, Ferreira MSR, Victoria M, Miranda V, Silva MFM, Strina A, Barreto ML, Miagostovicht MP, Leite JPG.** 2009. Detection of caliciviruses associated with acute infantile gastroenteritis in Salvador, an urban center in Northeast Brazil. *Brazilian Journal of Medical and Biological Research = Revista Brasileira De Pesquisas Médicas E Biológicas / Sociedade Brasileira De Biofísica [et Al]* **42**:438-444.
 72. **Wolffs PF, Bruggeman CA, van Well GT, van Loo IH.** 2011. Replacing traditional diagnostics of fecal viral pathogens by a comprehensive panel of real-time PCRs. *J Clin Microbiol* **49**:1926-1931.
 73. **Shigemoto N, Fukuda S, Tanizawa Y, Kuwayama M, Ohara S, Seno M.** 2011. Detection of norovirus, sapovirus, and human astrovirus in fecal specimens using a multiplex reverse transcription-PCR with fluorescent dye-labeled primers. *Microbiol Immunol* **55**:369-372.
 74. **Cantalupo PG, Calgua B, Zhao G, Hundesa A, Wier AD, Katz JP, Grabe M, Hendrix RW, Girones R, Wang D, Pipas JM.** 2011. Raw sewage harbors diverse viral populations. *MBio* **2**.
 75. **Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SW, Hibberd ML, Liu ET, Rohwer F, Ruan Y.** 2006. RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biol* **4**:e3.
 76. **Fauquet CM, Fargette D.** 2005. International Committee on Taxonomy of Viruses and the 3,142 unassigned species. *Virol J* **2**:64.
 77. **Alavandi SV, Poornima M.** 2012. Viral metagenomics: a tool for virus discovery and diversity in aquaculture. *Indian J Virol* **23**:88-98.
 78. **Bibby K, Viau E, Peccia J.** 2011. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. *Lett Appl Microbiol* **52**:386-392.
 79. **Li Y, Wang H, Nie K, Zhang C, Zhang Y, Wang J, Niu P, Ma X.** 2016. VIP: an integrated pipeline for metagenomics of virus identification and discovery. *Sci Rep* **6**:23774.
 80. **Edwards RA, Rohwer F.** 2005. Viral metagenomics. *Nat Rev Microbiol* **3**:504-510.
 81. **Berg G.** 1973. Removal of viruses from sewage, effluents and waters. 2. Present and future trends. *Bull World Health Organ* **49**:461-469.
 82. **Gerba CP.** 2007. Virus Occurrence and Survival in the Environmental Waters. In Bosch A (ed), *Human Viruses in Water* doi:10.1016/S0168-7069(07)17005-1. Elsevier B. V. All rights.
 83. **Bibby K, Peccia J.** 2013. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. *Environ Sci Technol* **47**:1945-1951.
 84. **Hrynyszyn A, Skonieczna M, Wiszniowski J.** 2013. Methods for Detection of Viruses in Water and Wastewater. *Advances in Microbiology* doi:10.4236/aim.2013.35060:7.

85. **Iaconelli M, Muscillo M, Della Libera S, Fratini M, Meucci L, De Ceglia M, Giacosa D, La Rosa G.** 2017. One-year Surveillance of Human Enteric Viruses in Raw and Treated Wastewaters, Downstream River Waters, and Drinking Waters. *Food Environ Virol* **9**:79-88.
86. **Iturriza-Gomara M, O'Brien SJ.** 2016. Foodborne viral infections. *Curr Opin Infect Dis* **29**:495-501.
87. **da Silva Polo T, Peiro JR, Mendes LC, Ludwig LF, de Oliveira-Filho EF, Bucardo F, Huynen P, Melin P, Thiry E, Mauroy A.** 2016. Human norovirus infection in Latin America. *J Clin Virol* **78**:111-119.
88. **Teixeira DM, Hernandez JM, Silva LD, Oliveira Dde S, Spada PK, Gurjao TC, Mascarenhas JD, Linhares AC, Morais LL, Gabbay YB.** 2016. Occurrence of Norovirus GIV in Environmental Water Samples from Belem City, Amazon Region, Brazil. *Food Environ Virol* **8**:101-104.
89. **Vieira CB, de Abreu Correa A, de Jesus MS, Luz SL, Wyn-Jones P, Kay D, Vargha M, Miagostovich MP.** 2016. Viruses Surveillance Under Different Season Scenarios of the Negro River Basin, Amazonia, Brazil. *Food Environ Virol* **8**:57-69.
90. **Gallimore CI, Appleton H, Lewis D, Green J, Brown DW.** 1995. Detection and characterisation of bisegmented double-stranded RNA viruses (picobirnaviruses) in human faecal specimens. *J Med Virol* **45**:135-140.
91. **Oude Munnink BB, Cotten M, Deijs M, Jebbink MF, Bakker M, Farsani SM, Canuti M, Kellam P, van der Hoek L.** 2015. A novel genus in the order Picornavirales detected in human stool. *J Gen Virol* **96**:3440-3443.
92. **Estes MK, Kapikian AZ.** 2007. Fields Virology, p 1917-1974. *In* Knipe RA, Howley DM, Griffin DE, Martin Lamb MA (ed), *Fields Virology*. Lippincott, Williams and Wilkins, Philadelphia.
93. **O'Ryan M, Lucero Y, Linhares AC.** 2011. Rotarix(R): vaccine performance 6 years postlicensure. *Expert Rev Vaccines* **10**:1645-1659.
94. **Kurtz JB, Lee TW, Pickering D.** 1977. Astrovirus associated gastroenteritis in a children's ward. *J Clin Pathol* **30**:948-952.
95. **Lee TW, Kurtz JB.** 1982. Human astrovirus serotypes. *J Hyg (Lond)* **89**:539-540.
96. **Kapikian AZ, Wyatt RG, Dolin R, Thornhill Ts, Kalica AR, Chanock RM.** 1972. VISUALIZATION BY IMMUNE ELECTRON-MICROSCOPY OF A 27-NM PARTICLE ASSOCIATED WITH ACUTE INFECTIOUS NONBACTERIAL GASTROENTERITIS. *Journal of Virology* **10**:1075-1081.
97. **Siebenga JJ, Lemey P, Kosakovsky Pond SL, Rambaut A, Vennema H, Koopmans M.** 2010. Phylodynamic reconstruction reveals norovirus GII.4 epidemic expansions and their molecular determinants. *PLoS Pathog* **6**:e1000884.
98. **Han TH, Kim SC, Kim ST, Chung CH, Chung JY.** 2014. Detection of norovirus genogroup IV, klassevirus, and pepper mild mottle virus in sewage samples in South Korea. *Arch Virol* **159**:457-463.
99. **Hamza IA, Jurzik L, Uberla K, Wilhelm M.** 2011. Evaluation of pepper mild mottle virus, human picobirnavirus and Torque teno virus as indicators of fecal contamination in river water. *Water Res* **45**:1358-1368.
100. **Kuroda K, Nakada N, Hanamoto S, Inaba M, Katayama H, Do AT, Nga TT, Oguma K, Hayashi T, Takizawa S.** 2015. Pepper mild mottle virus as

- an indicator and a tracer of fecal pollution in water environments: comparative evaluation with wastewater-tracer pharmaceuticals in Hanoi, Vietnam. *Sci Total Environ* **506-507**:287-298.
101. **Symonds EM, Sinigalliano C, Gidley M, Ahmed W, McQuaig-Ulrich SM, Breitbart M.** 2016. Faecal pollution along the southeastern coast of Florida and insight into the use of pepper mild mottle virus as an indicator. *J Appl Microbiol* **121**:1469-1481.
 102. **Song SI, Choi JN, Song JT, Ahn JH, Lee JS, Kim M, Cheong JJ, Choi YD.** 2002. Complete genome sequence of garlic latent virus, a member of the carlavirus family. *Mol Cells* **14**:205-213.
 103. **Wetter C, Conti M, Altshuh D, Tabillion R, Van Regenmortel MH.** 1984. Pepper mild mottle virus, a tobamovirus infecting pepper cultivars in Sicily. *Phytopathology*:5.
 104. **Rodriguez-Cerezo E, Moya A, Garcia-Arenal F.** 1989. Variability and evolution of the plant RNA virus pepper mild mottle virus. *J Virol* **63**:2198-2203.
 105. **Bonning BC.** 2009. The Dicistroviridae: An Emerging Family of Invertebrate Viruses. *Virologica Sinica* **24**:12.
 106. **Liu S, Vijayendran D, Carrillo-Tripp J, Miller WA, Bonning BC.** 2014. Analysis of new aphid lethal paralysis virus (ALPV) isolates suggests evolution of two ALPV species. *J Gen Virol* **95**:2809-2819.
 107. **de Miranda JR, Cordoni G, Budge G.** 2010. The Acute bee paralysis virus-Kashmir bee virus-Israeli acute paralysis virus complex. *J Invertebr Pathol* **103 Suppl 1**:S30-47.
 108. **Bosch A, Pinto RM, Guix S.** 2014. Human astroviruses. *Clin Microbiol Rev* **27**:1048-1074.
 109. **Xavier Mda P, Carvalho Costa FA, Rocha MS, Andrade Jda S, Diniz FK, Andrade TR, Miagostovich MP, Leite JP, Volotao Ede M.** 2015. Surveillance of Human Astrovirus Infection in Brazil: The First Report of MLB1 Astrovirus. *PLoS One* **10**:e0135687.
 110. **Patel MM, Hall AJ, Vinje J, Parashar UD.** 2009. Noroviruses: a comprehensive review. *J Clin Virol* **44**:1-8.
 111. **Miagostovich MP, Ferreira FF, Guimaraes FR, Fumian TM, Diniz-Mendes L, Luz SL, Silva LA, Leite JP.** 2008. Molecular detection and characterization of gastroenteritis viruses occurring naturally in the stream waters of Manaus, central Amazonia, Brazil. *Appl Environ Microbiol* **74**:375-382.
 112. **Zhang XF, Huang Q, Long Y, Jiang X, Zhang T, Tan M, Zhang QL, Huang ZY, Li YH, Ding YQ, Hu GF, Tang S, Dai YC.** 2015. An outbreak caused by GII.17 norovirus with a wide spectrum of HBGA-associated susceptibility. *Sci Rep* **5**:17687.
 113. **Silva LD, Bandeira RD, Junior EC, Lima IC, da Penha Junior ET, Teixeira DM, Siqueira JA, Resque HR, de Abreu Campos EM, Justino MC, Linhares AC, Gabbay YB.** 2016. Detection and genetic characterization of the emergent GII.17_2014 norovirus genotype among children with gastroenteritis from Northern Brazil. *Infect Genet Evol* **48**:1-3.
 114. **Di Bartolo I, Ponterio E, Battistone A, Bonomo P, Cicala A, Mercurio P, Triassi M, Pennino F, Fiore L, Ruggeri FM.** 2013. Identification and

- Genotyping of Human Sapoviruses Collected from Sewage Water in Naples and Palermo, Italy, in 2011. *Food Environ Virol* doi:10.1007/s12560-013-9124-2.
115. **Kitajima M, Oka T, Haramoto E, Katayama H, Takeda N, Katayama K, Ohgaki S.** 2010. Detection and genetic analysis of human sapoviruses in river water in Japan. *Appl Environ Microbiol* **76**:2461-2467.
 116. **Sano D, Perez-Sautu U, Guix S, Pinto RM, Miura T, Okabe S, Bosch A.** 2011. Quantification and genotyping of human sapoviruses in the Llobregat river catchment, Spain. *Appl Environ Microbiol* **77**:1111-1114.
 117. **Teixeira EW, Chen Y, Message D, Pettis J, Evans JD.** 2008. Virus infections in Brazilian honey bees. *J Invertebr Pathol* **99**:117-119.
 118. **Mendez E, Arias CF.** 2007. Astroviruses, p 981-1000. *In* Howley PM, Knipe DM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (ed), *Fields Virology*, vol I. Lippincott Williams & Wilkins, Philadelphia.
 119. **Oude Munnink BB, Cotten M, Canuti M, Deijs M, Jebbink MF, van Hemert FJ, Phan MV, Bakker M, Jazaeri Farsani SM, Kellam P, van der Hoek L.** 2016. A Novel Astrovirus-Like RNA Virus Detected in Human Stool. *Virus Evol* **2**:vew005.
 120. **Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ.** 2016. Redefining the invertebrate RNA virosphere. *Nature* doi:10.1038/nature20167.
 121. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**:3389-3402.
 122. **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A.** 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647-1649.
 123. **Conley M, Emmott E, Orton R, Taylor D, Carneiro DG, Murata K, Goodfellow IG, Hansman GS, Bhella D.** 2017. Vesivirus 2117 capsids more closely resemble sapovirus and lagovirus particles than other known vesivirus structures. *J Gen Virol* **98**:68-76.
 124. **Zeitler CE, Estes MK, Venkataram Prasad BV.** 2006. X-ray crystallographic structure of the Norwalk virus protease at 1.5-Å resolution. *J Virol* **80**:5050-5058.
 125. **Hogbom M, Jager K, Robel I, Unge T, Rohayem J.** 2009. The active form of the norovirus RNA-dependent RNA polymerase is a homodimer with cooperative activity. *J Gen Virol* **90**:281-291.
 126. **Zamyatkin DF, Parra F, Alonso JM, Harki DA, Peterson BR, Grochulski P, Ng KK.** 2008. Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J Biol Chem* **283**:7705-7712.
 127. **Tan M, Jiang X.** 2005. Norovirus and its histo-blood group antigen receptors: an answer to a historical puzzle. *Trends Microbiol* **13**:285-293.
 128. **Singh BK, Leuthold MM, Hansman GS.** 2015. Human noroviruses' fondness for histo-blood group antigens. *J Virol* **89**:2024-2040.
 129. **Fullerton SWB, Blaschke M, Coutard B, Gebhardt J, Gorbalenya A, Canard B, Tucker PA, Rohayem J.** 2007. Structural and functional

- characterization of sapovirus RNA-dependent RNA polymerase. *Journal of Virology* **81**:1858-1871.
130. **Kabsch W.** 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66**:125-132.
 131. **McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.** 2007. Phaser crystallographic software. *J Appl Crystallogr* **40**:658-674.
 132. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**:486-501.
 133. **Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**:213-221.
 134. **Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC.** 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**:12-21.
 135. **Muhire BM, Varsani A, Martin DP.** 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* **9**:e108277.
 136. **Liu W, Chen Y, Jiang X, Xia M, Yang Y, Tan M, Li X, Rao Z.** 2015. A Unique Human Norovirus Lineage with a Distinct HBGA Binding Interface. *PLoS Pathog* **11**:e1005025.
 137. **Chen R, Neill JD, Noel JS, Hutson AM, Glass RI, Estes MK, Prasad BV.** 2004. Inter- and intragenus structural variations in caliciviruses and their functional implications. *J Virol* **78**:6469-6479.
 138. **Miyazaki N, Taylor DW, Hansman GS, Murata K.** 2015. Antigenic and Cryo-Electron Microscopy Structure Analysis of a Chimeric Sapovirus Capsid. *J Virol* **90**:2664-2675.
 139. **Numata K, Hardy ME, Nakata S, Chiba S, Estes MK.** 1997. Molecular characterization of morphologically typical human calicivirus Sapporo. *Arch Virol* **142**:1537-1552.
 140. **Hansman GS, Katayama K, Oka T, Natori K, Takeda N.** 2005. Mutational study of sapovirus expression in insect cells. *Virology Journal* **2**:13-13.
 141. **Ravn V, Dabelsteen E.** 2000. Tissue distribution of histo-blood group antigens. *APMIS* **108**:1-28.
 142. **Zakhour M, Ruvoen-Clouet N, Charpilienne A, Langpap B, Poncet D, Peters T, Bovin N, Le Pendu J.** 2009. The alphaGal epitope of the histo-blood group antigen family is a ligand for bovine norovirus Newbury2 expected to prevent cross-species transmission. *PLoS Pathog* **5**:e1000504.
 143. **Zakhour M, Maalouf H, Di Bartolo I, Haugarreau L, Le Guyader FS, Ruvoen-Clouet N, Le Saux JC, Ruggeri FM, Pommepuy M, Le Pendu J.** 2010. Bovine norovirus: carbohydrate ligand, environmental contamination, and potential cross-species transmission via oysters. *Appl Environ Microbiol* **76**:6404-6411.

144. **Farkas T, Cross RW, Hargitt E, 3rd, Lerche NW, Morrow AL, Sestak K.** 2010. Genetic diversity and histo-blood group antigen interactions of rhesus enteric caliciviruses. *J Virol* **84**:8617-8625.
145. **Leuthold MM, Dalton KP, Hansman GS.** 2015. Structural analysis of a rabbit hemorrhagic disease virus binding to histo-blood group antigens. *J Virol* **89**:2378-2387.
146. **Tan M, Jiang X.** 2011. Norovirus-host interaction: multi-selections by human histo-blood group antigens. *Trends Microbiol* **19**:382-388.
147. **Neumann G, Kawaoka Y.** 2004. Reverse genetics systems for the generation of segmented negative-sense RNA viruses entirely from cloned cDNA. *Curr Top Microbiol Immunol* **283**:43-60.
148. **Sosnovtsev S, Green KY.** 1995. RNA transcripts derived from a cloned full-length copy of the feline calicivirus genome do not require VpG for infectivity. *Virology* **210**:383-390.
149. **Love DN.** 1973. The effect of DEAE-dextran on the infectivity of a feline calicivirus and its RNA. *Arch Gesamte Virusforsch* **41**:52-58.
150. **Flynn WT, Saif LJ.** 1988. Serial propagation of porcine enteric calicivirus-like virus in primary porcine kidney cell cultures. *J Clin Microbiol* **26**:206-212.
151. **Mitra T, Sosnovtsev SV, Green KY.** 2004. Mutagenesis of tyrosine 24 in the VPg protein is lethal for feline calicivirus. *J Virol* **78**:4931-4935.
152. **Sosnovtsev SV, Sosnovtseva SA, Green KY.** 1998. Cleavage of the feline calicivirus capsid precursor is mediated by a virus-encoded proteinase. *J Virol* **72**:3051-3059.
153. **Neill JD, Sosnovtsev SV, Green KY.** 2000. Recovery and altered neutralization specificities of chimeric viruses containing capsid protein domain exchanges from antigenically distinct strains of feline calicivirus. *J Virol* **74**:1079-1084.
154. **Sosnovtsev SV, Garfield M, Green KY.** 2002. Processing map and essential cleavage sites of the nonstructural polyprotein encoded by ORF1 of the feline calicivirus genome. *J Virol* **76**:7060-7072.
155. **Liu G, Zhang Y, Ni Z, Yun T, Sheng Z, Liang H, Hua J, Li S, Du Q, Chen J.** 2006. Recovery of infectious rabbit hemorrhagic disease virus from rabbits after direct inoculation with in vitro-transcribed RNA. *J Virol* **80**:6597-6602.
156. **Liu GQ, Ni Z, Yun T, Yu B, Zhu JM, Hua JG, Chen JP.** 2008. Rabbit hemorrhagic disease virus poly(A) tail is not essential for the infectivity of the virus and can be restored in vivo. *Arch Virol* **153**:939-944.
157. **Liu G, Ni Z, Yun T, Yu B, Chen L, Zhao W, Hua J, Chen J.** 2008. A DNA-launched reverse genetics system for rabbit hemorrhagic disease virus reveals that the VP2 protein is not essential for virus infectivity. *J Gen Virol* **89**:3080-3085.
158. **Chaudhry Y, Skinner MA, Goodfellow IG.** 2007. Recovery of genetically defined murine norovirus in tissue culture by using a fowlpox virus expressing T7 RNA polymerase. *J Gen Virol* **88**:2091-2100.
159. **Yunus MA, Chung LM, Chaudhry Y, Bailey D, Goodfellow I.** 2010. Development of an optimized RNA-based murine norovirus reverse genetics system. *J Virol Methods* **169**:112-118.

160. **Simmonds P, Karakasiliotis I, Bailey D, Chaudhry Y, Evans DJ, Goodfellow IG.** 2008. Bioinformatic and functional analysis of RNA secondary structure elements among different genera of human and animal caliciviruses. *Nucleic Acids Res* **36**:2530-2546.
161. **Chang KO, Sosnovtsev SV, Belliot G, King AD, Green KY.** 2006. Stable expression of a Norwalk virus RNA replicon in a human hepatoma cell line. *Virology* **353**:463-473.
162. **Chang KO, George DW.** 2007. Interferons and ribavirin effectively inhibit Norwalk virus replication in replicon-bearing cells. *J Virol* **81**:12111-12118.
163. **Chang KO.** 2009. Role of cholesterol pathways in norovirus replication. *J Virol* **83**:8587-8595.
164. **Katayama K, Hansman GS, Oka T, Ogawa S, Takeda N.** 2006. Investigation of norovirus replication in a human cell line. *Arch Virol* **151**:1291-1308.
165. **dos Anjos K, Lima LMP, Silva PA, Inoue-Nagata AK, Nagata T.** 2011. The possible molecular evolution of sapoviruses by inter- and intra-genogroup recombination. *Archives of Virology* **156**:1953-1959.
166. **Ward VK, McCormick CJ, Clarke IN, Salim O, Wobus CE, Thackray LB, Virgin HWt, Lambden PR.** 2007. Recovery of infectious murine norovirus using pol II-driven expression of full-length cDNA. *Proc Natl Acad Sci U S A* **104**:11050-11055.
167. **Liu GQ, Ni Z, Yun T, Yu B, Zhu JM, Hua JG, Chen JP.** 2007. [Construction of rabbit hemorrhagic disease virus replicons and its replication in RK-13 cells]. *Bing Du Xue Bao* **23**:481-484.
168. **Wei C, Farkas T, Sestak K, Jiang X.** 2008. Recovery of infectious virus by transfection of in vitro-generated RNA from tulane calicivirus cDNA. *J Virol* **82**:11429-11436.
169. **Abente EJ, Sosnovtsev SV, Bok K, Green KY.** 2010. Visualization of feline calicivirus replication in real-time with recombinant viruses engineered to express fluorescent reporter proteins. *Virology* **400**:18-31.
170. **Chiba S, Nakata S, Numata-Kinoshita K, Honma S.** 2000. Sapporo virus: history and recent findings. *The Journal of Infectious Diseases* **181 Suppl 2**:S303-308-S303-308.
171. **Stovesand P, Behrens F, Majer M, Maass G.** 1977. [Seroepidemiological Investigations on the epidemiology of human rotavirus infections (author's transl)]. *Zentralbl Bakteriol Orig A* **238**:16-19.
172. **Lee TW, Kurtz JB.** 1977. Astroviruses detected by immunofluorescence. *Lancet* **2**:406.
173. **Ferreira MS, Xavier MP, Fumian TM, Victoria M, Oliveira SA, Pena LH, Leite JP, Miagostovich MP.** 2008. Acute gastroenteritis cases associated with noroviruses infection in the state of Rio de Janeiro. *J Med Virol* **80**:338-344.
174. **Victoria M, Guimaraes FR, Fumian TM, Ferreira FF, Vieira CB, Shubo T, Leite JP, Miagostovich MP.** 2010. One year monitoring of norovirus in a sewage treatment plant in Rio de Janeiro, Brazil. *J Water Health* **8**:158-165.
175. **Vieira CB, Mendes AC, Guimaraes FR, Fumian TM, Leite JP, Gaspar AM, Miagostovich MP.** 2012. Detection of enteric viruses in recreational

- waters of an urban lagoon in the city of Rio de Janeiro, Brazil. *Mem Inst Oswaldo Cruz* **107**:778-784.
176. **Fumian TM, Vieira CB, Leite JP, Miagostovich MP.** 2013. Assessment of burden of virus agents in an urban sewage treatment plant in Rio de Janeiro, Brazil. *J Water Health* **11**:110-119.
177. **Chen H, Qian F, Xu J, Chan M, Shen Z, Zai S, Shan M, Cai J, Zhang W, He J, Liu Y, Zhang J, Yuan Z, Zhu Z, Hu Y.** 2015. A novel norovirus GII.17 lineage contributed to adult gastroenteritis in Shanghai, China, during the winter of 2014-2015. *Emerg Microbes Infect* **4**:e67.
178. **Gao Z, Liu B, Huo D, Yan H, Jia L, Du Y, Qian H, Yang Y, Wang X, Li J, Wang Q.** 2015. Increased norovirus activity was associated with a novel norovirus GII.17 variant in Beijing, China during winter 2014-2015. *BMC Infect Dis* **15**:574.
179. **Koromyslova AD, Leuthold MM, Bowler MW, Hansman GS.** 2015. The sweet quartet: Binding of fucose to the norovirus capsid. *Virology* **483**:203-208.

