

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE VETERINÁRIA
TRABALHO DE CONCLUSÃO EM MEDICINA VETERINÁRIA**

**SEQUENCIAMENTO DE ÚLTIMA GERAÇÃO APLICADO À VIROLOGIA
VETERINÁRIA**

Elaborado por Mariana Soares da Silva
Acadêmica em Medicina Veterinária

**PORTO ALEGRE
2015**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE VETERINÁRIA
TRABALHO DE CONCLUSÃO EM MEDICINA VETERINÁRIA**

**SEQUENCIAMENTO DE ÚLTIMA GERAÇÃO APLICADO À VIROLOGIA
VETERINÁRIA**

Autora: Mariana Soares da Silva

**Trabalho apresentado à Faculdade de
Veterinária como requisito parcial para
a obtenção da graduação em
Medicina Veterinária.**

Orientador: Cláudio Wageck Canal

Co-orientadora: Renata da Fontoura Budaszewski

PORTO ALEGRE

2015

RESUMO

Os vírus são microrganismos que ocorrem em todos os seres vivos e possuem altas taxas de mutação, sendo assim, necessitam de técnicas adequadas para sua detecção. A metagenômica viral permite a investigação da flora viral de animais saudáveis e doentes através de uma série de plataformas, denominadas “*Next-Generation Sequencing*” (NGS), que estão disponíveis no mercado e utilizam diferentes processos de sequenciamento do genoma. Através dessas metodologias, pode-se melhorar o entendimento da etiologia de doenças, assim como aprofundar o conhecimento sobre a circulação viral na natureza e sua interação com os hospedeiros. O objetivo do presente trabalho é realizar uma revisão bibliográfica sobre as diferentes plataformas de NGS e suas aplicações na virologia, assim como relatar a detecção de novas espécies e cepas virais na medicina veterinária através dessas plataformas.

Palavras-chave: vírus, metagenômica, *next-generation sequencing*, veterinária, sequenciamento.

ABSTRACT

Viruses are microorganisms that affect all living beings, with high mutation rate, therefore need appropriate detection techniques. Viral metagenomics allows the viral flora research of healthy and sick animals with a variety approaches, called “Next-Generation Sequencing” (NGS), which are available in the market and use different genome sequencing procedures. Through these methods is possible to improve diseases etiology, increasing viral circulation knowledge in the nature and host interaction. The aim of the present study is to perform a bibliographic review about different NGS platforms and their virology applications, as well as report the detection of new viruses or viral strains in veterinary medicine through these platforms.

Keywords: viruses, metagenomic, next-generation sequencing, veterinary, sequencing.

AGRADECIMENTOS

Agradeço, primeiramente, ao meu pai, que apesar de não estar mais aqui, foi o meu maior incentivador, nunca permitindo que eu desistisse dos meus objetivos e me ensinando que eu deveria fazer aquilo que me fizesse feliz. Pai, dedico esse momento à ti... Saudades.

À minha mãe, minha companheira, minha amiga, que sempre fez tudo por mim. Mãe, tu és a razão da minha caminhada, obrigada por ser essa mãe excepcional. Eu te amo demais.

Ao Bruno, por ter percorrido essa jornada comigo durante esses quase 6 anos, obrigada por me apoiar, me incentivar e compreender minhas ausências. Te amo.

À minha família, obrigada pelo incentivo e por acreditarem em mim. Em especial aos meus dindos, e a minha irmã Daniela, sempre presentes.

Aos presentes que a FAVET me deu, minhas amigas Ananda e Cibele. Ananda, obrigada por toda a dedicação e amizade, por todos os momentos em que estive ao meu lado, inclusive nos mais difíceis da minha vida. Cibele, te agradeço imensamente por todas as conversas, conselhos, companhia, aprendi muito contigo e te guardo no meu coração.

Às minhas amigas de longa data, Karen e Aline. Agradeço pela amizade e pelos momentos que compartilhamos, obrigada por compreenderem que nem sempre pude estar com vocês.

Aos colegas do Laboratório de Virologia. Não tenho palavras para agradecer pelos ensinamentos, por me incentivarem profissionalmente e também pela amizade diária. Em especial, agradeço à minha co-orientadora Renata, e ao meu orientador Cláudio Canal. Professor, muito obrigada pelos ensinamentos e pelas oportunidades.

E a todos que de alguma forma fizeram parte dessa jornada!

LISTA DE FIGURAS

Figura 1 – Custo médio do sequenciamento de 1.000 bases.....	10
Figura 2 – Etapas do método de sequenciamento da plataforma Roche/454.....	12
Figura 3 – Cascata de eventos da técnica de pirosequenciamento.....	13
Figura 4 – Amplificação em ponte utilizada na plataforma Illumina.....	14
Figura 5 – Técnica de sequenciamento por síntese utilizada na plataforma Illumina.....	15
Figura 6 – Metodologia de sequenciamento de ácidos nucleicos da plataforma Ion Torrent.....	17

ABREVIATURAS

ddNTP dideoxynucleotídeos Trifosfatados

DNA Ácido desoxirribonucleico

dNTP deoxynucleotídeos Trifosfatados

NGS Next-Generation Sequencing

pb pares de bases

PCR reação em cadeia da polimerase

PPi fosfato inorgânico

PPT placa de picotitulação

gDNA DNA genômico

ssDNA DNA fita simples

SUMÁRIO

1	INTRODUÇÃO	8
2	REVISÃO BIBLIOGRÁFICA	10
2.1	Plataformas de sequenciamento de DNA de última geração	10
2.1.1	<i>Roche 454 sequencing</i>	11
2.1.2	<i>Illumina Sequencers</i>	13
2.1.3	<i>SOLiD (Sequencing by Oligonucleotide Ligation and Detection)</i>	15
2.1.4	<i>Ion Torrent</i>	16
2.2	Preparo das amostras	17
2.3	Aplicações na virologia	18
2.3.1	Metagenômica viral.....	19
2.3.2	Sequenciamento de genoma viral completo.....	19
2.3.3	Detecção de novos vírus ou cepas virais na medicina veterinária.....	20
2.3.3.1	Caninos.....	20
2.3.3.2	Felinos.....	22
2.3.3.3	Suínos.....	22
2.3.3.4	Bovinos.....	22
2.3.3.5	Animais silvestres.....	23
2.3.3.6	Morcegos.....	24
3	CONCLUSÕES	25
	REFERÊNCIAS	26

1 INTRODUÇÃO

A emergência de novos vírus e a descoberta de novas doenças virais aumentaram consideravelmente nas últimas décadas devido a diversos fatores, como intensa globalização, mudança climática, evolução viral, entre outros. Muitos vírus, em particular os vírus de genoma RNA e os de DNA fita simples, possuem tempos de geração curtos e alta taxa de mutação, que ultrapassa em seis ordens de grandeza as mutações existentes em humanos (DUFFY; SHACKELTON; HOLMES, 2008). Sendo assim, é crucial o emprego de métodos adequados para detecção e identificação dessa ampla gama de agentes (BELÁK *et al.*, 2013).

Historicamente, as técnicas clássicas de detecção viral eram baseadas na inoculação intracerebral de animais de laboratório ou em isolamento viral em cultivo celular e visualização do efeito citopático. Após, os métodos imunológicos, como a hemaglutinação e a soroneutralização, começaram a ser usados para a detecção de antígenos virais nas diferentes amostras. Essas técnicas eram baseadas no isolamento dos agentes virais. Com o progresso da biologia molecular, a reação em cadeia da polimerase (PCR) tornou-se a principal técnica na descoberta de vírus e permitiu a detecção de vírus não cultiváveis (TEMMAN *et al.*, 2014).

Em 1977, Frederick Sanger (SANGER *et al.*, 1977) descreveu a primeira técnica de sequenciamento de DNA com o desenvolvimento de um método manual, também conhecido como método da cadeia interrompida, em função da utilização de dNTPs (deoxinucleotídeos Trifosfatados) modificados, os ddNTPs (dideoxinucleotídeos Trifosfatados), os quais, quando incorporados pela enzima *Taq* polimerase, não permitem a inclusão do nucleotídeo seguinte. Nos anos 1990, houve a automatização da técnica, com a possibilidade de utilizar ddNTPs marcados com diferentes fluorescências para cada um deles, produzindo sequências de até 750 pares de bases (pb). Com o passar dos anos, desenvolveram-se as técnicas de metagenômica, conhecidas como “Sequenciadores de DNA de Última Geração” ou “*Next-Generation Sequencing*” (NGS), que tornaram possível sequenciar todo o genoma viral de uma amostra, produzindo uma verdadeira revolução e abrindo novas perspectivas para a pesquisa e aplicação na medicina clínica (CAPOBIANCHI; GIOMBINI; ROZERA, 2012).

O primeiro sistema de NGS disponível comercialmente, desenvolvida pela Life Sciences, foi lançado em 2005. Desde então, em um período de tempo relativamente curto, várias tecnologias NGS foram desenvolvidas. Os sistemas podem ser referidos como: sequenciadores de alta capacidade, tais como o Genoma Analyzer e HiSeq sequenciadores (Illumina, San Diego, CA, EUA), a plataforma Heliscope (Helicos BioSciences Corporation, Cambridge, MA, USA) e 5500 Series Genetic Analysis Systems, que usa a tecnologia SOLiD

(Applied Biosystems, Carlsbad, CA, EUA), e os sequenciadores de leitura longa: tais como o Genoma Sequencer FLX (454 Life Sciences, Roche Diagnostics, Branford, CT, USA), Ion Torrent (Applied Biosystems) e PacBio RS (Pacific Biosciences, Gen-Probe, Menlo Park, CA, EUA) (CAPOBIANCHI; GIOMBINI; ROZERA, 2012).

O termo metagenômica designa a análise de todo ácido nucleico presente em uma dada amostra permitindo a exploração de comunidades microbianas e evitando a necessidade de isolamento e culturas individuais, além de não haver a necessidade de conhecimento prévio das sequências (RIESENFELD; SCHLOS; HANDELSMAN, 2004; CAPOBIANCHI; GIOMBINI; ROZERA, 2012.). A metagenômica viral avalia o genoma de todos os vírus existentes em amostras ambientais, como em água de lagos e água de reuso (DJIKENG, 2009; ROSÁRIO, 2009) ou biológicas, como em fezes e secreções respiratórias (SHAN, 2011; LYSHOLM, 2012).

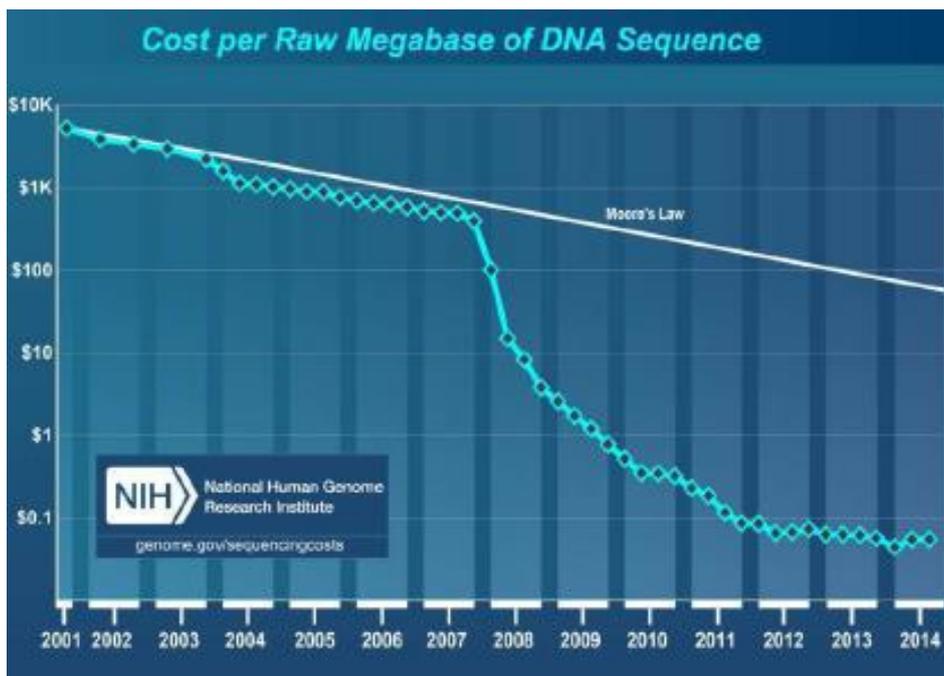
2 REVISÃO BIBLIOGRÁFICA

2.1 Plataformas de Sequenciamento de DNA de última geração

As plataformas NGS fornecem uma alta taxa de transferência, gerando centenas de gigabases (GB) de dados em um único experimento, além disso também são conhecidas como “sequenciamento intenso” em que muitas vezes o número total de *reads* (sequências curtas de leitura) é muito maior que o número de *reads* necessários para cobrir o comprimento da sequência de DNA estudada, gerando uma sobreposição. Embora o investimento inicial e o custo por corrida sejam altos atualmente, o valor por unidade de informação (nucleotídeo) foi drasticamente reduzido quando comparado ao sequenciamento de primeira geração (Figura 1) (TOON, 2015). NGS são subdividas em sequenciadores de segunda (Roche 454, Illumina, Ion Torrent, SOLiD) e terceira geração (PacBio RS, HeliScope).

As plataformas de sequenciamento incluem etapas como a preparação da fita molde, sequenciamento, geração de imagem e análise de dados. A exclusiva combinação de protocolos específicos distingue as tecnologias e determina os tipos de dados produzidos a partir de cada plataforma. Estas diferenças estão relacionadas com a qualidade desses dados e com o custo de cada técnica (METZKER, 2010).

Figura 1 - Custo médio de sequenciamento de 1000 bases (dólar americano, escala logarítmica).



Fonte: Wetterstrand (2015).

Ao longo dos últimos 5-7 anos, todas as principais plataformas fizeram melhorias significativas, com avanços notáveis em termos de complexidade do protocolo, desempenho e melhor relação custo-benefício (TOON, 2015). O desenvolvimento de sequenciadores menores (de bancada) tornaram a tecnologia mais acessível para o uso em laboratórios de microbiologia (LOMAN, 2012).

2.1.1 Roche 454 *sequencing*

O sequenciador 454 de última geração da empresa Roche foi o primeiro a ser introduzido comercialmente (em 2004) e utiliza uma tecnologia de sequenciamento alternativa, conhecida como pirosequenciamento (RONAGHI *et al.*, 1998). Esse sequenciador possui uma metodologia baseada em três etapas (**Erro! Fonte de referência não encontrada.**2), sendo que a primeira compreende a preparação da biblioteca de DNA, em que o DNA genômico (gDNA) é fracionado em fragmentos menores que posteriormente são desnaturados, gerando fitas simples de DNA (ssDNA). Nesses fragmentos, dois pequenos adaptadores (A e B) são ligados em suas extremidades atuando como *primers* para a amplificação e sequenciamento dos mesmos. O adaptador B possui biotina ligada à extremidade 5', o que permite o isolamento dos fragmentos ligados ao adaptador A na extremidade 3' (MARGULIES *et al.*, 2005). O adaptador A é utilizado no anelamento do *primer* que inicia a reação de sequenciamento (DRESSMAN *et al.*, 2003).

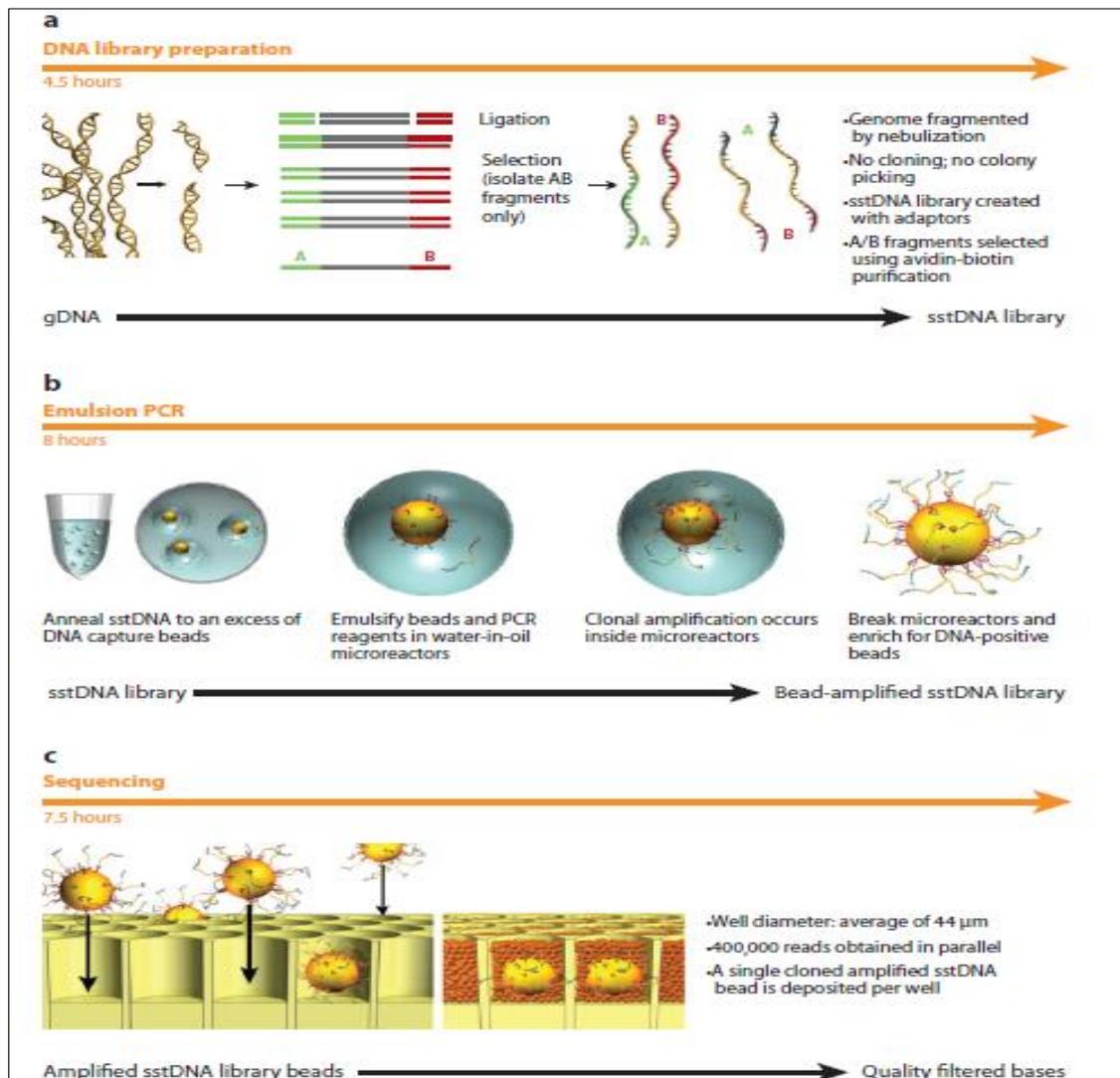
Os fragmentos da biblioteca de DNA são misturados à *beads* de agarose cuja superfície carrega oligonucleotídeos complementares às sequências de adaptadores 454-específicos e cada fragmento é associado à uma *bead*. Cada um desses complexos *bead*:fragmento é isolado individualmente em micelas de água e óleo que contém os reagentes da PCR, e os ciclos térmicos nas micelas produzem aproximadamente 1 milhão de cópias de cada fragmento de DNA na superfície de cada *bead*. As *beads* são arranjadas em uma placa de picotitulação (PPT), de forma que cada uma permaneça individualmente em um poço, permitindo o monitoramento das reações (MARDIS, 2008).

A placa de sequenciamento é inserida junto ao sistema óptico de leitura no equipamento. As soluções e os reagentes de sequenciamento são então distribuídos por toda a placa a cada ciclo para obtenção do sequenciamento paralelo dos 1,6 milhões de poços. O sequenciamento é realizado em ciclos e, a cada ciclo, um tipo determinado de nucleotídeo é adicionado à reação. No pirosequenciamento, a cascata de reações inicia com a polimerização dos ácidos nucleicos

em que o pirofosfato inorgânico (PPi) é liberado como resultado da incorporação de nucleotídeos. A liberação de PPi é subsequentemente convertida em ATP pela ATP sulfúrilase que fornece energia para a luciferase oxidar a luciferina e gerar fluorescência (Figura 3). Uma vez que o nucleotídeo adicionado é reconhecido, a sequência é revelada (RONAGHI, 2001).

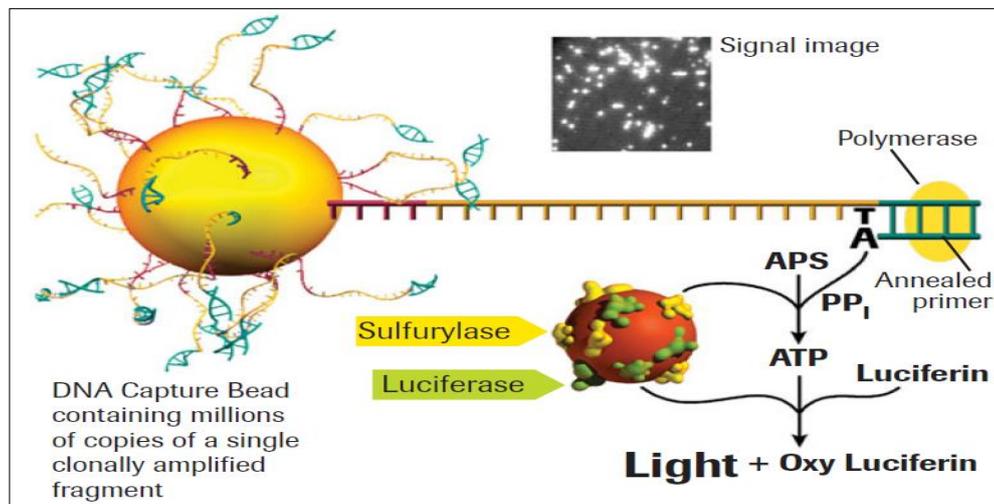
Atualmente, existem duas plataformas no mercado usando essa tecnologia, os sistemas GS FLX+ e o GS Júnior+. O GS FLX+ pode gerar aproximadamente 700 Mb de dados de sequência em um dia, com comprimentos de *reads* de até 1.000 pb. O GS Júnior+ é essencialmente uma versão de bancada por ser menor, gera cerca de 70 Mb por corrida com *reads* entre 700-800 pb de comprimento.

Figura 2 - Etapas do método de sequenciamento da plataforma Roche/454. (a) Preparação da biblioteca de DNA; (b) PCR em emulsão; (c) Sequenciamento.



Fonte: Mardis (2008).

Figura 3 - Cascata de eventos da técnica de pirosequenciamento.



Fonte: Mardis (2008).

2.1.2 Illumina Sequencing

As técnicas de NGS da empresa Illumina (formalmente conhecida como Solexa) foram introduzidas no ano de 2007. Devido a sua alta capacidade, logo tornou-se amplamente utilizada nos projetos de sequenciamento de genoma total, incluindo humanos e demais organismos. As plataformas HiSeq 2000, HiSeq 1000 e *Genome Analyzer IIx* geram resultados de sequenciamento de 600, 300 e 95 Gb, respectivamente. A plataforma MiSeq pode gerar até 150 pb de sequenciamento, com um rendimento total de 1,5-2 Gb por corrida. Em 2012, a Illumina apresentou a plataforma HiSeq2500 como uma atualização de HiSeq2000. Esta nova plataforma pode gerar até 120 Gb de dados em 27 h, permitindo aos pesquisadores a sequência de um genoma inteiro em 24 h (ou seja, *Genome in a Day*) (SHOKRALLA *et al.*, 2008). A NextSeq é uma plataforma que comporta uma quantidade média de dados (até 120 Gb). Até o momento, a tecnologia NGS Illumina é dominante no mercado, com baixa taxa de erro e custos operacionais aceitáveis (TOON, 2015).

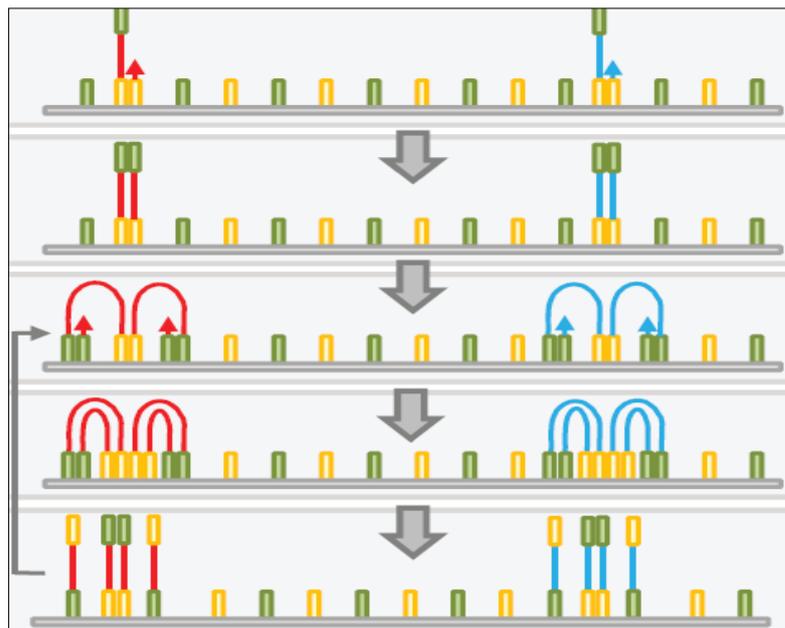
A síntese é realizada através da enzima DNA polimerase e nucleotídeos terminadores marcados com diferentes fluoróforos. A inovação dessa plataforma consiste na clonagem *in vitro* dos fragmentos em uma plataforma sólida de vidro, processo também conhecido como PCR de fase sólida (FEDURCO, 2006). Nessa tecnologia, o DNA é fragmentado aleatoriamente e, após a fragmentação, dois tipos diferentes de adaptadores são ligados às extremidades dos fragmentos. Em seguida, esses fragmentos são desnaturados, distribuídos e imobilizados (por uma das extremidades), também aleatoriamente, em uma superfície sólida

que utiliza a tecnologia denominada *flow cell*, a qual é revestida por uma camada densa de oligonucleotídeos complementares aos dois tipos de adaptadores dos fragmentos (ANSORGE, 2009).

Esses oligonucleotídeos atuam na captura da fita molde de DNA e também como *primers* para a amplificação subsequente que ocorre por um processo denominado de amplificação em ponte (

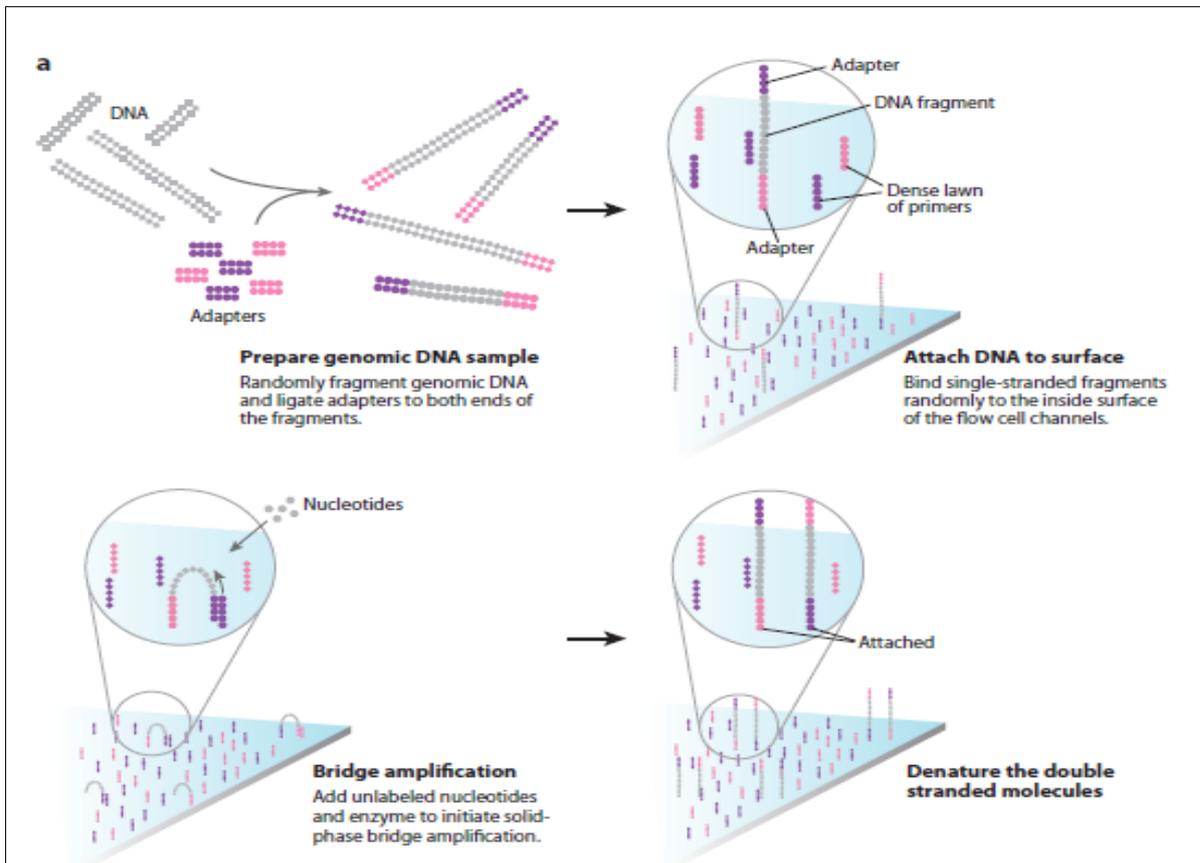
Figura 44), em que ambas extremidades das fitas simples de DNA ligam-se aos oligonucleotídeos fixados na *flow cell*. As sucessivas etapas de PCR geram pequenas ilhas ou *clusters* que servem como clones para o subsequente sequenciamento. A plataforma utiliza o método de sequenciamento por síntese (Figura 5), em que para cada *cluster* é fornecida a polimerase e os quatro nucleotídeos marcados com diferentes fluorescências que tem sua terminação 3'OH quimicamente inativada, garantindo que apenas uma base seja incorporada por ciclo. Depois que cada nucleotídeo é incorporado, há a geração de uma imagem que permite sua identificação em cada *cluster*. Um tratamento de desbloqueio químico remove o radical fluorescente e a incorporação do nucleotídeo seguinte é realizada no próximo ciclo. A sequência de cada *cluster* é computada e submetida a filtração para eliminar os *reads* de baixa qualidade (SHENDURE, 2008).

Figura 4 - Amplificação em ponte utilizada na plataforma Illumina.



Fonte: Radford *et al.* (2012).

Figura 5 - Técnica de sequenciamento por síntese utilizada na plataforma Illumina.



Fonte: Mardis (2008).

Essa plataforma pode produzir leituras de até 150 pb e possui como principal vantagem a elevada qualidade dos dados gerados e como limitação o tamanho reduzido das leituras (METZKER, 2010).

2.1.3 SOLiD (*Sequencing by Oligonucleotide Ligation and Detection*)

Diferentemente dos demais processos, no sistema SOLiD, a reação de sequenciamento é realizada por uma DNA ligase e não por uma polimerase. Os fragmentos de 60-90 pb são diretamente ligados a adaptadores universais (P1 e P2) em ambas as extremidades. Uma vez selecionados, os fragmentos são circularizados e clivados com uma enzima de restrição que reconhece seu sítio no adaptador e cliva adiante, liberando fragmentos formados por: 27 bases de uma região, a sequência dos adaptadores e 27 bases adicionais de outra região que está separada da primeira pela distância utilizada no intervalo de seleção dos fragmentos. O

adaptador P1 é utilizado no anelamento do *primer* da PCR de emulsão (CARVALHO; SILVA, 2010).

Os fragmentos de DNA ligados aos adaptadores são imobilizados em uma superfície magnética de 1 mm (*beads*). Essas *beads* são amplificadas individualmente através de PCR em emulsão. Após a amplificação, elas são covalentemente ligadas à uma superfície de vidro especialmente tratada que dará início ao sequenciamento. O sistema SOLiD™, portanto, gera o resultado do sequenciamento no espaço de cores e não no espaço de bases, sendo necessária uma decodificação dos sinais das leituras pela combinação dos dados. Como as bases do adaptador P1 são conhecidas, é possível fazer a identificação da primeira base do fragmento, quando da segunda etapa do processo de sequenciamento. Daí em diante, os demais sinais são especificados pela única combinação possível de cores, levando-se em consideração a base conhecida (CARVALHO; SILVA, 2010).

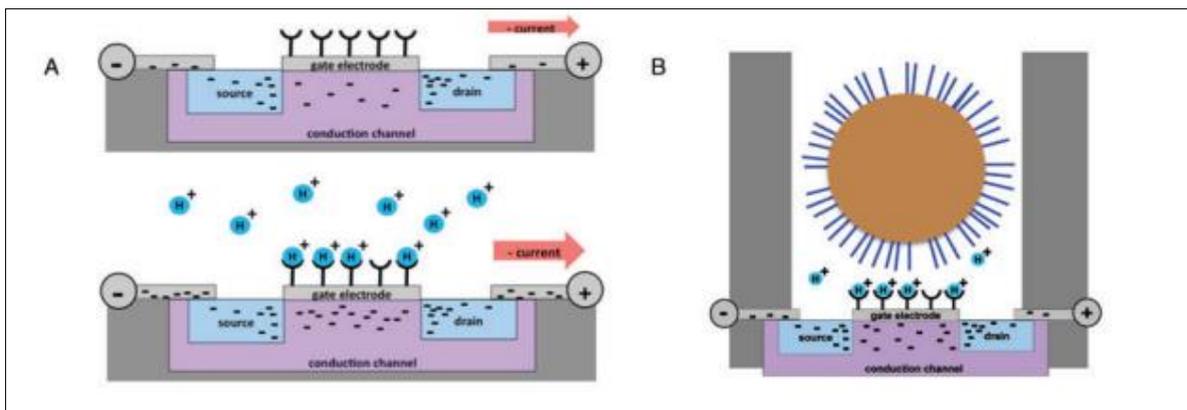
Essa plataforma promove alta acurácia de dados, com comprimentos máximos de *reads* relativamente curtos (75 pb). Por isso SOLiD tem sido utilizada principalmente para aplicações em que não há necessidade de uma nova montagem dos *reads* (*de novo assembly*). Atualmente, existem duas plataformas SOLiD disponíveis, a 5500 e a 5500x1 (TOON, 2015).

2.1.4 Ion Torrent

Ion Torrent é uma plataforma de NGS semicondutora que detecta, através de um *microchip*, a liberação de íons hidrogênio emitida durante a polimerização do DNA (DAUM *et al.*, 2012), sendo que a maioria dos outros sequenciadores utiliza bases marcadas com fluoróforos. O conceito básico é o de sequenciamento por detecção eletroquímica da síntese de DNA e que consiste na química fundamental de liberação de um hidrogênio (H^+) no sítio 3'OH a cada incorporação de nucleotídeos. Cada reação é acoplada em seu próprio sensor, que por sua vez são organizadas em uma matriz de sensores paralelos massivos em um *CMOSchip*. Esse sensor é um transistor de H^+ , conhecido como pHFET (transistor de efeito de campo sensível ao pH), que converte essa liberação em um sinal elétrico (MERRIMAN *et al.*, 2012). Embasando essas considerações, os elementos fundamentais do processo de sequenciamento estão ilustrados na Figura 6.

A concentração local de H^+ cria uma voltagem positiva perto da região de entrada do transistor, que resulta em uma mudança de corrente que flui através dele. Para o sequenciamento, um sensor é implantado para servir de pHFET na porção inferior de um micro poço e uma *bead* é usada para depositar o fragmento de DNA específico da amostra a ser sequenciado em cada poço. A *bead* possui uma população clonal de seqüências de DNA amplificada, produzida por um processo de PCR em emulsão. Essa *bead* normalmente transporta centenas de milhares da mesma cópia de fragmento de DNA de modo a amplificar o sinal químico produzido pela reação de sequenciamento (MERRIMAN *et al.*, 2012).

Figura 6 - Metodologia da plataforma Ion Torrent. Em (A) a ilustração do princípio básico de um transistor usado para detectar os íons de hidrogênio (H^+). Em (B) o sensor integrado ao sequenciamento de DNA, registra o H^+ produzido pela incorporação dos nucleotídeos pela ação da enzima polimerase, medindo assim, um sinal de incorporação.



Fonte: Merriman *et al.* (2012, Ilustração Modificada)

2.2 Preparo das Amostras

O benefício do uso de NGS para estimar a diversidade viral depende principalmente da qualidade da amostra e do seu processo de preparação. As seqüências de *reads* refletem o acúmulo de possíveis erros, alguns deles inevitáveis, no entanto, outros podem ser minimizados. Para diminuir a taxa de erro, as amostras requerem manuseio cuidadoso, desde a recuperação da amostra biológica até os últimos passos do procedimento de NGS (BEERENWINKEL *et al.*, 2012).

Para todos os métodos de NSG mencionados, a seleção da amostra e sua preparação é crucial. Um dos maiores obstáculos enfrentados na detecção de novos vírus utilizando as

plataformas NGS em amostras clínicas é a falta de sensibilidade e a contaminação com ácidos nucleicos do hospedeiro (BLOMSTÖM, 2010). Os ácidos nucleicos virais são frequentemente constituídos apenas de uma pequena proporção de ácidos nucleicos totais de uma amostra extraída. Sendo assim, é necessário concentrar as partículas virais e/ou degradar os ácidos nucleicos do hospedeiro. O método mais comum é através de centrifugação, filtração e tratamento com nucleases (THURBER *et al.*, 2009; DELWART, 2007).

Um caminho possível para aumentar a probabilidade de identificar os vírus é realizando um *pool* de um grande número de amostras suspeitas de conter um agente. Esse *pool* pode ser feito com facilidade quando se trabalha com amostras de fluidos, como soro e sangue. A ultracentrifugação pode ser usada subsequentemente para concentrar os vírus. As bactérias e complexos celulares maiores podem ser removidos por filtração. Os filtros (0,22 µm) são frequentemente usados para remover bactérias, pois acredita-se que os vírus podem passar através deles (DELWART *et al.*, 2007). No entanto, recentemente, descobriu-se que alguns vírus são tão grandes quanto bactérias, como por exemplo, os mimivírus que possuem um genoma de aproximadamente 1,2 Mb e capsídeo de aproximadamente 500 nm (VAN ETTEN; LANE; DUNIGAN, 2010). As partículas virais também podem ser separadas de outras partículas baseado em sua densidade, utilizando colchões de sacarose ou separação por gradiente de cloreto de cério (THURBER *et al.*, 2009).

As nucleases também são utilizadas para reduzir os ácidos nucleicos do hospedeiro. O princípio por trás dessa técnica é o de que o capsídeo viral protege os ácidos nucleicos virais da ação das nucleases e apenas os ácidos nucleicos do hospedeiro é degradado (ALLANDER *et al.*, 2001). Esta noção é verdadeira e atualmente o tratamento com DNase ou RNase está incluído na maioria das plataformas quando a abordagem trata de metagenômica viral. Ainda assim, deve-se levar em conta que diferentes vírus possuem estabilidades diferentes e podem ter sensibilidade variável a esse tipo de tratamento (BLOMSTÖM, 2010).

2.3 Aplicações na Virologia

As aplicações da metagenômica na virologia são diversas e incluem a descoberta de novos vírus a partir de amostras clínicas de humanos e animais doentes, como por exemplo, o

novo vírus Ebola Bundiubugyo associado a doença hemorrágica que foi identificado através da plataforma 454 Life Sciences (TOWNER *et al.*, 2008), a identificação da etiologia viral de um surto de uma doença em abelhas (COX-FOSTER *et al.*, 2007), além da caracterização do viroma, ou seja, da comunidade viral no ambiente (DJIKENG *et al.*, 2009), em animais (GE *et al.*, 2012) e em seres humanos (NAKAMURA *et al.*, 2009; BIBBY *et al.*, 2011).

2.3.1 Metagenômica viral

As técnicas de metagenômica viral abrem novas possibilidades para análises comparativas diretas entre várias amostras clínicas e para detecção de novos vírus emergentes. Um exemplo prático recente foi a detecção do vírus Schmallenberg, um novo ortobuniavírus, na Europa (HOFFMANN *et al.*, 2012). A metagenômica tem a capacidade de detectar vírus como agentes isolados ou como componentes de infecções complexas multifatoriais. Considerando essa gama de usos, é importante compreender que eles não são necessariamente causadores de doença. Isso significa que não é suficiente informar sobre os achados de metagenômica, mas sim acoplar a uma gama de investigações complementares como a recuperação do agente viral por isolamento ou por genética reversa e através de estudos de patogenicidade e epidemiologia, considerando os postulados de Koch (BELÀK, *et al.* 2013).

Com o desenvolvimento das diferentes plataformas de NGS, tornou-se possível sequenciar todos os genomas virais em uma determinada amostra, sem a necessidade de conhecimento prévio sobre a sua natureza. Dessa forma, a metagenômica viral permite a descoberta de novas espécies virais dentro de uma amostra complexa e, devido aos custos decrescentes, o uso dessas técnicas hoje em dia vem aumentando exponencialmente (TEMMAM *et al.*, 2014). Esta nova ciência é um dos campos que avança de forma mais rápida na biologia e expande a nossa compreensão da diversidade, ecologia, evolução e funcionamento do mundo microbiano, bem como contribui para o surgimento de novas aplicações em diferentes áreas. O desenvolvimento contínuo e dinâmico de NGS representa uma oportunidade sem precedentes para aplicações de metagenômica e está expandindo a nossa capacidade de analisar comunidades microbianas de uma variedade de habitats e ambientes (ZHANG *et al.*, 2011)

2.3.2 Sequenciamentos de genoma viral completo

Uma das principais aplicações do uso dessas tecnologias é o sequenciamento do genoma viral completo que, embora sejam relativamente pequenos, possuem alto valor acadêmico. A necessidade de sequenciamento desses genomas completos permite compreender a evolução e epidemiologia molecular desses vírus, que são dificultadas quando a análise filogenética é realizada com sequências parciais (RADFORD *et al.*, 2012). Além disso, permite comparar múltiplos genomas afim de identificar similaridades e diferenças genéticas entre elas (SZPARA *et al.*, 2010).

Ao permitir o sequenciamento de genomas completos, NGS tem promovido a compreensão do surgimento e transmissão de importantes vírus (BAILLIE *et al.*, 2012). O desafio restante é reduzir o tempo de produção dessas sequências a um ponto em que a informação possa ser utilizada rotineiramente em tempo real no auxílio de controle de surtos para agências de proteção da saúde (RADFORD *et al.*, 2012). Estas abordagens têm sido usadas na medicina humana para sequenciamento do vírus da gripe pandêmica H5N1 (CROVILLE *et al.*, 2012), vírus da imunodeficiência humana (HIV) (HOPER *et al.*, 2011), herpesvírus humanos (SZPARA *et al.*, 2010) e outros vírus.

2.3.3 Detecção de novos vírus ou cepas virais na Medicina Veterinária

O conhecimento de novos vírus em animais expandiu nosso conhecimento em relação a diversidade viral e quanto a emergência de potenciais zoonoses. O advento de NGS tem permitido a descoberta de muitos novos vírus que acometem animais como, por exemplo, novas espécies de circovírus, astrovírus e bocavírus suíno (BAYLIS *et al.*, 2011; SHAN *et al.*, 2011), novos sapovírus, norovírus e dependovírus em leões marinhos (LI *et al.*, 2011), novos hepacivírus em cães afetados por um surto respiratório (KAPOORA *et al.*, 2011), novos anelovírus em focas (NG, T.F. *et al.*, 2011), novos astrovírus em tecidos cerebrais de martas com *Shaking Syndrome* (BLÖMSTROM *et al.*, 2010) e muitas outras famílias virais em humanos (GRENINGER *et al.*, 2010), patos, morcegos e insetos (DAY *et al.*, 2010).

2.3.3.1 Caninos

Os cães são animais de companhia que mantém contato íntimo com humanos, incluindo crianças e, frequentemente, também estão presentes no meio silvestre. Essa interação promove

a possibilidade de transmissão de zoonoses, como o vírus da raiva e rotavírus (COOK, 1989; LACKAY; KUANG; FUNG, 2008). Estudo realizado por Linlin Li e colaboradores, analisou, através de pirosequenciamento (*Roche 454*), amostras de fezes de 18 cães com diarreia e aproximadamente 10% dos *reads* obtidos eram de origem viral. Como resultado, identificaram coronavírus canino (CCoV), parvovírus canino tipo 2 (CPV 2) e rotavírus, além de uma nova espécie de vírus do gênero *Kobuvirus*, família *Picornaviridae*, provisoriamente chamado de kobuvírus canino (CKoV). Até o momento, há três espécies de kobuvírus reconhecidas pelo ICTV: Aishi vírus humano, kobuvírus bovino e kobuvírus suíno. O CKoV possui maior similaridade filogenética com o Aishi vírus que está associado com gastroenterite em humanos, sugerindo uma origem comum recente e uma possível transmissão interespecíes. Uma nova espécie de vírus do gênero *Sapovirus*, Família *Sapoviridae*, denominados provisoriamente de sapovírus canino tipo 1 e 2 (CaSaV-1/2) também foi identificada. Esses vírus são reportados por causarem diarreia em humanos, suínos e martas (CHIBA *et al.*, 1979; MARTELLA *et al.*, 2008c; GUO *et al.*, 2001).

Outro estudo realizado através de NGS detectou em amostras de suabe nasal de cães com um surto de doença respiratória, um novo vírus com alta similaridade ao vírus da hepatite C (HCV), denominado hepacivírus canino (CHV). Após, realizaram RT-PCR com *primers* específicos em 33 amostras de cães provenientes de cinco surtos de doenças respiratórias. Obtiveram, em um dos surtos, seis amostras positivas de um total de nove e, do outro surto, três positivas de um total de cinco amostras. Através da análise filogenética, esse estudo indicou que a presença de hepacivírus não está restrita a primatas e sugere a possibilidade de que HCV possa ter sido introduzido em humanos através do contato com cães ou outras espécies não primatas (KAPOORA *et al.*, 2011).

Uma nova cepa de bocavírus canino também foi detectada através de NGS. Uma ninhada de cães com onze dias de vida apresentava sinais clínicos clássicos de doença entérica, incluindo vômito e diarreia, sendo que dois filhotes morreram. A necropsia revelou uma enterite muito semelhante a desencadeada pela infecção por parvovírus canino, no entanto, os testes laboratoriais convencionais descartaram a infecção por parvovírus e por outros agentes comuns, como rotavírus. Dessa forma, as amostras foram processadas por PCR randômica e submetidas a NGS, revelando 10 *reads* de alta similaridade com bocavírus canino tipo 2 (CaBov-2), chamado canine bocavirus F13000791S. Até a realização desse estudo, não há relatos anteriores de bocavírus causando enterite severa, portanto, mais estudos devem ser realizados para confirmar sua implicação na etiologia da doença (BODEWES *et al.*, 2014).

2.3.3.2 Felinos

Em 2014, Terry Fei Fan Ng *et al.* analisaram o viroma fecal de um felino saudável através da plataforma Illumina MiSeq. Cinco vírus foram identificados na seguinte ordem decrescente de quantidade de *reads*: picornavírus > astrovírus > bocavírus > picobirnavírus > rotavírus. Foram detectados quatro novos vírus que, através das análises filogenéticas, foram propostos como sakobuvírus A (SakoV-A cepa FFUP1; GenBank KF387721) da Família *Picornaviridae*; bocavírus-2 cepa POR1 (FBoV2-POR1) da Família *Parvoviridae*; rotavírus felino Viseu (GenBank KF792839) da Família *Reoviridae* e picobirnavírus felino pertencente à família *Picobirnaviridae*. Até o momento, os picobirnavírus têm sido relatados em amostras fecais e respiratórias de humanos, mamíferos não-humanos, répteis e pássaros, mas sua patogenicidade permanece desconhecida, no entanto, nunca haviam sido detectados em gatos ou na família *Felidae* (KING *et al.*, 2011).

2.3.3.3 Suínos

Em estudo realizado em 2013, por BELAK e colaboradores, foram coletados os gânglios linfáticos de 36 suínos com sintomatologia de Síndrome Multissistêmica do Definhamento (SMDS) e 24 indivíduos que não apresentavam alterações clínicas. Além do circovírus suíno tipo 2 (PCV-2), sabidamente envolvido na etiologia da doença, também foi detectado Torque Teno vírus (TTV) e um novo parvovírus suíno relacionado geneticamente com bocavírus. Essa co-infecção foi mais comum em animais afetados (70%) do que em não afetados (33%). A hipótese de que haja um sinergismo entre diferentes agentes virais na composição da SMDS tem sido suportada por estudos experimentais e epidemiológicos (BLÖMSTROM *et al.*, 2010).

2.3.3.3 Bovinos

O rotavírus do grupo A (RVA) é um dos patógenos de maior importância na causa de diarreia em terneiros neonatos (PAPP *et al.*, 2013). O genoma dos rotavírus consiste em 11 segmentos codificantes, seis proteínas estruturais (VP1, VP2, VP3, VP4, VP6 e VP7) e seis proteínas não estruturais (NS1-NS6). Através da plataforma Illumina Mi-Seq, foi detectado um novo RVA em um bovino leiteiro que apresentava diarreia líquida e diminuição na produção de leite. RVAs são classificados em múltiplos tipos G e P definidos por duas proteínas de capsídeo, VP7 e VP4, respectivamente. Até o momento, 27 tipos G e 35 tipos P de genótipos

foram determinados, muitos dos quais foram identificados nos últimos cinco anos (MATTHIJNSSENS *et al.*, 2011). Uma combinação muito rara G/P, G15P[14] foi encontrada nesse trabalho (MASUDA *et al.*, 2011).

Em outubro de 2011, em uma fazenda de gado leiteiro próxima a cidade de Schmallingenberg na Alemanha, animais apresentavam sintomas de uma doença desconhecida que cursava com febre, diminuição da produção de leite e diarreia. Nessa ocasião, três amostras de sangue de diferentes animais com a sintomatologia foram coletadas e através de um *pool* foram analisadas por metagenômica utilizando a plataforma 454 Genome Sequencer FLX (Roche). Como resultado, a análise filogenética preliminar identificou um novo vírus, que foi denominado de vírus Schmallingenberg (SBV), pertencente à família *Bunyaviridae* e ao gênero *Orthobunyavirus* que possui muitos vírus transmitidos por mosquitos e culicídeos. As sequências encontradas foram intimamente relacionadas com os vírus Akabane, Ainoa e Shamonda pertencentes ao Sorogrupo Simbu detectado na Europa (HOFFMANN *et al.*, 2011). Além dos sintomas citados, SBV também pode causar defeitos genéticos se ocorrer infecção fetal durante o estágio suscetível da gestação (GARIGLIANY *et al.*, 2012a). Esse vírus se espalhou através dos rebanhos de ruminantes em toda a Europa durante 2011-2013, atingindo, por exemplo, a Finlândia, no norte, a Irlanda, no oeste, a Turquia ao leste, e a Espanha, no sul (ELBERT *et al.*, 2014), além de ter sido detectado não apenas em bovinos, mas também em outros ruminantes, como ovinos, caprinos e veados (EFSA, 2012). Atualmente a vacinação é a melhor e única forma efetiva de reduzir a doença clínica e, conseqüentemente, diminuir as perdas econômicas causadas pelos surtos. Além disso, repelentes e piretróides podem auxiliar no controle do vetor (KAUFFOLD; VAHLENKAMO; HOOPS, 2014).

2.3.3.4 Animais Silvestres

Os carnívoros silvestres, sabidamente, são carreadores de patógenos virais que podem infectar animais domésticos e humanos, incluindo o vírus da raiva e o vírus da cinomose (DE BENEDICTIS *et al.*, 2008; VIANA *et al.*, 2015). Estudo publicado por Rogier Bodewes *et al.* (2014) realizou PCR randômico associado a NGS (*454 GS Junior - Roche*) de 42 amostras de fezes coletadas de 10 espécies de pequenos carnívoros silvestres habitantes da região norte da Espanha, como vison-americano (*Neovison vison*), vison-europeu (*Mustela lutreola*), marta-europeia (*Martes martes*), lontra-europeia (*Lutra lutra*) da família *Mustelidae*; geneta (*Genetta genetta*) da família *Viverridae*; raposa vermelha (*Vulpes vulpes*) da família *Canidae* e gato selvagem europeu (*Felis silvestris*) da família *Felidae*. Em uma das amostras da raposa

vermelha e da lontra-europeia, foi detectada uma sequência de alta similaridade com vírus do gênero *Plebovirus* pertencente à família *Bunyaviridae* (BODEWES *et al.*, 2014). Alguns membros dessa família foram identificados como importantes patógenos em humanos e animais domésticos, incluindo o vírus da febre do Vale do Rift e vírus da síndrome da febre severa com trombocitopenia (ZHANF *et al.*, 2012).

2.3.3.5 Morcegos

É sabido que os morcegos são potenciais transmissores de uma série de patógenos além de serem reservatórios ideais de novos agentes infecciosos emergentes (WONG *et al.*, 2007). Um total de 96 suabes retais foram coletados e armazenados em meio de transporte viral, provenientes de 10 espécies, sendo elas, *Rhinolophus sinicus*, *Rhinolophus affinis*, *Hipposideros pomona*, *Miniopterus pusillus*, *Miniopterus schreibersii*, *Pipistrellus abramos*, *Pipistrellus spp*, *Myotis ricketti*, *Myotis chinensis*, e *Nyctalus noctula*. Através da NGS, utilizando a plataforma 454 – Roche, 22 sequências de *reads* e uma *contig* estavam relacionados com papilomavírus em que a identidade dos aminoácidos variou entre 42% a 73%. O novo vírus foi proposto como *Miniopterus schreibersii* papilomavírus tipo 1 (MscPV1).

3 CONCLUSÕES

Ao longo dos últimos anos, diversas plataformas NGS que são baseadas em metodologias distintas foram lançadas no mercado e seu uso disseminou-se nas diferentes áreas da biologia através do sequenciamento massivo de ácidos nucleicos de amostras biológicas ou ambientais. No ramo da virologia, essas tecnologias potencializaram a descoberta de novos vírus e cepas virais através da metagenômica viral, aprofundando as pesquisas relacionadas à interação e presença desses vírus em diferentes hospedeiros.

A capacidade das plataformas em gerar inúmeros dados de genomas virais pode ser apresentada através da detecção de agentes isolados ou como componentes de infecções complexas, como no caso de doenças multifatoriais. Quando aplicada ao diagnóstico clínico, a metagenômica viral utiliza uma abordagem sistemática de amostras de seres humanos ou de animais que apresentem um quadro clínico sem etiologia conhecida. No entanto, é importante lembrar que a presença desses vírus não significa necessariamente que eles causem doença, dessa forma, essas técnicas devem ser acopladas à investigações complementares, como a tentativa de isolamento viral ou ainda por estudos de patogenicidade, considerando os postulados de Koch. Tendo em vista que as técnicas de NGS são relativamente recentes e muitas questões devem ser abordadas antes de ser utilizada mais amplamente a nível de laboratório clínico.

Na virologia veterinária, NGS permitiu que uma série de pesquisas estabelecesse a presença de vírus, já conhecidos ou não, em diferentes espécies animais. O viroma ou flora viral de animais saudáveis e doentes foi caracterizado em felinos, suínos, carnívoros domésticos, animais silvestres, entre outros. Os resultados de tais estudos têm enriquecido o conhecimento a respeito de reservatórios virais e, inclusive, estudos de zoonoses, tendo em vista que muitos vírus já identificados ou não, podem circular entre seres humanos, animais silvestres, domésticos e de produção.

NGS oferece muitas novas oportunidades e é cada vez mais utilizado para estudar a etiologia, genomas, evolução e surto de doenças infecciosas humanas e animais além das interações entre patógeno e hospedeiro. Essas técnicas foram impulsionadas por avanços tecnológicos que levaram a uma diminuição substancial dos custos de sequenciamento de ácidos nucleicos. Ainda assim, o elevado investimento na compra e funcionamento desses equipamentos é um dos fatores limitantes na disseminação de seu uso.

REFERÊNCIAS

- ANSORGE, W. J. Next-generation DNA sequencing techniques. **New Biotechnology**, Amsterdam, v. 25, n. 4, p. 195-203, Apr. 2009
- BAILLIE, G. J. *et al.* Evolutionary dynamics of local pandemic H1N1/2009 influenza virus lineages revealed by whole-genome analysis. **Journal of Virology**, Baltimore, v. 86, n. 1, p. 11–18, Oct. 2012.
- BAYLIS, S. A. *et al.* Analysis of porcine circovirus type 1 detected in Rotarix vaccine. **Vaccine**, Kidlington, v. 29, n.4, p. 690–697, Jan. 2011.
- BEERENWINKEL, N. *et al.* Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. **Frontiers in microbiology**, Lausanne, v. 3, n.1, p. 313-324, Sept. 2012.
- BELÀK, S. *et al.* New viruses in veterinary medicine, detected by metagenomic approaches. **Veterinary Microbiology**, Amsterdam, v. 165, n.1-2, p. 95-101, Feb. 2013.
- BIBBY, K.; VIAU, E.; PECCIA, J. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. **Letters in Applied Microbiology**, Oxford, v. 52, n. 4, p. 386-392, Apr. 2011.
- BLOMSTRÖM, A-L *et al.* Detection of a novel astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral metagenomics. **Journal of Clinical Microbiology**, Washington, v. 48, n. 12, p. 4392–396, Dec. 2010.
- BLOMSTRÖM, A-L. **Applications of viral metagenomics in the veterinary field: looking for the unknown.** 2010. Thesis (**Phd in Veterinaty Medicine**) - Faculty of Veterinary Medicine and Animal Sciences, Swedish University of Agricultural Sciences, Uppsala, 2010.
- BLOMSTRÖM, A-L. *et al.* Studies of porcine circovirus type 2, porcine boca-like virus and torque teno virus indicate the presence of multiple viral infections in postweaning multisystemic wasting syndrome pigs. **Virus Research**, Amsterdam, v. 152, n. 1-2, p. 59-64, Sep. 2010.
- BODEWES, R. *et al.* Novel canine bocavirus strain associated with severe enteritis in a dog litter. **Veterinary Microbiology**, Amsterdam, v. 174, n. 1-2, p. 1-8, Sep. 2014.
- BODEWES, R. Viral metagenomic analysis of feces of wild small carnivores. **Virology Journal**, London, v. 11, n. 1, p. 89-102 May 2014.
- CAPOBIANCHI, M. R.; GIOMBINI, E.; ROZERA, G. Next-generation sequencing technology in clinical virology. **Clinical Microbiology and Infection**, Malden, v. 19, n. 1, p. 15-22, Jan. 2012.
- CHIBA, S. *et al.* An outbreak of gastroenteritis associated with calicivirus in an infant home. **Journal of Medical Virology**, Malden, v. 4, n.4, p. 249–254, Apr. 1979.

COOK, G. C. *et al.* Canine-associated zoonoses: an unacceptable hazard to human health. **The Quarterly Journal of Medicine**, Oxford, v. 70, n. 261, p. 5–26, June 1989.

COX-FOSTER, D. L. *et al.* A metagenomic survey of microbes in honey bee colony collapse disorder. **Science**, New York, v. 318, n. 5848, p. 283-287, Oct. 2007.

CROVILLE, G. *et al.* Field monitoring of avian influenza viruses: whole genome sequencing and tracking of neuraminidase evolution using 454 pyrosequencing. **Journal of Clinical Microbiology**, Washington, v. 50, n. 9, p. 2881–2887, Sept. 2012.

DAUM, S. L. T. *et al.* Next-Generation Ion Torrent sequencing of drug resistance mutations in mycobacterium tuberculosis strains. **Journal of Clinical Microbiology**, Washington, v. 50, n. 12, p. 3831–3837, Dec. 2012.

DAY, J. M. *et al.* Metagenomic analysis of the turkey gut RNA virus community. **Virology Journal**, London, v. 7, n. 7, p. 313-321, Nov. 2010.

DELWART, E. L. Viral metagenomics. **Reviews in Medical Virology**, v. 17, n. 2, p. 115-31, Apr. 2007

DJIKENG, A. Metagenomic analysis of RNA viruses in a fresh water lake. **PLoS One**, San Francisco, v. 4, n. 9, p. e7264, Sep. 2009.

DUFFY, S.; SHACKELTON, L. A.; HOLMES, E. C. Rates of evolutionary change in viruses: patterns and determinants. **Nature Reviews Genetics**, New York, v. 9, n. 4, p. 267–276, Apr. 2008.

ELBERS, A. R. W. *et al.* Schmallenberg Virus Antibodies in Adult Cows and Maternal Antibodies in Calves. **Emerging Infectious Diseases**, Atlanta, v. 20, n. 5, p. 901-902, May 2014.

ETTEN, V. *et al.* DNA viruses: the really big ones (giruses). **Annual Review of Microbiology**, Palo Alto, v. 13, n. 64, p. 83-99, 2010.

EUROPEAN FOOD SAFETY AUTHORITY. “Schmallenberg” virus: analysis of the epidemiological data. **Supporting Publications**, v. 12, n. 5, Nov. 2012.

FEDURCO, M. *et al.* BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. **Nucleic acids research**, Coinsins, v. 34, n. 3, p. 22-24, Jan. 2006.

GARIGLIANY, M.M. *et al.* Schmallenberg virus in calf born at term with porencephaly, Belgium. **Emerging Infectious Diseases**, Atlanta, v.18, n.6, p.1005–1006, June 2012.

GE, X. *et al.* Metagenomic analysis of viruses from bat fecal samples reveals many novel viruses in insectivorous bats in China. **Journal of Virology**, Baltimore, v. 86, n. 8, p. 4620–4630, Apr. 2012.

GRENINGER, A. L. A metagenomic analysis of pandemic influenza A (2009 H1N1) infection in patients from North America. **PLoS One**, San Francisco, v. 5, n. 10, p e13381, Oct. 2010.

GUO, M.; EVERMANN, J. F.; SAIF, L. J. Detection and molecular characterization of cultivable caliciviruses from clinically normal mink and enteric caliciviruses associated with diarrhea in mink. **Archives of Virology**, New York, v. 1, n. 146, p. 479–493, 2001.

HARVELL, C. D. *et al.* Emerging marine diseases – climate links and anthropogenic factors. **Science**, New York, v. 285, n. 5433, p. 1505–1510, Sep. 1999.

HOFFMANN, B. *et al.* Novel Orthobunyavirus in Cattle, Europe. **Emerging Infectious Disease**, Atlanta, v. 18, n. 3, p. 469-472, Mar. 2012.

HOPER, D.; HOFFMANN, B.; BEER, M. A comprehensive deep sequencing strategy for full-length genomes of influenza A. **PLoS ONE**, San Francisco, v. 6, n. 4, p. e19075, Apr. 2011

KAPOORA, A. *et al.* Characterization of a canine homolog of hepatitis C virus. **Proceeding of the Natural Academy of Sciences of the USA**, Washington, v. 108, n. 28, p. 11608-11613, July 2012.

KING, A. M. *et al.* **Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses**. San Diego: Elsevier, 2012.

LACKAY, S. N.; KUANG, Y., FU, Z. F. Rabies in small animals. **Veterinary Clinics of North America: Small Animal Practice**, v. 38, n. 4, p. 851–861, July 2008.

LI, L. *et al.* The fecal viral flora of California sea lions. **Journal of Virology**, Baltimore, v. 85, n. 19, p. 9909–9917, Oct. 2011.

Li, L. *et al.* Viruses in diarrhoeic dogs include novel kobuviruses and sapoviruses. **Journal of General Virology**, London, v. 92, p. 2534–2541, Nov. 2011.

LOMAN, N. J. Performance comparison of benchtop high-throughput sequencing platforms. **Nature Biotechnology**, New York, v. 30, n. 5, p. 434-439, May. 2012.

LYSHOLM, F. *et al.* Characterization of the viral microbiome in patients with severe lower respiratory tract infections, using metagenomic sequencing. **PLoS One**, San Francisco, v. 7, n. 2, p. e30875, Feb. 2012.

MAGI, A. *et al.* Bioinformatics for next generation sequencing data. **Genes**, New York, v. 1, n. 2, p. 294-307, Sep. 2010.

MARDIS, R. M. Next-generation DNA sequencing methods. **Annual Review of Genomics and Human Genetics**, London, v. 9, p. 387-402, June 2008.

MARGULIES, M. *et al.* Genome sequencing in open microfabricated high density picoliter reactors. **Nature**, Brandford, v.437, n. 7057, p. 376-380, July 2005.

MARTELLA, V. Canine distemper virus. **Veterinary Clinics of North America: Small Animal Practice**, v. 38, n.4, p. 787–797, July 2008a.

MASUDA, T. *et al.* Identification of novel bovine group A rotavirus G15P[14] strain from epizootic diarrhea of adult cows by de novo sequencing using a next-generation sequencer. **Veterinary Microbiology**, Amsterdam, v. 171, n. 1-2, p. 66-73, June 2014.

MATTHIJNSSENS, J. *et al.* Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). **Archives of Virology**, New York, v. 156, p. 1397–1413, Aug. 2011.

MERRIMAN, B. *et al.* Progress in Ion Torrent semiconductor chip based sequencing. **Electroforesis**, Weinheim, v. 33, n. 23, p. 3397-3417, Dec. 2012.

MIHALOV-KOVA, E. *et al.* The fecal virome of domesticated animals. **Virus Disease**, New Delhe, v. 25, n. 2, p. 150-157, June 2014.

NAKAMURA, S. *et al.* Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased highthroughput sequencing approach. **PLoS One**, São Francisco, v. 4, n. 1, p. e4219, Jan. 2009.

NG, T. F. F. *et al.* Metagenomic identification of a novel anellovirus in Pacific harbor seal (*Phoca vitulina richardsii*) lung samples and its detection in samples from multiple years. **Journal of General Virology**, London, v. 92, p. 1318–323, June 2011.

NG, T. F. F. *et al.* Feline fecal virome reveals novel and prevalent enteric viroses. **Veterinary Microbiology**, Amsterdam, v. 171, n. 1-2, p. 102-111, June 2014.

PAPP, H. *et al.* Review of group A rotavirus strains reported in swine and cattle. **Veterinary Microbiology**, Amsterdam, v. 165, n 3-4, p. 190–199, Aug. 2013.

RADFORD, A. D. *et al.* Application of next generation sequencing technologies in virology. **The Journal of General Virology**, London, v. 93, n. 9, p. 1853-1868, Sep. 2012.

RIESENFELD, C.; SCHLOSS, P.D.; HANDELSMAN, J. Metagenomics: genomic analysis of microbial communities. **Annual Review of Genetics**, London, v. 38, p. 525-552, July 2004.

RONAGHI, M. Pyrosequencing sheds light on DNA sequencing. **Genome research**, New York, v.11, n.1, p. 03-11, Jan. 2001.

RONAGHI, M.; UHLÉN, M.; NYRÉN, P. A sequencing method based on real-time pyrophosphate. **Science**, New York, v. 281, n. 5375 p. 363–365, July 1998.

ROSARIO, K. *et al.* Metagenomic analysis of viruses in reclaimed water. **Environment Microbiology**, Petersburg, v. 11, n. 11, p. 2086-2090, June 2009.

SCHOWALTER, R. M. *et al.* Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. **Cell Host & Microbe**, Bethesda, v. 7, n. 6, p. 509–515, June 2009.

SHAN, T. *et al.* The fecal virome of pigs on a high-density farm. **Journal of Virology**, Baltimore, v. 85, n. 22, p. 11697–11708, Nov. 2011.

SHOKRALLA, S. *et al.* Next-generation sequencing technologies for environmental DNA research. **Molecular Ecology**, Ontario, v. 21, n. 8, p. 1794-1805, Apr. 2012.

SZPARA, M. L.; PARSONS, L.; ENQUIST, L. W. Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. **Journal of Virology**, Baltimore, v. 84, n.10, p. 5303–5313, May 2010.

TEMMAN, S. *et al.* Viral Metagenomics on Animals as a Tool for the Detection of Zoonoses Prior to Human Infection. **International Journal of Molecular Sciences**, Basel, v. 15, n. 6, p. 10377-10397, June 2014.

THURBER, R. V. *et al.* Laboratory procedures to generate viral metagenomes. **Nature Protocols**, London, v.4, n. 4, p. 470-83, Mar. 2009.

TOWNER, J. S.; SEALY, T. K.; KHRISTOVA, M. L. *et al.* Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. **PLoS Pathogens**, San Francisco, v. 4, n. 11, p. e1000212, Nov. 2008.

VIANA, M. *et al.* Dynamics of a morbillivirus at the domestic–wildlife interface: Canine distemper virus in domestic dogs and lions. **Proceedings of the National Academy of Sciences**, New York, v. 112, n. 5, p. 1464-1469, Feb. 2015.

WETTERSTRAND, K. DNA sequencing costs: data from the NHKRI Genome Sequencing Program Research Institute, 2015. Disponível em <<http://www.genome.gov/sequencingcosts>>. Acesso em 01 jun. 2015>.

WONG, S. *et al.* Bats as a continuing source of emerging infections in humans. **Reviews in Medical Virology**, Malden, v. 17, n. 2, p. 67–91, Apr. 2007.

WONK, K. *et al.* Application of enteric viruses for fecal pollution source tracking in environmental waters. **Environment International**, New York, v. 45, p. 151–164, Sep. 2012.

ZHANG, Y. *et al.* The ecology, genetic diversity, and phylogeny of Huaiyangshan virus in China. **Journal of Virology**, Baltimore, v. 86, n. 5, p. 2864–2868, Dec. 2012.