

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
DEPARTAMENTO DE BIOLOGIA

Adriana de Oliveira Fernandes

**IDENTIFICAÇÃO E CARACTERIZAÇÃO FUNCIONAL DE GENES VIRAIS
RELACIONADOS COM A PAPILOMATOSE BOVINA**

SÃO CRISTÓVÃO - SE

2018.2

ADRIANA DE OLIVEIRA FERNANDES

IDENTIFICAÇÃO E CARACTERIZAÇÃO FUNCIONAL DE GENES VIRAIS
RELACIONADOS COM A PAPILOMATOSE BOVINA

Trabalho de conclusão de curso apresentado ao curso de Ciências Biológicas Bacharelado da Universidade Federal de Sergipe como parte dos requisitos exigidos para obtenção do título de Graduação em Ciências Biológicas.

Orientador: Prof. Dr. Marcus Vinicius de Aragão Batista

SÃO CRISTÓVÃO - SE

2018.2

RESUMO

IDENTIFICAÇÃO E CARACTERIZAÇÃO FUNCIONAL DE GENES VIRAIS RELACIONADOS COM A PAPILOMATOSA BOVINA

O papilomavírus bovino (BPV) é o vírus causador da doença infectocontagiosa conhecida como papilomatose, que acomete principalmente bovinos. A doença é conhecida por formar verrugas benignas em tecidos epiteliais, porém, a depender das condições, tais formações podem se tornar cancerosas. É uma enfermidade de distribuição mundial e que causa sérios danos, especialmente aos rebanhos do Brasil, pois pode levar à perda de centenas de cabeças de gado num país onde a produção de carne e leite é o cerne da economia agropecuária. Sabe-se que um mesmo animal pode estar acometido não só pelo BPV, mas também por outros microrganismos, e isto pode influenciar na evolução da doença. Sendo assim, este trabalho busca identificar e caracterizar as funções e processos biológicos de genes virais expressos em lesões papilomatosas. Para tanto, foi utilizada a metodologia de RNA-seq para sequenciar RNA mensageiro de seis animais (bovinos) jovens, sendo três infectados e três não-infectados por BPV. Com o uso da plataforma Galaxy, foram separadas aquelas sequências que não pertenciam ao hospedeiro, e com o programa BLASTn foram identificados os genes virais expressos nas lesões. A anotação funcional dos genes destes vírus foi feita com os bancos de dados do Gene Ontology, UniProtKB, KEGG e ViruSite, bem como buscas em artigos científicos. No total foram encontradas 18 famílias virais, sendo oito de relevância clínica, tendo vários genes relacionados com a progressão oncogênica, supressão do sistema imune e proliferação celular. Devido à falta de dados anotados para a maior parte dos genes, são necessários mais esforços da comunidade acadêmica para organizar as informações funcionais de genes virais, facilitando os estudos de anotação. Entretanto, este é o primeiro trabalho que buscou averiguar a possível influência de outros genes virais no desenvolvimento das lesões papilomatosas como um co-fator associado à infecção por BPV.

Palavras-chave: Papilomavírus bovino, RNA-seq, Anotação funcional, Coinfecção

ABSTRACT

IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF VIRUS GENES RELATED TO BOVINE PAPILOMATOSIS

Bovine papillomavirus (BPV) is the virus that causes the infectious-contagious disease known as papillomatosis, which mainly affects cattle. The disease is known to form benign warts in epithelial tissues, however, depending on the conditions, such formations may become cancerous. It is a disease of worldwide distribution and it causes serious damage, especially to Brazil's herds, as it can lead to the loss of hundreds of cattle in a country where meat and milk production is at the heart of the agricultural and livestock economy. It is known that the same animal may be affected not only by BPV, but also by other microorganisms, and this may influence the evolution of the disease. Thus, this work seeks to identify and characterize the biological functions and processes of viral genes expressed in papillomatous lesions. For this, the RNA-seq methodology was used to sequence messenger RNA from six young animals (bovine), three infected and three non-infected by BPV. With the use of the Galaxy platform, those non-host sequences were separated, and the viral genes expressed in the lesions were identified with the BLASTn program. The functional annotation of the genes of these viruses was done with the databases of Gene Ontology, UniProtKB, KEGG and ViruSite, as well as searches in scientific articles. In total eighteen viral families were found, eight of them being of clinical relevance, with several genes related to oncogenic progression, immune system suppression and cell proliferation. Due to the lack of annotated data for most genes, more efforts are required from the academic community to organize the functional information of viral genes, facilitating annotation studies. However, this is the first study to investigate the possible influence of other viral genes on the development of papillomatous lesions as a cofactor associated with BPV infection.

Keywords: Bovine papillomavirus, RNA-seq, Functional annotation, Coinfection

LISTA DE ABREVIATURAS

pb	Pares de bases
PV	Papilomavírus
BPV	Papilomavírus bovino
HPV	Papilomavírus humano
ORF	Matriz de leitura aberta (<i>Open Reading Frame</i>)
DNA	Ácido desoxirribonucleico (<i>Deoxyribonucleic Acid</i>)
RNA	Ácido ribonucleico (<i>Ribonucleic Acid</i>)
RNA-seq	Tecnologia de sequenciamento de RNA de nova geração
cDNA	Ácido desoxirribonucleico complementar (<i>Complementary DNA</i>)
GO	<i>Gene Ontology</i>
UniProtKB	<i>Universal Protein Knowledgebase</i>
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
RefSeq	Sequência de referência (<i>Reference Sequence</i>)
NCBI	<i>National Center for Biotechnology Information</i>
SRC	Proto-oncogene da família tirosina-quinase
EGF-R	Receptor de fator de crescimento epidérmico (<i>Epidermal Growth Factor Receptor</i>)
mb	Megabyte
kb	Kilobyte
GnRH	Hormônio liberador de gonadotrofina (<i>Gonadotrofin-releasing Hormone</i>)
ErbB	Receptor de tirosina quinase, estruturalmente relacionado ao EGF-R
MAPK	Proteína-quinases ativadas por mitógenos (<i>Mitogen Activated Protein Kinases</i>)
Ras1	Proteína relacionada a Ras 1 (<i>Ras-related protein 1</i>)
P53	Proteína supressora de tumor
mTOR	Complexo proteico que contém TOR (<i>Target of Rapamycin</i>) e que media a fosforilação e ativação de sinalizadores
STAT	Transdutores de sinal e fatores de transcrição ativados por citocinas e fatores de crescimento
BHV	Herpesvírus bovino
dsDNA	DNA de fita dupla

LISTA DE FIGURAS

Figura 1. Representação esquemática do genoma do BPV22 (Fonte: Bauermann <i>et al.</i> , 2017).....	11
Figura 2. Árvore filogenética do papilomavírus baseada no gene L1 (Fonte: Bauermann <i>et al.</i> , 2017).....	13
Figura 3. Quantidade, em porcentagem, de organismos encontrados por família viral em ambos os tipos de bibliotecas.....	26
Figura 4. Quantificação dos processos biológicos anotados pelo GO.....	31
Figura 5. Quantificação dos componente celulares anotados pelo GO.....	32
Figura 6. Quantificação das funções moleculares anotadas pelo GO.....	33
Figura 7. Adaptação de parte do mapa metabólico fornecido pelo KEGG, evidenciando a relação entre o gene SRC com o receptor EGF para promover a proliferação celular durante o câncer de bexiga.....	34
Figura 8. Adaptação de recorte da via de infecção do citomegalovírus humano, ilustrando a participação do gene SRC na regulação do citoesqueleto de actina e recorte adaptado do mesmo mapa de infecção do citomegalovírus humano, desta vez mostrando como o gene UL38 interage indiretamente com a mTOR para promover a replicação viral.....	34 e 35
Figura 9. Mapa esquemático simplificado (adaptado) fornecido pelo KEGG, ilustrando as interações da proteína E6 com a proteína p53 e o fator regulador de interferon 3 (IRF3).....	35

LISTA DE TABELAS

Tabela 1. Quantidade de <i>reads</i> totais e após o corte do Trimmomatic.....	22
Tabela 2. Tamanho dos arquivos FASTA e quantidade de sequências após separação.....	23
Tabela 3. Resultados do BLASTn, mostrando os melhores <i>hits</i> para cada organismo viral com base no índice de identidade e <i>e-value</i>	24
Tabela 4. Símbolos dos genes obtidos para cada organismo viral dentre as famílias encontradas pelo BLASTn.....	27
Tabela 5. Símbolos dos genes obtidos e suas respectivas funções segundo os bancos e artigos.....	28

SUMÁRIO

1 INTRODUÇÃO.....	9
2 OBJETIVOS.....	17
2.1 Objetivo geral.....	17
2.2 Objetivos específicos.....	17
3 MATERIAL E MÉTODOS.....	18
3.1 Coleta do material e controle de qualidade.....	18
3.2 Identificação das sequências do hospedeiro.....	19
3.3 Identificação das sequências de RNA de outros organismos.....	19
3.4 Anotação funcional.....	20
4 RESULTADOS.....	22
4.1 Controle de qualidade, quantificação dos <i>reads</i> e mapeamento dos transcritos.....	22
4.2 Alinhamento local usando o BLASTn.....	24
4.3 Anotação funcional dos genes identificados.....	27
5 DISCUSSÃO.....	36
6 CONSIDERAÇÕES FINAIS	41
7 REFERÊNCIAS.....	42

1 INTRODUÇÃO

A família *Papillomaviridae* é conhecida por agrupar organismos virais não envelopados, com DNA circular de fita dupla, compostos por aproximadamente 8000 pb (da Silva *et al.*, 2015). Estes organismos são conhecidos por vírus do papiloma ou apenas papilomavírus (PV), sendo capazes de infectar uma ampla gama de amniotas no mundo inteiro, desde humanos e outros mamíferos até aves, répteis e peixes (López-Bueno *et al.*, 2016; Rector e Van Ranst, 2013). Estes vírus possuem esse nome em decorrência do principal sintoma que causam no hospedeiro, que é o crescimento de verrugas, pólipos, calos e/ou condilomas, que, em termos técnicos, são caracterizados como tumores epiteliais benignos constituídos por hipertrofia ou neoformação de papilas (Daudt *et al.*, 2017; Araldi *et al.*, 2015).

O papilomavírus bovino (BPV) é um tipo de PV que infecta preferencialmente indivíduos da subfamília *Bovidae*, que compreende os gêneros *Bubalus*, *Bison* e *Bos*, este último sendo o gênero dos bois domésticos, no entanto, segundo Daudt *et al.* (2017), já foram detectadas infecções cruzadas entre o BPV e outros animais além dos bovinos, como antílopes, cavalos e girafas. A infecção por BPV pode muitas vezes ser assintomática, porém, quando há a manifestação de sintomas, geralmente há o crescimento de tumores benignos no epitélio cutâneo ou mucoso e que tendem a regredir à normalidade (Claus *et al.*, 2007). Entretanto, a depender do tipo de BPV e se houver a interferência de cofatores ambientais (ingestão de broto de samambaia do campo, por exemplo) (Sharma *et al.*, 2012; Roperto *et al.*, 2015), tais crescimentos podem se tornar malignos, passando a ser chamados, então, de papilomatoses ou fibropapilomas malignos (Claus *et al.*, 2007; Hamad *et al.*, 2018).

A samambaia do campo, nome popular de *Pteridium aquilinum* v. *arachnoideum*, é conhecida por seus efeitos imunossupressores, mutagênicos e

carcinogênicos devido ao flavonóide quercetina, que permite que os papilomas neoplásicos benignos se tornem malignos (Sharma *et al.*, 2013). Por ser uma enfermidade transmitida preferencialmente por contato físico, ordenha e/ou amamentação, as zonas mais atingidas pelas lesões são a cabeça, vulva, úbere e tetos (Savini *et al.*, 2016).

A evolução destes tumores até um quadro canceroso pode agravar as perdas físicas e econômicas na agropecuária não somente no Brasil, mas em várias outras regiões como Itália (Savini *et al.*, 2016), Iraque (Hamad *et al.*, 2018), Estados Unidos (Bauermann *et al.*, 2017), Japão (Hatama *et al.*, 2008) e Alemanha (Schmitt *et al.*, 2010), onde o vírus foi encontrado e documentado. No Brasil já foram relatados casos de infecção por BPV em diversos estados, como São Paulo (Módolo *et al.*, 2017), Paraná (Claus *et al.*, 2007), Rio de Janeiro (Torres *et al.*, 2007), Espírito Santo (Dias *et al.*, 2012), Pernambuco (Diniz *et al.*, 2009; Batista *et al.*, 2013), Santa Catarina (Silva *et al.*, 2010), Goiás (Torres *et al.*, 2007), Rondônia (Torres *et al.*, 2007) e na região da Amazônia (Daudt *et al.*, 2016; da Silva *et al.*, 2015; da Silva *et al.*, 2016).

Entre os prejuízos acarretados pela papilomatose bovina, pode-se citar a depreciação estética do animal, bem como de seu couro, o que dificulta sua comercialização em feiras e exposições (Torres *et al.*, 2007). A enfermidade também altera a capacidade reprodutiva dos indivíduos quando infecta o pênis ou a vulva, interferindo no acasalamento e disseminando o vírus (Batista *et al.*, 2013).

No Japão, o BPV-6 tem sido o maior causador de papilomatose de tetos, levando a perdas severas de gado leiteiro (Watanabe *et al.*, 2017). No Brasil a situação é a mesma, onde o animal que é acometido pela doença, principalmente nos tetos, perderá grande parte da sua produção leiteira e ficará suscetível a

ordenhas dolorosas e possível contração de mastite (Torres *et al.*, 2007). A presença de feridas abertas também pode levar a um quadro de miíase, que não só é uma condição estressante para o animal, mas ainda o desvaloriza comercialmente e o deixa vulnerável a outras doenças (Torres *et al.*, 2007).

O genoma do BPV é dividido em duas partes: precoce, apresentando cinco ou seis matrizes abertas de leitura (em inglês: *open reading frame* - ORF) – E1, E2, E4, E5, E6 e E7, e tardio, apresentando duas ORF – L1 e L2 (Daudt *et al.*, 2016) (Figura 1). Os genes precoces são responsáveis por mediar todos os estágios do ciclo de replicação dentro da célula hospedeira, enquanto que L1 e L2 são proteínas de capsídeo que estão envolvidas com a entrada ou saída do vírus da célula hospedeira (Culleton *et al.*, 2016).

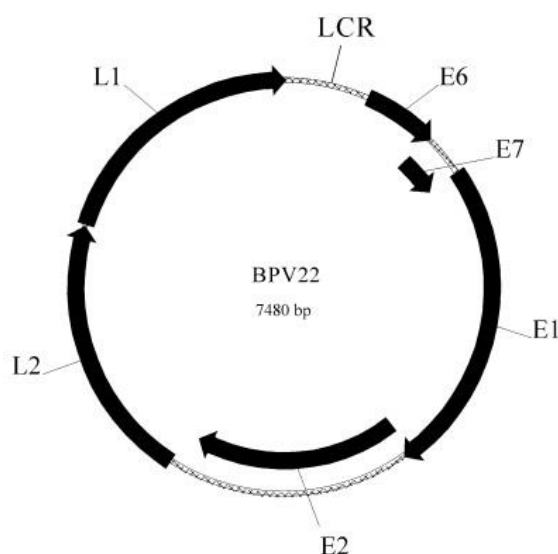


Figura 1. Representação esquemática do genoma do BPV22. Fonte: Bauermann et al (2017)

O gene E1 é responsável por codificar a maior proteína viral, que possui função de helicase (Bergvall *et al.*, 2013; Daudt *et al.*, 2016). A proteína E2 tem papel importante na transcrição e replicação do genoma viral (Culleton *et al.*, 2016), enquanto que a E4, apesar de precoce, se expressa de forma tardia nos estágios da infecção e está relacionada com a transmissão do vírus pela formação de um

envelope cornificado, em outras palavras atuando como um facilitador do brotamento do vírus na célula (Jareborg *et al.*, 1991; Doorbar, 2013). A proteína E5 é conhecida pela indução de defeitos no ciclo celular devido à obstrução dos mecanismos de supressão (Dimaio *et al.*, 2013; Karabadzhak *et al.*, 2017). E6 e E7 agem de forma similar, sendo proteínas que estimulam a proliferação e sobrevivência celular e modulam a diferenciação de queratinócitos, logo, tanto a E5 quanto E6 e E7 possuem potencial oncogênico (Pol *et al.*, 2013).

L1 e L2 são proteínas formadoras de capsídeo e são de suma importância no ciclo infeccioso, pois interagem diretamente com as proteoglicanas da membrana plasmática de células próximas à membrana basal (Buck *et al.*, 2013; Culp *et al.*, 2005; McKinney *et al.*, 2015).

O ciclo de infecção começa com a abertura de uma porta de entrada para o vírus, que seria uma lesão ao epitélio cutâneo ou mucoso escamoso, que irá expor a camada de células basais (Pol *et al.*, 2013). Diferentemente do papilomavírus humano (HPV), a contaminação por BPV ocorre principalmente por meio de contato físico com um indivíduo afetado, logo, o vírus contido nas verrugas ou papilas entrará pela lesão exposta. Uma vez dentro das células, inicia-se a replicação do DNA viral como plasmídeo (Pol *et al.*, 2013). Quando as células infectadas se dividem, elas empurram as demais lateralmente na membrana basal ou para cima em direção à camada de células espinhosas (Pol *et al.*, 2013). Quando afetadas pelo vírus, estas células irão reentrar, incomumente, no ciclo celular para amplificar o DNA viral até um alto número de cópias (Pol *et al.*, 2013). Em pouco tempo haverá a formação de condilomas epiteliais de aspecto rugoso e de tamanho variável (Torres *et al.*, 2007).

Em geral se utiliza a ORF L1 para classificar os PV, e estes são considerados de um tipo diferente se sua ORF L1 tiver menos de 90% de similaridade com outro PV (Munday, 2014; de Villiers *et al.*, 2004; Bernard *et al.*, 2010; de Villiers, 2013). Dentro da família em questão, cinco gêneros estão ligados ao BPV (*Deltapapillomavirus*, *Xipapillomavirus*, *Epsilonpapillomavirus*, *Dyokappapapillomavirus* e *Dyoxipapillomavirus*), sendo que a similaridade da ORF L1 entre eles é maior que 60% (Bauermann *et al.*, 2017) (Figura 2).

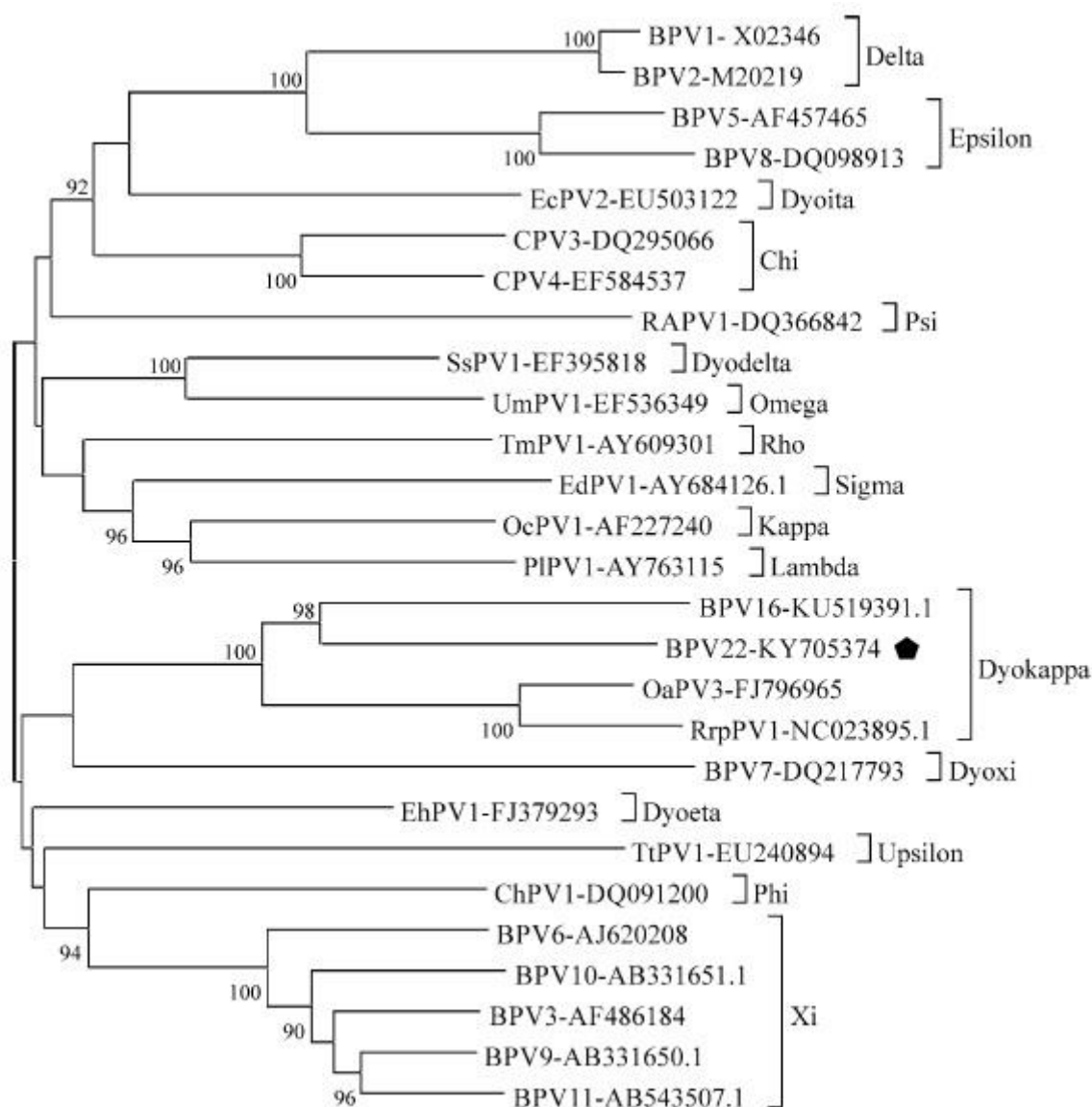


Figura 2. Árvore filogenética do papilomavírus baseada no gene L1. Fonte: Bauermann et al (2017).

Apesar do crescente número de estudos feitos acerca dos BPV, ainda há muito que não foi respondido, especialmente no que tange suas vias de infecção e a razão de alguns tipos induzirem mutações cancerosas, enquanto outros restringem-se à tumores benignos. Ainda não se conhece muito bem as funções de todas as proteínas virais, nem quais são os fatores que poderiam interferir na infecção, como a expressão de tais proteínas, as interações destas com o hospedeiro e possíveis coinfeções. Nesse sentido, a tecnologia do RNA-seq tem se mostrado relevante para a identificação de possíveis marcadores envolvidos em infecções virais.

O RNA-seq é uma ferramenta de sequenciamento que oferece inúmeras vantagens sobre outras tecnologias existentes. Seu funcionamento consiste em converter uma população de RNA a um conjunto de fragmentos de cDNA com adaptadores ligados a um ou ambos os lados. Cada molécula é, então, sequenciada para se obter sequências mais curtas de cada final. Em seguida, qualquer tecnologia de sequenciamento pode ser utilizada, a exemplo da Illumina ou SOLiD. Após o sequenciamento, os resultados podem ser alinhados ao transcrito ou genoma de referência ou ainda ser agregados novamente sem a sequência genômica para produzir um mapa de transcrição em escala genômica, que consiste na estrutura transcricional ou no nível de expressão de cada gene (Wang *et al.*, 2009). O RNA-seq pode ser uma técnica atrativa para ser utilizada com organismos não-modelo, onde as sequências genômicas ainda não foram determinadas. Ela pode indicar a localização exata dos limites de transcrição em até 1 pb, e além disso pode mostrar a conectividade entre múltiplos éxons e informações sobre como estes estão conectados, a depender do tamanho da leitura (Wang *et al.*, 2009).

Numa lesão por BPV, ainda mais considerando o fato de os hospedeiros serem animais que, em geral, vivem em regime de confinamento e nem sempre se

pode ter total controle sanitário sobre eles, é de se esperar que outros microrganismos oportunistas estejam presentes nela, às vezes sendo mascarados pela patologia viral. É possível encontrar microrganismos como os do gênero *Staphylococcus*, causadores da mastite bovina (Sartori *et al.*, 2018), interagindo com o vírus numa coinfeção. Em humanos, num caso de infecção por HPV, já foram reportadas coinfeções no tecido intraepitelial cervical com *Chlamydia* (Ssedyabane *et al.*, 2019), tumores de faringe e laringe com os poliomavírus BK e *Epstein-Barr* (Drop *et al.*, 2017) ou até em adenocarcinomas gástricos com o vírus *Epstein-Barr* e o procaríoto *Helicobacter pylori* (de Souza *et al.*, 2018).

Sabe-se que existem relações entre o HPV e os patógenos citados, como afirma Pietro *et al.* (2017) em relação à *Chlamydia trachomatis*, em que a infecção por clamídia pode facilitar a entrada do HPV ou ainda debilitar o sistema imune, favorecendo a persistência do vírus.

Adicionalmente, segundo Drop *et al.* (2017) e de Souza *et al.* (2018), percebe-se que o HPV também pode interagir com outros vírus, como é o caso do vírus *Epstein-Barr* e citomegalovírus (Ali *et al.*, 2019), que são classificados como pertencentes à família *Herpesviridae*, e do poliomavírus de células de Merkel, pertencente à família *Polyomaviridae* (Vazquez-Guillen *et al.*, 2018). Sabendo-se que os PV compartilham diversas características morfofuncionais, é possível estabelecer relações entre o HPV e o BPV. Além disso, como afirma Yagui *et al.* (2008) e Carvalho *et al.* (2013), já foram detectadas coinfeções não só de um, mas de diversos tipos de BPV, com relatos de até cinco tipos virais circulando na corrente sanguínea de um mesmo animal, além da presença de outros papilomavírus, a exemplo do felino (da Silva *et al.*, 2012) em verrugas cutâneas de bovinos. Já foi relatado também, num estudo mais antigo de Goldstein *et al.* (1987), que o

citomegalovírus humano (um herpesvírus) pode aumentar o potencial carcinogênico do BPV em um ensaio *in vitro*. O próprio herpesvírus bovino (BHV) já foi relacionado com diversas enfermidades em ruminantes, como vulvovaginite, encefalite, pneumonia e mastite (Cvetojevic *et al.*, 2016), esta última sendo um ponto de interesse para estudos com o BPV, pois, como já foi dito, o papilomavírus pode atingir e causar lesões nos tetos de bovinos fêmeas, o que poderia favorecer uma coinfeção entre os dois vírus. Porém, poucos estudos até agora foram realizados a fim de estabelecer uma relação firme entre a infecção por BPV e outros vírus que, possivelmente, viriam a realçar a manifestação papilomatosa.

Tendo isso em mente, é de grande importância que se possa traçar uma relação entre o BPV e outros vírus que possam estar correlacionados com ele, alterando de alguma maneira o desenvolvimento e a progressão da doença, até mesmo pelo fato de haverem poucos estudos que abordem esse tema. Nesse sentido, a tecnologia de RNA-seq pode ser de suma importância, pois não apenas serve para mostrar quais organismos estão presentes nas lesões, mas também pode identificar quais genes do hospedeiro estão sendo expressos e regulados e de que forma interagem com o BPV, sendo assim pode-se evidenciar como esses vírus oportunistas participam do ciclo infeccioso do papilomavírus, alterando ou não a manifestação patológica.

2 OBJETIVOS

2.1 Objetivo geral

O objetivo deste estudo foi identificar e caracterizar funcionalmente transcritos de genes virais que podem estar envolvidos no desenvolvimento da papilomatose bovina.

2.2 Objetivos específicos

- Obter os transcritos de lesões papilomatosas causadas por BPV e de tecidos não infectados;
- Isolar os transcritos dos vírus;
- Identificar os genes encontrados nas lesões papilomatosas e em tecidos não infectados pelo BPV;
- Realizar a anotação funcional dos genes virais identificados.

3 MATERIAL E MÉTODOS

3.1 Coleta do material e controle de qualidade

O presente estudo se utilizou de um banco de dados local do Laboratório de Genética Molecular e Biotecnologia (GMBio) com sequências oriundas do transcriptoma comparativo de bovinos infectados ou não com BPV. A coleta foi realizada em seis bovinos fêmeas da raça Girolando, um gado de tipo leiteiro de até 2 anos de idade (Barreto *et al.*, 2018). Foram obtidas três amostras infectadas por BPV e três não infectadas por BPV, cada uma proveniente de um animal diferente. O projeto foi aprovado no Comitê de Ética em Pesquisa com Animais de Produção da Universidade Federal de Sergipe (Protocolo: 05/14).

O sequenciamento foi realizado utilizando-se diversas ferramentas para cada etapa, tendo a confecção de bibliotecas feita a partir de 250 ng de RNA total usando o TruSeq Stranded mRNA Sample Preparation LS Protocol Kit (Illumina Inc.), a quantificação delas com o Library Quantification Kit (Illumina Inc.) e a avaliação do tamanho do fragmento com o Bioanalyzer 2100 (Agilent Technologies) (Barreto *et al.*, 2018). Também foi feita a formação de clusters na célula de fluxo com o cBot Instrument (Illumina Inc.) e as bibliotecas foram sequenciadas de modo multiplex no HiSeq 2000 System (Illumina Inc.) (Barreto *et al.*, 2018).

As sequências de RNA obtidas pelo sequenciamento foram submetidas a uma análise de qualidade utilizando-se a ferramenta FastQC versão 0.11.7 (Andrews, 2010) que avaliou a qualidade das sequências a depender do valor de Phred. O FastQC funciona seguindo o referido algoritmo, Phred, que é uma unidade de medida da probabilidade de uma dada base estar incorreta. O valor de Phred é determinado utilizando vários preditores de possíveis erros como o espaçamento, a taxa e a resolução do pico (Liao *et al.*, 2017). Em termos equacionais, o valor de Phred é definido como $Q = -10\log_{10}Pr$ (alelo observado \neq alelo verdadeiro). Logo, um valor $Q = 30$ corresponde a uma taxa de erro de 0.1% (Liao *et al.*, 2017). Portanto, neste estudo, foram aceitas sequências com valores de Phred iguais ou maiores que 30.

3.2 Identificação das sequências do hospedeiro

Tendo em mãos as sequências obtidas após o FastQC, estas foram submetidas ao software Trimmomatic versão 0.38 (Bolger *et al.*, 2014), que serviu para cortar as sequências de adaptadores, homopolímeros e *reads* de baixa qualidade, que por ventura poderiam interferir no sequenciamento e também para diminuir o peso dos arquivos. Isto é feito pois, em geral, dados de sequenciamento de nova geração (NGS) estão na forma de sequências pareadas, onde a *forward* e a *reverse* estão em dois arquivos separados de formato FastQ (Bolger *et al.*, 2014). Muitas outras ferramentas utilizam essa relação de posição entre as sequências, então é adequado que ela seja mantida (Bolger *et al.*, 2014). Após os cortes e tendo em mãos as sequências de boa qualidade, utilizando-se a ferramenta TopHat versão 2.1.1 (Trapnell *et al.*, 2013) as sequências foram alinhadas e mapeadas com o genoma de referência (UMD3.1 *Bos taurus* 8), onde foi possível diferenciar o RNA bovino de outras sequências ali presentes.

3.3 Identificação das sequências de RNA de outros organismos

Tendo em mãos os arquivos alinhados pelo TopHat, estes foram utilizados na plataforma Galaxy (<https://usegalaxy.org/>), onde, com o pacote ChIP-seq e a ferramenta BAM filter versão 0.5.9 que pertence a ele (Mendonza-Parra *et al.*, 2016), foram excluídas todas as sequências provenientes do bovino, restando apenas aquelas que não foram mapeadas, ou seja, que pertencem a outros organismos desconhecidos. Foram gerados seis arquivos em formato BAM (formato padrão do TopHat), nomeados simplificarmente de query1, query2, query5, query6, query7 e query10, sendo os três primeiros infectados com BPV, e os três últimos os não infectados. Cada arquivo tem entre 146 a 204 mb de tamanho, e contém cerca de quatro milhões de sequências.

Sendo assim, os seis arquivos em formato BAM advindos do TopHat foram convertidos para FASTA utilizando a ferramenta *Convert, Merge, Randomize*, contido no pacote BAMtools versão 2.4.0 (Masella *et al.*, 2016) da plataforma Galaxy e assim utilizados no BLASTn local, pois é o único formato aceito pelo programa. Mesmo após a conversão, o número de sequências foi mantido, porém, o tamanho

do arquivo diminuiu consideravelmente, já que se trata de uma extensão compatível com .txt.

A identificação dos organismos foi realizada a partir de buscas utilizando a plataforma BLAST (*Basic Local Alignment Search Tool*) (Altschul *et al.*, 1990), criada e mantida pelo *National Center for Biotechnology Information* (NCBI). O algoritmo Nucleotídeo BLAST compara as sequências de nucleotídeos com as depositadas no GenBank do NCBI (<https://www.ncbi.nlm.nih.gov/>) e calcula a significância estatística dos correspondentes com base na similaridade (Altschul *et al.*, 1990).

A plataforma BLASTn na web tem um limite de aproximadamente 100 sequências por busca e, sabendo que as bibliotecas geradas pelo TopHat possuem cerca de quatro milhões delas, optou-se pelo uso do Standalone BLASTn (versão 2.7.1+), programa de função idêntica ao BLASTn web, porém com o benefício de poder ser realizado por meio do *prompt* de comando do Windows, logo podendo aceitar arquivos maiores e sem dependência de conexão com a internet. A principal diferença entre os dois métodos é que, no programa local, é necessário ter também os bancos de dados com sequências de referência específicos para a sua busca. Neste caso foi utilizado o banco de dados RefSeq (*Reference Sequence*, em português, Sequências de Referência) para vírus, próprio para uso no BLASTn, encontrado na página FTP do NCBI (<ftp://ftp.ncbi.nlm.nih.gov/blast/db/>).

Assim, o trabalho de identificação foi feito em duas etapas: primeiro fez-se o BLASTn com o banco de dados de RefSeq viral, e depois foram identificadas as famílias de organismos prevalentes para serem analisadas na próxima fase. Vale ressaltar que a configuração foi para que o limite máximo de *hits* fosse 50.

3.4 Anotação funcional

Para identificar os transcritos dos genes virais encontrados nas lesões papilomatosas, foi feito um alinhamento entre as sequências correspondentes a cada organismo e seu genoma completo, obtido no NCBI. Esta tarefa foi realizada utilizando-se o programa MEGA versão 7.0 (Kumar *et al.*, 2016), e, ao saber em que

posição o alinhamento foi feito, pôde-se saber qual era o gene correspondente, também tendo como base o genoma completo do vírus no NCBI.

Após a identificação dos genes expressos, foi realizada a anotação funcional dos mesmos. Para tanto, os códigos de acesso obtidos pelo NCBI após o alinhamento com o MEGA foram utilizados para fazer buscas nas plataformas Gene Ontology (<http://www.geneontology.org/>), Kyoto Encyclopedia of Genes and Genomes (<https://www.genome.jp/kegg/kegg2.html>), ViruSite (<http://www.virusite.org/>) e Universal Protein Knowledgebase (<https://www.uniprot.org/>), onde foram anotados os componentes celulares, processos biológicos e funções moleculares dos genes expressos. Paralelamente, também foi feita uma busca pelas informações dos genes na literatura, em artigos científicos.

4 RESULTADOS

4.1 Controle de qualidade, quantificação dos *reads* e mapeamento dos transcritos

Tendo sido extraído o RNA das lesões de três animais não infectados e três animais infectados (contidas num banco de dados local), foram construídas seis bibliotecas de cDNA em duplicata: para cada amostra, há um arquivo denominado R1 (*Forward*) e outro denominado R2 (*Reverse*), totalizando 12 arquivos em formato FastQ. No total, o sequenciamento gerou 121.722.238 *reads* pareados, cada um com aproximadamente 101 pares de bases (pb), sendo considerados curtos (Biederstedt *et al.*, 2018), e cada arquivo individual possui cerca de 20 milhões de *reads*.

Após o corte das sequências de baixa qualidade, homopolímeros e adaptadores com o programa Trimmomatic, o número de *reads* em cada arquivo R1 e R2 diminuiu em torno de dois milhões, restando aproximadamente 17 milhões de *reads* por biblioteca (Tabela 1).

Tabela 1. Quantidade de *reads* totais e após o corte do Trimmomatic

Biblioteca	<i>Reads</i> totais	<i>Reads</i> após o corte
Infectados por BPV		
1	20335461	17917849
2	19950125	17165233
5	20407539	17796330
Não infectados por BPV		
6	19266702	17389959
7	21908947	19081327
10	19853464	17389655

Na avaliação de qualidade, o FastQC mostrou que todos os *reads* apresentaram um valor de Phred médio acima de 32, e, sabendo que um valor de Phred igual a 30 corresponde a uma taxa de erro de 0.1%, equivalente a uma base errada em mil (Liao *et al.*, 2017), pode-se afirmar que as sequências possuem uma boa qualidade e poucos erros são esperados.

Após serem alinhados com o genoma de referência, cada arquivo R1 e R2 retornou unido, resultando em apenas seis arquivos (uma para cada amostra), e estes também precisaram ser convertidos para o formato BAM, padrão do software TopHat. Ao alinhar os transcritos com o genoma bovino, obteve-se 87,56%, equivalente a 93.261.478 de *reads*, mapeados com o genoma de referência (UMD3.1 *Bos taurus* 8) (Barreto *et al.*, 2018). Os *reads* restantes, pertencentes a organismos desconhecidos e que somam 28.460.760 (equivalente a 12,44%), foram obtidos ao se separar todos os que não foram mapeados dos que foram mapeados, utilizando o programa BAM filter na plataforma Galaxy. Após a separação, cada arquivo ficou com cerca de quatro a cinco milhões de sequências e um tamanho, em mb, em torno de 146 a 200.

Como o programa BLASTn necessita de arquivos em formato FASTA para começar as análises, todos os seis arquivos em formato BAM tiveram que ser convertidos, também com o uso do plataforma Galaxy. Com esta conversão, o número de sequências foi mantido, como pode ser visto na Tabela 2, porém o tamanho diminuiu consideravelmente, já que trata-se de um formato compatível com arquivos de texto.

Tabela 2. Tamanho dos arquivos FASTA e quantidade de sequências após separação

Biblioteca	Tamanho (em KB)	Quantidade de sequências não mapeadas
Infectados por BPV		
Query1	430.5	4554602
Query2	585.4	5928512
Query5	366.7	3932626
Não infectados por BPV		
Query6	384.3	4165352
Query7	343.4	4031091
Query10	333.6	3686303

4.2 Alinhamento local usando o BLASTn

O programa BLASTn gerou arquivos de extensão FASTA em formatação tabular, onde foi possível ver o nome genérico das sequências, o código de acesso do NCBI, geralmente composto pelas iniciais “NC_” e seguido de sete dígitos, o valor de identidade daquele *hit* e o *e-value*. Vale ressaltar que a maior parte dos *hits* em todos os arquivos teve um índice de identidade considerado alto, nunca ficando abaixo de 85%. Os melhores *hits* obtidos pelo BLASTn, juntamente com seus respectivos códigos de acesso, *e-values* e índices de identidade podem ser vistos na Tabela 3.

Tabela 3. Resultados do BLASTn, mostrando os melhores *hits* para cada organismo viral com base no índice de identidade e *e-value*

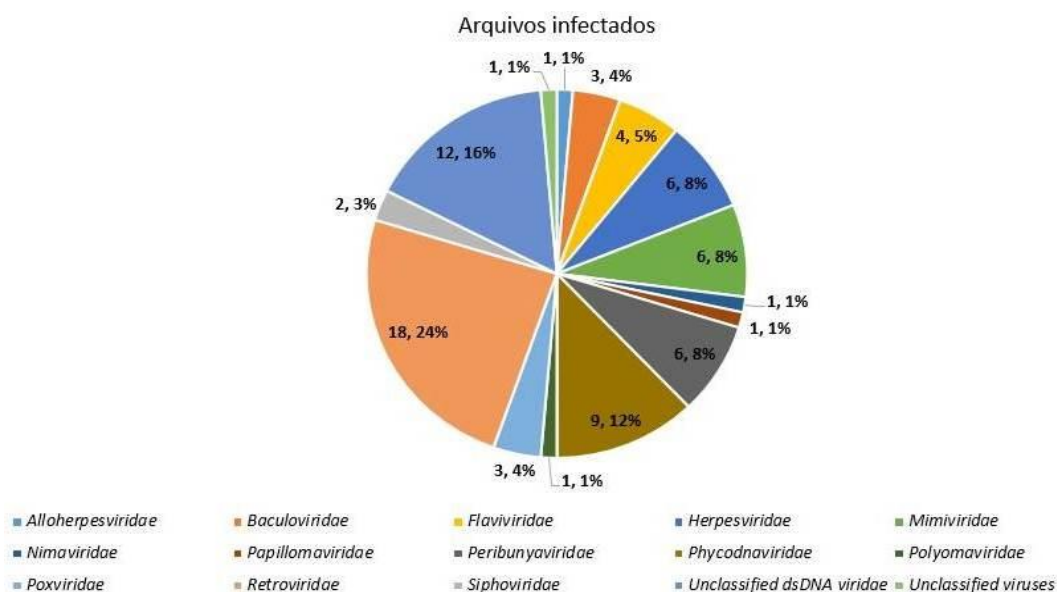
Família	Organismo	Código de acesso (NCBI)	Identidade	<i>E-value</i>
<i>Alloherpesviridae</i>	<i>Cyprinid herpesvirus 2</i>	NC_019495.1	100%	2.44E-17
<i>Baculoviridae</i>	<i>Choristoneura fumiferana granulovirus</i>	NC_008168.1	90%	2.56E-11
<i>Circoviridae</i>	<i>Bat associated circovirus 1</i>	NC_038385.1	100%	1.30E-06
	<i>Tadarida brasiliensis circovirus 1</i>	NC_028045.1	100%	2.54E-06
<i>Flaviviridae</i>	<i>Bovine viral diarrhea virus 1</i>	NC_001461.1	100%	1.67E-11
	<i>Pestivirus giraffe-1</i>	NC_003678.1	100%	1.19E-38
<i>Herpesviridae</i>	<i>Alcelaphine herpesvirus 2</i>	NC_024382.1	100%	1.97E-07
	<i>Bovine herpesvirus 6</i>	NC_024303.1	100%	1.99E-07
	<i>Canid herpesvirus 1</i>	NC_030117.1	89%	1.60E-07
	<i>Cercopithecine herpesvirus 2</i>	NC_006560.1	100%	1.74E-06
	<i>Felis catus gammaherpesvirus 1</i>	NC_028099.1	100%	1.77E-07
<i>Mimiviridae</i>	<i>Murid herpesvirus 1</i>	NC_004065.1	100%	1.14E-07
	<i>Acanthamoeba polyphaga mimivirus</i>	NC_014649.1	92%	1.99E-07
	<i>Mimivirus terra2</i>	NC_023639.1	92%	1.99E-07
<i>Nimaviridae</i>	<i>Yellowstone lake mimivirus</i>	NC_028104.1	100%	1.89E-07
	<i>White spot syndrome virus</i>	NC_003225.3	100%	2.68E-07
<i>Papillomaviridae</i>	<i>Bos taurus papillomavirus 13</i>	NC_030795.1	100%	2.34E-07
<i>Peribunyaviridae</i>	<i>Shamonda virus</i>	NC_018464.1	100%	1.34E-08
	<i>Simbu virus</i>	NC_018476.1	100%	1.53E-13
<i>Phycodnaviridae</i>	<i>Acanthocystis turfacea Chlorella virus 1</i>	NC_008724.1	97%	1.54E-08
	<i>Aureococcus anophagefferens virus</i>	NC_024697.1	97%	1.54E-08

	<i>Chrysochromulina ericina virus</i>	NC_028094.1	94%	2.28E-09
	<i>Ectocarpus siliculosus virus 1</i>	NC_002687.1	100%	2.58E-06
	<i>Orpheovirus IHUMI-LCC2</i>	NC_036594.1	100%	1.99E-07
<i>Podoviridae</i>	<i>Escherichia phage vB_EcoP_PhAPEC5</i>	NC_024786.1	100%	2.16E-07
<i>Polyomaviridae</i>	<i>Adelie penguin polyomavirus</i>	NC_026141.2	92%	2.88E-08
<i>Poxviridae</i>	<i>BeAn 58058 virus</i>	NC_032111.1	100%	2.60E-09
	<i>Turkeypox virus</i>	NC_028238.1	91%	1.59E-10
<i>Retroviridae</i>	<i>Abelson murine leukemia virus</i>	NC_001499.1	92%	1.64E-27
	<i>Avian myelocytomatosis virus</i>	NC_001866.1	92%	1.62E-08
	<i>Avian sarcoma virus CT10</i>	NC_038922.1	89%	1.75E-10
	<i>FBR murine osteosarcoma</i>	NC_038858.1	95%	1.49E-14
	<i>Hardy-Zuckerman 4 feline sarcoma virus</i>	NC_038923.1	96%	1.14E-39
	<i>Murine osteosarcoma virus</i>	NC_001506.1	100%	1.58E-12
	<i>Rous sarcoma virus</i>	NC_001407.1	91%	1.48E-28
	<i>Woolly monkey sarcoma virus</i>	NC_009424.4	96%	1.14E-39
	<i>Y73 sarcoma virus</i>	NC_008094.1	92%	1.49E-33
<i>Siphoviridae</i>	<i>Achromobacter phage JWF</i>	NC_029075.1	100%	2.58E-06
	<i>Gordonia phage Kvothe</i>	NC_030914.1	100%	1.57E-06
<i>Unclassified archaeal dsDNA viridae</i>	<i>Halorubrum phage HF2</i>	NC_003345.1	97%	1.54E-08
	<i>Halovirus HF1</i>	NC_004927.1	97%	1.54E-08
<i>Unclassified dsDNA viridae</i>	<i>Pandoravirus dulcis</i>	NC_021858.1	100%	2.34E-06
	<i>Pandoravirus inopinatum</i>	NC_026440.1	96%	2.34E-06
	<i>Pandoravirus macleodensis</i>	NC_037665.1	100%	2.34E-06
	<i>Pandoravirus neocaledonia</i>	NC_037666.1	100%	2.34E-06
	<i>Pandoravirus quercus</i>	NC_037667.1	100%	2.34E-06
	<i>Pandoravirus salinus</i>	NC_022098.1	100%	2.68E-06
<i>Unclassified viruses</i>	<i>Mollivirus sibericum</i>	NC_027867.1	100%	2.58E-06

Todas as 18 famílias de vírus encontradas e suas respectivas porcentagens estão descritas na Figura 3 (A e B). As famílias mais representativas, ou seja,

aquelas que estiveram presentes em maior quantidade de *hits* e em ambas as classificações de bibliotecas (infectadas por BPV e não infectadas por BPV) foram *Retroviridae*, *Unclassified dsDNA viridae*, *Phycodnaviridae* e *Herpesviridae*, nessa ordem.

A)



B)

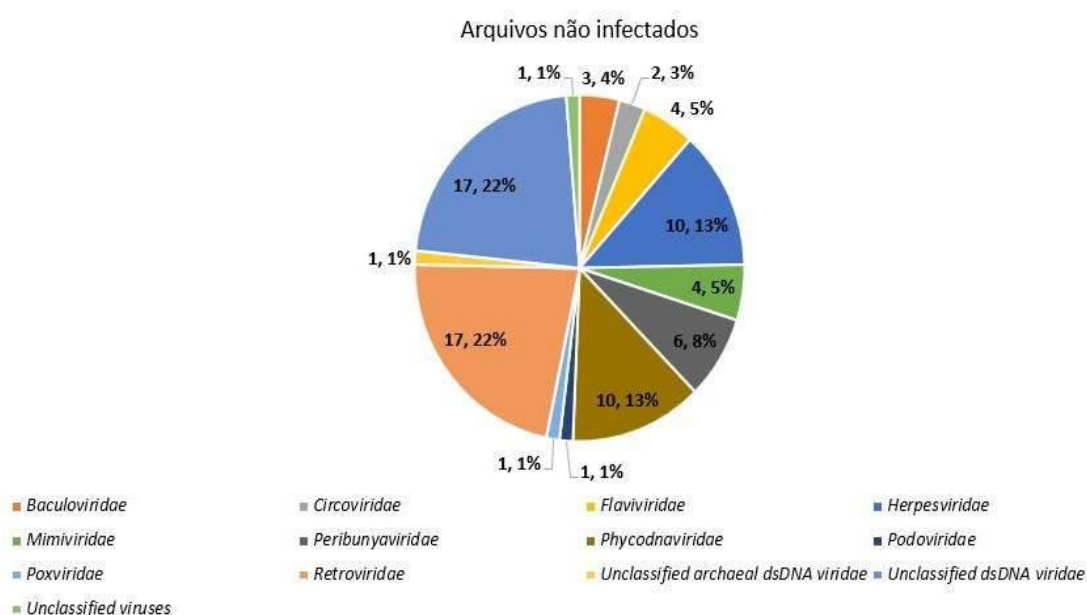


Figura 3. Quantidade, em porcentagem, de organismos encontrados por família viral em ambos os tipos de bibliotecas. Percebe-se uma prevalência das famílias *Retroviridae*, *Unclassified dsDNA viridae*, *Phycodnaviridae* e *Herpesviridae*. **A)** Porcentagem de organismos encontrados por

família viral nas bibliotecas infectadas por BPV (Query1, Query2 e Query5). **B)** Porcentagem de organismos encontrados por família viral nas bibliotecas não infectadas por BPV (Query5, Query7 e Query10).

4.3 Anotação funcional dos genes identificados

Para realizar a anotação funcional dos genes, foram feitas buscas nos diversos bancos de dados internacionais citados anteriormente, tendo como referência as classificações de processos biológicos, componentes celulares e funções moleculares do GO. No entanto, a fim de preencher as lacunas observadas pela falta de informações em vários deles, a maior parte do que foi descrito à respeito dos genes foi obtida em artigos científicos.

Cada vírus e seu respectivo gene foi discriminado na Tabela 4. É importante ressaltar que não foi possível encontrar os genes expressos dos organismos *Acanthamoeba polyphaga mimivirus*, *Mimivirus terra2*, *Aureococcus anophagefferens virus*, *Chrysomulina ericina virus*, *Ectocarpus siliculosus virus 1*, *Orpheovirus IHUMI-LCC2*, *Pandoravirus dulcis*, *Pandoravirus inopiatum*, *Pandoravirus macleodensis*, *Pandoravirus neocaledonia*, *Pandoravirus quercus*, *Pandoravirus salinus* e *Mollivirus sibericum* pois seus respectivos genomas ultrapassam a faixa de 300.000 pares de bases, impossibilitando o alinhamento pelo programa MEGA.

Tabela 4. Símbolos dos genes obtidos para cada organismo viral dentre as famílias encontradas pelo BLASTn.

Família	Espécie	Símbolo do gene
<i>Alloherpesviridae</i>	<i>Cyprinid herpesvirus 2</i>	CyHV2_ORF130
<i>Baculoviridae</i>	<i>Choristoneura fumiferana granulovirus</i>	COGV_gp022
<i>Circoviridae</i>	<i>Bat associated circovirus 1</i>	rep
	<i>Tadarida brasiliensis circovirus 1</i>	rep
<i>Flaviviridae</i>	<i>Bovine viral diarrhea virus 1</i>	PestiV1gp1
	<i>Pestivirus giraffe-1 H138</i>	PGV1gp1
<i>Herpesviridae</i>	<i>Alcelaphine herpesvirus 2</i>	ALHV2gp68
	<i>Bovine herpesvirus 6</i>	BoHV6ORF6
	<i>Canid herpesvirus 1</i>	UL48
	<i>Cercopithecine herpesvirus 2</i>	UL1
	<i>Felis catus gammaherpesvirus 1</i>	AQ624_gp08

	<i>Murid herpesvirus 1</i>	MuHV1_gp018
	<i>Papiine herpesvirus 2</i>	RL2
Mimiviridae	<i>Yellowstone lake mimivirus</i>	AR680_gp073
Nimaviridae	<i>White spot syndrome virus</i>	SWSSV_gp010
Papillomaviridae	<i>Bos taurus papillomavirus 13</i>	E6
Peribunyaviridae	<i>Shamonda virus</i>	N
	<i>Simbu virus</i>	RdRp
Phycodnaviridae	<i>Acanthocystis turfacea Chlorella virus 1</i>	Z771L
Podoviridae	<i>Escherichia phage vB_EcoP_PhAPEC5</i>	LD33_gp79
Polyomaviridae	<i>Adelie penguin polyomavirus</i>	SP79_gp3
Poxviridae	<i>BeAn 58058 virus</i>	ASN15_gp026
	<i>Turkeypox virus</i>	BTW14_gp014
Retroviridae	<i>Abelson murine leukemia virus</i>	gag-abl
	<i>Avian myelocytomatosis virus</i>	Amvgp1
	<i>Avian sarcoma virus CT10</i>	D1T39_gp1
	<i>FBR murine osteosarcoma</i>	D1S92_gp1
	<i>Hardy-Zuckerman 4 feline sarcoma virus</i>	D1T40_gp1
	<i>Murine osteosarcoma virus</i>	Movgp2
	<i>Rous sarcoma virus</i>	SRC
	<i>Woolly monkey sarcoma virus</i>	WMSV_gp6
Siphoviridae	<i>Y73 sarcoma virus</i>	SRC
	<i>Achromobacter phage JWF</i>	AXJ13_gp071
Unclassified archaeal dsDNA viridae	<i>Gordonia phage Kvothe</i>	9
	<i>Halorubrum phage HF2</i>	HF2p104
	<i>Halovirus HF1</i>	HalHV1gp091

A tabela 5 descreve de forma sucinta as principais funções dos genes obtidos, tendo como base as informações fornecidas pelos bancos e trabalhos científicos. Não foram descritas as funções de genes repetidos, como é o caso dos genes *SRC* e *rep* vistos na tabela anterior, pois, mesmo pertencendo a organismos distintos, pode-se dizer que sua função é similar. Evidentemente, devido ao não reconhecimento dos genes dos organismos citados no parágrafo anterior, estes não tiveram suas funções anotadas.

Tabela 5. Símbolos dos genes obtidos e suas respectivas funções segundo os bancos e artigos.

Símbolo	Função
CyHV2_ORF130	Possui função de molécula estrutural.
COGV_gp022	Possui função de molécula estrutural.

rep	Participa do processo de replicação; possui atividade de helicase no RNA.
PestiV1gp1	Suprime a atividade do IRF3 do hospedeiro; faz poros na membrana celular do hospedeiro e promove a ligação do vírion a ela.
PGV1gp1	Participa do processo de supressão imunológica do hospedeiro, além de promover a entrada do vírus, sua replicação e processamento proteico; promove autofagia nas células do hospedeiro.
ALHV2gp68	Pode estar relacionado com a progressão celular no núcleo e na regulação negativa da ubiquitinação de proteínas.
BoHV6ORF6	Participa da replicação bidirecional do DNA viral.
UL48	Regula a transcrição com molde de DNA 53.
UL1	Permite a fusão do vírus com a membrana plasmática do hospedeiro.
AQ624_gp08	Possivelmente regula o processo apoptótico.
MuHV1_gp018	Possível transdução de sinal mediada por GTPase.
RL2	Possivelmente promove a degradação de diversas proteínas celulares.
AR680_gp073	Atua como um componente integral da membrana do vírus.
SWSSV_gp010	Possui atividade de helicase.
E6	Modula o processo apoptótico do hospedeiro, bem como suprime a via do Interferon tipo I.
N	Possui ligação com o RNA.
RdRp	Participa da replicação de RNA do genoma viral, bem como na transcrição.
Z771L	Tem função de polissacarídeo-liase.
LD33_gp79	Possui atividade estrutural.
SP79_gp3	Possui atividade estrutural.
ASN15_gp026	Biossíntese de ácidos graxos.
BTW14_gp014	Possível sinalização de estiramento em músculos esqueléticos.
gag-abl	Possui atividade estrutural; ligação ao ATP.
Amvgp1	Apresenta atividade de dimerização proteica, bem como atividade do fator transcricional de ligação ao DNA.
D1T39_gp1	Possível atividade de molécula estrutural.
D1S92_gp1	Possui atividade de ligação ao DNA.
D1T40_gp1	Possível atividade de brotamento viral e

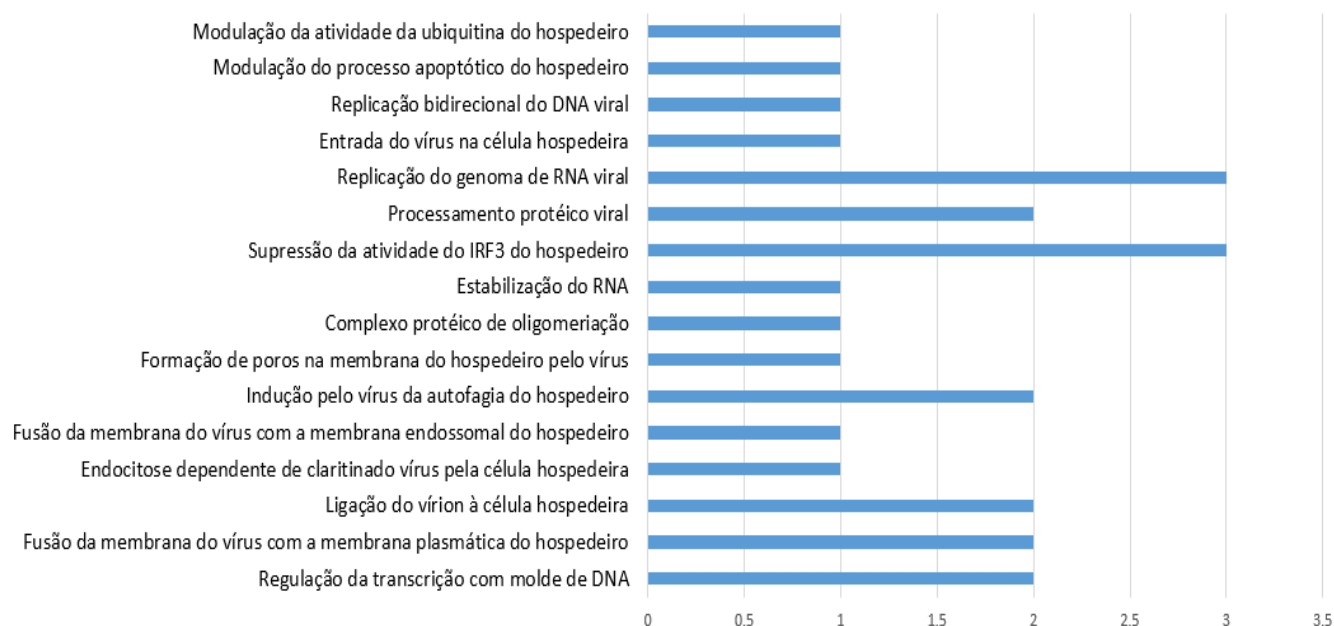
	ligação ao RNA.
Movgp2	Possivelmente envolvida na ubiquitinação de proteínas.
SRC	Apresenta atividade da proteína tirosina-quinase.
WMSV_gp6	Promove a ligação do vírion à membrana celular do hospedeiro, assim como a fusão de suas membranas.
AXJ13_gp071	Atua como molécula estrutural.
9	Apresenta ligação a íons metálicos.
HF2p104	Possui função de helicase.
HalHV1gp091	Possui função de helicase.

A figura 4 (A e B) quantifica os processos biológicos anotados pelo GO, separados por infectados por BPV e não infectados por BPV. Como já foi dito, nem todos os genes possuem todas as características anotadas nos bancos de dados buscados, tendo que, na maioria dos casos, se recorrer à genes similares descritos na literatura para se obter essas informações.

Percebe-se que os processos mais registrados nas bibliotecas infectadas por BPV foram os de replicação do genoma viral (RNA) e supressão do sistema imune do hospedeiro. Em segundo lugar, também pode-se citar as funções de processamento proteico viral, indução de autofagia, ligação do vírion à célula hospedeira, fusão das membranas do vírus e do hospedeiro e regulação da transcrição com molde de DNA.

Já nas bibliotecas não infectadas por BPV, os processos foram mais bem distribuídos, ficando em primeiro lugar os processos de replicação do genoma viral e ligação do vírion ao hospedeiro. Outros processos que merecem destaque são os de replicação bidirecional do DNA, processamento proteico viral, supressão do sistema imune, indução de autofagia no hospedeiro, replicação de DNA, fusão das membranas do vírus e do hospedeiro e regulação da transcrição com molde de DNA.

A)



B)

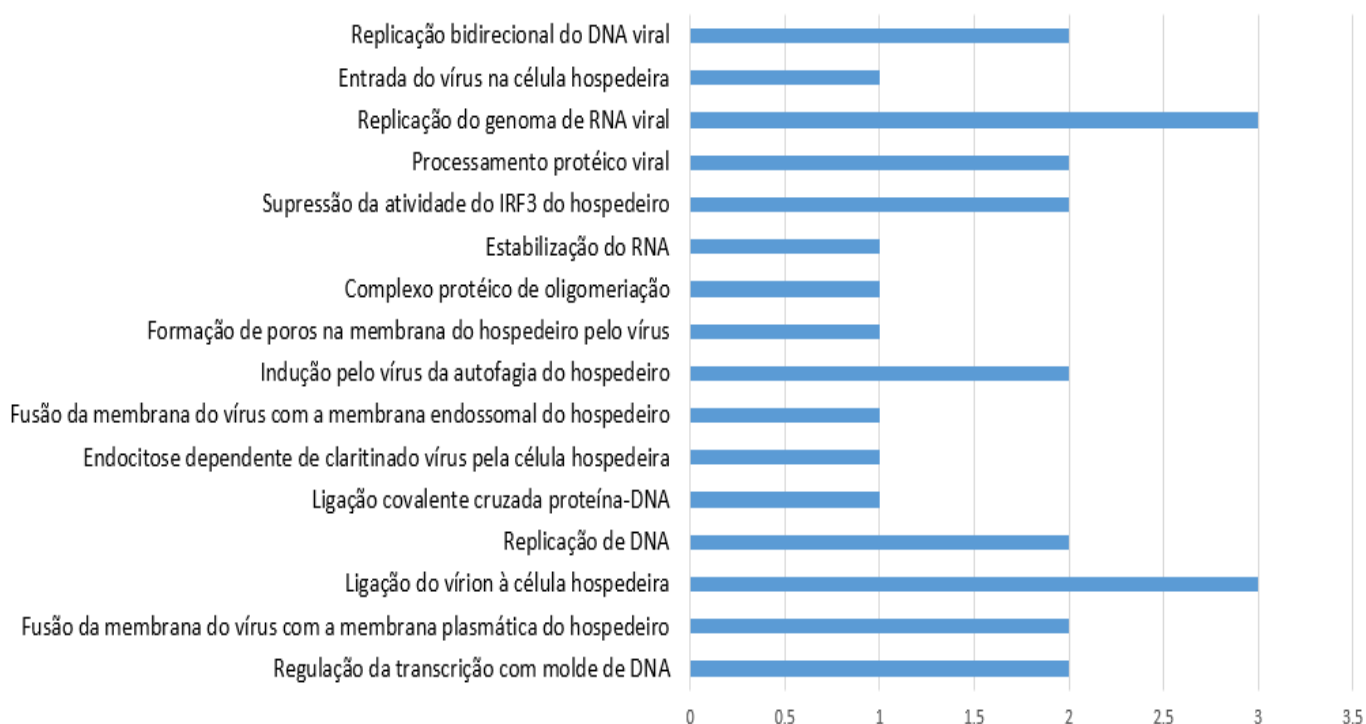
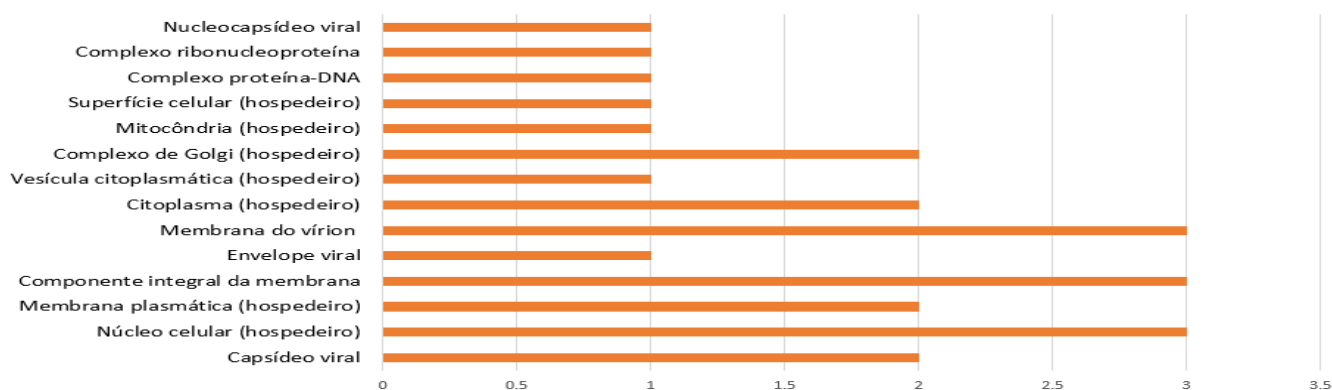


Figura 4. A) Quantificação dos processos biológicos anotados pelo GO das bibliotecas infectadas por BPV. No eixo X tem-se a quantidade de genes em valores absolutos. **B)** Quantificação dos processos biológicos anotados pelo GO das bibliotecas não infectadas por BPV. No eixo X tem-se a quantidade de genes em valores absolutos.

Os componentes celulares descritos pelos bancos de dados foram discriminados na figura 5 (A e B). Nas bibliotecas infectadas por BPV, nota-se que os componentes mais repetitivos foram a membrana do vírion, como componente integral da membrana do vírus e no núcleo celular do hospedeiro, com destaque para o complexo de Golgi, citoplasma e membrana plasmática, bem como para o capsídeo viral.

De forma semelhante ocorre nas bibliotecas não infectadas por BPV, com a diferença que os componentes de maiores quantidades foram a membrana do vírion e como componente integral da membrana. Abaixo destes, ficaram o complexo de Golgi, envelope viral, membrana plasmática e núcleo do hospedeiro.

A)



B)

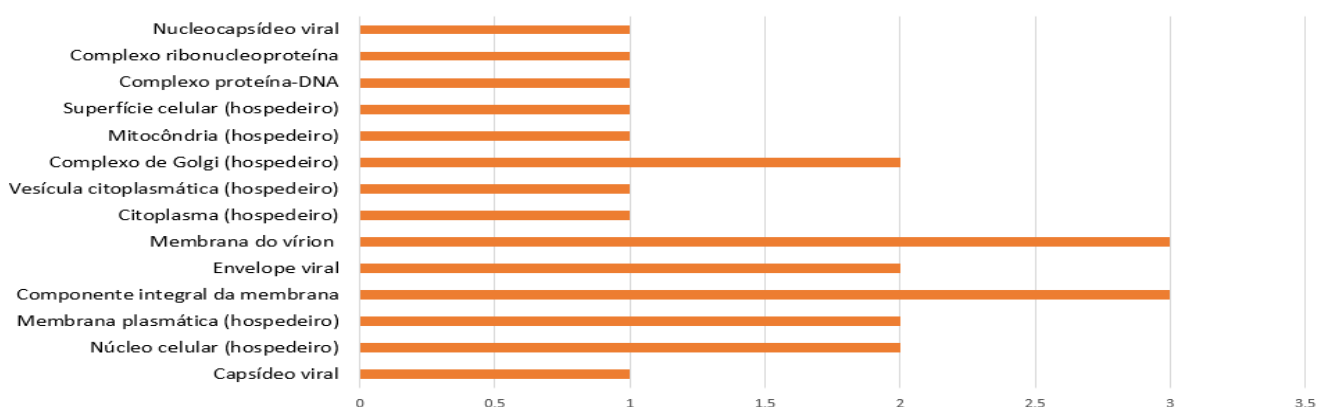
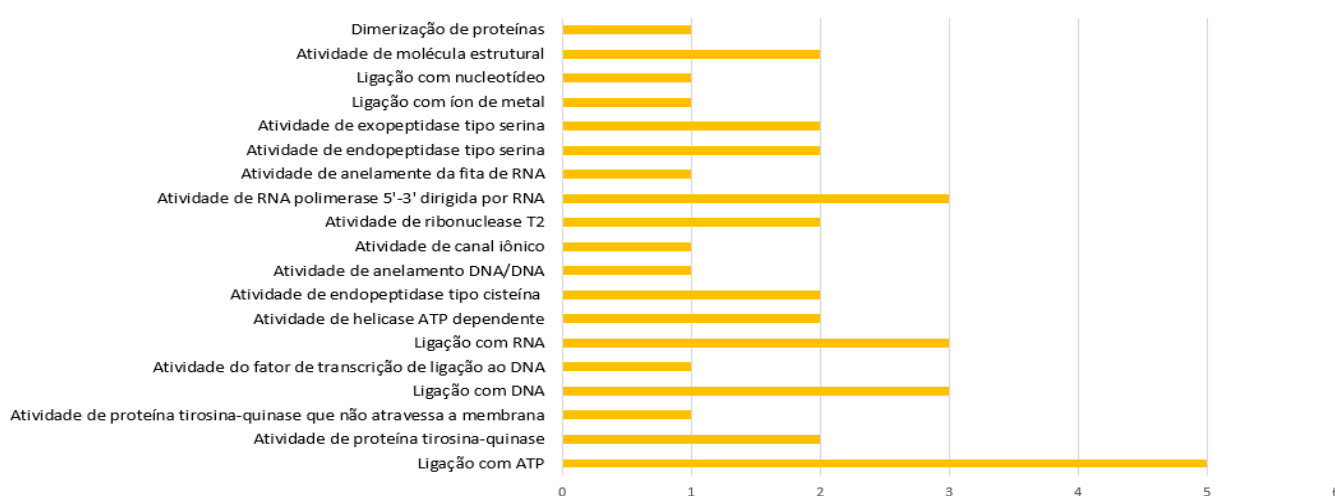


Figura 5. A) Quantificação dos componente celulares anotados pelo GO das bibliotecas infectadas por BPV. No eixo X tem-se a quantidade de genes em valores absolutos. **B)** Quantificação

dos componente celulares anotados pelo GO das bibliotecas não infectadas por BPV. No eixo X tem-se a quantidade de genes em valores absolutos.

No quesito de funções moleculares, a função que se destacou de forma unânime foi a de ligação com o ATP, em ambos os tipos de bibliotecas. Até mesmo nas funções em segundo lugar não houve grandes diferenças, podendo-se observar que as funções de destaque em ambas as situações (infectadas ou não por BPV), foram ligação com RNA, DNA e atividade de RNA polimerase (Figura 6 A e B).

A)



B)

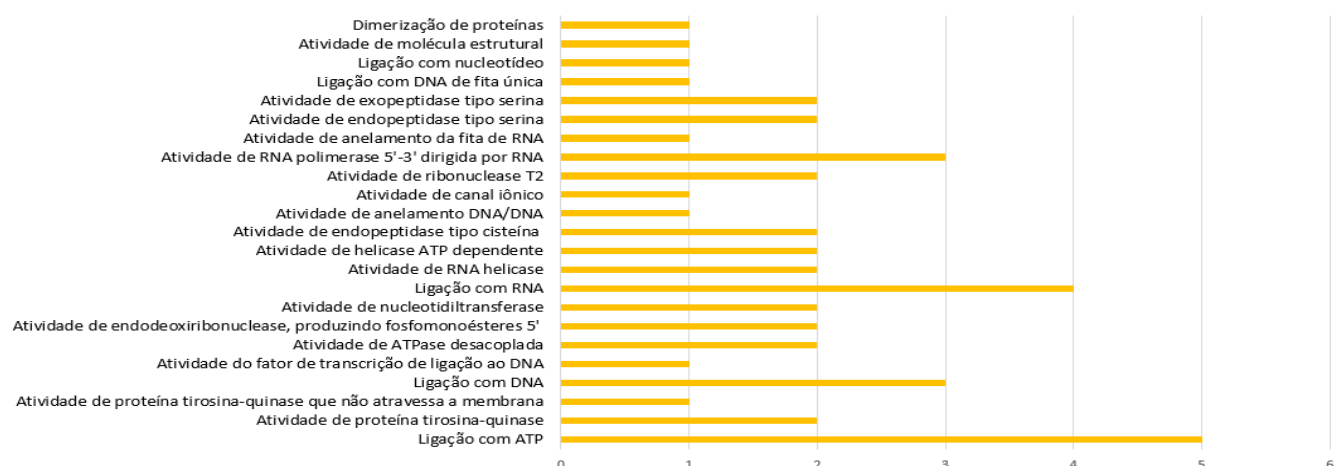


Figura 6. **A)** Quantificação das funções moleculares anotadas pelo GO das bibliotecas infectadas por BPV. No eixo X tem-se a quantidade de genes em valores absolutos. **B)** Quantificação das funções moleculares anotadas pelo GO das bibliotecas não infectadas por BPV. No eixo X tem-se a quantidade de genes em valores absolutos.

Da mesma forma que os bancos de dados se mostraram insuficientes para a realização da anotação funcional, a busca por mapas de vias metabólicas foi igualmente difícil. Entretanto, nota-se que houve uma repetição de genes *SRC* na forma de homólogos (como o *gag-ab1*), principalmente presentes na família *Retroviridae*. Esses genes, que fazem parte da família *SRC* de proteínas serina/treonina quinases (Roskoski Jr., 2004), foram observados no mapa de regulação do ciclo celular, geralmente estimulando o receptor do fator de crescimento epidermal (EGF-R), na transição entre as fases G2 e M, bem como nas vias de infecção do vírus da Hepatite B e do câncer de bexiga (Figura 7), também promovendo a diferenciação e proliferação celular.

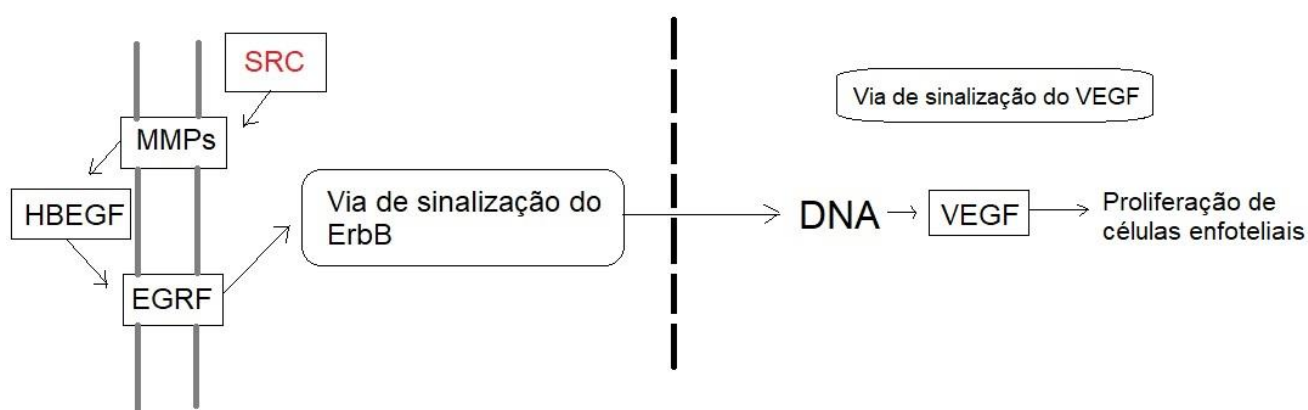
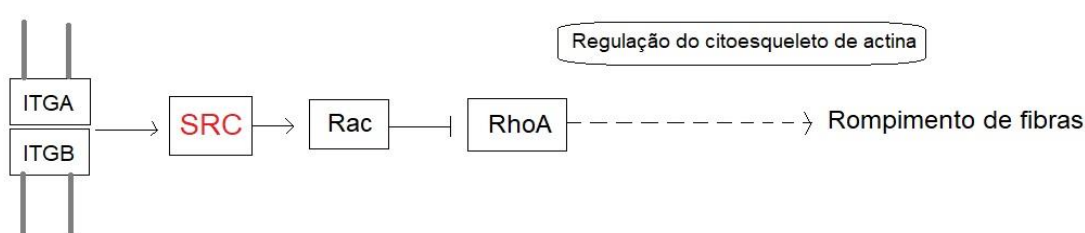


Figura 7. Adaptação de parte do mapa metabólico fornecido pelo KEGG, evidenciando a relação entre o gene *SRC* com o receptor EGF para promover a proliferação celular durante o câncer de bexiga. KEGG: ko05219.

Na via de infecção do citomegalovírus humano, o gene *SRC* foi observado sendo modulado pelas integrinas A ou B para promover o rompimento das fibras do citoesqueleto. Também, na mesma via, foi encontrado o gene *UL38*, homólogo dos genes obtidos nesse estudo *UL48* e *UL1*, interagindo com a mTOR (proteína serina/treonina quinase) para promover a replicação viral (Figura 8 A e 6B).

A)



B)

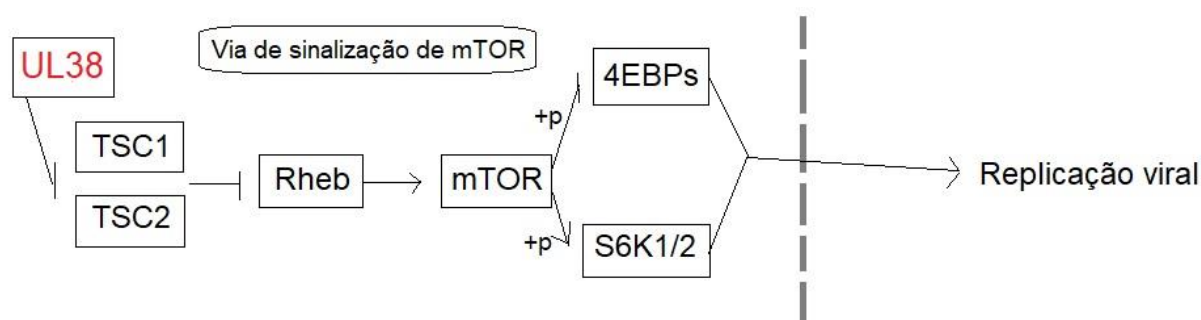


Figura 8. A) Adaptação de recorte da via de infecção do citomegalovírus humano, ilustrando a participação do gene *SRC* na regulação do citoesqueleto de actina. **B)** Recorte adaptado do mesmo mapa de infecção do citomegalovírus humano, desta vez mostrando como o gene *UL38* interage indiretamente com a *mTOR* para promover a replicação viral. KEGG: ko05163.

Evidentemente, não se pode deixar de mencionar a proteína E6 que foi encontrada na família *Papillomaviridae*. A E6 é citada nas vias metabólicas de infecção de papilomavírus humano e no mapa de carcinogêneses virais, modulando diretamente a p53 para promover a proliferação celular ou inibindo o IRF3, para suprimir a resposta imune no hospedeiro (Figura 9).

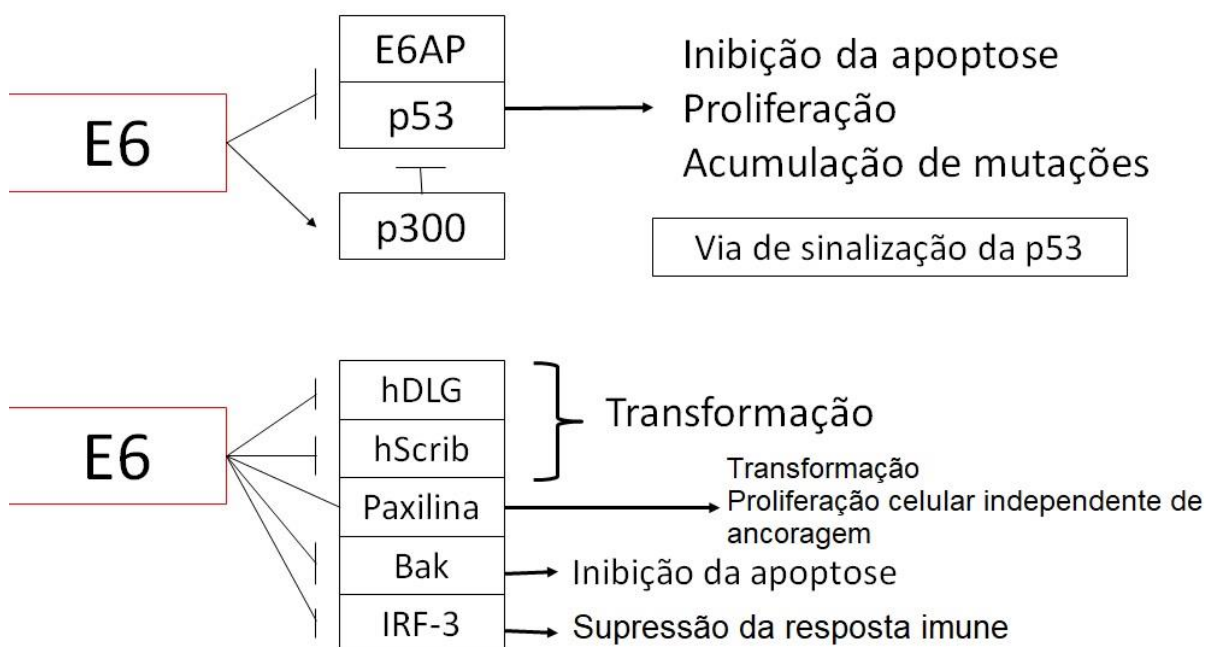


Figura 9. Mapa esquemático simplificado (adaptado) fornecido pelo KEGG, ilustrando as interações da proteína E6 com a proteína p53 e o fator regulador de interferon 3 (IRF3). KEGG: ko05203.

5 DISCUSSÃO

Neste estudo, o uso das tecnologias de sequenciamento de nova geração possibilitou a obtenção, tratamento e apresentação de dados provenientes de lesões papilomatosas no tecido epitelial de três bovinos para a possível análise de interações entre o BPV e outros agentes virais ali contidos. Com o uso do RNA-seq, foi possível realizar uma análise metagenômica, ou seja, com a extração de amostras provenientes diretamente do ambiente, sem a necessidade de clonagens, meios de cultura e/ou conhecimento prévio à respeito dos organismos a serem identificados (Willner *et al.*, 2009). Este método de pesquisa abre um leque de oportunidades para a criação de conhecimentos de base, pois, à partir deste tipo de trabalho é que surgem pesquisas mais aprofundadas acerca da identificação, classificação, descrição populacional, anotação funcional, e busca de possíveis tratamentos para microrganismos patogênicos. Foi com este intuito que este estudo foi realizado, para que se possa entender como é composta a microbiota viral em lesões desse tipo e como ela se comporta face à infecção papilomatosa.

Com os alinhamentos visualizados pelo BLASTn, não foram encontradas somente famílias virais relacionadas com infecções em vertebrados, mas diversas outras que variam desde vírus que infectam arqueias e algas, e até mesmo organismos virais ainda não classificados nos bancos de dados. É provável que isto tenha acontecido por ser um dos vieses de se trabalhar com a metagenômica: amostras retiradas diretamente do ambiente, apesar de terem suas vantagens, não são puras e podem apresentar contaminantes durante o sequenciamento (Willner *et al.*, 2009). Existem diversas razões para estas contaminações, podendo-se citar o ar, que por si só já é um meio rico em microrganismos, a pele do animal, que, como qualquer outro, não é estéril e pode apresentar células de outros organismos ou até mesmo de outros tecidos, além de outros fatores como o contato com outros bovinos, humanos tratadores, entre outros (Willner *et al.*, 2009).

Sendo assim, os vírus foram selecionados com base num critério de potencial patogênico, para evitar levar em conta organismos que não estavam efetivamente expressando nenhum gene no tecido do animal. Um exemplo disso foi o vírus *Choristoneura fumiferana granulovirus*, um vírus da família *Baculoviridae*, ou seja, específicos de insetos (neste caso, de uma espécie de mariposa). Apesar de estar

presente em grande quantidade em todas as bibliotecas, é possível que isso tenha acontecido apenas por homologia às sequências do bovino, mas que não estava de fato presente na amostra, ou, se estava presente, poderia estar apenas inativo. Outra razão para a presença de outros vírus que fogem ao nosso interesse é que, devido às amostras terem sido retiradas de animais que viviam em propriedades rurais, é possível que houvessem ali outros vírus latentes, no caso daqueles que apareceram em nossas bibliotecas não infectadas por BPV.

A maior parte dos vírus presentes nas lesões não eram específicos de bovinos, mas infectavam preferencialmente animais modelo (famílias *Cricetidae* e *Muridae*) (Santos *et al.*, 2017), animais taxonomicamente relacionados ao *Bos taurus* (famílias *Giraffidae* e *Bovidae*) (Daudt *et al.*, 2017), animais aparentados aos humanos (famílias *Cercopithecidae* e *Atelidae*) (Brush *et al.*, 2014; Goldstein *et al.*, 1987) e animais que possivelmente podem ter contato com bovinos no ambiente em que habitam ou que apresentaram vírus que possuíam a capacidade de infectar outras classes de vertebrados (famílias *Molossidae*, *Felidae*, *Canidae*, *Spheniscidae* e *Phasianidae*) (da Silva *et al.*, 2012).

Após a anotação dos genes expressos, foi possível observar que há uma predominância de processos de regulação de DNA e ligação do vírion à célula hospedeira, bem como fosforilação de proteínas e supressão da resposta do sistema imune. Segundo a literatura (Pagliarini *et al.*, 2015), estas constituem características típicas de oncogenes, ou seja, genes que estão em contato direto com o DNA e induzem mutações nele, principalmente alterações no ciclo celular e regulação da apoptose da célula hospedeira. O fato de os genes encontrados estarem situados na membrana do vírus ou no núcleo celular do hospedeiro corrobora com este fato, além de eles terem como principais funções moleculares a ligação com o DNA, RNA ou ATP.

A anotação de todos os genes obtidos não foi uniforme, havendo ausência de informações consistentes nos bancos verificados em várias ocasiões. Casos como os genes RL2, ALHV2gp68, AQ624_gp08, BTW14_gp014, Movgp2, D1T39_gp1, D1S62_gp1, D1T40_gp1 se mostraram um grande desafio, pois tanto no UniProtKB, KEGG, GO ou ViruSite as informações eram escassas. Além disso, por serem genes muito específicos ou de organismos pouco estudados, até mesmo na literatura suas

descrições eram raras e sucintas, o que impossibilitou que encontrássemos descrições a nível metabólico para eles.

De forma geral, os genes mais bem descritos nos bancos de anotação foram os das famílias *Herpesviridae*, *Peribunyaviridae*, *Flaviviridae* e *Retroviridae*, que coincidiram de ser os mais representativos em termos numéricos durante o nosso levantamento, além do próprio BPV na forma da proteína E6. Os genes de cada uma destas famílias, em geral, compartilhavam similaridades em uma ou mais das características anotadas (processo biológico, componente celular e função molecular). Até mesmo entre famílias, existiu uma semelhança de funcionamento viral, tendo, por exemplo, os genes PestiV1gp1, PGV1gp1, SRC e gag-abl a mesma função molecular de ligação ao ATP. Também foi observada a mesma similaridade entre os genes PestiV1gp1, PGV1gp1, N e RdRp na função de ligação ao RNA e atividade de RNA polimerase. Os genes E6, UL48, BoHV6ORF6 e Amvgp1 também apresentaram a mesma função de ligação ao DNA.

Tendo em mente estas informações, é possível traçar alguns paralelos, como por exemplo o fato de os genes UL48 e BoHV6ORF6 compartilharem funções com a proteína E6, implicando que poderiam interagir de alguma forma na progressão da doença papilomatosa. Isto acontece pois, ambos os genes pertencem ao gênero *Varicellavirus* (Larsen *et al.*, 2015), conhecido por acometer diversos grupos de mamíferos, atingindo o preferencialmente o epitélio mucoso e reprogramando vias de sinalização induzindo fatores celulares (por exemplo, os fatores de transcrição citoplasmáticos – STAT) para suportar sua replicação (Zerboni *et al.*, 2014). Com isso, fica estabelecida uma infecção latente, onde seria uma boa oportunidade para a proteína E6 modular o processo apoptótico do hospedeiro.

Os genes das famílias *Shamonda* e *Simbu virus* são conhecidos por causar a “síndrome do bezerro fraco”, condição caracterizada por má formações congênitas em ruminantes (Behar *et al.*, 2019). Já foi relatado que tais vírus conseguem atravessar a placenta do animal, causando situações de abortos e fetos natimortos, quando não causam danos cerebrais severos no filhote (Behar *et al.*, 2019). Num estudo realizado na China, foi observado que mulheres infectadas com HPV tinham frequentes ocorrências de abortos espontâneos (Xiong *et al.*, 2018). Tendo em

mente que o BPV também já foi detectado nas vias sexuais de ruminantes, uma coinfeção entre os dois vírus poderia agravar essas ocorrências durante a gestação.

Ambos os Flavivírus encontrados neste estudo pertencem ao grupo dos *Pestivirus*, conhecidos por possuírem estratégias de subversão ou uso da autofagia do hospedeiro para melhorar a eficiência de sua replicação (Rajput *et al.*, 2017). Evidentemente, os dois genes obtidos, PestiV1gp1 e PGVgp1, possuem como característica a indução da autofagia no hospedeiro e supressão da resposta imune inata. Já foi demonstrado, no estudo de Barreto *et al.* (2018), que a infecção por BPV causa mudanças significativas na expressão de genes envolvidos com o sistema imune. Possivelmente, uma interação entre ambos os vírus possa deixar o bovino mais debilitado e suscetível a uma progressão no quadro patogênico.

O gene *SRC* e homólogos, provenientes da família *Retroviridae*, foram os que apresentaram as melhores informações nos bancos de dados e esteve presente em 36 mapas metabólicos do KEGG, a maioria estando relacionada a infecções bacterianas (shigelose, tuberculose, sinalização na via da *H. pylori* e invasão de bactérias nas células do hospedeiro), vias de sinalização de hormônios (relaxina, oxitocina, hormônios tireoidianos, estrogênio e GnRH), infecções virais (Hepatite B, citomegalovírus humano e herpesvírus associado ao sarcoma de Kaposi), câncer (carcinogênese viral, câncer de bexiga e proteoglicanas no câncer), funções celulares (endocitose, mitofagia em animais, ativação de plaquetas e regulação do citoesqueleto de actina) e na sinalização de fatores celulares (EGF-R, ErbB, MAPK, Rap1), entre outros.

Percebe-se que a maioria destas funções está relacionada com a promoção da invasão de microrganismos no hospedeiro, manutenção do quadro infeccioso, supressão da atividade do sistema imune e proliferação e sobrevivência celular. Apenas com estas informações é possível afirmar que este gene tem características típicas de um oncogene (Pagliarini *et al.*, 2015), ou seja, possui uma alta capacidade transformante a nível genético.

Segundo a literatura (Hbib *et al.*, 2008), os genes *SRC* detém a capacidade de causar cânceres dos mais diversos, a exemplo dos colorretais, ovarianos, de cabeça e pescoço, pulmonares ou até esofágicos. No mesmo estudo, de Hbib *et al.* (2008),

foi detectado que a progressão tumoral e posterior metástase está diretamente relacionada com a super expressão do fator de crescimento epidermal (EGF-R), que interage com o receptor de tirosina quinase p60-SRC para aumentar a síntese de DNA em fibroblastos. Acredita-se, então, que o mesmo possa ocorrer com os queratinócitos durante uma coinfeção entre o BPV e um vírus que expresse o gene SRC (Xie *et al.*, 2017; Barreto *et al.*, 2018).

Apesar de o SRC ser um gene que atinge com maior frequência os tecidos internos do animal, este pode servir como uma porta entrada oportuna para o BPV, pois já se foi reportada a presença do vírus ou mesmo de verrugas em outros tecidos que não são epiteliais, como sangue (Lindsey *et al.*, 2009; Santos *et al.*, 2016), sêmen (Lindsey *et al.*, 2009) e até mesmo no canal digestivo. Devido às suas atividades de fusão das membranas do vírus e do hospedeiro e imunossupressão, pode-se inferir que os genes SRC podem ser de suma importância para a manutenção da patologia papilomatosa, fazendo-a perdurar ou até transformando-a num quadro canceroso, tendo em vista que grande parte dos Retrovírus presentes neste estudo têm capacidade para isso (Wang *et al.*, 2016; Rosenberg e Witte, 1980; Aaronson *et al.*, 1975; Majumder *et al.*, 1990).

6 CONSIDERAÇÕES FINAIS

A análise metagenômica dos transcritos de RNA não mapeados com os bovinos nos possibilitou identificar vírus que podem estar relacionados com a infecção por BPV e alterar de alguma forma sua manifestação patológica, geralmente para casos mais severos da doença. Devido à escassa anotação funcional de vários genes pesquisados nesse trabalho, não foi possível estabelecer uma relação firme entre o papilomavírus e as demais famílias que não apresentaram genes e/ou proteínas com funções e processos definidos. Entretanto, para aqueles genes que possuíam anotações satisfatórias, foi possível perceber que pode sim haver alguma interação à nível molecular entre eles e as principais proteínas oncogênicas do BPV, fazendo o papilomavírus perdurar no organismo ou até leva-lo a um quadro maligno. Logo, mais estudos nessa área podem ser necessários para realizar a anotação destes genes, e, posteriormente, inferir com mais certeza qual seria o elo entre estas famílias virais e o BPV. Contudo, este trabalho mostra seu valor na formação de conhecimento de base sobre o tema e por contribuir para a abordagem de um assunto que não é tão corriqueiro na comunidade científica, que é a coinfeção entre BPV/BPV e BPV/outros microrganismos.

7 REFERÊNCIAS

AARONSON, S. A.; STEPHENSON, J. R.; TRONICK, S. R.; HINO, S. Immunological analysis of structural polypeptides of woolly monkey-gibbon ape type C viruses. Expression of woolly monkey helper viral antigens in cells transformed by woolly monkey sarcoma virus. **Bibliotheca Haematologica**, v. 43, out. 1975.

ALI, S. M. A.; MIRZA, Y.; AHMAD, Z.; ZAHID, N.; ENAM, S. A. Human papillomavirus and human cytomegalovirus infection and association with prognosis in patients with primary glioblastoma in Pakistan. **World Neurosurgery**, v. 121, jan. 2019.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic Local Alignment Search Tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.

ANDREWS, S. FastQC: a quality control tool for high throughput sequence data. Disponível em: <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>>, 2010.

ARALDI, R. P.; MAZZUCHELI-DE-SOUZA, J.; MODOLO, D. G.; SOUZA, E. B.; DE MELO, T. C.; SPADACCI-MORENA, D. D.; MAGNELLI, R. F.; DE CARVALHO, M. A. C. R.; DE SÁ JÚNIOR, P. L.; DE CARVALHO, R. F.; BEÇAK, W.; STOCCO, R. C. Mutagenic potential of *Bos taurus* papillomavirus type 1 E6 recombinant protein: First description. **BioMed Research International**, London, v. 2015, 15 p., out. 2015.

BARRETO, D. M.; BARROS, G. S.; SANTOS, L. A. B. O.; SOARES, R. C.; BATISTA, M. V. A. Comparative transcriptomic analysis of bovine papillomatosis. **BMC Genomics**, v. 19 (949), 2018.

BATISTA, M. V. A.; FREITAS, A. C.; BALBINO, V. Q. Entropy-based approach for selecting informative regions in the L1 gene of bovine papillomavirus for phylogenetic inference and primer design. **Genetics and Molecular Research**, Ribeirão Preto, v. 12 (1), p. 400-407, fev. 2013.

BATISTA, M. V. A.; SILVA, M. A. R.; PONTES, N. E.; REIS, M. C.; CORTEGGIO, A.; CASTRO, R. S.; BORZACHIELLO, G.; BALBINO, V. Q.; FREITAS, A. C. Molecular epidemiology of bovine papillomatosis and the identification of a putative new virus type in Brazilian cattle. **The Veterinary Journal**, London, v. 197 (2), p. 368-373, jan. 2013.

BAUERMANN, F. V.; JOSHI, L. R.; MOHR, K. A.; KUTISH, G. F.; MEIER, P.; CHASE, C.; CHRISTOPHER-HENNINGS, J.; DIEL, D. G. A novel bovine papillomavirus type in the genus Dyokappapapillomavirus. **Archives of Virology**, Vienna, v. 162 (10), p. 3225-3228, jun. 2017.

BEHAR, A.; LEIBOVICH, B.; EDERY, B.; YANASE, T.; BRENNER, J. First genomic detection of Peaton virus in a calf with hydranencephaly in Israel. **Veterinary Medicine and Science**, v. 5, p. 87-92, 2019.

BERGVALL, M.; MELENDY, T.; ARCHAMBAULT, J. The E1 proteins. **Virology**, London, v. 445 (1-2), p. 35-56, set. 2013.

BERNARD, H. U.; BURK, R. D.; CHEN, Z.; DOORSLAER, K. V.; ZUR HAUSEN, H.; DE VILLIERS, E. M. Classification of papillomaviruses (PVs) based on 189 PV

types and proposal of taxonomic amendments. **Virology**, London, v. 401 (1), p. 70-79, mai. 2010.

BIEDERSTEDT, E.; OLIVER, J. C.; HANSEN, N. F.; JAJOO, A.; DUNN, N.; OLSON, A.; BUSBY, B.; DILTNEY, A. T. NovoGraph: Human genome graph construction from multiple long-read de novo assemblies. **F1000 Research**, v. 7:1391, set. 2018.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, Oxford, v. 30 (15), p. 2114-2120, abr. 2014.

BRUSH, L. A.; BLACK, G. H.; MCCORMACK, K. A.; MAXWELL, L. K.; WRIGHT, G.; RITCHEY, J. W.; PAYTON, M. E.; EBERLE, R. Papiine herpesvirus 2 as a predictive model for drug sensitivity of Macacine herpesvirus 1 (monkey B virus). **Comparative Medicine**, v. 64 (5), out. 2014.

BUCK, C. B.; DAY, P. M.; TRUS, B. L. The papillomavirus major capsid protein L1. **Virology**, London, v. 445 (1-2), p.169-174, abr. 2013.

CARVALHO, R. F.; SAKATA, S. T.; GIOVANNI, D. N. S.; MORI, E.; BRANDÃO, P. E.; RICHTZENHAIN, L. J.; POZZI, C. R.; ARCARO, J. R. P.; MIRANDA, M. S.; MASSUZHELLI-DE-SOUZA, J.; MELO, T. C.; GOMENALE, G.; ASSAF, S. L. M. R.; BEÇAK, W.; STOCCO, R. C. Bovine papillomavirus in Brazil: detection of coinfection of unusual types by a PCR-RFLP method. **BioMed Research International**, v. 2013, mai. 2013.

CLAUS, M. P.; VIVIAN, D.; LUNARDI, M.; ALFIERI, A. F.; ALFIERI, A. A. Análise filogenética de papilomavírus bovino associado com lesões cutâneas em

rebanhos no Estado do Paraná. **Pesquisa Veterinária Brasileira**, Rio de Janeiro, v. 27 (7), p. 314-318, jul. 2007.

CULLETON, S. P.; KANGINAKUDRU, S.; DESMET, M.; GILSON, T.; XIE, F.; WU, S.; CHIANG, C.; QI, G.; WANG, M.; ANDROPHY, E. J. Phosphorylation of the bovine papillomavirus E3 protein on tyrosine regulates its transcription and replication functions. **Journal of Virology**, Washington DC, v. 91 (16), 1854 p., nov. 2016.

CULP, T. D.; BUDGEON, L. R.; CHRISTENSEN, N. D. Human papillomaviruses bind a basal extracellular matrix component secreted by keratinocytes which is distinct from a membrane-associated receptor. **Virology**, v. 341 (2006), p. 147-159, dez. 2005.

CVETOJEVIC, D.; SAVIC, B.; MILICEVIC, V.; KURELJUSIC, B.; JEZDIMIROVIC, N.; JAKIC-DIMIC, D.; PAVLOVIC, M.; SPALEVIC, L. J. Prevalence of bovine herpesvirus type 4 in aborting dairy cows. **Polish Journal of Veterinary Sciences**, v. 19 (4), p. 731-736, dez. 2016.

DA SILVA, F. R. C.; DAUDT, C.; CIBULSKI, S. P.; WEBER, M. N.; VARELA, A. P. M.; MAYER, F. Q.; ROEHE, P. M.; CANAL, C. W. Genome characterization of a bovine papillomavirus type 5 from cattle in the Amazon region, Brazil. **Virus Genes**, Vienna, v. 53, p. 130-133, nov. 2016.

DA SILVA, F. R. C.; DAUDT, C.; STRECK, A. F.; WEBER, M. N.; LEITE FILHO, R. V.; DRIEMEIER, D.; CANAL, C. W. Genetic characterization of Amazonian bovine papillomavirus reveals the existence of four new putative types. **Virus Genes**, Vienna, v. 51, p. 77-84, jun. 2015.

DA SILVA, M. A. R.; CARVALHO, C. C. R.; COUTINHO, L. C. A.; REIS, M. C.; BATISTA, M. V. A.; DE CASTRO, R. S.; DOS ANJOS, F. B. R.; DE FREITAS, A. C. Co-infection of bovine papillomavirus and feline-associated papillomavirus in bovine cutaneous warts. **Transboundary and Emerging Diseases**, v. 59 (2012), p. 539-543, aug. 2012.

DAUDT, C.; DA SILVA, F. R. C.; LUNARDI, M.; ALVES, C. B. D. T.; WEBER, M. N.; CIBULSKI, S. P.; ALFIERI, A. F.; ALFIERI, A. A.; CANAL, C. W. Papillomaviruses in ruminants: An update. **Transboundary and Emerging Diseases**, Medford, v. 00, p. 1-15, out. 2017.

DAUDT, C.; DA SILVA, F. R.C.; CIBULSKI, S. P.; WEBER, M. N.; MAYER, F. Q.; VARELA, A. P. M.; ROEHE, P. M.; CANAL, C. W. Complete genome sequence of Deltapapillomavirus 4 (bovine papillomavirus 2) from a bovine papillomavirus lesion in Amazon Region, Brazil. **Memórias do Instituto Oswaldo Cruz**, Rio de Janeiro, v. 111 (4), p. 277-279, abr. 2016.

DE JESUS, S. P.; DA COSTA, A. C. M.; BARCELLOS, R. B.; MEDEIROS, R. M.; DA SILVA, C. M. D.; ROSSETTI, M. L. A high prevalence of human papillomavirus 16 and 18 co-infections in cervical biopsies from southern Brazil. **Brazilian Journal of Microbiology**, v. 49 (1), p. 220-223, nov. 2018.

DE SOUZA, C. R. T.; ALMEIDA, M. C. A.; KHAYAT, A. S.; DA SILVA, E. L.; SOARES, P. C.; CHAVES, L. C.; BURBANO, R. M. R. Association between *Helicobacter pylori*, Epstein-Barr virus, human papillomavirus and gastric adenocarcinomas. **World Journal of Gastroenterology**, v. 24 (43), p. 4928-4938, nov. 2018.

DE VILLIERS, E. M. Cross-roads in the classification of papillomaviruses. **Virology**, London, v. 445, p. 2-10, mai. 2013.

DE VILLIERS, E. M.; FAUQUET, C.; BROKER, T. R.; BERNARD, H. U.; ZUR HAUSEN, H. Classification of papillomaviruses. **Virology**, London, v. 324, p. 17-27, mar. 2004.

DIAS, J. D. C.; IGNACCHITI, M. D. C.; GIURIATO, P. G. G.; NUNES, L. de C.; PEREIRA JÚNIOR, O. dos S. Detecção do papillomavirus bovino tipo 2 em bexigas de bovinos com hematúria enzoótica pela técnica de reação em cadeia de polimerase no sul do Espírito Santo, Brasil. **Revista Brasileira de Medicina Veterinária**, v. 34(2), p. 146-151, jun. 2012.

DIMAIO, D.; PETTI, L. M. The E5 Proteins. **Virology**, London, v. 445 (1-2), p. 99-114, mai. 2013.

DINIZ, N.; MELO, T. C.; SANTOS, J. F.; MORI, E.; BRANDÃO, P. E.; RITCHZENHAIN, L. J.; FREITAS, A. C.; BEÇAK, W.; CARVALHO, R. F.; STOCCO, R. C. Simultaneous presence of bovine papillomavirus in blood and in short-term lymphocyte cultures from dairy cattle in Pernambuco, Brazil. **Genetics and Molecular Research**, Ribeirão Preto, v. 8 (4), p. 1474-1480, dez. 2009.

DOORBAR, J. The E4 protein; structure, function and patterns of expression. **Virology**, London, v. 445 (1-2), p. 80-98, set. 2013.

DROP, B.; STRYCHARZ-DUDZIAK, M.; KLISZCZEWSKA, E.; POLZ-DACEWICZ, M. Coinfection with Epstein-Barr virus (EBV), human papillomavirus (HPV) and polyoma BK virus (BKPyV) in laryngeal, oropharyngeal and oral cavity cancer. **International Journal of Molecular Sciences**, v. 18, dez. 2017.

ENAM, S. A. Human papillomavirus and human cytomegalovirus infection and association with prognosis in patients with primary glioblastoma in Pakistan.

World Neurosurgery, v. 121, jan. 2019.

FERREIRA, H. C. C.; CAMPOS, M. G.; VIDIGAL, P. M. P.; SANTOS, M. R.; DE CARVALHO, O. V.; BRESSAN, G. C.; FIETTO, J. L. R.; COSTA, E. P.; ALMEIDA, M. R.; SILVA JÚNIOR, A. Latent bovine herpesvirus 1 and 5 in milk from naturally infected dairy cattle. **The Journal of Veterinary Medical Science**, v. 80 (11), p. 1787-1790, out. 2018.

GOLDSTEIN, S. C.; BYRNE, J. C.; RABSON, A. S. Huma cytomegalovirus (HCMV) enhances bovine papillomavirus (BPV) transformation in vitro. **Journal of Medical Virology**, v. 23, p. 157-164, fev. 1987.

HAMAD, M. A.; AL-SHAMMARI, A. M.; ODISHO, S. M.; YASEEN, N. Y. Molecular epidemiology of bovine papillomatosis and identification of three genotypes in central Iraq. **Intervirology**, Basel, v. 60, p. 156-164, feb. 2018.

HATAMA, S.; NOBUTOMO, K.; KANNO, T. Genomic and phylogenetic analysis of two novel bovine papillomavirus, BPV-9 and BPV-10. **Journal of General Virology**, London, v. 89, p.158-163, jan. 2008.

HBIBI, A. T.; LAGORCE, C.; WIND, P.; SPANO, J. P.; GUETZ, G.; MILANO, G.; BENAMOUZIG, R.; RIXE, O.; MORERE, J. F.; BREAU, J. L.; MARTIN, A.; FAGARD, R. Identification of a functional EGF-R/p60c-src/STAT3 pathway in colorectal carcinoma: Analysis of its long-term prognostic value. **Cancer Biomarkers**, v. 4, p. 83-91, 2008.

JAREBORG, N.; BURNETT, S. Immunofluorescent detection of Bovine Papillomavirus E4 antigen in the cytoplasm of cells permissive to in vitro for viral DNA amplification. **Journal of General Virology**, London, v. 72, p. 2269-2274, set. 1991.

KARABADZHAK, A. G.; PETTI, L. M.; BARRERA, F. N.; EDWARDS, A. P. B.; MOYA-RODRÍGUEZ, A.; POLIKANOV, Y. S.; FREITES, J. A.; TOBIAS, D. J.; ENGELMAN, D. M.; DIMAIO, D. Two transmembrane dimers of the bovine papillomavirus E5 oncoprotein clamp the PDGF beta receptor in an active dimeric conformation. **Proceedings of the National Academy of Sciences of the United States of America**, Washington DC, v. 114 (35), p. 7262-7271, ago. 2017.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. **Biology and Evolution**, v. 33 (7), p. 1870-1874, 2016.

LARSEN, R. W.; KIUPEL, M.; BALZER, H. J.; AGERHOLM, J. S. Prevalence of canid herpesvirus-1 infection in stillborn and dead neonatal puppies in Denmark. **Acta Veterinaria Scandinavica**, v. 2015 (57), 2015.

LIAO, P.; SATTEN, G. A.; HU, Y. PhredEM: A phred-score-informed genotype-calling approach for next-generation sequencing studies. **Genetic Epidemiology**, Medford, v. 41 (5), p. 375-387, jul. 2017.

LINDSEY, C.J.; ALMEIDA, M.E.; VICARI, C.F.; CARVALHO, C.; YAGUIU, A.; FREITAS, A.C. Bovine papillomavirus DNA in milk, blood, urine, semen, and

spermatozoa of bovine papillomavirus-infected animals. **Genetics and Molecular Research**, v. 8 (1):310–8, 2009.

LÓPEZ-BUENO, A.; MAVIAN, C.; LABELLA, A. M.; CASTRO, D.; BORREGO, J. J.; ALCAMI, A.; ALEJO, A. Concurrence of iridovirus, polyomavirus and an unique member of a new group of fish papillomaviruses in lymphocystis disease-affected gilthead sea bream. **Journal of Virology**, v. 90, p. 8768-8779.

MAJUMDER, S.; RAY, P.; BESMER, P. Tyrosine protein kinase activity of the HZ4-feline sarcoma virus p80-gag-kit-transforming protein. **Oncogene Research**, v. 5 (4), 1990.

MARINS, R. S. de Q. S. Phylogeny of papillomavirus and its correlation with human papillomavirus (HPV) and animals – review. **Publicações em Medicina Veterinária e Zootecnia**, Londrina, v. 1 (7), 20 p., nov. 2015.

MASELLA, A. P.; LALANSINGH, C. M.; SIVASUNDARAM, P.; FRASER, M.; BRISTOW, R. G.; BOUTROS, P. C. BAMQL: a query language for extracting reads from BAM files. **BMC Bioinformatics**, v. 2016, (17), 2016.

MCKINNEY, C. C.; HUSSMANN, K. L.; MCBRIDE, A. A.; The role of the DNA damage response throughout the papillomavirus life cycle. **Viruses**, v. 7, p. 2450-2469, 2015.

MENDONZA-PARRA, M. A.; SALEEM, M. A. M.; BLUM, M.; CHOLLEY, P. E.; GRONEMEYER, H. NGS-QC Generator: a quality control system for ChIP-seq and related deep sequencing-generated datasets. **Statistical Genomics: Methods and Protocols**, New York, v. 1418, 2016.

MÓDOLO, D. G.; ARALDI, R. P.; MAZZUCHELLI-DE-SOUZA, J.; PEREIRA, A.; PIMENTA, D. C.; ZANPHORLIN, L. M.; BEÇAK, W.; MENOSSI, M.; STOCCO, R. de C.; CARVALHO, R. de F. Integrated analysis of recombinant BPV-1 L1 protein for the production of a bovine papillomavirus VLP vaccine. **Vaccine**, New York, v. 35, p. 1590-1593, fev. 2017.

MUNDAY, J. S. Bovine and human papillomavirus: a comparative review. **Veterinary Pathology**, Los Angeles, v. 51 (6), p. 1063-1075, jun. 2014.

OLUWAYELU, D.; WERNIKE, K.; ADEBIYI, A.; CADMUS, S.; BEER, M. Neutralizing antibodies against Simbu serogroup viruses in cattle and sheep, Nigeria, 2012-2014. **BMC Veterinary Research**, v. 14 (277), 2018.

PAGLIARINI, R.; SHAO, W.; SELLERS, W. R. Oncogene addiction: pathways of therapeutic response, resistance, and road maps toward a cure. **EMBO Reports**, v. 16, p. 280-296, fev. 2015.

PATEL, K. K.; STANISLAWEK, W. L.; BURROWS, E.; HEUER, C.; ASHER, G. W.; WILSON, P. R.; HOWE, L. Investigation of association between bovine viral diarrhoea virus and cervid herpesvirus type-1, and abortion in New Zealand farmed deer. **Veterinary Microbiology**, v. 228, p. 1-6, nov. 2018.

PIETRO, M. Di; FILARDO, S.; PORPORA, M. G.; RECINE, N.; LATINO, M. A.; SESSA, R. HPV/Chlamydia trachomatis co-infection: metagenomic analysis of cervical microbiota in asymptomatic women. **New Microbiologica**, Pavia, v. 41 (1), p. 34-41, jun. 2017.

POL, S. B. V.; KLINGELHUTZ, A. J. Papillomavirus E6 oncoproteins. **Virology**, London, v. 445, p. 115-137, mai. 2013.

RAJPUT, M. K. S.; ABDELSALAM, K.; DARWEESH, M. F.; BRAUN, L. J.; KERKVLIT, J.; HOPPE, A. D.; CHASE, C. C. L. Both cytopathic and non-cytopathic bovine viral diarrhoea virus (BVDV) induced autophagy at a similar rate.

Veterinary Immunology and Immunopathology, 2017.

RECTOR, A.; VAN RANST, M. Animal papillomaviruses. **Virology**, London, v. 445, p. 213-223, mai. 2013.

ROPERTO, S.; RUSSO, V.; LEONARDI, L.; MARTANO, M.; CORRADO, F.; RICCARDI, M. G.; ROPERTO, F. Bovine papillomavirus type 13 expression in the urothelial bladder tumors of cattle. **Transboundary and Emerging Diseases**, Medford, v. 63 (6), p. 628-634, jan. 2015.

ROSENBERG, N.; WITTE, O. N. Abelson murine leukemia virus mutants with alterations in the virus-specific P120 molecule. **Journal of Virology**, v. 33 (1), p. 340-348, jan. 1980.

ROSKOSKI JR., R. Src protein-tyrosine kinase structure and regulation. **Biochemical and Biophysical Research Communications**, v. 324, p. 1155-1164, out. 2004.

SALEM, N. B.; DE VAL, B. P.; MARTIN, M.; MOENS, U.; EHLERS, B. Genome sequence of bovine polyomavirus 1 detected in a salers cow (*Bos Taurus*) from Catalonia, Spain. **Genome Announcements**, v. 4 (1), jan. 2016.

SANTOS, C.; VILANOVA, M.; MEDEIROS, R.; GIL DA COSTA, R. M. HPV-transgenic mouse models: tools for studying the cancer-associated immune response. **Virus Research**, v. 235, p. 49-57, mai. 2017.

SANTOS, E.U.D.; SILVA, M.A.R.; PONTES, N.E.; COUTINHO, L.C.A.; PAIVA, S.S.L.; CASTRO, R.S. Detection of Different Bovine Papillomavirus Types and Co-infection in Bloodstream of Cattle. **Transboundary and Emerging Diseases**, v. 63(1):e103–8, 2016.

SARTORI, C.; BOSS, R.; BODMER, M.; LEUENBERGEN, A.; IVANOVIC, I.; GRABER, H. U. Sanitation of Staphylococcus aureus genotype B-positive dairy herds: A field study. **Journal of Dairy Science**, Urbana, v. 101 (8), p. 1-18, ago. 2018.

SAVINI, F.; MANCINI, S.; GALLINA, L.; DONATI, G.; CASÀ, G.; PELI, A.; SCAGLIARINI, A. Bovine papillomatosis: first detection of bovine papillomavirus types 6, 7, 8, 10 and 12 in Italian cattle herds. **The Veterinary Journal**, London, v. 210, p. 82-84, fev. 2016.

SCAGLIARINI, A.; CASÀ, G.; TRENTIN, B.; GALLINA, L.; SAVINI, F.; MORENT, M.; LAVAZZA, A.; PULEIO, R.; BUTTACI, C.; CANNELLA, V.; PURPARI, G.; DI MARCO, P.; PIQUEMAL, D.; GUERCIO, A. Evidence of zoonotic Poxviridae coinfections in clinically diagnosed papillomas using a newly developed mini-array test. **Journal of Veterinary Diagnostic Investigation**, v. 28 (1), 2015.

SCHMITT, M.; FIEDLER, V.; MÜLLER, M. Prevalence of BPV genotypes in a German cowshed determined by a novel multiplex BPV genotyping assay. **Journal of Virological Methods**, London, v. 170 (1-2), p. 67-72, dez. 2010.

SHARMA, R.; BHAT, T. K.; SHARMA, O. P. The environmental and human effects of ptaquiloside-induced enzootic bovine hematuria: a tumorous disease in

cattle. **Reviews of Environmental Contamination and Toxicology**, New York, v. 224, p. 53-95, nov. 2012.

SILVA, M. S. e; WEISS, M.; BRUM, M. C. S.; DOS ANJOS, B. L.; TORRES, F. D.; WEIBLEN, R.; FLORES, E. F. Molecular identification of bovine papillomaviruses associated with cutaneous warts in southern Brazil. **Journal of Veterinary Diagnostic Investigation**, Guelph, v. 22, p. 603-606, jul. 2010.

SSEDYABANE, F.; AMNIA, D. A.; MAYANJA, R.; OMONIGHO, A.; SSUUNA, C.; NAJJUMA, J. N.; FREDDIE, B. HPV-Chlamydial coinfection, prevalence and association with cervical intraepithelial lesions: a pilot study at Mbarara Regional Referral Hospital. **Journal of Cancer Epidemiology**, v. 2019, jan. 2019.

STEIN, L. Genome annotation: from sequence to biology. **Nature Reviews**, New York, v. 2, p. 493-505, jul. 2001.

TORRES, R. de A.; DA SILVA, A. A.; DIAS, E. R. G. D. Anais do 4º Rio Leite Serrano. **Embrapa Gado de Leite**, Juíz de Fora, v. 1, 172 p., nov. 2007.

TRAPNELL, C.; ROBERTS, A.; GOFF, L.; PERTEA, G.; KIM, D.; KELLEY, D. R.; PIMENTEL, H.; SALZBERG, S. L.; RINN, J. L.; PACHTER, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nature Protocols**, London, v. 7 (3), p. 562-578, mar. 2013.

VAZQUEZ-GUILLEN, J. M.; PALACIOS-SAUCEDO, G. C.; RIVERA-MORALES, L. G.; ALONZO-MORADO, M. V.; BURCIAGA-BERNAL, S. B.; MONTUFAR-MARTINEZ, M.; ORTIZ-LOPEZ, R.; GONZALEZ-VILLASANA, V.; MARTINEZ-TORRES, A. C.; SERNA-HERNANDEZ, J. C.; HERNANDEZ-MARTINEZ, S. J.; CASTELAN-MALDONADO, E. E.; ZAVALA-POMPA, A.; MONTALVO-

BAÑUELOS, M. S.; GARCIA-CABELLO, R.; SANCHEZ-FRESNO, E. C.; RODRIGUEZ-PADILLA, C. Infection and coinfection by human papillomavirus, Epstein-Barr virus and Merkel cell polyomavirus in patients with squamous cell carcinoma of the larynx: a retrospective study. **PeerJ**, out. 2018.

WANG, Y.; LI, J.; LI, Y.; FANG, L.; SUN, X.; CHANG, S.; ZHAO, P.; CUI, Z. Identification of avian leukosis virus subgroup J-associated acutely transforming viruses carrying the v-src oncogene in layer chickens. **Journal of General Virology**, v. 97, p. 1240-1248, fev. 2016.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-seq: a revolutionary tool for transcriptomics. **Nature Review Genetics**, London, v. 10 (1), p. 57-63, jan. 2009.

WATANABE, S.; IIZUKA, T.; HATAMA, S.; KANNO, T.; MASE, M.; SHIBAHARA, T. Production on highly immunogenic virus-like particles of bovine papillomavirus type 6 in silkworm pupae. **Vaccine**, New York, v. 35 (43), p. 5878-5882, out. 2017.

WEN, L.; MAO, A.; FAN, Z.; LI, W.; XIAO, Q.; LIU, Q.; XIE, J.; HE, K. Porcine circovirus-like virus P1 in cattle, goats and rabbits in China. **Transboundary and Emerging Diseases**, mar. 2017.

WHITE, E. A.; KRAMER, R. E.; HWANG, J. H.; FERNANDO, A. T. P.; NAETAR, N.; HAHN, W. C.; ROBERTS, T. M.; SCHAFFHAUSEN, B. S.; LIVINGSTON, D. M.; HOWLEY, P. M. Papillomavirus E7 oncoproteins share functions with polyomavirus small T antigens. **Journal of Virology**, v. 89 (5), p. 2857-2865, dez. 2014.

WILLNER, D.; FURLAN, M.; HAYNES, M.; SCHMIEDER, R.; ANGLY, F. E.; SILVA, J.; TAMMADONI, S.; NOSRAT, B.; CONRAD, D.; ROHWER, F. Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. **Plos One**, v. 4 (10), out. 2009.

XIE, F.; DESEMET, M.; KANGINAKUDRU, S.; JOSE, L.; CULLETON, S. P.; GILSON, T.; LI, C.; ANDROPHY, E. J. Kinase activity of fibroblast growth factor receptor 3 regulates activity of the papillomavirus E2 protein. **Journal of Virology**, v. 91 (20), ago. 2017.

XIONG, Y. Q.; MO, Y.; LUO, Q. M.; HUO, S. T.; HE, W. Q.; CHEN, Q. The risk of human papillomavirus infection for spontaneous abortion, spontaneous preterm birth, and pregnancy rate of assisted reproductive technologies: a systematic review and meta-analysis. **Gynecologic and Obstetric Investigation**, v. 2018.

YAGUIU, A.; DAGLI, M. L. Z.; BIRGEL JR, E. H.; ALVES REIS, B. C. A.; FERRAZ, O. P.; PITUCO, E. M.; FREITAS, A. C.; BEÇAK, W.; STOCCO, R. C. Simultaneous presence of bovine papillomavirus and bovine leukemia virus in different bovine tissues: in situ hybridization and cytogenetic analysis. **Genetics and Molecular Research**, Ribeirão Preto, v. 7 (2), p. 487-497, jun. 2008.

ZERBONI, L.; SEN, N.; OLIVER, S. L.; ARVIN, A. M. Molecular mechanisms of varicella zoster virus pathogenesis. **Nature Reviews Microbiology**, v. 12 (3), p. 197-210, jun. 2014.