

**Ministério da Saúde
Fundação Oswaldo Cruz
Centro de Pesquisas René Rachou
Programa de Pós-graduação em Ciências da Saúde**

**ISOLAMENTO E CARACTERIZAÇÃO GENÔMICA DE
BACTERIÓFAGOS QUANTO AO SEU POTENCIAL DE USO
TERAPÊUTICO EM INFECÇÕES CAUSADAS POR
ENTEROBACTÉRIAS**

por
Assmaa El Khal

Belo Horizonte
2016

DISSERTAÇÃO MCS-CPqRR

A. E. KHAL

2016

ASSMAA EL KHAL

**ISOLAMENTO E CARACTERIZAÇÃO GENÔMICA DE BACTERIÓFAGOS
QUANTO AO SEU POTENCIAL DE USO TERAPÊUTICO EM INFECÇÕES
CAUSADAS POR ENTEROBACTÉRIAS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Saúde do Centro de Pesquisas René Rachou, como requisito parcial para obtenção do título de Mestre em Ciências - área de concentração Biologia Celular e Molecular.

Orientação: Guilherme Corrêa de Oliveira

Coorientação: Sara Cuadros Orellana

Belo Horizonte

2016

Catálogo-na-fonte
Rede de Bibliotecas da FIOCRUZ
Biblioteca do CPqRR
Segemar Oliveira Magalhães CRB/6 1975

K45i El Khal, Assmaa.
2016

Isolamento e caracterização genômica de bacteriófagos quanto ao seu potencial de uso terapêutico em infecções causadas por enterobactérias / Assmaa El Khal. – Belo Horizonte, 2016.

XVI, 80 f: il.: 210 x 297 mm.

Bibliografia: 89 - 96

Dissertação (mestrado) – Dissertação para obtenção do título de Mestre em Ciências pelo Programa de Pós-Graduação em Ciências da Saúde do Centro de Pesquisas René Rachou. Área de concentração: Biologia Celular e Molecular.

1. Diarreia/terapia 2. *Escherichia coli*/genética 3. Bacteriófagos/isolamento & purificação 4. Genômica/métodos I. Título. II. Oliveira, Guilherme Corrêa de (Orientação). III. Orellana, Sara Cuadros (Coorientação).

CDD – 22. ed. – 616.342 7

ASSMAA EL KHAL

**ISOLAMENTO E CARACTERIZAÇÃO GENÔMICA DE BACTERÍÓFAGOS
QUANTO AO SEU POTENCIAL DE USO TERAPÊUTICO EM INFECÇÕES
CAUSADAS POR ENTEROBACTÉRIAS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Saúde do Centro de Pesquisas René Rachou, como requisito parcial para obtenção do título de Mestre em Ciências - área de concentração Biologia Celular e Molecular.

Banca Examinadora:

Dr. Guilherme Corrêa de Oliveira (CPqRR/FIOCRUZ) Orientador/Presidente

Dr. Carlos Eduardo Calzavara Silva (CPqRR/FIOCRUZ) Titular

Dr. Flávio Guimarães da Fonseca (UFMG) Titular

Dra. Marina de Moraes Mourão (CPqRR/FIOCRUZ) Suplente

Dissertação defendida e aprovada em Belo Horizonte, 27/06/2016.

Dedico este trabalho aos meus pais, minha base, incentivo, exemplo de vida,
conduta e amor.

AGRADECIMENTOS

Gostaria de agradecer a Deus: porque sem a fé que tudo daria certo eu não teria conseguido.

Aos colegas e amigos da FIOCRUZ (Anderson, Angela, Carol, Daniel, Fabiano, Fausto, Fran, Gabriel, Juliana, Julliane, Larissa, Laura, Sara, Victor e Wagner): pelo acolhimento e pelo carinho, seja em forma de sorrisos, olhares, palavras carinhosas, guloseimas, ajuda e até puxões de orelha. Sem vocês e todo o apoio e energia positiva que me passam, não sei se teria conseguido passar pelas dificuldades do mestrado. Muito obrigada por me ajudarem e me fazerem tão bem!

Guilherme e Sara, por aceitarem me orientar e terem me dado a oportunidade de ser desse grupo tão especial.

Julliane: pela amizade, pelos ensinamentos e confiança depositada em mim e em “meus fagos”, pelas várias vezes que corrigiu meu trabalho.

Gabriel e Laura: pela ajuda na montagem e análise dos genomas, pela orientação contínua, pelo incentivo e amizade.

Eric: pela ajuda no início das análises e pela paciência para me ensinar como rodar os comandos.

Aos colegas e amigos do LPCM: pela convivência agradável, divisão das áreas de uso comum e pela torcida. Principalmente à Carol, Paulinha, Isabela e Erick, vocês eu levarei para a vida!

Vivi: pela ajuda na extração do DNA, fazendo as tardes serem mais leves e felizes.

Anna e Flávio: pelo sequenciamento do DNA, além do carinho e a preocupação constante comigo e com o meu trabalho.

Aos meus amigos, que estiveram sempre presentes me apoiando e distraíndo nos momentos de tensão e também aos que entenderam a ausência nos de dedicação exclusiva. Obrigada pelo apoio e confi Alessandra, Aninha, Camilinha, Nai, Nath, Sarah, Thaiana. Amo vocês!

Família: pelo amor e confiança plena de que eu seria capaz, pela companhia nas noites mal dormidas, pelos cafés e pelas flores nas manhãs, pelas

horas acordados fazendo meu almoço do dia seguinte, pelas caronas para o laboratório e pelo jeito único de demonstrarem apoio. Vocês são tudo o que poderia pedir a Deus. Meu amor por vocês é imensurável.

À minha irmã caçula, Imane, que sempre foi meu porto seguro, me dando os conselhos mais maduros que já ouvi, suportando meus dias de estresse.

À banca que avaliou esse trabalho, pelas sugestões, colaborando para o aprimoramento deste.

À Plataforma de Bioinformática do Centro de Pesquisas René Rachou.

À Plataforma de Sequenciamento do Centro de Pesquisas René Rachou.

À COPASA, pelo fornecimento das amostras que serviram como fonte de bacteriófagos.

À CAPES e à Pós-Graduação do Centro de Pesquisas René Rachou, pelo financiamento e suporte a esse trabalho.

RESUMO

O crescente surgimento de resistência bacteriana aos antibióticos convencionais é um grave problema que precisa ser enfrentado, seja pela descoberta de novas substâncias antimicrobianas, naturais ou sintéticas, ou através da pesquisa de terapias alternativas que sejam economicamente acessíveis. A terapia de fagos é uma dessas alternativas. Trata-se de uma forma de controle biológico, baseado em vírus específicos que infectam e destroem células bacterianas: os bacteriófagos. No entanto, esta é uma fonte terapêutica ainda pouco explorada. Esse trabalho utilizou o cultivo, isolamento e sequenciamento do genoma, além de técnicas de genômica de alto desempenho para isolar e caracterizar o genoma de bacteriófagos específicos para a linhagem enteroinvasiva de *Escherichia coli* ATCC 43893, visando o entendimento e a definição do ciclo de infecção desses vírus (líticos ou lisogênicos). A metodologia utilizada nessa pesquisa possibilitou o isolamento de 12 vírus. 8 diferentes linhagens virais tiveram seu material genético extraído e purificado, apresentando bom rendimento e quantidade reduzida de DNA bacteriano contaminante. O sequenciamento do genoma desses 8 vírus foi realizado usando a plataforma de nova geração *MiSeq*. Foi analisada a diversidade genética desses bacteriófagos e verificou-se que são vírus da ordem *Caudovirales*, sendo 2 da família *Siphoviridae* e 6 da família *Myoviridae*. Apenas um deles mostrou potencial de ter ciclo lisogênico, os outros sete vírus não continham nenhum gene que sugerisse isso. Entretanto, apesar dos bacteriófagos isolados não terem apresentado genes relacionados ao ciclo lisogênico, análises mais aprofundadas devem ser realizadas para comprovar que são realmente exclusivamente líticos, já que muitos não apresentam seu genoma completo e mais de 50% dos genes anotados não têm função definida.

Palavras-chave: bacteriófagos, *Escherichia coli*, genômica.

ABSTRACT

The increasing emergence of bacterial resistance to conventional antibiotics is a serious problem that needs to be faced, either through the discovery of new antimicrobial substances, natural or synthetic, or by searching for alternative therapies that are affordable. The phage therapy is one of those alternatives. It is a form of biological control based on specific viruses that infect and kill bacterial cells: the bacteriophages. However, this therapeutic source is still poorly explored. This study used the cultivation, isolation and sequencing of the genome, as well as high-performance genomic techniques to isolate and characterize the genome of specific bacteriophages for enteroinvasive *Escherichia coli* ATCC 43893, for the understanding and the definition of the infection cycle (lytic or lysogenic) of these viruses. The methodology used in this study allowed the isolation of 12 viruses. 8 different viral strains had their genetic material extracted and purified, with good yield and reduced amount of contaminating bacterial DNA. The sequencing of the genome of these 8 viruses was conducted using the new generation *MiSeq* platform. The the genetical diversity of these bacteriophages was analyzed and it was found that the viruses belong to the *Caudovirales* order, which 2 belong to the *Siphoviridae* family and 6 to the *Myoviridae* family. Only one of them showed the potential to have lysogenic cycle, the other seven viruses contained no gene to suggest that. However, despite the isolated bacteriophages have not presented genes related to lysogenic cycle, further analysis should be conducted to demonstrate that they are really exclusively lytic, since many do not have their genome completed and more than 50% of the annotated genes have no defined function.

Keywords: bacteriophages, *Escherichia coli*, genomics.

LISTA DE FIGURAS

Figura 1 – Morfotipos dos vírus de Procariotos.....	23
Figura 2 – Estrutura do vírion do bacteriófago T4.....	24
Figura 3 – Montagem dos vírions em bacteriófagos caudados.....	27
Figura 4 – Esquema ilustrativo do ciclo de infecção dos bacteriófagos.....	28
Figura 5 – Etapas do ciclo lítico de infecção e respostas bacterianas à invasão viral.....	31
Figura 6 – Integração e excisão do genoma de bacteriófagos ao cromossomo bacteriano.....	36
Figura 7 – Placas de lise geradas pelo bacteriófago F6.....	47
Figura 8 – Diluição de 1:100 do bacteriófago F1.....	48
Figura 9 – Eletroforese em gel de agarose a 0,5%.....	49
Figura 10 – Aspecto turvo das placas de lise produzidas pelo bacteriófago F2....	85

LISTA DE GRÁFICOS

Gráfico 1 - Comparação dos genes possivelmente ortólogos dos bacteriófagos utilizando a ferramenta BRIG.....	52
Gráfico 2 - Origem taxonômica dos genes usados para anotação do bacteriófago F1.....	84
Gráfico 3 - Origem taxonômica dos genes usados para anotação do bacteriófago F10.....	84

LISTA DE QUADROS

Quadro 1 - Padrão de aderência e fatores de virulência das linhagens de <i>E. coli</i>	19
Quadro 2 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F1.....	53
Quadro 3 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F2.....	54
Quadro 4 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F3.....	55
Quadro 5 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F6.....	58
Quadro 6 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F8.....	60
Quadro 7 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F9.....	63
Quadro 8 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F10.....	65
Quadro 9 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F12.....	66

LISTA DE TABELAS

Tabela 1 – Reagentes para a reação de PCR.....	42
Tabela 2 – Quantidade de placas de lise verificadas nas diferentes coletas.....	48
Tabela 3 – Rendimento do DNA isolado de bacteriófagos.....	50
Tabela 4 – Quantidade de sequências antes e após a filtragem.....	50
Tabela 5 – Características dos genomas montados.....	51
Tabela 6 – Resultado da anotação dos genes do bacteriófago F1.....	68
Tabela 7 – Resultado da anotação dos genes do bacteriófago F2.....	69
Tabela 8 – Resultado da anotação dos genes do bacteriófago F3.....	71
Tabela 9 – Resultado da anotação dos genes do bacteriófago F6.....	73
Tabela 10 – Resultado da anotação dos genes do bacteriófago F8.....	76
Tabela 11 – Resultado da anotação dos genes do bacteriófago F9.....	78
Tabela 12 – Resultado da anotação dos genes do bacteriófago F10.....	80
Tabela 13 – Resultado da anotação dos genes do bacteriófago F12.....	81

LISTA DE SIGLAS E ABREVIATURAS

ASM: Sociedade Americana de Microbiologia
BFP: bundle-forming pili
CDC: Centers for Disease Control and Prevention
COPASA: Companhia de Saneamento de Minas Gerais
CPqRR: Centro de Pesquisas René Rachou
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats
crRNAs: CRISPR ribonucleic acids
DAEC: *E. coli* difusivamente aderente
DALY: Disability-adjusted life year
dNTPs: Desoxirribonucleotídeos fosfatados
dsDNA: double stranded deoxyribonucleic acid
DVH: Doenças de Veiculação Hídrica
EAEC: *E. coli* enteroagregativa
EAggEC: *E. coli* enteroagregativa
EHEC: *E. coli* entero-hemorrágica
EIEC: *E. coli* enteroinvasiva
EPEC: *E. coli* enteropatogênica
ESBL: Beta-lactamases de espectro estendido
ETE-Arrudas: Estação de Tratamento de Esgoto-Arrudas
FDA: Food and Drug Administration
INCQS: Instituto Nacional de Controle de Qualidade em Saúde
Kbp: Kilo-base pair
LEE: locus for enterocyte effacement
NCBI: National Center for Biotechnology Information
OMS: Organização Mundial da Saúde
pb: pares de bases
RM: Restrição-modificação
TSA: Tryptic Soy Agar
TSB: Tryptic Soy Broth

SUMÁRIO

1 INTRODUÇÃO E JUSTIFICATIVA	17
1.1 Contexto epidemiológico	17
1.2 Enterobactérias	17
1.3 <i>E. coli</i> , suas características e importância médica	18
1.4 Resistências a drogas antimicrobianas	20
1.5 Bacteriófagos: definição e características	21
1.6 Classificação dos bacteriófagos	21
1.7 Mecanismos de infecção dos bacteriófagos.....	25
1.8 Ciclos de bacteriófagos	27
1.9 Mecanismos de defesa bacteriana à invasão por bacteriófagos	29
1.10 Evasão da defesa bacteriana pelos bacteriófagos	32
1.11 Bacteriófagos e a fagoterapia	32
1.12 Emprego de bacteriófagos como bactericidas na indústria de alimentos.....	33
1.13 Utilização de proteínas codificadas por bacteriófagos como antimicrobianos.....	34
1.14 Genômica de bacteriófagos	35
1.15 Fagos geneticamente modificados	37
1.16 Identificação de novos bacteriófagos	38
2 OBJETIVOS.....	39
2.1 Objetivo Geral	39
2.2 Objetivos Específicos	39
3 METODOLOGIA	40
3.1 Coleta de amostras	40

3.2 Linhagens bacterianas e condições de cultivo	40
3.3 Indução de infecção bacteriana <i>in vitro</i>	41
3.4 Isolamento, purificação e armazenamento dos fagos das unidades formadoras de placa (UFP)	41
3.5 Diluição seriada de fagos para utilização no isolamento de bacteriófagos	41
3.6 Titulação de fagos e determinação de virulência.....	42
3.7 Análise da pureza das amostras virais	42
3.8 Eletroforese em gel de agarose	43
3.9 Extração do DNA de bacteriófagos	43
3.10 Sequenciamento e análise do material genético dos bacteriófagos	44
3.11 Montagem e análise dos genomas virais	44
4 RESULTADOS.....	47
4.1 Diluição seriada dos vírus	47
4.2 Comparação da carga de bacteriófagos da estação de tratamento de esgoto em diferentes coletas	48
4.3 Verificação da pureza do extrato viral	49
4.4 Quantificação do DNA	50
4.5 Caracterização dos genomas virais	50
5 DISCUSSÃO	83
6 CONCLUSÃO	87
7 PERSPECTIVAS.....	88
REFERÊNCIAS.....	89

1 INTRODUÇÃO E JUSTIFICATIVA

1.1 Contexto epidemiológico

O continente sul-americano abriga diversos países em desenvolvimento com grandes problemas na área de Saúde Pública. Um dos maiores problemas crônicos é a dificuldade de acesso à água potável e ao saneamento de qualidade. Em 2015, 2,4 bilhões de pessoas não tinham acesso a sistemas de saneamento adequados (WHO, 2015). Segundo a UNESCO, a falta de saneamento básico faz com que muitos reservatórios de água destinados ao consumo humano sejam contaminados com micro-organismos patogênicos, o que representa um problema potencial de Saúde Pública. Doenças de veiculação hídrica (DVH) são responsáveis por aproximadamente 5 milhões de mortes anuais, metade das quais ocorre como resultado de complicações derivadas de infecções intestinais causadas por enterobactérias (CABRAL, 2010).

As diarreias são os exemplos mais frequentes de DVH. Elas podem ser causadas por bactérias, protozoários ou vírus (WHO, 2013). Anualmente, ocorrem aproximadamente 1,7 bilhões de casos de diarreia, o que resulta em cerca de 1,8 milhões de óbitos e contribui com 3,6% do DALY (Disability-adjusted life year) global (WHO, 2013). Elas são a principal causa de desnutrição e a segunda maior causa de mortalidade em crianças, causando 760 000 óbitos/ano (WHO, 2013). Cerca de 50% dos casos de mortalidade infantil no continente africano são causados por essa doença (WHO, 2009).

1.2 Enterobactérias

As enterobactérias, responsáveis pelas diarreias bacterianas, são um grupo de organismos Gram-negativos, anaeróbios facultativos, capazes de colonizar o intestino de mamíferos (ENVIRONMENTAL PROTECTION AGENCY, 1997). Muitos de seus representantes, tais como alguns sorotipos de *Escherichia coli*, são simbioses e compõem a microbiota intestinal normal, porém outras linhagens de *Escherichia coli*, *Salmonella*, *Yersinia pestis*, *Klebsiella* e *Shigella*, podem ser potenciais patógenos, geralmente oportunistas, podendo causar infecção no trato digestivo, infecção urinária, sepse, pneumonia, entre outras (ANVISA, 2004).

Apesar de estarem frequentemente associadas a um hospedeiro, bactérias dessa família também podem ser de vida livre, sendo encontradas em ecossistemas aquáticos, principalmente após a contaminação desses ambientes com esgoto (BEAUDOIN, 2007).

1.3 *E. coli*, suas características e importância médica

Escherichia é o gênero tipo da família Enterobacteriaceae. *E. coli* é a espécie mais estudada dessa família, e embora a maior parte das linhagens de *E. coli* sejam inócuas, alguns sorotipos podem ser altamente patogênicos, sendo responsáveis por infecções hospitalares, do trato urinário e da corrente sanguínea, assim como infecções de origem alimentar em todo o mundo (WHO, 2014). *E. coli* também é a principal causadora de diarreias no homem, sendo os sintomas da doença considerados linhagem-dependentes (OKEKE, 2009; NATARO; KAPER, 1998).

As linhagens patogênicas de *E. coli* são classificadas de acordo com suas manifestação clínicas em humanos e com o seu sítio de infecção, podendo ser: enteropatogênicas (EPEC), uropatogênicas (UPEC) e patogênicas extra-intestinais. Essas bactérias também podem ser agrupadas com base em sua sorologia. Além disso, podem ser classificadas, de acordo com o mecanismo de virulência utilizado, em: enterohemorrágica (EHEC), enterotoxigênica (ETEC), enteroinvasiva (EIEC), enteroagregativa (EAaggEC) e difusamente-aderente (DAEC) (LUKJANCENKO et al., 2010).

Os mecanismos de virulência em *E. coli* são codificados por genes cromossômicos, plasmidiais ou de bacteriófagos. Os genes envolvidos nesses mecanismos são: *eae*, envolvido em lesões por adesão ou comprometimento das microvilosidades intestinais; *bfpA*, relacionado a aderência localizada; *ipaH*, que se relaciona a mecanismos enteroinvasivos. Genes que codificam as toxinas Shiga (*stx1* e *stx2*), a toxina termo-lábil (LT) e toxina termo-estável (ST) também são associados à virulência (ARANDA; FAGUNDES-NETO; SCALETISKY, 2004).

Os padrões de aderência e fatores de virulência das linhagens de *E. coli* estão apresentados no **Quadro 1**. As EPEC podem causar lesão por adesão, através da síntese de fatores de virulência, que é regulada pela ativação de genes

localizados no lócus cromossômico *LEE* (*Locus of Enterocyte Effacement*). O *LEE* é considerado uma ilha de patogenicidade e está presente em linhagens EPEC e EHEC. (MCDANIEL et al., 1995; CLEARY et al., 2004)

As linhagens de EIEC são capazes de causar disenteria bacilar. Após a ingestão de alimentos contaminados com fezes de humanos doentes, essas bactérias invadem as células do epitélio do intestino. A doença resultante, cujos sintomas aparecem após 12 a 72 horas, caracteriza-se pelo aparecimento de sangue e muco nas fezes de pacientes infectados, podendo também ser acompanhada de cólicas abdominais, diarreia, vômitos, febre, calafrios e mal-estar generalizado. Uma complicação comumente associada com a infecção é a síndrome hemolítico-urêmica (FDA, 2012).

QUADRO 1 - Padrão de aderência e fatores de virulência das linhagens de *E. coli*.

Linhagem de <i>E. coli</i>	Padrão de aderência	Principais fatores de virulência conhecidos
<i>E. coli</i> enteropatogênica (EPEC)	Localizado	<i>LEE</i> , enterotoxina EspC, BFP (<i>bundle forming pili</i>), Sistema de secreção Tipo III, Proteínas: EspABD, Intimina
<i>E. coli</i> enterotoxigênica (ETEC)	Difuso fraco	Toxina termo-estável ao calor (ST), Toxina termo-lábil ao calor (LT), fímbrias
<i>E. coli</i> enteroinvasiva (EIEC) e <i>Shigella</i>	Variável	Enterotoxinas de <i>Shigella</i> , toxinas <i>Shiga</i> (STX; Sh dysenteriae), exclusão em genes <i>housekeeping</i>
<i>E. coli</i> enterohemorrhagica (EHEC)	Variável	Enterohemolisina, <i>LEE</i> , Shiga toxinas (Stx), toxina termo-estável (EAST-1), intimina, adesinas acessórias
<i>E. coli</i> enteroagregativa (EAEC ou EAggEC)	Agregativo	Dispersina, toxina termo-estável (EAST-1), adesinas acessórias
<i>E. coli</i> difusamente-aderente (DAEC)	Difuso	Adesina difusa, outras adesinas

FONTE: elaborado pelo autor a partir de dados de CLEARY et al., 2004.

1.4 Resistência a drogas antimicrobianas

A primeira droga antimicrobiana, a penicilina, foi descoberta por Alexander Fleming, no ano de 1928. Esse acontecimento foi um marco na medicina, tornando possível o tratamento de sérias infecções bacterianas. Entretanto, o uso descontrolado de antibióticos fez com que bactérias resistentes fossem selecionadas, causando um grande problema, que hoje é ameaça à saúde pública: a resistência a antibióticos (DAVIES; DAVIES, 2010).

Sabe-se que as bactérias podem ser intrinsecamente resistentes a certos antibióticos, mas também podem adquirir resistência através de mutações em genes cromossômicos ou através da transferência gênica horizontal (BLAIR, et al., 2015). Alguns dos mecanismos de resistência descritos em *E. coli* são as mutações (ex: resistência às fluoroquinolonas) e a aquisição de elementos genéticos móveis (ex: resistência às penicilinas de largo espectro, como ampicilina e amoxicilina, e cefalosporinas de terceira geração). Esta última é conferida principalmente por enzimas conhecidas como beta-lactamases de espectro estendido (ESBL), as quais destroem muitos antibióticos beta-lactâmicos (ALEKSHUN; LEVY, 2007).

O relatório mais recente da Organização mundial da Saúde (OMS) deixa claro que a resistência bacteriana tem atingido níveis alarmantes em várias partes do mundo (WHO, 2014), incluindo o surgimento de linhagens resistentes a múltiplos antibióticos (superbactérias), contra as quais poucos tratamentos permanecem eficazes. De acordo com esse relatório, que teve como fonte dados fornecidos pelos programas de vigilância de países de todo o mundo, assim como artigos científicos, a porcentagem de *Escherichia coli*, *Klebsiella pneumoniae* e *Staphylococcus aureus* resistentes a antibióticos comumente utilizados ultrapassou os 50% (WHO, 2014). Em se tratando de resistência de *Staphylococcus aureus* a meticilinas foi verificada uma porcentagem de 90% nas Américas e 60% na Europa. A Organização Pan-Americana da Saúde cita altos níveis de resistência de *Escherichia coli* às cefalosporinas de terceira geração e às fluoroquinolonas (WHO, 2014).

Foi descrito por Losee e colaboradores, em janeiro de 2015, um novo antibiótico, a teixobactina, que mostrou ter excelente atividade contra patógenos

Gram-positivos, incluindo cepas resistentes, como enterococos, *Mycobacterium tuberculosis*, *Clostridium difficile* e *Bacillus anthracis*. Essa nova droga também teve atividade bactericida contra *Staphylococcus aureus*, porém mostrou ser ineficaz contra a maioria das bactérias Gram-negativas, como é o caso das *Escherichia coli* e outras enterobactérias. Nesse contexto, torna-se importante a pesquisa e o desenvolvimento de métodos alternativos e complementares para o tratamento de infecções por bactérias resistentes aos medicamentos convencionais.

1.5 Bacteriófagos: definição e características

Bacteriófagos, também denominados fagos, são vírus capazes de infectar células bacterianas, dentro das quais eles podem se replicar (D'HÉRELLE, 1917). São considerados as entidades biológicas mais abundantes do planeta (ROHWER; EDWARDS, 2002). Por serem incapazes de se replicar de forma independente, são normalmente encontrados em maior número em ambientes onde seu hospedeiro está presente (WEINBAUER, 2004). Além disso, por apresentarem alta especificidade em relação ao hospedeiro, não são nocivos a organismos eucariotos (SUMMERS, 2001). Além disso, os fagos geram significativo impacto na estrutura e na função de comunidades de bactérias, seja causando a sua morte ou realizando transferência horizontal de genes, podendo alterar os fenótipos bacterianos e conduzir à expressão de genes de virulência (PAUL, 1999).

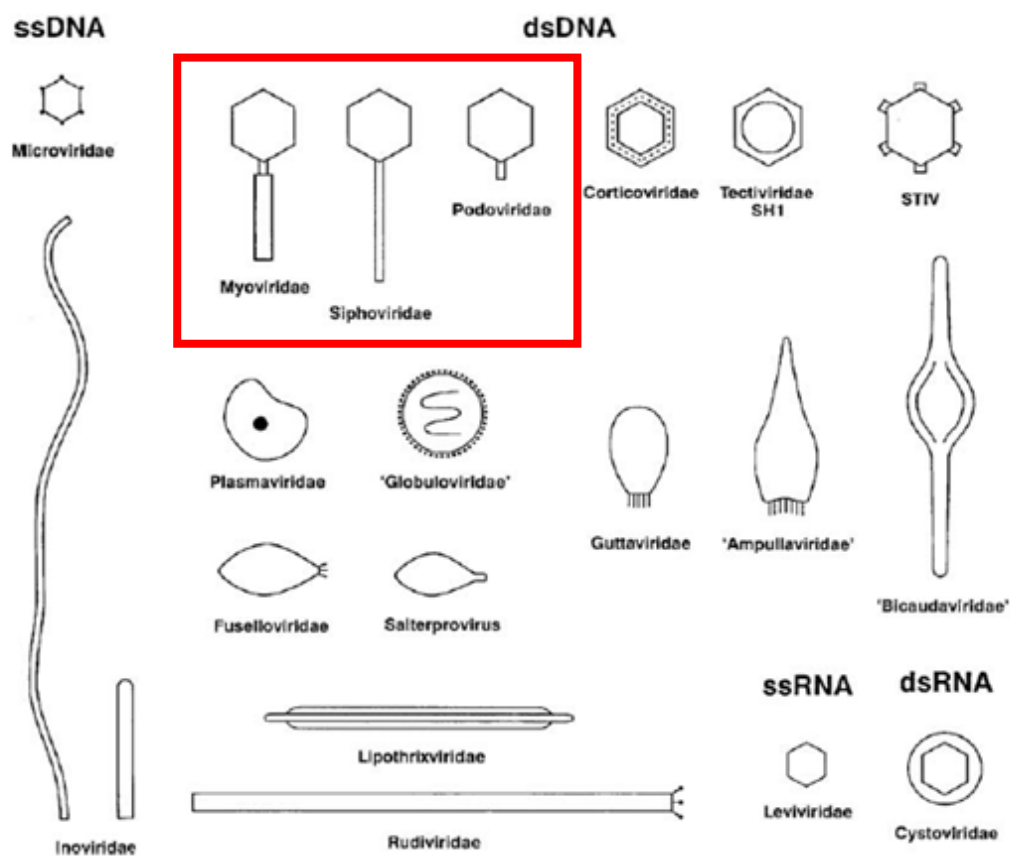
1.6 Classificação dos bacteriófagos

A classificação dos bacteriófagos é ainda uma questão pouco definida. Aproximadamente 70 propriedades de bacteriófagos são utilizadas com fins taxonômicos, porém, as mais importantes são: natureza dos ácidos nucleicos, morfologia da partícula viral (vírion), suas propriedades físico-químicas e dados genômicos (ACKERMANN; EISENSTARK, 1974).

Inicialmente, em 1943, Ruska propôs um método de classificação viral baseado em microscopia eletrônica. Mais tarde, em 1962, Lwoff, Horne e Tournier estabeleceram que a classificação desses vírus deveria ser baseada em

propriedades do vírion e do material genético e propuseram um novo sistema de nomenclatura (LWOFF et al., 1962). O *Provisional Committee on Nomenclature of Viruses (PCNV)*, fundado em 1965, posteriormente se tornou o *International Committee on Taxonomy of Viruses (ICTV)*, que é, atualmente, o único órgão internacional responsável por taxonomia viral (MURPHY et al., 1995). Em 1974, os fagos foram separados por morfotipos (**Figura 1**), para facilitar a identificação por microscopia eletrônica. (ACKERMANN; EISENSTARK, 1974).

FIGURA 1 - Morfotipos dos vírus de Procariotos.



Acima estão representados os diversos morfotipos de bacteriófagos conhecidos. Os bacteriófagos da ordem *Caudovirales* (famílias *Myoviridae*, *Siphoviridae* e *Podoviridae*) estão localizados entre os vírus de DNA de fita dupla (dsDNA), como indicado na figura.

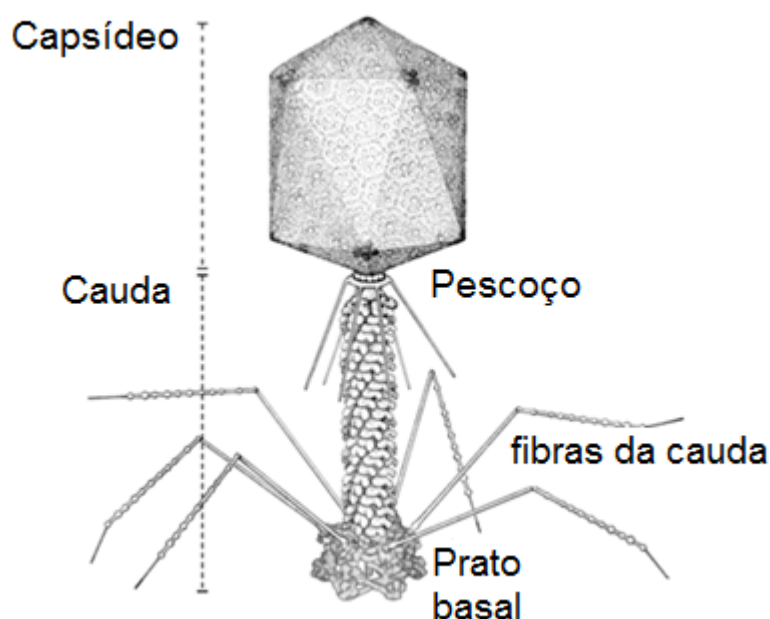
FONTE: CLOKIE; KROPINSKI, 2009.

Apesar de ainda não existir um método universal para a classificação desses fagos, eles são mais comumente agrupados de acordo com sua morfologia e com a composição de seu material genético (www.ictvonline.org/).

A ordem *Caudovirales*, é formada de vírus caudados, que constituem mais de 96% da população de bacteriófagos. São fagos que apresentam genoma de DNA de fita dupla, com tamanho variando de 18 a 500 mil pares de bases. Os vírions formados são únicos e apresentam um capsídeo icosaédrico, contendo o DNA e uma cauda, que é responsável por fazer a ligação à célula hospedeira alvo e pelo transporte do material genético para o interior da célula (**Figura 2**) (CASJENS, 2005; ACKERMANN, 2007). A ordem *Caudovirales* é composta por

três famílias virais: *Myoviridae*, com o maior capsídeo (150 nm) e uma cauda contrátil; *Siphoviridae*, com um capsídeo relativamente pequeno (50-60 nm) e apresentando cauda longa, flexível e não-contrátil; e *Podoviridae*, caracterizada por apresentar pequeno capsídeo (50-60 nm) e cauda curta (DRULIS-KAWA et al., 2012). As famílias, por sua vez, são subdivididas em subfamílias, gêneros e subgêneros utilizando-se o tamanho e as características do genoma, assim como a gama de hospedeiros dos vírus. (NELSON, 2004). Estima-se que o número de espécies de *Caudovirales* seja aproximadamente dez vezes superior ao número espécies bacterianas. Dessa forma, uma estimativa conservadora é a de que existam, no mínimo, cem milhões de espécies de bacteriófagos dessa ordem, embora apenas uma pequena fração dessa diversidade tenha sido isolada e caracterizada (ROHWER, 2003).

FIGURA 2 – Estrutura do vírion do bacteriófago T4.



Na figura acima está representada a morfologia usual de um bacteriófago da ordem *Caudovirales*. O material genético desses vírus encontra-se protegido pelo capsídeo viral. O pescoço faz a ligação entre o capsídeo e a cauda e as fibras da cauda são responsáveis pela ligação com receptores específicos da célula bacteriana.

FONTE: MILLER et al., 2003.

1.7 Mecanismos de infecção dos bacteriófagos

A infecção por bacteriófagos ocorre quando há a ligação entre a partícula viral e uma célula hospedeira suscetível. Inicialmente, ocorre a adsorção da partícula viral a receptores de superfície do hospedeiro. Esse processo é específico e pode ser reversível ou irreversível. A adsorção reversível é influenciada por ações eletrostáticas. Já a irreversível tem ligações mais estáveis e ação de enzimas hidrolíticas (DRULIS-KAWA et al., 2012; GREGORACCI, 2006).

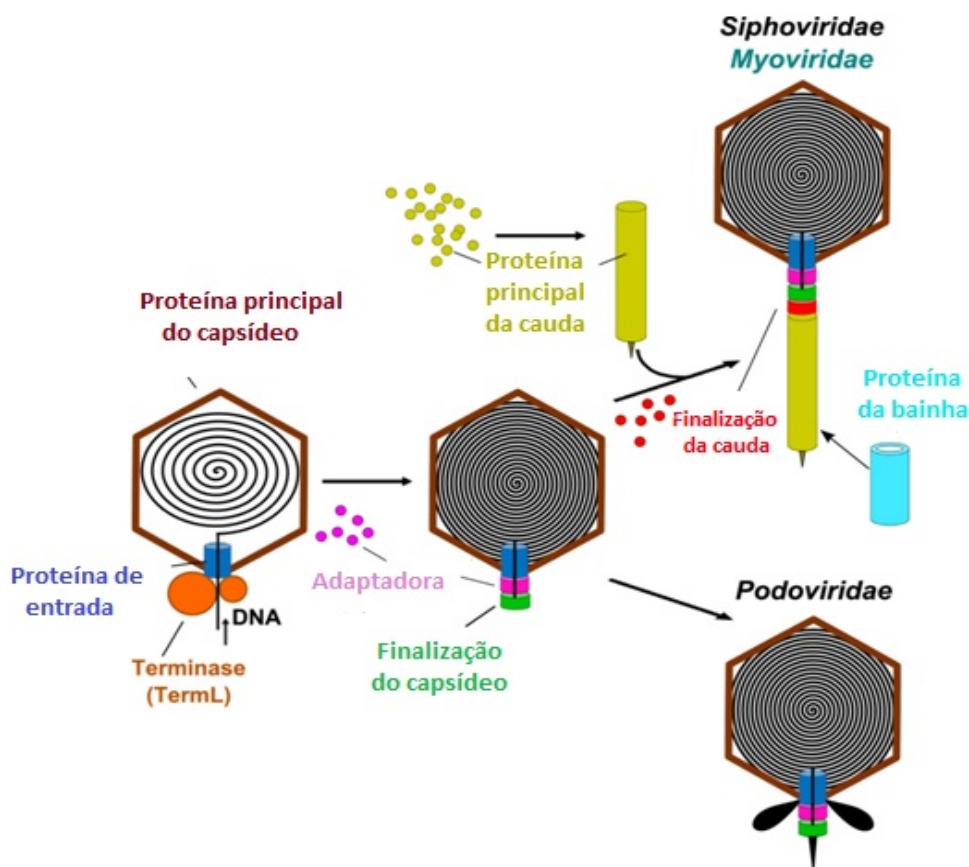
Existem dois tipos principais de receptores bacterianos aos quais os fagos se ligam: componentes da superfície e pílus sexual. Assim, são classificados como bacteriófagos somáticos os fagos que têm como receptores os componentes da superfície bacteriana (proteínas de membrana, lipopolissacarídeos, peptidoglicanos, ácidos teicóicos, oligossacarídeos, cápsula e fímbrias do tipo IV e o flagelo bacteriano). Já quando a ligação ocorre no pílus sexual da bactéria os vírus são classificados como fagos RNA F⁺ específicos ou “male-specific” (DRULIS-KAWA et al., 2012). No caso de vírus da ordem *Caudovirales*, fibras da cauda viral geralmente são responsáveis pela ligação com os receptores bacterianos, sendo a gama de hospedeiros desse bacteriófago definida pelo tipo de fibras da cauda desse vírus (CHATURONGAKUL; OUNJAI, 2014; MILLER et al., 2003).

A etapa seguinte é a injeção do material genético dos bacteriófagos no interior da célula bacteriana (DRULIS-KAWA et al., 2012). Depois que o DNA penetra o citoplasma, os genes precoces dos fagos redirecionam a maquinaria de síntese bacteriana para que sejam produzidos os ácidos nucleicos e as proteínas virais. Em seguida, ocorre a montagem e o empacotamento dos vírions (**Figura 3**) (MAURICE et al., 2013).

A primeira etapa é a montagem do capsídeo viral, que é iniciada ao se construir o procapsídeo, uma proteína estrutural composta pela proteína principal do capsídeo (MCP). Posteriormente, a proteína de entrada (portal protein) forma um canal em um dos vértices do procapsídeo. Um complexo ATPase se acopla a esse canal. Esse complexo é denominado terminase e é composto por diversas cópias da grande subunidade da terminase (TermL) e uma pequena subunidade

de ligação que reconhece o DNA do bacteriófago e, usando o canal da proteína de entrada (portal protein), o transfere para dentro do procapsídeo, realizando o empacotamento do DNA. Ao final deste processo, o complexo terminase é desmontado. Para evitar que o material genético se disperse, a proteína de entrada reúne outras, como a proteína adaptadora (Ad), que se liga a ela. Em seguida há a suplementação do processo pela proteína finalizadora do capsídeo (Hc), já que a MCP sozinha não é capaz de garantir a geometria correta do mesmo (AKSYUK; ROSSMANN, 2011). Esta conclui a montagem das pequenas caudas de vírus da família *Podoviridae* e possibilita a ligação de caudas longas dos membros da família *Myoviridae* e *Siphoviridae*. As proteínas finalizadoras da cauda, ou *tail completion proteins* (Tc) fazem a ligação entre a cauda e o capsídeo. O pescoço dos bacteriófagos é formado pelas proteínas Hc, Tc e a proteína de entrada, conectando a cabeça à cauda. A cauda é formada principalmente pela proteína principal da cauda (MTP). Vírus da família *Myoviridae* possui um envoltório ao redor da cauda, formado pela proteína da bainha (sheath protein). Essa bainha se contrai no momento da infecção, para que o DNA seja injetado na célula bacteriana (LOPES et al., 2014). As etapas seguintes dependem do tipo de ciclo biológico de cada bacteriófago.

FIGURA 3 - Montagem dos vírions em bacteriófagos caudados.



A montagem dos vírions é iniciada pela proteína principal do capsídeo e é continuada com o auxílio de diversas proteínas, como a proteína de entrada, a proteína adaptadora, terminasse, proteína finalizadora do capsídeo, proteína principal da cauda, proteína de finalização da cauda e, em Myoviridae, pela proteína da bainha, que envolve a cauda viral.

FONTE: LOPES et al., 2014.

1.8 Ciclos de bacteriófagos

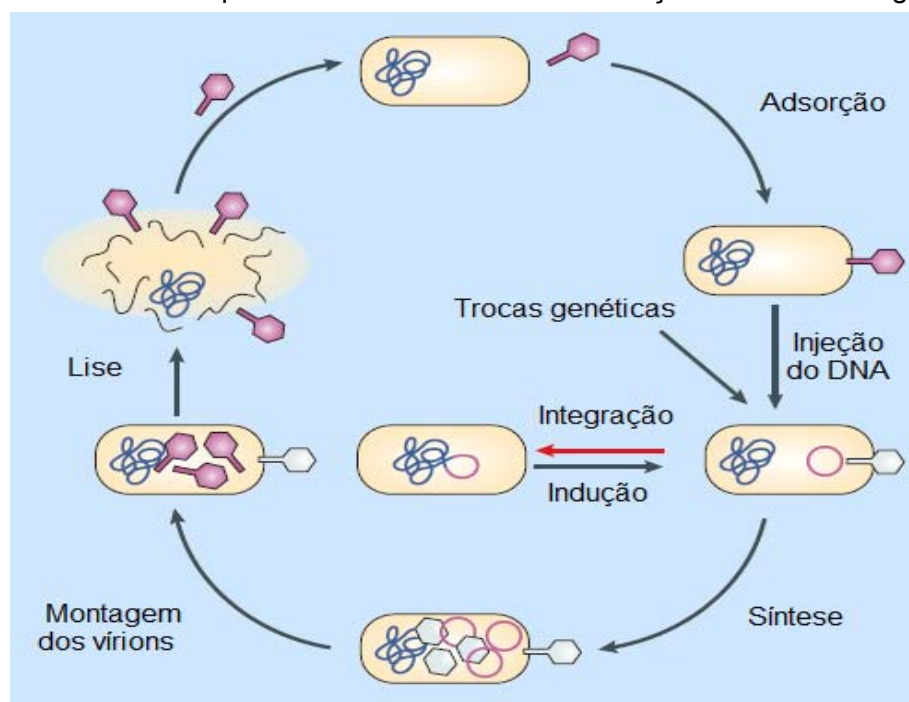
Bacteriófagos podem apresentar diferentes ciclos biológicos. Os dois principais são: lítico e lisogênico (DRULIS-KAWA et al., 2012).

Os fagos líticos, ou virulentos, têm altas taxas de replicação, ocorrendo acúmulo de grande quantidade de vírus e de proteínas expressas tardiamente por eles (endolisinas, holinas ou inibidoras da síntese de mureína), que provocam a lise da célula bacteriana e liberação desses bacteriófagos (**Figura 4**) (MAURICE et al., 2013).

Quando se trata de fagos lisogênicos (ou temperados), o material genético do bacteriófago pode ser integrado ao da bactéria, porém esse processo é reversível. Nele, há silenciamento da expressão de genes do fago, a partir de

proteínas que conferem à bactéria imunidade a esses vírus. Após ser integrado ao genoma bacteriano, o vírus passa a ser denominado profago. Este pode sofrer indução se a célula for submetida à ação de raios ultravioleta ou outro estresse físico-químico. Assim, o vírus, que antes estava inserido no genoma bacteriano, passa para o ciclo lítico. Através da lisogenia, o hospedeiro adquire imunidade à superinfecção por fagos relacionados ao invasor, pois o genoma de alguns fagos lisogênicos possui genes capazes de modificarem os receptores da célula bacteriana infectada, evitando que outros vírus se adsorvam (MARKINEGORIAYNOFF et al., 2004). Pode também haver a aquisição de resistência a antibióticos e de sequências que influenciem na expressão de toxinas (conversão lisogênica), como a Shiga toxina de *E. coli*, aumentando a aptidão (*fitness*) da bactéria. O DNA bacteriano também pode ser danificado por inserção de elementos móveis e deleção. Esse processo é denominado transdução especializada (MAURICE et al., 2013).

FIGURA 4 – Esquema ilustrativo do ciclo de infecção dos bacteriófagos.



O ciclo é iniciado com a adsorção do vírus à célula bacteriana. Posteriormente ele injeta seu DNA; realiza a síntese de suas proteínas; as partículas virais são montadas, produzem endolisinas e lisam a célula infectada, liberando novas partículas virais. A seta vermelha indica o ciclo lisogênico, no qual o DNA do fago é integrado ao genoma bacteriano.

FONTE: STURINO; KLAENHAMMER, 2006.

1.9 Mecanismos de defesa bacteriana à invasão por bacteriófagos

Os bacteriófagos apresentam capacidade de infectar as células bacterianas, porém, as bactérias, por sua vez, possuem diversos mecanismos de defesa contra partículas ou DNA exógeno (**Figura 5**). Entre esses mecanismos estão: inibição da adsorção, bloqueio da injeção do material genético, restrição-modificação (RM), infecção abortiva, o mecanismo de CRISPRs (*Clustered Regularly Interspaced Short Palindromic Repeats*) (KESIK- SZELOCH, et al., 2013).

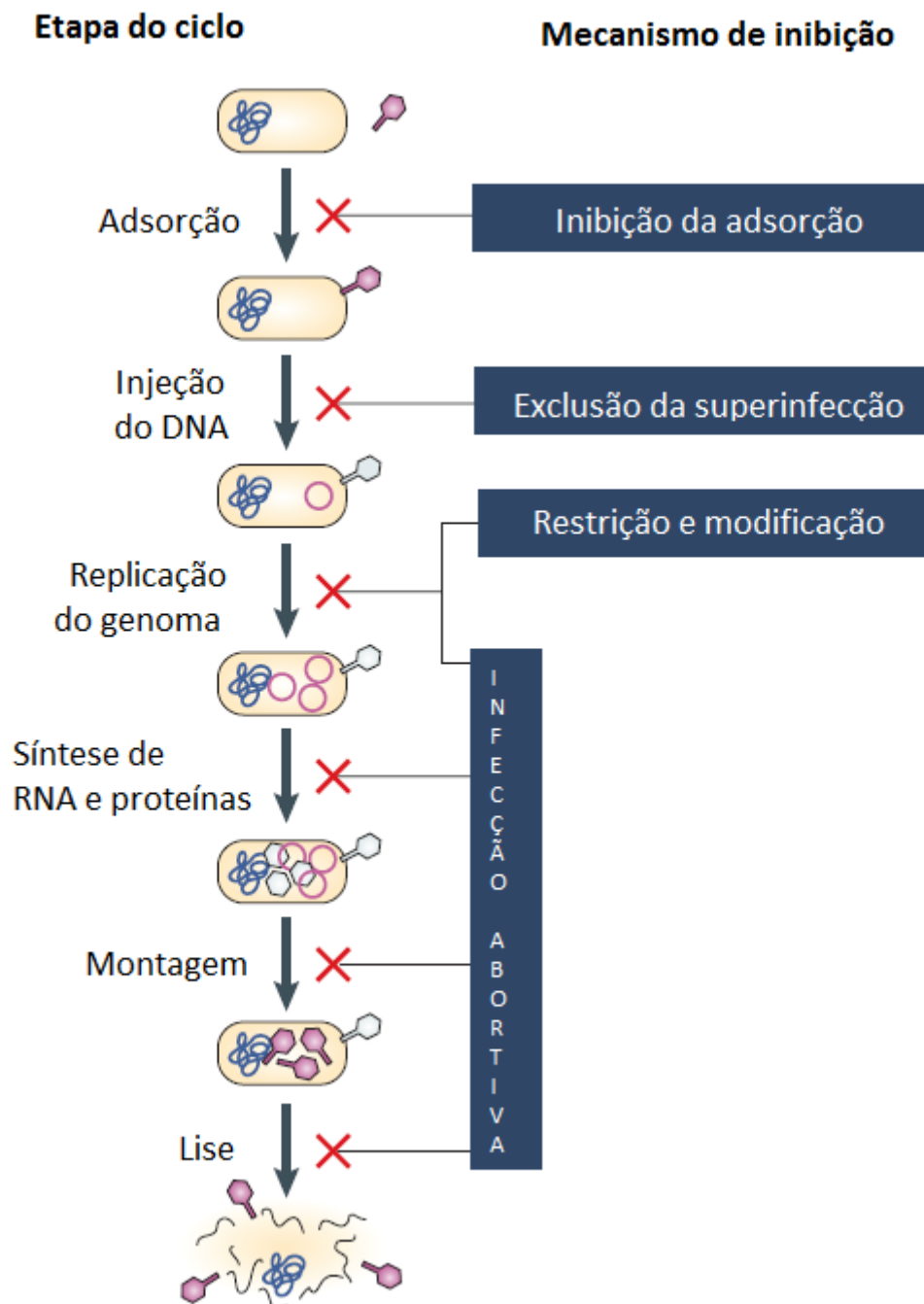
A inibição da adsorção, que torna inviável a infecção pelos bacteriófagos, pode ocorrer por meio de mascaramento dos receptores da superfície bacteriana ou modificação na sua estrutura. Porém, é mais comum a resistência à infecção por fagos devido à ausência desses receptores. Tanto o cromossomo bacteriano, como plasmídeos podem mediar os mecanismos que previnem a adsorção (RAKHUBA et al., 2010).

O mecanismo mais bem estudado é o de restrição-modificação (RM) e está presente em mais de 90% das bactérias. Consiste na ação de endonucleases de restrição, que são enzimas capazes de reconhecer e clivar sequências específicas nas moléculas de DNA (KESIK-SZELOCH, et al., 2013). As endonucleases participam do sistema de restrição, fazendo a degradação do DNA invasor, que não é metilado. Já as metil-transferases, responsáveis pela modificação, introduzem um grupo metil em um sítio específico do DNA do hospedeiro, protegendo-o de ser degradado. As sequências não metiladas são degradadas pelas endonucleases. Os sistemas de restrição-modificação são classificados em quatro grupos (I, II, III e IV), dependendo de sua estrutura molecular, cofatores necessários, reconhecimento de sequências e sítio de clivagem (KRÜGER; BICKLE, 1983).

A infecção abortiva (Abi) é o processo responsável pela defesa da célula quando o sistema RM falha. Assim, os mecanismos de infecção abortiva têm diferentes alvos na célula, atuando em distintas fases do ciclo (replicação, transcrição, síntese proteica, montagem) e limitando a proliferação dos fagos (LABRIE et al., 2010).

Além disso, também existe o CRISPR. O *locus* CRISPR e seus genes associados (*cas*) codificam um mecanismo de defesa sequência-específico contra bacteriófagos e plasmídeos. O CRISPR consiste em conjuntos de sequências repetitivas curtas (cerca de 40 pb) que são separadas por sequências espaçadoras derivadas de DNA exógeno. Após a clivagem, pela endoribonuclease Cas, em sequências repetidas, conjuntos de CRISPR são transcritos como um longo precursor. Este é processado em pequenos CRISPR RNAs (crRNAs) (BIKARD; MARRAFFINI, 2012). Esse sistema é encontrado em 48% das Bactérias e em 95% das Arqueas. Utilizando os crRNAs, os complexos crRNA/Cas clivam DNA exógeno em sítios específicos, conferindo às bactérias resistência contra o DNA invasor. A incorporação de novas sequências ao locus CRISPR é realizada também pela ação das proteínas Cas. (BONDY-DENOMY et al., 2013).

FIGURA 5 - Etapas do ciclo lítico de infecção e respostas bacterianas à invasão viral.



À esquerda estão representadas as etapas do ciclo de infecção dos bacteriófagos e à direita os mecanismos de defesa bacteriana contra o agente invasor.

FONTE: STURINO; KLAENHAMMER, 2006.

1.10 Evasão da defesa bacteriana pelos bacteriófagos

Para obter sucesso ao infectar o hospedeiro, o fago necessita evadir os diferentes mecanismos de defesa da célula bacteriana. Entre as estratégias de evasão mais conhecidas estão a ação de inibidores de enzimas de restrição e a hidrólise de cofatores do sistema de RM (KESIK- SZELOCH, et al., 2013).

Outro mecanismo de evasão da defesa bacteriana por bacteriófagos envolve cinco distintos genes “anti-CRISPR” que foram identificados nos genomas de fagos. Enquanto mutações no gene anti-CRISPR do fago o tornam incapaz de infectar uma bactéria com um sistema CRISPR/Cas funcional, a adição do mesmo gene ao seu genoma permite que esse vírus evada a defesa bacteriana (BONDY-DENOMY et al., 2013).

Os fagos também têm a capacidade de adquirir novas mutações em genes que codificam proteínas que se ligam aos receptores bacterianos (RBPs) ou fibras da cauda viral em resposta às modificações que esses receptores sofrem, mantendo assim a capacidade de infectar o hospedeiro. Para lidar com o mascaramento de receptores da membrana bacteriana por algumas moléculas o fago pode expressar despolimerases, que degradam a molécula mascaradora, possibilitando a ligação do bacteriófago ao receptor-alvo (SAMSON et al., 2013).

1.11 Bacteriófagos e a fagoterapia

A terapia de fagos, também conhecida como fagoterapia é uma forma de controle biológico, baseada em vírus específicos de bactérias. Ela foi inicialmente proposta em 1917, por Félix d’Herelle, no Instituto Pasteur (Paris, França), mas o interesse nessa aplicação foi reduzido significativamente quando os antibióticos foram descobertos. Porém, com a ameaça emergente de infecções causadas por bactérias multirresistentes e poucas perspectivas para o surgimento de novos antibióticos, a terapia de fagos passou a ser alvo de novos estudos visando a uma possível aplicação terapêutica (FRUCIANO; BOURNE, 2007).

Como agentes terapêuticos, os fagos poderiam substituir ou atuar como coadjuvantes de drogas no combate a infecções bacterianas (VIERTEL et al., 2014). A fagoterapia já tem aplicação em alguns países, como a Rússia, onde o uso do coquetel de fagos, medicamento composto por bacteriófagos, já é uma realidade.

A companhia russa *Microgen* desenvolve preparações de fagos em suspensão ou em comprimidos para diversas infecções, por exemplo disenterias, infecções por *Staphylococcus*, *Streptococcus*, *Pseudomonas*, além de coquetéis de bacteriófagos (en.microgen.ru/) (MCCALLIN et al., 2013). As pesquisas nessa área têm atraído maior interesse global e a incorporação da fagoterapia no tratamento de infecções tem sido apoiada pelo governo de vários países. O Instituto Nacional de Alergia e Doenças Infecciosas dos Estados Unidos, por exemplo, listou a fagoterapia como um dos sete pilares no seu plano de pesquisa para combater a resistência aos antibióticos. Além disso, a Sociedade Americana de Microbiologia (ASM) apoia o *Phagoburn*: o primeiro centro multi-núcleo geral de clínica da fagoterapia para infecções humanas, financiado pela Comissão Europeia, no âmbito do *7th Framework Programme for Research and Development*. O *Phagoburn* é um projeto inicialmente focado no tratamento de pacientes vítimas de queimaduras e, ao mesmo tempo, portadores de infecções por *Escherichia coli* ou *Pseudomonas aeruginosa* (phagoburn.eu/).

Algumas das vantagens observadas na fagoterapia são: melhor atividade profilática e terapêutica se comparada aos medicamentos convencionais; ausência de destruição da microbiota bacteriana normal dos indivíduos (tendo em vista a especificidade dos bacteriófagos em infectar a célula de interesse) e a possibilidade de utilização em dose única (PAISANO; BOMBONA, 2010). Além disso, a fagoterapia não oferece perigo ao ser humano, pois esses vírus não têm a capacidade de infectar ou causar danos a células eucarióticas (KESIK-SZELOCH et al., 2013). Alguns trabalhos como o de McCallin e colaboradores (2013) demonstram a eficácia e segurança da terapia de fagos.

Apesar da possibilidade de surgimento de resistência bacteriana a fagos, essa é muito menos frequente do que a resistência a antibióticos convencionais (hu, 2006). Dessa forma, a fagoterapia é uma alternativa eficiente e segura, podendo ser forte aliada ao tratamento no caso de resistência bacteriana aos antibióticos usuais (REARDON, 2014).

1.12 Emprego de bacteriófagos como bactericidas na indústria de alimentos

Além da fagoterapia, os bacteriófagos possuem diversas outras aplicações.

Entre elas está o uso como agentes bactericidas em alimentos (HUDSON et al., 2005). O FSIS (*Department of Agriculture's Food Safety and Inspection Service*) aprovou o uso de uma mistura de seis bacteriófagos específicos para o patógeno alimentar *Listeria monocytogenes* (*ListShieldTM*, fabricado pela empresa *Intralytix Inc.*). Esse “coquetel de fagos” é utilizado em carnes e produtos avícolas. O *ListShieldTM* é usado como aditivo alimentar, assim como outros aditivos normalmente empregados pela indústria para manter a qualidade dos alimentos, além de não oferecer risco à saúde humana (SHARMA, 2013).

Além disso, existem diversos outros produtos baseados em um ou em uma combinação de bacteriófagos, como o *Listex P100* (*Micreos Food Safety*), esse também é reconhecido como seguro pela *Food and Drug Administration* (FDA), não havendo objeções científicas para o seu uso. Esses produtos são utilizados para inibir o crescimento de *L. monocytogenes* em queijos e em alimentos prontos para o consumo (SHARMA, 2013).

Em um trabalho de Jianxiong e colaboradores, um coquetel de seis bacteriófagos específicos para *Salmonella spp.*, aplicado em associação com *Enterobacter asburiae* reduziu a concentração de *Salmonella spp.* em 6 log CFU/mL em feijões, se comparado a brotos de feijão sem o coquetel (YE et al., 2010). As soluções de bacteriófagos são aplicadas de diferentes formas, como spray ou imersão do alimento, sempre gerando diminuição significativa da carga bacteriana nos produtos alimentícios (SHARMA, 2013).

1.13 Utilização de proteínas codificadas por bacteriófagos como antimicrobianos

Outra opção ao uso direto de fagos como antimicrobianos é o uso de proteínas codificadas por esses agentes: as endolisinas, que degradam o peptídeoglicano; e as holinas, que rompem a membrana celular da célula infectada. Essas proteínas demonstraram ter importantes ações antibacterianas (YOUNG; BLÄSI, 1995; DRULIS-KAWA et al., 2015).

As lisinas não são tóxicas para a célula eucariótica e são ainda mais promissoras que fagos isolados. Elas penetram melhor nos tecidos-alvo, têm menor imunogenicidade, a probabilidade de desenvolvimento de resistência

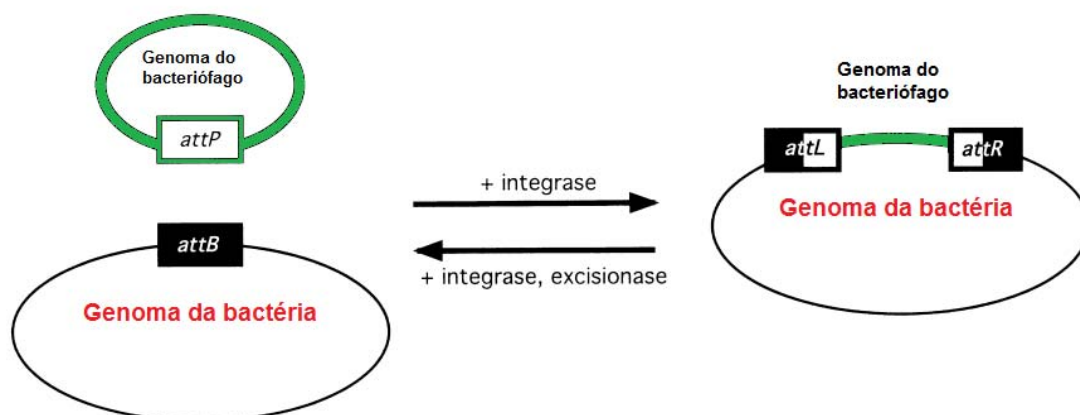
bacteriana é menor, apresentam maior espectro de atividade, além de serem altamente eficientes na destruição dos patógenos-alvo (DRULIS-KAWA et al., 2012).

1.14 Genômica de bacteriófagos

Fagos com potencial de uso terapêutico necessitam ser caracterizados biologicamente e geneticamente. O uso de vírus que não sejam exclusivamente líticos pode levar à expressão de fatores de virulência bacterianos como resposta adaptativa à infecção por fagos. Um exemplo é a cepa O104:H4 de *E. coli*, que produz uma toxina capaz de causar diarreia e lesões no rim, a Shiga toxina. A Shiga toxina é responsável pela síndrome hemolítico-urêmica. Os genes envolvidos na sua produção são transferidos por bacteriófagos. Logo, após essa transferência gênica, a bactéria torna-se patogênica (TURNER, 2011; MUNIESA et al., 2012).

Dessa forma, é essencial a caracterização genômica dos vírus bacteriófagos antes de qualquer ensaio clínico e uso como terapia. A caracterização genômica possibilita saber se o bacteriófago é unicamente lítico ou se apresenta a capacidade de se integrar ao genoma bacteriano, causando possíveis problemas. Uma informação importante que caracteriza bacteriófagos lisogênicos é a presença de genes que codificam a proteína integrase (*Int*). Essa enzima é mediadora da recombinação entre o DNA do bacteriófago e o de seu hospedeiro em sítios específicos (*attP*, de bacteriófagos e *attB*, no caso das bactérias) (**Figura 6**). Logo fagos que apresentam essa enzima não são indicados para uso na fagoterapia (GROTH; CARLOS, 2004; CHO et al., 2002; GRAINGE; JAYARAM, 1999).

FIGURA 6 – Integração e excisão do genoma de bacteriófagos ao cromossomo bacteriano.



O Genoma da bactéria está representado em forma de elipse de cor preta e em verde está indicado o genoma do bacteriófago e seu sítio de ligação (*attP*). Com a ação da enzima integrase ocorre a integração do genoma viral no cromossomo bacteriano, flanqueado pelos sítios híbridos *attL* e *attR*. A excisão desse genoma pode ocorrer com o auxílio das enzimas excisionase e integrase.

FONTE: GROTH; CARLOS, 2004.

O avanço das plataformas de sequenciamento, bem como a redução no custo dessa técnica tem possibilitado o avanço no campo da genômica de bacteriófagos. Em 2012, o banco de dados de genomas do NCBI possuía aproximadamente 600 genomas de *Caudovirales* (KLUMPP et al., 2012). Com apenas 3 anos de pesquisas, esse número dobrou. Em maio de 2016 havia 1.863 genomas de bacteriófagos depositados no NCBI, sendo que para a ordem *Caudovirales* havia 1.748 (www.ncbi.nlm.nih.gov/).

A bioinformática possibilita obter informações sobre genes de bacteriófagos que ainda não foram caracterizados, o que é importante para a compreensão da diversidade e função desses vírus. Apesar de relativamente recentes, os estudos genômicos dos bacteriófagos têm sido essenciais para aprimorar o conhecimento da diversidade molecular e estratégias empregadas pelos fagos caudados (CASJENS, 2005).

A genômica funcional consiste em sequenciar o material genético dos bacteriófagos, realizar a sua montagem, anotar o genoma e, finalmente, atribuir funções aos produtos dos genes anotados. Portanto, a genômica funcional, a partir da abordagem *in silico* (fazendo a predição de genes) ou acoplada à experimental (analisando a função deles *in vitro*), tem grande potencial em

aperfeiçoar a anotação de genoma de fagos e, assim, melhor caracterizá-los (KLUMPP et al., 2013).

Devido à ausência de dados experimentais complementares, mais de 50% dos produtos gênicos de fagos são hipotéticos e não têm uma função definida. Além disso, aproximadamente 50% das cepas bacterianas que já tiveram seu material genético sequenciado apresentam prófagos em seu genoma, chegando a compor cerca de 20% do genoma (KLUMPP et al., 2013).

Em *Caudovirales* a especificidade de um fago pelo hospedeiro é determinada por genes que codificam proteínas da cauda e sabe-se que a fagoterapia é limitada pela gama de hospedeiros que um bacteriófago isolado pode infectar. Portanto, o estudo comparativo de genomas permite detectar o possível hospedeiro pela manipulação de genes da cauda dos fagos (BRÜSSOW; HENDRIX, 2012; SHIN et al., 2012).

Vários membros da ordem *Caudovirales* têm genes conservados, como o gene que codifica terminases, proteínas de entrada e proteínas do capsídeo viral, assim como as proteínas da cauda (proteína principal da cauda e da bainha). Porém, a intensa recombinação genética que ocorre dá origem a mosaicismos no genoma e dificulta a caracterização filogenética desses vírus (REYES et al., 2012). Isso explica a ausência de uma definição exata de espécies de bacteriófagos e a dificuldade de padronizar um esquema de classificação (CASJENS, 2005; LOPES, 2014).

Portanto, utilizando-se de estudos de genômica, podem-se esclarecer questões essenciais a respeito de mecanismos de infecção dos bacteriófagos, evasão da defesa bacteriana por esses vírus, assim como sobre seu ciclo e o procedimento de replicação do material genético (KLUMPP, et al., 2013).

1.15 Fagos geneticamente modificados

As técnicas contemporâneas de engenharia genética tornam possível desenvolver fagos com funções adicionais, desenhados para a solução de problemas específicos como, por exemplo, degradar diferentes componentes estruturais em biofilmes de bactérias. Sabe-se que o biofilme de *E. coli* é composto principalmente por celulose, logo, fagos que expressem a enzima

celulase são capazes de degradá-lo. *Lu e Collins*, por exemplo, desenvolveram um fago geneticamente modificado capaz de realizar a degradação do biofilme de *E. coli*: o bacteriófago T7 modificado, sendo ele cem vezes mais eficiente na remoção desse biofilme do que o fago original. (LU; COLLINS, 2007).

1.16 Identificação de novos bacteriófagos

É estimado que menos de 0,0002% da diversidade global de bacteriófagos foi amostrada. Entretanto, apesar dos fagos serem as entidades biológicas mais abundantes e diversas no planeta, pouco se sabe sobre este universo (IGNACIO-ESPINOZA et al, 2013; ROHWER et al., 2009).

No entanto, apesar de haver busca ativa por bacteriófagos que infectam bactérias patogênicas, essa é uma fonte terapêutica ainda pouco explorada. Além disso, é importante que se busque bacteriófagos específicos para as bactérias de interesse local, tendo em vista que para o sucesso desse método terapêutico é necessária a utilização de fagos capazes de infectar e destruir especificamente o organismo causador da infecção, sem afetar o restante da microbiota. Nesse contexto, estudos que visam a identificação de novos fagos, bem como o sequenciamento do genoma desses vírus são essenciais para o conhecimento da sua diversidade, ainda desconhecida (ROHWER, 2003).

2 OBJETIVOS

2.1 OBJETIVO GERAL

Identificar e caracterizar o genoma de bacteriófagos capazes de infectar enterobactérias de interesse médico.

2.2 OBJETIVOS ESPECÍFICOS

- Isolar bacteriófagos capazes de infectar a linhagem enteroinvasiva de *Escherichia coli* (EIEC) ATCC 43893.
- Sequenciar, montar e anotar os genomas dos bacteriófagos isolados.
- Descrever o repertório de genes virais possivelmente envolvidos na infecção e na evasão dos sistemas de defesa bacteriana.

3 METODOLOGIA

3.1 Coleta de amostras

Amostras de água residual foram coletadas na ETE-Arrudas (COPASA), a maior estação de tratamento de esgoto da região metropolitana de Belo Horizonte. Foram realizadas duas coletas: uma no mês de abril de 2015 e outra no mês de agosto de 2015. A primeira coleta foi realizada tomando-se amostras da fase intermediária e final de tratamento, estágios em que a contagem de células de *E. coli* ainda é superior a 10^5 células/mL, e que a matéria orgânica já é reduzida. Foram coletados 10 tubos Falcon de 15 mL cheios. As amostras foram transportadas até o laboratório, onde foram centrifugadas a 4410 *g*, por 20 minutos, a 4°C para decantação de sólidos em suspensão. Posteriormente, as amostras foram filtradas em filtros de acetato de celulose com poros de 0,22 μ m, para remoção de células bacterianas. O filtrado recuperado, parte que contém a fração viral (cerca de 10 mL), foi armazenado em geladeira até o momento do processamento.

A segunda coleta foi realizada tomando-se amostras apenas da fase intermediária do tratamento. Os procedimentos seguintes foram idênticos aos da primeira coleta.

3.2 Linhagens bacterianas e condições de cultivo

O cultivo bacteriano foi realizado no laboratório do Grupo de Genômica e Biologia Computacional do Centro de Pesquisas René Rachou (CPqRR). Devido à sua importância clínica, foi utilizada como modelo a linhagem enteroinvasiva de *Escherichia coli* (EIEC) ATCC 43893, provinda da Coleção de Culturas do Instituto Nacional de Controle de Qualidade em Saúde (INCQS/Fiocruz). A linhagem foi cultivada em Caldo Trípico de Soja (Sigma-Aldrich) para sua reativação. Alíquotas dessa cultura foram conservadas em meio TSB contendo glicerol na concentração de 20%, a -70°C, para uso como inóculo, quando necessário.

3.3 Indução de infecção bacteriofágica *in vitro*

Para o isolamento de bacteriófagos presentes nas amostras ambientais, foram aplicados, em meio de cultivo semissólido aquecido (0,35% de Agar), 100 µL do filtrado resultante da coleta (que contém os fagos) e 100 µL da cultura de *E. coli* em fase exponencial de crescimento (OD_{600nm} = 0,6 a 0,9). As placas foram incubadas em estufa a 37°C, por 24 horas. Após esse período, foi verificado o aparecimento de placas de lise (zonas claras dentro da camada de crescimento bacteriano), que indicaram as áreas de ocorrência de infecção e lise celular (KESIK-SZELOCH et al., 2013; ADAMS, 1959).

3.4 Isolamento, purificação e armazenamento dos fagos das Unidades Formadoras de Placa (UFP)

Para a purificação dos bacteriófagos foram realizadas sete passagens sucessivas. Para cada passagem foi retirado, com o auxílio de uma ponteira, um pequeno pedaço de ágar do centro da placa de lise. Essa placa foi transferida para um tubo Eppendof de 1,5 mL, contendo 1 mL de tampão SM (20 mg/L Tris-HCl, 10 mg/L $MgSO_4$, 10 mg/L $CaCl_2$, 100 mg/L NaCl, pH 7,5) (ANDREWS-PFANNKUCH et al., 2010). Esse conteúdo foi levado ao vórtex para desprendimento dos vírus presentes nas placas de ágar e posteriormente centrifugou-se os tubos a 4410 g, por 15 minutos, a 4°C para decantação do ágar e das células bacterianas. O sobrenadante foi filtrado em filtros de acetato de celulose com poros de 0,22 µm para obtenção de partículas virais purificadas, que foram armazenadas a 4°C, em tampão SM. Esse processo foi realizado para possibilitar o isolamento de uma linhagem única de bacteriófagos.

3.5 Diluição seriada de fagos para utilização no isolamento de bacteriófagos

A fim de obter uma concentração ideal de fagos para a utilização no isolamento, foram realizadas diluições seriadas da amostra final contendo os vírus coletados anteriormente de cada placa de lise. Dessa forma, foram transferidos 100 µL dessa amostra para um microtubo estéril contendo 900 µL de tampão SM. A partir dessa diluição foram tomados outros 100 µL e transferidos a um novo

microtubo, também contendo 900 µL de tampão. Foram realizadas sete diluições idênticas a essa e testadas para padronização da melhor diluição.

3.6 Titulação de fagos e determinação de virulência

A contagem das placas de lise observáveis a olho nu foi feita e o número de unidades virais infecciosas (título de fagos) foi expresso como unidades formadoras de placa por mL (UFP/mL) (KESIK- SZELOCH, et al., 2013; ADAMS, 1959).

$$\text{UFP/mL} = n \times \text{FC} \times 10^*$$

Tendo em vista que:

n é igual ao número de placa de lises contadas;

FC é o fator de correção (numero pelo qual o valor do inoculo utilizado deve ser multiplicado para que se obtenha 1 mL);

10* é o inverso da diluição na qual foi encontrado o valor de n.

3.7 Análise da pureza das amostras virais

Visando analisar se havia presença de contaminação das amostras virais com DNA bacteriano, foi realizada uma reação de PCR usando os *primers* 515F (5'-GTGCCAGCMGCCGCGGTAA-3') e 806R (5'-GGACTACHVGGGTWTCTAAT-3') para a região V4 do gene do 16S rRNA bacteriano. Como controle positivo, foi utilizado DNA de *E. coli*. A eletroforese em gel de agarose foi feita para avaliar os resultados da PCR.

TABELA 1 - Reagentes para a reação de PCR

Reagente	Volume
DNA molde (5 ng/µL)	2,5 µl
<i>Primer forward</i> (1µM)	5,0 µl
<i>Primer reverse</i> (1µM)	5,0 µl
2x KAPA HiFi HotStart ReadyMix	12,5 µl
Total	25,0 µl

FONTE: elaborado pelo autor.

A reação de PCR ocorreu em um termociclador Veriti® 96-Well (Applied Biosystems®), utilizando como programação: uma desnaturação inicial a 95°C por 5 minutos, seguida de 25 ciclos de desnaturação a 95°C por 30 segundos, anelamento dos *primers* a 55°C por 30 segundos e extensão a 72°C, também por 30 segundos. A extensão final à temperatura de 70°C teve duração de 5 minutos.

3.8 Eletroforese em gel de agarose

Os produtos da PCR foram analisados por eletroforese em gel de agarose a 0,5%. Foram aplicados 25 µl dos produtos de PCR, acrescidos de 5 µl de tampão de amostra corado com GelRed™ (Biotium). Usou-se padrão de peso molecular de 1 kb. O gel foi submetido a eletroforese a 100 V e os resultados foram analisados no equipamento ImageQuant LAS 4000 (GE Healthcare Bio-Sciences).

3.9 Extração do DNA de bacteriófagos

Foram incubadas placas da passagem 7, diluição 1:10 (diluição em que foi verificado alto título de bacteriófagos, porém foi possível a visualização de placas de lise) com 3 mL de tampão SM, por 24 horas, a 4°C. Após esse período, a superfície das placas foi raspada com uma alça de Drigalsky e o tampão, contendo uma suspensão de células bacterianas íntegras, lisadas e partículas virais, foi recolhido com uma seringa e submetido à filtração em filtros com poros de 0,22 µm. O filtrado (aproximadamente 2 mL), aqui denominado lisado bacteriano, foi transferido para um tubo Falcon de 50 mL, tratado com 5 µl de DNase I e 5 µl de RNase A, e incubado a 37°C *overnight*. Uma alíquota dessa solução foi reservada para a realização de PCR, e o restante foi tratado para a lise das partículas virais. Para isso, a cada mL de lisado bacteriano, foram adicionados 12,5 µL de MgCl₂, 40 µL EDTA a 0,5 M, 5 µL de proteinase K (10 mg/mL) e 50 µL de SDS a 10%.

A mistura foi agitada vigorosamente, utilizando um vórtex, e incubada a 55°C por uma hora, agitando novamente os tubos a cada intervalo de 20 minutos. Em seguida, foram feitas alíquotas de 500 µl, em tubos *Eppendorf* de 1,5 mL, e a esse conteúdo foi adicionado igual volume de uma mistura de fenol:clorofórmio:álcool isoamílico (25:24:1) e os tubos foram invertidos e agitados

para misturar homogeneamente o conteúdo. Os tubos foram centrifugados a 21.000 *g* por 30 minutos, a 4°C.

A fase aquosa superior foi coletada e transferida para um novo tubo e novamente foi adicionado igual volume de fenol:clorofórmio:álcool isoamílico (25:24:1) e os tubos foram invertidos e agitados para misturar homogeneamente o conteúdo.

A fase aquosa superior foi coletada e transferida para um novo tubo e o DNA foi precipitado com o dobro do volume de isopropanol gelado e 1/10 do volume de acetato de sódio 3M. Esse material foi armazenado *overnight* a 4°C.

No dia seguinte, centrifugou-se por uma hora, a 21.000 *g*, a 4°C. A fase líquida foi removida cuidadosamente sem desfazer o *pellet*. Ao tubo, foram adicionados 500 µl de etanol 70% e centrifugou-se a 21.000 *g* por 20 minutos, a 4°C. A fase líquida foi removida e deixou-se o *pellet* secar por alguns minutos. O DNA foi dissolvido em 25 µl de 5 mM Tris-HCl (pH 8,0 a 8,5). Os tubos foram armazenados em geladeira, *overnight*, e no dia seguinte, o DNA foi quantificado em fluorímetro (Qubit) e armazenado em freezer a -20°C.

3.10 Sequenciamento e análise do material genético dos bacteriófagos

O DNA gênomico dos fagos passou por uma etapa de construção de bibliotecas, no qual se deu sua fragmentação, ligação de adaptadores e índice, utilizando o kit Nextera Library Preparation Kit (Illumina, USA). Posteriormente foi sequenciado de maneira *paired-end* (2x300bp) na plataforma MiSeq (Illumina, USA) seguindo as recomendações dos fabricantes. Essa etapa foi realizada na plataforma de sequenciamento de nova geração do Centro de Pesquisas René Rachou (FIOCRUZ-MG).

3.11 Montagem e análise dos genomas virais

As sequências geradas foram pré-processadas com objetivo de remover regiões de baixa qualidade. Para tal, foi utilizada a ferramenta Trimmomatic com os parâmetros para exclusão dos nucleotídeos próximos às extremidades 5' e 3' com qualidade menor que 5 na escala Phred (Q) (BOLGER et al., 2014). Também

foram excluídos fragmentos de sequências em que uma média do valor Q em uma janela de 4 nucleotídeos fosse menor que 20.

As sequências remanescentes após a filtragem foram mapeadas ao genoma da linhagem enterohemorrágica de *E. coli*, O157:H7 Sakai, obtido no banco de dados do Ensembl (HERRERO et al., 2016). Esse mapeamento, realizado pela ferramenta Bowtie2, teve como objetivo retirar as sequências do DNA bacteriano contaminante (LANGMEAD; SALZBERG, 2012). Fragmentos não mapeados ao genoma referência da bactéria foram selecionados aleatoriamente de maneira a permitir uma cobertura mínima de 120 vezes. A partir dessa seleção foram montados os genomas utilizando a ferramenta SPADES, versão 3.6, e o parâmetro “careful” para reduzir a probabilidade de erros (BANKEVICH et al., 2012). O procedimento foi repetido por cem vezes com o objetivo de reduzir os erros de sequenciamento que pudessem estar super-representados devido à alta cobertura inicial. Os cem genomas montados foram utilizados para a construção de um consenso através da ferramenta NUCmer (DELCHER et al., 2002). A predição de genes foi realizada pela ferramenta utilizando o modelo de predição para genes codificados por *E. coli* (LUKASHIN; BORODOVSKY, 1998). Tal modelo foi selecionado pelo fato de se tratar do hospedeiro primário dos bacteriófagos estudados, e é esperado que os genes a serem identificados sigam seu uso de códons.

As proteínas produzidas pelos genes preditos na etapa anterior foram classificadas funcionalmente utilizando uma busca ao banco de dados UniProt, executado pela ferramenta BLAST (www.ncbi.nlm.nih.gov/BLAST/). Os melhores resultados foram selecionados de acordo com um critério de cobertura maior que 50% da proteína referência do banco de dados, e melhor pontuação do BLAST (BATEMAN et al., 2015). Isso confere uma classificação mais relacionada ao contexto geral da proteína referência do que a um domínio funcional conservado. Cada uma das proteínas anotadas foi associada a um grupo KEGG Orthology (KO) através de ferramenta disponível na página do UniProt. (KANEHISA et al., 2016).

Para análise comparativa da origem dos genes, as proteínas codificadas foram agrupadas em ortólogos utilizando a ferramenta OrthoMCL (FISCHER et al.,

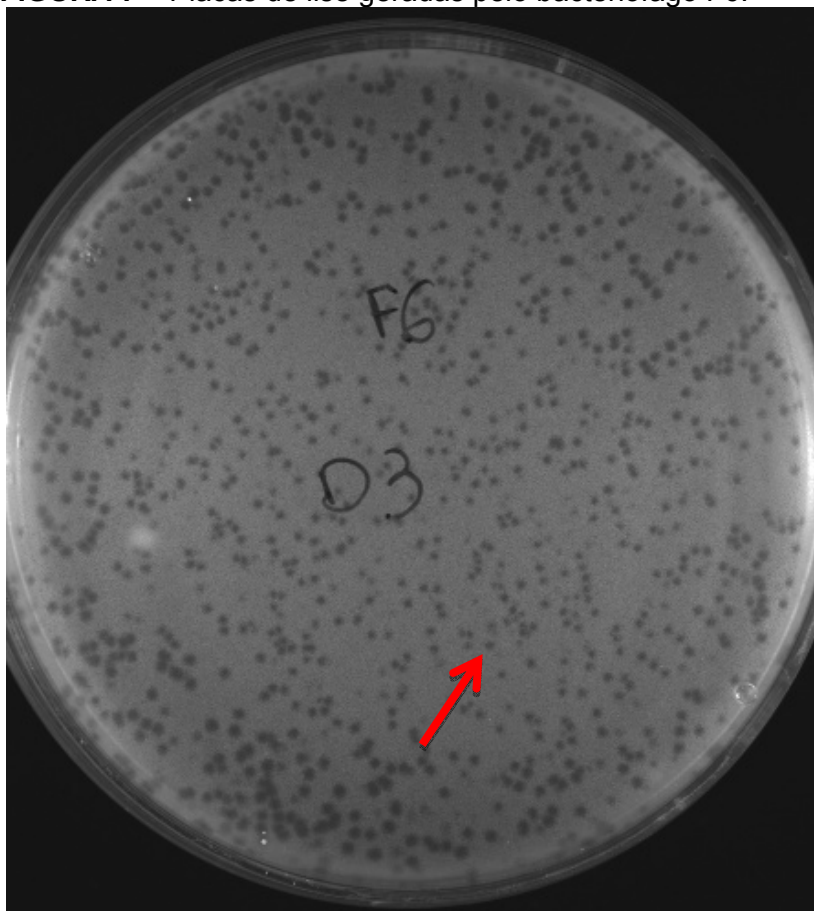
2011). Baseado nos resultados obtidos com o OrthMCL, foi feita a comparação das sequências dos genes ortólogos com a ferramenta BRIG visando analisar o nível de identidade dos órtologos em cada um dos genomas (ALIKHAN et al., 2011).

4 RESULTADOS

4.1 Diluição seriada dos vírus

Através da diluição seriada foi possível isolar 12 bacteriófagos, os quais foram denominados F1 a F12. Um exemplo de placas de lise geradas por esses bacteriófagos pode ser visto na **Figura 7**. Todos os bacteriófagos tiveram seu DNA extraído.

FIGURA 7 – Placas de lise geradas pelo bacteriófago F6.

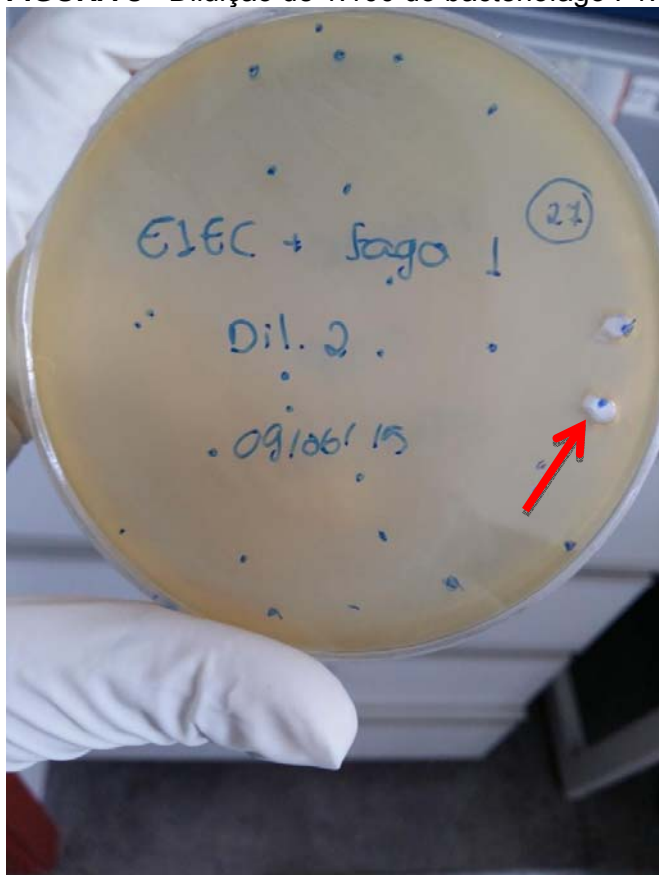


Acima está representada a placa de Ptri contendo o cultivo bacteriano juntamente com o bacteriófago denominado F6. As placas de lise geradas por esse vírus são identificadas pelas áreas claras, livres de crescimento bacteriano, como o exemplo indicado pela seta.

FONTE: elaborado pelo autor.

Ao final da etapa de diluição seriada, foram determinadas as diluições 1:100 e 1:1.000 como as ideais para o isolamento de placas de lise individuais, já que as placas geradas a partir dessas diluições não se sobrepuseram, reduzindo assim a possibilidade de isolamento de mais de uma linhagem viral (**Figura 8**).

FIGURA 8 - Diluição de 1:100 do bacteriófago F1.



As placas de lise estão identificadas como pontos. O número de placas de lise está circulado à direita (27). As placas de lise individuais foram retiradas para uso na próxima passagem, como apontado pela seta.

FONTE: elaborado pelo autor.

4.2 Comparação da carga de bacteriófagos da estação de tratamento de esgoto em diferentes coletas

As amostras coletadas no mês de abril apresentaram maior número de placas de lise do que as coletadas no mês de agosto (**Tabela 2**). Isso pode ter ocorrido pela possível influência das variações ambientais (como a temperatura) sobre as comunidades bacterianas nessas amostras (ADRIAENSSENS; COWAN, 2014).

TABELA 2 – Quantidade de placas de lise verificadas nas diferentes coletas.

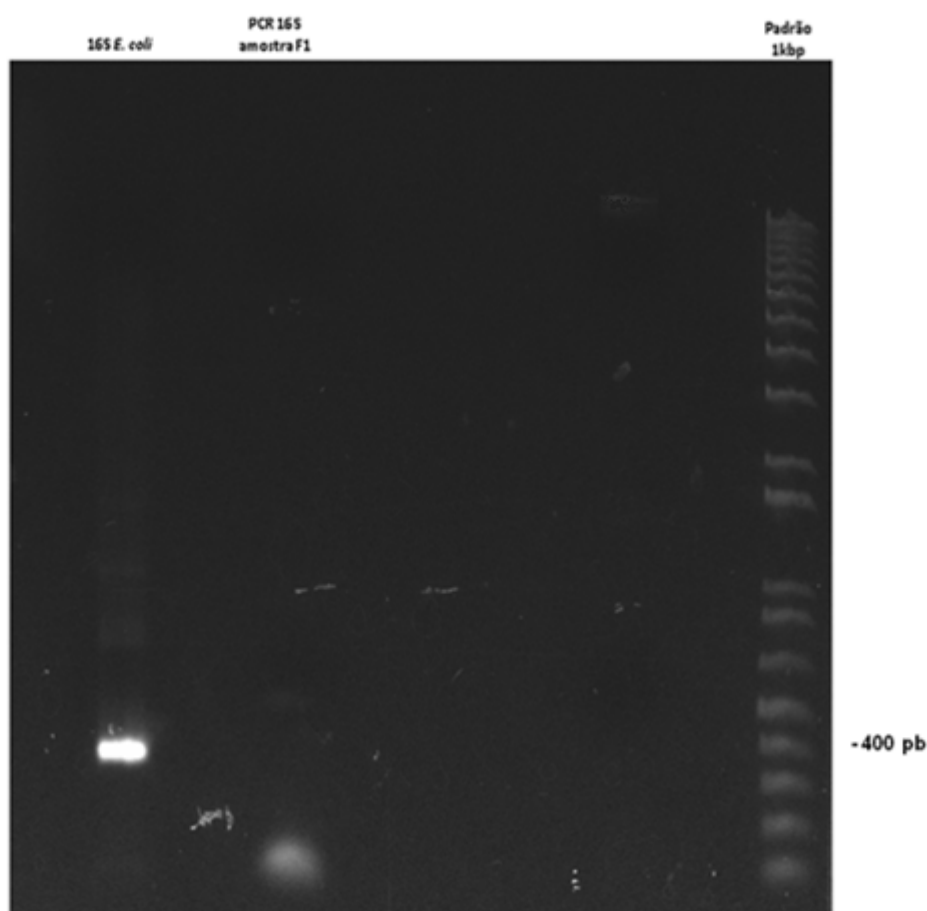
Mês de Coleta	Abril	Abril	Agosto	Agosto
Número de placas de lise	11	9	0	2

FONTE: elaborado pelo autor.

4.3 Verificação da pureza do extrato viral

A verificação da pureza dos lisados virais por PCR indicou que o DNA extraído estava livre de contaminação por DNA genômico bacteriano. A eletroforese do DNA extraído em gel de agarose a 0,5% indicou a presença de uma banda fraca de tamanho superior a 12 kpb (**Figura 9**), que possivelmente representa o genoma viral.

FIGURA 9 – Eletroforese em gel de agarose a 0,5%.



1 - Amplicon do gene 16S rRNA de *E. coli* (aproximadamente 400 pb); 2 - Produto de PCR com *primers* para o gene 16S rRNA bacteriano da amostra do lisado do fago F1 (ausência de amplicon).

FONTE: elaborado pelo autor.

4.4 Quantificação do DNA

Após a extração do material genético dos bacteriófagos seguindo os protocolos descritos na metodologia desse trabalho, o rendimento obtido pode ser visualizado na **Tabela 3**.

TABELA 3 - Rendimento do DNA isolado de bacteriófagos.

Identificação da amostra	Rendimento do DNA (ng)
F1	10.020
F2	3.180
F3	2.840
F4	400
F5	0
F6	1.710
F7	0
F8	1.860
F9	6.200
F10	2.250
F11	0
F12	6.000

FONTE: elaborado pelo autor.

Devido à baixa concentração do DNA extraído para F4, F5, F7 e F11 o material genético de apenas 8 dos 12 bacteriófagos isolados (F1, F2, F3, F6, F8, F9, F10 e F12) foi submetido a sequenciamento e posterior montagem e análise funcional.

4.5 Caracterização dos genomas virais

O sequenciamento deu origem a mais de 60.000 sequências por bacteriófago. O número exato de sequências geradas, filtradas e utilizadas na etapa de montagem dos genomas está indicado na **Tabela 4**.

TABELA 4 – Quantidade de sequências antes e após a filtragem.

	Sequências ao final do sequenciamento	Sequências após filtro qualidade	Sequências após filtro <i>E. coli</i>
F1	1.097.008	958.383	945.841
F2	279.072	237.293	148.717
F3	2.820.550	2.330.143	2.324.383
F6	732.676	668.44	666.68

F8	3.127.184	2.997.162	2.993.555
F9	332.393	306.321	305.284
F10	518.159	404.552	403.014
F12	34.911	31.932	31.763

A primeira coluna da tabela contém o número total de sequências geradas ao final do sequenciamento. A segunda coluna contém o número de sequências remanescentes após ser utilizado o filtro de qualidade pela ferramenta Trimmomatic. A terceira coluna contém o número final de sequências após a filtragem visando remoção de possível DNA bacteriano.

FONTE: elaborado pelo autor.

O genoma do bacteriófago F1 resultou em 38 contigs, sendo o seu tamanho aproximado igual a 93.557 pb e o conteúdo GC igual a 42%; F2 apresentou genoma formado por 522 contigs, com tamanho aproximado de 526.978 pb, sendo seu conteúdo GC igual a 40,4%; F3 resultou em 1 contig, com tamanho de 334.163 pb e conteúdo GC igual a 35,45%; F6 resultou em 1 contig de 677.814 pb e conteúdo GC igual a 37,64%; F8 deu origem a 1 contig com 339.235 pb e %GC equivalente a 37,64%; F9 resultou em 39 contigs, seu conteúdo GC foi igual a 48,90% e o tamanho do seu genoma é 191.788 pb. O bacteriófago F10 apresenta genoma com tamanho aproximado de 58.863 pb e conteúdo GC de 43.6%; F12 é formado por 108 contigs, tendo %GC = 47,80% e genoma com 128.523 pb. O tamanho médio dos contigs variou de 1.199 pb a 677.814 pb. O número de genes preditos nos genomas variou de 72 a 465. Todos os bacteriófagos tiveram mais de 50% dos seus genes sem função definida, como pode ser visto na **tabela 5**.

TABELA 5 – Características dos genomas montados.

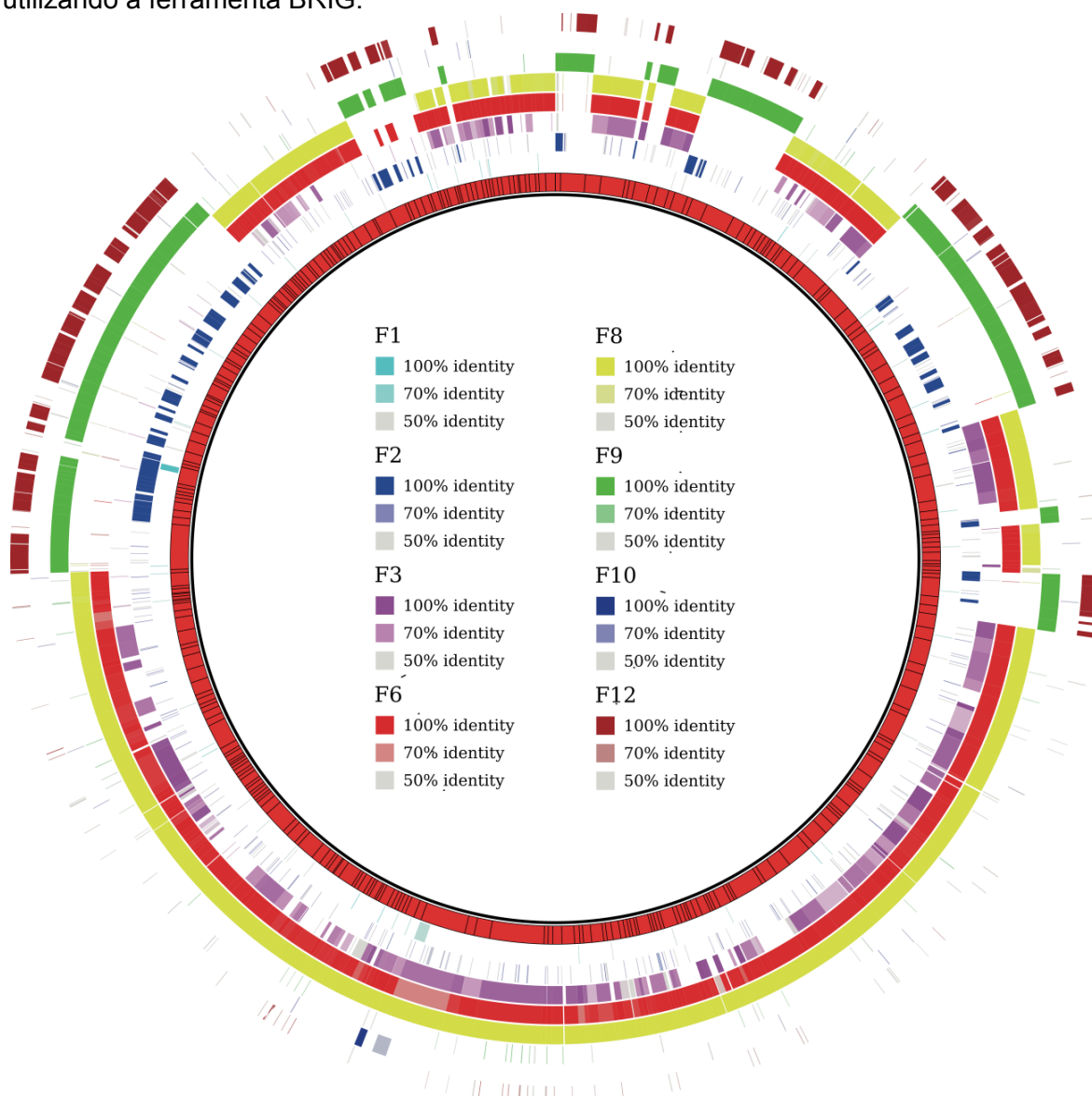
Bacteriófago	Contigs	Conteúdo G+C	Tamanho total	N50	Genes	Genes sem função descrita
F1	38	42%	93.557 pb	55.286 pb	105	67
F2	522	40,40%	526.978 pb	1.199 pb	465	375
F3	1	35,45%	334.163 pb	334.163 pb	254	155
F6	1	37,66%	677.814 pb	677.814 pb	256	153
F8	1	37,64%	339.235 pb	339.235 pb	253	147
F9	39	48,90%	191.788 pb	30.599 pb	208	154
F10	3	43.6%	58.863 pb	57.900 pb	72	49
F12	108	47,80%	128.523 pb	1.366 pb	183	125

Coluna 1: identificação dos fagos; 2: número de contigs montados; 3: conteúdo G+C dos genomas; 4: tamanho total dos genomas; 5: tamanho médio dos contigs; 6: número total de genes; 7: número de genes sem função descrita na literatura.

FONTE: elaborado pelo autor.

A comparação dos genes classificados como ortólogos pela ferramenta OrthoMCL sugere que os ortólogos dos vírus F3, F6 e F8 possuem alta identidade entre si, assim como os genes ortólogos dos bacteriófagos F9 e F12 (**Gráfico 1**).

GRÁFICO 1 – Comparação dos genes possivelmente ortólogos dos bacteriófagos utilizando a ferramenta BRIG.



A figura acima ilustra o resultado da comparação entre os genes possivelmente ortólogos. O círculo interno, em vermelho, é constituído por todos os genes dos bacteriófagos e é utilizado como referência para a comparação. Os possíveis ortólogos estão ilustrados exteriormente a ele. As diferentes cores estão relacionadas ao nome do bacteriófago e ao nível de identidade entre os genes, como pode ser visto na legenda do interior da figura.

FONTE: elaborado pelo autor.

A classificação funcional dos genomas dos bacteriófagos mostrou que todos os vírus sequenciados tiveram a presença de genes produtores de proteínas

líticas e apenas F2 teve a presença de genes relacionados ao ciclo lisogênico (**Quadro 2 a Quadro 9**). A relação de genes encontrados em cada genoma está na **Tabela 6 à Tabela 13**.

Quadro 2 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F1.

Metabolismo
6-carboxy-5,6,7,8-tetrahydropterin synthase
Deoxyuridine 5'-triphosphate nucleotide hydrolase
GTP cyclohydrolase 1
7-cyano-7-deazaguanine synthase
Organic radical activating enzyme
Síntese, empacotamento, recombinação e reparo de DNA
DNA polymerase beta subunit
DNA polymerase elongation subunit 1
RNAse H
ATP-dependent DNA ligase
Superfamily II DNA/RNA helicase
DNA primase
ATPase
Estruturais
Tail fiber assembly protein
Side tail fiber
Tape measure protein
Terminase large subunit
Tail subunit
Tail protein
Transferase
Queuosine tRNA-ribosyltransferase
Lise
Lysis protein

FONTE: elaborado pelo autor.

Quadro 3 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F2.

Metabolismo
Lipopolysaccharide core biosynthesis protein
Biotin sulfoxide reductase 2 (Fragment)
Bifunctional acetaldehyde-CoA/alcohol dehydrogenase, nonfunctional (Fragment)
Chain length determinant protein
Phospho-2-dehydro-3-deoxyheptonate aldolase (Fragment)
Síntese e empacotamento, recombinação e reparo de DNA
Gp61 DNA primase subunit
RNA-DNA + DNA-DNA helicase
DNA primase-helicase subunit
DNA binding protein
Gp33 T4-like late promoter transcription accessory protein
IS1 protein
Transcriptional regulatory domain protein
Transposase family protein
Resolvase, N terminal domain protein (Fragment)
Phage-related DNA-binding protein
Transposase family protein (Fragment)
VirF
Initiator RepB protein (Replication protein)
ATPase subunit of terminase family protein
Gp59 T4-like loader of gp41 DNA helicase
Transferase
Glycosyl transferase, group 2 family protein
Lipopolysaccharide glucosyltransferase I
UDP-galactose:(Glucosyl) LPS alpha-1,3-galactosyltransferase

Estruturais
Head completion protein
Tailspike protein
Gp15 proximal tail sheath stabilization protein
Gp59 T4-like loader of gp41 DNA helicase
Lise
Lyase
Lisogenia
Integrase
Imunidade
Colicin E1 immunity protein

FONTE: elaborado pelo autor.

Quadro 4 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F3.

Metabolismo
DN 3'phosphatase
Endonuclease II
NrdB aerobic NDP reductase, small subunit
Putative homing endonuclease
Putative thymidylate synthase
Dihydrofolate reductase, phage-associated
Anaerobic NTP reductase small subunit
Síntese, empacotamento, recombinação e reparo de DNA
RNA ligase 2 (Rnl2)
RNA-DNA and DNA-DNA helicase ATPase
DNA ligase
RNA ligase 1
Ribonucleoside-diphosphate reductase
Single-stranded DNA binding protein
DNA helicase loader protein

Topoisomerase II medium subunit
Gp15
Gp18
DNA helicase
RNA polymerase ADP-ribosylase
Mobile nuclease
DNA primase-helicase subunit
DNA polymerase
Clamp loader subunit DNA polymerase accessory protein
RNA polymerase binding protein
Recombination endonuclease subunit
AAA domain-containing protein
Thymidine kinase
Valyl-tRNA synthetase modifier
DNA end protector during packaging
RNA ligase 2
Estruturais
Major head protein
Capsid vertex protein gp24
Putative large head outer capsid protein
Base plate protein gp25
Baseplate hub subunit
Baseplate hub assembly protein
Base plate hub subunit
Baseplate distal hub subunit
Baseplate hub subunit tail length determinant
Capsid and scaffold protein
Tail fiber protein
Hypothetical membrane protein
Phage long tail fiber proximal subunit
Tail fiber protein p36

Long tail fiber distal subunit
Tail fiber adhesin
Small outer capsid protein
Tail completion and sheath stabilizer protein
Head completion protein
Phage baseplate wedge
Phage baseplate wedge initiator
Gp8 base plate wedge
Phage baseplate wedge
Baseplate wedge subunit and tail pin
Neck protein
Neck protein
Tail sheath stabilizer and completion protein
Terminase small subunit
Tail sheath protein
Tail tube protein
Portal vertex of the head
Prohead core protein
Prohead core protein
Prohead core scaffold protein
Gp24 of head vertex subunit
Transferase
Deoxycytidylate 5-hydroxymethyltransferase
Alpha-glucosyl-transferase
Alpha-glucosyl-transferase
Lise
Holin lysis protein
Inhibitor of host Lon protease
Lysozyme (EC 3.2.1.17)
Nudix hydrolase
Lysozyme (EC 3.2.1.17)

FONTE: elaborado pelo autor.

Quadro 5 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F6.

Metabolismo
Aerobic NDP reductase small subunit
Anaerobic NTP reductase large subunit
Deoxycytidylate deaminase
Deoxynucleoside monophosphate kinase
DNA endonuclease IV
dTMP thymidylate synthase
Frd (Frd dihydrofolate reductase)
Inh inhibitor of prohead protease gp21
Inhibitor of host transcription
Nucleoid disruption protein
Pin protease inhibitor
Síntese, empacotamento, recombinação e reparo de DNA
Activator of middle period transcription
Aerobic NDP reductase small subunit
AroG
Clamp loader small subunit
DNA helicase
DNA ligase
DNA polymerase (EC 2.7.7.7)
DNA topoisomerase II medium subunit
DsDNA-binding, transcription factor
Endonuclease II
Endonuclease V, N-glycosylase UV repair enzyme
Exonuclease subunit 2 (EC 3.1.11.-) (Protein Gp46)
Gp49 EndoVII packaging and recombination endonuclease VII

Helicase loading protein
Late promoter transcription accessory protein
Portal protein
RecA-like recombination protein
Recombination, repair and ssDNA binding protein
Ribonucleoside-diphosphate reductase (EC 1.17.4.1)
RnaseH
Single-stranded DNA binding protein
Sliding clamp
Terminase DNA packaging enzyme, large subunit
Thymidine kinase (EC 2.7.1.21)
Topoisomerase II large subunit
Transcription modulator under heat shock
Translational regulator
UvsW RNA-DNA and DNA-DNA helicase, ATPase
Vs valyl-tRNA synthetase modifier
Estruturais
Baseplate hub assembly
Baseplate hub distal subunit
Baseplate hub subunit
Baseplate hub subunit, tail length determinator
Baseplate tail tube junction
Baseplate tail tube junction
Baseplate wedge subunit
Baseplate wedge subunit and tail pin
Baseplate wedge tail fiber connector
Chaperone for tail fiber formation
Gp24 head vertex protein
Head completion protein
Head scaffolding protein
Internal head protein with nuclease inhibitor activity

Large head outer capsid protein
Major capsid protein
Neck protein
Phage tail fibers
Prohead core protein
Sheath completion
Tail completion and sheath stabilizer protein
Tail sheath protein
Tail tube protein
Lise
Lysozyme (EC 3.2.1.17)
Nudix hydrolase
Phage holin

FONTE: elaborado pelo autor.

Quadro 6 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F8.

Metabolismo
AroG
Deoxycytidylate deaminase
NrdG
Nucleoid disruption protein
Pin protease inhibitor
Thioredoxin
Thymidylate synthase
Síntese, empacotamento, recombinação e reparo de DNA
Activator of middle period transcription
ADP-ribosylase
Aerobic NDP reductase small subunit
Clamp loader small subunit
DenV endonuclease

Deoxynucleoside monophosphate kinase
DNA endonuclease IV
DNA helicase
DNA ligase
DNA polymerase (EC 2.7.7.7)
DNA topoisomerase II large subunit
DsDNA-binding, transcription factor
Endonuclease II
Endonuclease subunit
Exonuclease
Exonuclease subunit 2 (EC 3.1.11.-) (Protein Gp46)
Gp49 EndoVII packaging and recombination endonuclease VII
Gp52 topoisomerase II, medium subunit
Gp61 DNA primase subunit
Gp59
Large terminase protein
RecA-like recombination protein
Recombination, repair and ssDNA binding protein
RegA translational repressor protein
Ribonuclease H
Ribonucleoside-diphosphate reductase (EC 1.17.4.1)
RNA ligase
RNA polymerase ADP-ribosylase
RNaseH ribonuclease
Single-stranded DNA binding protein
Site-specific RNA endonuclease
Sliding clamp
Terminase small subunit
Thymidine kinase (EC 2.7.1.21)
Transcription modulator under heat shock
UvsW RNA-DNA and DNA-DNA helicase, ATPase

Vs valyl-tRNA synthetase modifier
Estruturais
Baseplate hub assembly
Baseplate hub distal subunit
Baseplate hub subunit, tail length determinator
Baseplate hub subunit
Baseplate tail tube junction
Baseplate wedge subunit
Baseplate wedge subunit and tail pin
Baseplate wedge tail fiber connector
Capsid vertex protein
Chaperone for tail fiber formation
Head completion protein
Head scaffolding protein
Internal head protein
Large head outer capsid protein
Long tail fiber distal subunit
Major capsid protein
Neck protein
Phage tail fibers
Portal protein
Prohead core protein
Sheath completion
Sheath protein
Tail completion and sheath stabilizer protein
Tail tube
Lise
Lysozyme (EC 3.2.1.17)

NudE nudix hydrolase
Phage holin

FONTE: elaborado pelo autor.

Quadro 7 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F9.

Metabolismo
Nucleoside triphosphate pyrophosphohydrolase
Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)
Transferase
Nicotinamide phosphoribosyltransferase (EC 2.4.2.12)
Síntese, empacotamento, recombinação e reparo de DNA
DNA binding protein
DNA helicase
DNA ligase
DNA polymerase (EC 2.7.7.7)
DNA primase-helicase subunit
DNA topoisomerase
DNA topoisomerase II (Fragment)
DnaB-like helicase C terminal domain protein
Gp16 terminase DNA packaging enzyme small subunit
Gp17 terminase DNA packaging enzyme large subunit
Gp32 T4-like ssDNA binding protein
Gp33 T4-like late promoter transcription accessory protein
Gp44 clamp loader subunit DNA polymerase accessory protein
Gp45 sliding clamp DNA polymerase accessory protein

Gp52 topoisomerase II medium subunit
Gp61 DNA primase subunit
Gp62 clamp loader subunit DNA polymerase accessory protein
Helicase loader
Initiator RepB protein (Replication protein)
Recombination and repair protein UvsX
RegA translational repressor protein
Ribonucleoside-diphosphate reductase (EC 1.17.4.1)
Ribonucleotide reductase of class Ia (Aerobic) beta subunit (EC 1.17.4.1)
RNA-DNA and DNA-DNA helicase UvsW
RNAseH
UvsX RecA-like recombination protein
Estruturais
Base plate family protein
Gp14 neck protein
Gp15 proximal tail sheath stabilization protein
Gp20 portal vertex protein of head
Gp22 prohead core protein
Gp23 major capsid protein (Fragment)
Gp3 tail completion and sheath stabilizer protein
Gp4 head completion protein
Gp48 putative tail tube associated base plate protein
Gp5 baseplate hub subunit and tail lysozyme
Gp53 baseplate wedge subunit

Gp6 baseplate wedge subunit
Major capsid protein g23 (Fragment)
Neck protein gp14 (Gene product 14) (gp14)
Tail fiber protein
Tail tube protein
Tailspike protein
Lise
Lyase (Fragment)

FONTE: elaborado pelo autor.

Quadro 8 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F10.

Metabolismo
Deoxyuridine 5'-triphosphatenucleotidohydrolase
GTP cyclohydrolase 1
6-carboxy-5,6,7,8-tetrahydropterin synthase
7-cyano-7-deazaguanine synthase
Organic radical activating enzyme
Síntese, empacotamento, recombinação e reparo de DNA
Superfamily II DNA/RNA helicase
ATP-dependent DNA ligase
DNA primase
DNA polymerase elongation subunit 1
DNA polymerase beta subunit
RNAse H
ATPase
Estruturais
Terminase large subunit
Main coat protein

Tail protein
Tail subunit
Tape measure protein
Side tail fiber
Tail fiber assembly protein
Host specificity protein
Transferase
Queuosine tRNA-ribosyltransferase
Lise
Lysis protein
Phage holin

FONTE: elaborado pelo autor.

Quadro 9 - Classificação funcional de algumas das proteínas produzidas pelos genes do bacteriófago F12.

Síntese, empacotamento, recombinação e reparo de DNA
Topoisomerase II large subunit (Topoisomerase II large subunit C-terminal region)
Recombination endonuclease subunit
Gp43 DNA polymerase
RNA-DNA + DNA-DNA helicase
Terminase DNA packaging enzyme large subunit
DNA primase-helicase subunit
DNA helicase
RNAseH
PhoH-like protein
Gp33 T4-like late promoter transcription accessory protein
Large terminase subunit
DNA topoisomerase 2
Gp45 sliding clamp DNA polymerase accessory protein
Gp61 DNA primase subunit
Gp52 topoisomerase II medium subunit
DNA polymerase

Estruturais
Tail fiber protein
Tailspike protein
Tail tube protein
Gp48 baseplate tail tube cap
Gp3 tail completion and sheath stabilizer protein
Gp53 baseplate wedge subunit
Gp15 proximal tail sheath stabilization protein
Gp5 baseplate hub subunit and tail lysozyme
Gp14 neck protein
Viral capsid assembly protein
Gp23 major capsid protein
Lise
Lyase

FONTE: elaborado pelo autor.

TABELA 6 – Resultado da anotação dos genes do bacteriófago F1.

Identificador	Proteína	Gene	Tamanho
X2KTT1	6-carboxy-5,6,7,8-tetrahydropterin synthase	9g_00005	177
X2KU64	GTP cyclohydrolase 1	9g_00004	204
X2L008	Queuosine tRNA-ribosyltransferase	9g_00003	317
X2KSN7	DNA polymerase beta subunit	9g_00002	330
X2KP98	DNA polymerase elongation subunit 1	9g_00001	657
X2KPE1	Tail fiber assembly protein	9g_00071	192
X2KTY7	Side tail fiber	9g_00070	1211
X2L058	Host specificity protein	9g_00068	1029
X2KUA9	Tape measure protein	9g_00064	1086
X2L019	Deoxyuridine 5'-triphosphatenucleotidohydrolase	9g_00018	169
X2KPA7	RNAse H	9g_00016	196
X2KTT7	ATP-dependent DNA ligase	9g_00015	291
X2KU72	ATPase	9g_00014	256
X2KPA5	Superfamily II DNA/RNA helicase	9g_00011	667
X2KPD4	Tail subunit	9g_00061	393
X2KUA6	Tail protein	9g_00059	139
X2KTU7	DNA primase	9g_00025	755
X2KTW4	Lysis protein	9g_00045	166
D2AJH9	IpaJ, invasion plasmid antigen	ipaJ SFxv_4965	242
B3X714	Protein IpgB	Sd1012_4490	113
B3X756	OspC2	Sd1012_3244	484
I6CHT0	Invasin IpaB family protein	SFK315_3091	206
J2ZHD4	Spa47 domain protein	spa47 SSMOSELEY_1368	68
Q55285	ORF1		60
F5P4G8	Invasin ipaA domain protein	SFK227_5406 SFK227_5434	197
J2N7U0	Bacterial regulatory helix-turn-helix s, AraC family protein	SSMOSELEY_0002	234
J2G911	Bacterial type II/III secretion system short domain protein	SSMOSELEY_1375	199
Q8KXQ9	IpaB (Fragment)	ipaB	306
F3V4P5	NleE domain protein	SD15574_1405	44
F3V305	OspD1 domain protein	SD15574_0797	41
B2TTA0	Protease 7 (EC 3.4.23.-)	SbBS512_A0311	326
A0A1K9	Initiator RepB protein (Replication protein)	repE ERS085366_04868 UC41_06395	233
A4SH38	Colicin E1 immunity protein	imm SSON_PB02	113
D8BRD0	Colicin pore forming domain protein	HMPREF9551_00172	460
X2L013	Organic radical activating enzyme	9g_00008	254
X2KPA0	7-cyano-7-deazaguanine synthase	9g_00006	471
B2TSZ7	MxiC	SbBS512_A0180	355

FONTE: elaborado pelo autor.

TABELA 7 – Resultado da anotação dos genes do bacteriófago F2.

Identificador	Proteína	Gene	Tamanho
H9UWA1	Surface presentation of antigens protein SpaS	spaS P12B_c2957	243
K3KK56	Type III secretion apparatus protein, YscR/HrcR family	epaP ECN1_3018	221
H9UWA5	Type III secretion system apparatus protein YscQ/HrcQ	epaO P12B_c2961	315
H9UQQ1	IS1 protein	P12B_c0974	167
Q31SP3	Plasmid segregation protein	parA SBO_P043	399
Q31SM4	IS600 ORF1	SBO_P062	100
A0A070SGH2	Transcriptional regulatory domain protein	yqeI AC45_5537	139
I6FGP2	Transposase family protein	SD22575_4945	249
S1I8K1	Radical SAM additional 4Fe4S-binding SPASM domain-containing protein	A1WS_02075	203
Q7DC26	RhsC protein in rhs element	Z0848	73
A0A0B1MVP7	Surface exclusion protein	PU06_18645	165
A4SH38	Colicin E1 immunity protein	imm SSON_PB02	113
A0A069WWQ 5	Resolvase, N terminal domain protein (Fragment)	AD16_5641	78
N2QN37	NAD-dependent dihydropyrimidine dehydrogenase sunbunit PreT domain protein	EC2875150_2510	53
U5PUI9	DNA binding protein	Marshall_98	92
W1V5I2	Fimbrial-like protein ybgD (Fragment)	Q610_ECBC00376G0001	61
T6L1D4	TraT complement resistance protein	G749_05193	91
B7MDH1	Toxin of the YoeB-YefM toxin-antitoxin system	yoeB ECS88_2116	84
B7MDH2	Antitoxin	yefM ECS88_2117	83
B7MCP7	Probable transcriptional regulatory protein YeeN	yeeN ECS88_2051	238
I3VZV0	Molecular chaperone		125
H9UYH6	Glycosyl transferase, group 2 family protein	waaV P12B_c3752	327
E3PMY9	Lipopolysaccharide core biosynthesis protein	ETEC_3868	230
H9UYI0	Lipopolysaccharide glucosyltransferase I	rfaJ P12B_c3756	331
E3PMZ1	UDP-galactose:(Glucosyl) LPS alpha-1,3- galactosyltransferase	ETEC_3870	338
W1XHF8	Biotin sulfoxide reductase 2 (Fragment)	Q604_UNBC15760G0001	132
K8WUM7	Phage-related DNA-binding protein	OOA_11358	120
F5B3T4	Head completion protein	gp4 SCRM01_064c	139
H8ZNH7	Gp61 DNA primase subunit		104
Q9XJ82	Gp17v (Fragment)	17v	248
W1YL15	Bifunctional acetaldehyde-CoA/alcohol dehydrogenase, nonfunctional (Fragment)	Q604_UNBC04170G0001	60
K0X8D6	Transposase family protein (Fragment)	SF148580_1305	159
Q31SN4	VirF	virF SBO_P052	262
L2VDS0	Protein yiaL	WCM_00987	92
I6DSJ3	Mutated periplasmic disulfide isomerase	SB96558_0843	40
Q3YTS0	OspC3	ospC3 SSON_P072	484

Q327J0	Acid phosphatase (EC 3.1.3.2)	apy phoN2 SDY_P004	246
Q3YTY8	OspB	ospB SSON_P003	288
A0A1K9	Initiator RepB protein (Replication protein)	repE ERS085366_04868 UC41_06395	233
U9XZE6	Chain length determinant protein	HMPREF1589_03529	137
B2TT20	ISL3 family transposase	SbBS512_A0206	125
B2TSM6	Killing factor	SbBS512_A0024	239
Q31SS0	OspD2	ospD2 SBO_P016	569
A4SH43	Entry exclusion protein 2	exc SSON_PB07	139
B6IC08	ParB protein	ECSE_P1-0040	138
A0A0B1M885	Integrase (Uncharacterized protein)	PU06_28870 SM09_00423	75
U9ZSU3	Hemagglutinin (Fragment)	HMPREF1620_04091	126
K3KMD8	Sensor evgS domain protein (EC 2.7.3.-)	evgS ECN1_2506	1127
F5NW70	IbrB domain protein	SFK227_2383	112
C8URC9	Probable Replication regulatory protein RepA2	ECO111_p3-14	85
C8XUV3	SegD-like protein	orf00242	228
C8XUD1	Cd allosteric enzyme	orf00033	168
I6YMU9	DNA primase-helicase subunit		137
W1XLL6	Flagellar hook-basal body complex protein FlIE (Fragment)	Q604_UNBC14451G0001	52
K0XMD1	ATPase subunit of terminase family protein	SF148580_5100	156
W1V3P7	Bacterial transcription activator, effector binding protein (Fragment)	Q610_ECBC00105G0001	122
A0A0A6ZZY1	Transposase	Asd1617_04379	43
I6B9S2	Inner membrane protein CbrB	cbrB SF285071_4460	118
D2AJ59	IS3 orf2	SFxfv_4845	39
F3V580	Gifsy-2 prophage protein	SD15574_1593	111
F3V579	Exonuclease VIII domain protein	SD15574_1592	45
W1FTA4	Phosphoglycerate transporter protein PgtP		148
C8XUP1	RegA translational repressor protein	orf00162	154
W1Y4H9	Transcriptional activator feaR (Fragment)	Q604_UNBC09665G0001	87
W1XX30	Low affinity tryptophan permease protein (Fragment)	Q604_UNBC10698G0001	114
W1YC61	Tryptophanase protein (Fragment)	Q604_UNBC07235G0001	65
A0A029MPP8	DnaJ domain protein	AB99_1984	647
C8XUD9	Gp59 T4-like loader of gp41 DNA helicase	orf00041	221
S1D332	Protein rhsA (Fragment)	A1U1_03810	184
Q31SK3	TonB-like colicin Js sensitive putative membrane protein	cjrB SBO_P083	258
W1YE43	CjrA protein (Fragment)	Q604_UNBC05439G0001	63
I6D0J8	OspC3 domain protein	SFK315_0965	66
V0ADC6	Fimbrial usher protein	HMPREF1620_01518	273
C8UH79	Hcp-like protein	ECO111_0217	160
F4NI87	Surface composition regulator	glgS SSJG_01882	68
M9ENX1	Fimbrial family protein	EC174750_3347	163
W1X8I1	Phospho-2-dehydro-3-deoxyheptonate	Q609_ECAC00699G0001	57

	aldolase (Fragment)		
B7L5C1	DicA protein	dicA EC55989_1734	135
M2NV47	Transporter protein (Fragment)	A364_23868	196
Q31SK9	OspC1	ospC1 SBO_P077	470
C8XUS2	Gp15 proximal tail sheath stabilization protein	orf00199	231
E3SMV7	RNA-DNA + DNA-DNA helicase	uvsW PHM1_126	279
IOVY98	TRAP transporter, DctM subunit	ECW26_03730	310
F8X6A9	Lipopolysaccharide core biosynthesis glycosyl transferase RfaQ	HMPREF9349_00263	64
X4YGA2	Tailspike protein		185
M4MEQ6	Lyase (Fragment)	gp17	713
E1JES9	Transcriptional regulator, AlpA family	HMPREF9347_05485	111
C8XUH2	Gp33 T4-like late promoter transcription accessory protein	orf00076	81
Q5GFL8	Gnd (Fragment)	Gnd	86

FONTE: elaborado pelo autor.

TABELA 8 – Resultado da anotação dos genes do bacteriófago F3.

Identificador	Proteína	Gene	Tamanho
A0A0A0PU16	AAA domain-containing protein	HY01_0088	322
A0A097J4G8	Activator of host PrrC lysyl-tRNA endonuclease	stp RB9_261	26
G0X5W8	Adenosylribosyl-transferase packaged injected with DNA		697
A0A097J8M4	Alpha-glucosyl-transferase	a-gt RB68_063	400
A0A097J8Q6	Anaerobic NTP reductase small subunit	nrdG RB68_080	156
A0A097J922	Base plate hub subunit	RB68_190	390
Q712G0	Base plate protein gp25	25	132
D4Z9Z2	Baseplate distal hub subunit	28 AR1_196	191
A0A097J8Z0	Baseplate hub assembly protein	RB68_189	250
G0X5V9	Baseplate hub subunit		208
A0A097J905	Baseplate hub subunit tail length determinator	RB68_192	590
A0A097J7B8	Baseplate wedge subunit and tail pin	RB55_p155	602
I6ZYE7	Capsid and scaffold protein	ECML134_207	111
D4Z9X5	Capsid vertex protein gp24 (Gene product 24) (gp24)	24 AR1_179	427
A0A097J8M0	Clamp loader subunit DNA polymerase accessory protein	RB68_051	319
S5M332	dCTPase		172
G1FHP2	Deoxycytidylate 5-hydroxymethyltransferase	gp42	246
I7LF46	Dihydrofolate reductase, phage-associated	g241 BN81_241	193
C3V1T0	DN 3'phosphatase	pseT RB14ORF218	301
A0A097J8V4	DNA end protector during packaging	RB68_149	274
A0A097J641	DNA helicase	dda RB33_016	439

A0A0A0PSM9	DNA helicase loader protein	HY01_0233	217
A0A023ZW70	DNA ligase	e112_210	487
A0A097J8K9	DNA polymerase (EC 2.7.7.7)	RB68_048	898
K4FCG3	DNA primase-helicase subunit	ACG-C40_0039	475
A0A097J925	Endonuclease II	denA RB68_226	136
C3V1Y3	Endonuclease IV	denB RB14ORF271	185
Q8H9J6	Endoribonuclease	regB	153
A0A097BWM2	Gp15	pSs1_0014	312
A0A097BWL3	Gp18	pSs1_0017	605
C3V1N5	Gp24 of head vertex subunit	24 RB14ORF173	427
A0A097BWM5	Gp29	pSs1_0028	81
Q08I13	Gp30.1	30.1	89
A0A097BWS4	Gp79	pSs1_0078	108
C3V1L6	Gp8 base plate wedge	8 RB14ORF154	334
A0A097BWT1	Gp80	pSs1_0079	66
A0A097J8Y0	Head completion protein	RB68_150	150
D4ZA56	Holin lysis protein	t AR1_260	218
A0A097J4K1	Homing endonuclease	segG RB9_237	210
A0A097J3Z7	Inhibitor of host Lon protease	pin RB9_080	161
K4FC21	Inhibitor of prohead protease	ACG-C40_0178	226
A0A060BN58	Long tail fiber distal subunit	JS09_0201	1080
A0A0A0PUN6	Lysozyme (EC 3.2.1.17)	HY01_0140	575
I7B328	Lysozyme (EC 3.2.1.17)	ECML134_123	164
A0A0A0PUJ6	Macro domain-containing protein	HY01_0105	159
S5M3C4	Major head protein		521
A0A0A7HEC4	Mobile nuclease	mobE VR5_028	218
D4Z9G8	ModB ADP-ribosylase	modB AR1_022	207
A0A097J3T3	Modifier of transcription	motB RB9_011	157
A0A0A0PUH7	Mur ligase domain-containing protein	HY01_0085	135
S5M7B8	Neck protein		309
A0A023ZVH7	Neck protein	e112_173	256
D4ZA33	NrdB aerobic NDP reductase, small subunit	nrdB AR1_237	392
D4Z9M2	NrdH glutaredoxin	nrdH AR1_076	140
D4ZA69	Nuclear disruption protein	ndd AR1_273	147
D4Z9S4	Nudix hydrolase	nudE AR1_128	151
I7A903	Phage baseplate wedge	ECML134_145	196
I6ZYA9	Phage baseplate wedge	ECML134_152	288
I7KRS9	Phage baseplate wedge initiator	g159 BN81_159	1032
G1FHG6	Phage long tail fiber proximal subunit	gp34	1289
G1FH59	Phage protein	e.8	109
A0A097J901	Portal vertex of the head	RB68_170	524
D4Z9G6	Postulated decoy of host sigma70 or sigmaS	srd AR1_020	251
A0A097J8X6	Prohead core protein	RB68_171	78
A0A097J8Y3	Prohead core protein	RB68_172	141
A0A097J8X7	Prohead core scaffold protein	RB68_174	269
I7B359	Prohead core scaffold protein and protease	ECML134_168	212

A0A023ZUE0	Recombination endonuclease subunit	e112_059	560
Q0ILH2	Ribonucleoside-diphosphate reductase (EC 1.17.4.1)	nrdA	754
G0X5C4	RIIA protein		725
G1FHE7	RNA ligase 1	rnIA	374
P32277	RNA ligase 2 (Rnl2) (EC 6.5.1.3)	Y10A_24.1	334
S5M9D8	RNA ligase 2	rnIB	334
G0X5W7	RNA polymerase ADP-ribosylase		685
A0A023ZVR7	RNA polymerase ADP-ribosylase	e112_021	200
S5M716	RNA polymerase binding protein	rpbA	129
A0A023ZUU5	RNA-DNA and DNA-DNA helicase ATPase	e112_192	80
A0A097J8Z3	RNA-DNA and DNA-DNA helicase ATPase	uvsW RB68_182	503
G0X593	RNase H ribonuclease		305
A0A023ZU24	Sigma factor	e112_068	187
S5ML91	Single-stranded DNA binding protein		301
A0A097J5E2	Small outer capsid protein	soc RB27_030	85
A0A097J8W9	Tail completion and sheath stabilizer protein	RB68_148	176
A0A097J7P2	Tail fiber adhesin	RB55_p247	183
I7AU89	Tail fiber protein	ECML134_209	78
I7B392	Tail fiber protein	ECML134_208	102
G1FHG8	Tail fiber protein p36	gp36	221
A0A097J7A8	Tail sheath protein	RB55_p164	659
A0A097J7B7	Tail sheath stabilizer and completion protein	RB55_p161	272
A0A097J8X1	Tail tube protein	RB68_169	163
I7A916	Terminase small subunit	ECML134_160	164
A0A0A0PUJ1	Thymidine kinase (EC 2.7.1.21)	HY01_0100	193
G0X5B0	Topoisomerase II medium subunit		442
A0A097J8J7	Transcription modulator	srh RB68_026	67
A0A097J8H9	Transcription modulator under heat shock	mrh RB68_027	161
A0A097J8T5	Valyl-tRNA synthetase modifier	vs RB68_118	113

FONTE: elaborado pelo autor.

TABELA 9 – Resultado da anotação dos genes do bacteriófago F6.

Identificador	Proteína	Gene	Tamanho
A0A067ZJI3	Activator of middle period transcription	PhAPEC2_245	210
A0A060BN31	ADP-ribosyltransferase	JS09_0248	695
A0A067ZHI2	Aerobic NDP reductase small subunit	PhAPEC2_219	392
A0A067ZJA5	Anaerobic NTP reductase large subunit	PhAPEC2_75	605
C3V1D8	Anaerobic nucleotide reductase subunit	nrdG RB14ORF76	156
A0A088FVY9	AroG		330
A0A088FVM3	Baseplate hub assembly		250
A0A088FVX6	Baseplate hub distal subunit		177
A0A088FRD6	Baseplate hub subunit		208
A0A088FR65	Baseplate hub subunit		390

A0A060BHK9	Baseplate hub subunit, tail length determinant	JS09_0252	590
A0A088FRE0	Baseplate tail tube junction		369
A0A088FVM8	Baseplate tail tube junction		320
A0A060BDK5	Baseplate wedge subunit	JS09_018	657
A0A060BHJ3	Baseplate wedge subunit	JS09_017	1032
A0A060BHJ8	Baseplate wedge subunit	JS09_022	191
E1A1H9	Baseplate wedge subunit	phiAS4_ORF0131	304
A0A060BN47	Baseplate wedge subunit	JS09_016	334
A0A088FSA4	Baseplate wedge subunit		132
A0A060BDK0	Baseplate wedge subunit and tail pin	JS09_013	219
A0A060BGV9	Baseplate wedge subunit and tail pin	JS09_014	601
A0A060BMC1	Baseplate wedge tail fiber connector	JS09_015	290
A0A060BN74	Chaperone for tail fiber formation	JS09_027	76
A0A060BNF8	Clamp loader small subunit	JS09_0127	320
A0A088FRZ1	Clamp loader subunit		187
A0A088FRF3	Co-chaperone for GroEL		110
A0A088FQQ8	dCTPase		173
A0A067ZHH3	Deoxycytidylate deaminase	PhAPEC2_204	193
A0A060BME6	Deoxynucleoside monophosphate kinase	JS09_026	244
Q7Y599	Dmd discriminator of mRNA degradation	Dmd	63
A0A067ZK51	DNA endonuclease IV	PhAPEC2_252	185
A0A060BHA6	DNA helicase	JS09_0166	437
A0A088FR13	DNA helicase		480
A0A060BE80	DNA ligase	JS09_0246	497
A0A060BH71	DNA polymerase (EC 2.7.7.7)	JS09_0130	903
A0A060BNM5	DNA topoisomerase II medium subunit	JS09_0188	441
A0A088FRH2	DsDNA-binding, transcription factor		95
A0A060BNP6	dTMP thymidylate synthase	JS09_0214	286
A0A060BMZ5	Endonuclease II	JS09_0218	136
A0A067ZK98	Endonuclease subunit	PhAPEC2_58	339
A0A067ZKE0	Endonuclease V, N-glycosylase UV repair enzyme	PhAPEC2_113	137
Q7Y5S0	Endoribonuclease	regB	152
A0A060BMT9	Exonuclease	JS09_0167	225
O64299	Exonuclease subunit 2 (EC 3.1.11.-) (Protein Gp46)	46	562
A0A060BN42	Fibritin	JS09_011	482
Q7Y267	Frd (Frd dihydrofolate reductase)	Frd	195
Q7Y4V9	Gp24 head vertex protein	24	427
Q7Y557	Gp49 EndoVII packaging and recombination endonuclease VII	49	157
Q7Y5A8	Gp61 DNA primase subunit	58 ^a	340
A0A088FR25	Head completion protein		149
A0A088FRC5	Head scaffolding protein		270
A0A067ZK36	Helicase loading protein	PhAPEC2_227	217
Q7Y4V4	Inh inhibitor of prohead protease gp21	Inh	222

A0A067ZJ23	Inhibitor of host transcription	PhAPEC2_216	187
A0A067ZJD8	Internal head protein with nuclease inhibitor activity	PhAPEC2_140	95
Q06EE8	Large head outer capsid protein	hoc RB32ORF178c	376
A0A067ZKJ5	Late promoter transcription accessory protein	PhAPEC2_228	112
A0A088FQU5	Late sigma transcription factor		185
A0A067ZKF7	Lysozyme (EC 3.2.1.17)	PhAPEC2_148	577
A0A060BDQ1	Lysozyme (EC 3.2.1.17)	JS09_054	162
A0A088FR50	Major capsid protein		522
Q7Y5B9	ModB ADP-ribosylase	modB	193
A0A067ZIU2	Modifier of transcription	PhAPEC2_11	151
A0A060BNK9	mRNA metabolism modulator	JS09_0173	136
A0A088FRB6	Neck protein		308
A0A088FVK3	Neck protein		254
A0A060BHC7	Nucleoid disruption protein	JS09_0186	148
A0A067ZIY1	Nudix hydrolase	PhAPEC2_116	152
I7LF52	Phage holin	g256 BN81_256	219
I7K3Z1	Phage tail fibers	g255 BN81_255	183
I7KRW6	Phage tail fibers	g254 BN81_254	1038
Q7Y555	Pin protease inhibitor	Pin	137
A0A088FR45	Portal protein		523
X2KLC5	Prohead core protein	JS09_02	78
A0A088FS90	Prohead core protein		141
A0A067ZK06	Prohead core protein protease	PhAPEC2_167	213
A0A088FQS2	RecA-like recombination protein		390
A0A060BHI9	Recombination, repair and ssDNA binding protein	JS09_0258	156
A0A076G8C3	RI lysis inhibition regulator	Av05_00199	100
A0A060BE47	Ribonucleoside-diphosphate reductase (EC 1.17.4.1)	JS09_0216	751
A0A088FV76	RIIA protein		737
A0A060BE15	RIIB protector from prophage-induced early lysis	JS09_0180	315
A0A060BE70	RIII lysis inhibition accessory protein	JS09_0236	82
A0A067ZK08	RNA ligase	PhAPEC2_172	332
A0A060BNQ3	RNA ligase A	JS09_0219	374
A0A060BNK0	RNA polymerase ADP-ribosylase	JS09_0163	202
A0A088FQT7	RNA polymerase-binding protein		138
A0A060BI41	RnaseH	JS09_0205	305
A0A088FR40	Sheath completion		273
A0A067ZJ29	Single-stranded DNA binding protein	PhAPEC2_226	300
A0A088FVC1	Sliding clamp		228
A0A088FVK4	Spackle		97
A0A060BGY3	Tail completion and sheath stabilizer protein	JS09_025	194
X2KQU6	Tail sheath protein	JS09_05	660
A0A067ZKG3	Tail tube protein	PhAPEC2_163	163

A0A067ZIZ4	Terminase DNA packaging enzyme, large subunit	PhAPEC2_161	611
A0A088FVV8	Terminase small subunit		164
A0A060BHR2	Thymidine kinase (EC 2.7.1.21)	JS09_073	193
A0A067ZH10	Topoisomerase II large subunit	PhAPEC2_4	605
A0A067ZIU6	Transcription modulator under heat shock	PhAPEC2_21	154
A0A088FVL6	Translational regulator		125
Q7Y4V3	UvsW RNA-DNA and DNA-DNA helicase, ATPase	UvsW	504
Q7Y529	Vs valyl-tRNA synthetase modifier	VS	115

FONTE: elaborado pelo autor.

TABELA 10 – Resultado da anotação dos genes do bacteriófago F8.

Identificador	Proteína	Gene	Tamanho
A0A067ZIJ3	Activator of middle period transcription	PhAPEC2_245	210
A0A060BMT5	ADP-ribosylase	JS09_0162	193
A0A067ZHI2	Aerobic NDP reductase small subunit	PhAPEC2_219	392
A0A088FVY9	AroG		330
A0A088FVM3	Baseplate hub assembly		250
A0A088FVX6	Baseplate hub distal subunit		177
A0A088FRD6	Baseplate hub subunit		208
A0A088FR65	Baseplate hub subunit		390
A0A067ZJ06	Baseplate hub subunit, tail length determinator	PhAPEC2_186	590
A0A088FRE0	Baseplate tail tube junction		369
A0A088FVM8	Baseplate tail tube junction		320
A0A060BDK5	Baseplate wedge subunit	JS09_018	657
A0A060BHJ3	Baseplate wedge subunit	JS09_017	1032
A0A060BHJ8	Baseplate wedge subunit	JS09_022	191
A0A060BN47	Baseplate wedge subunit	JS09_016	334
A0A088FSA4	Baseplate wedge subunit		132
A0A060BDK0	Baseplate wedge subunit and tail pin	JS09_013	219
A0A060BGV9	Baseplate wedge subunit and tail pin	JS09_014	601
A0A060BMC1	Baseplate wedge tail fiber connector	JS09_015	290
A0A060BDM3	Capsid vertex protein	JS09_0269	427
A0A060BN74	Chaperone for tail fiber formation	JS09_027	76
A0A060BNF8	Clamp loader small subunit	JS09_0127	320
A0A060BHV9	Clamp loader small subunit	JS09_0128	187
A0A088FRF3	Co-chaperone for GroEL		110
A0A088FQQ8	dCTPase		173
A0A076G790	DenV endonuclease	Av05_00180	137
A0A067ZHH3	Deoxycytidylate deaminase	PhAPEC2_204	193
A0A060BME6	Deoxynucleoside monophosphate kinase	JS09_026	244
Q7Y599	Dmd discriminator of mRNA degradation	Dmd	63
A0A067ZK51	DNA endonuclease IV	PhAPEC2_252	185

A0A060BHA6	DNA helicase	JS09_0166	437
A0A088FR13	DNA helicase		480
A0A060BE80	DNA ligase	JS09_0246	497
A0A060BH71	DNA polymerase (EC 2.7.7.7)	JS09_0130	903
A0A060BHB9	DNA topoisomerase II large subunit	JS09_0176	605
A0A088FRH2	DsDNA-binding, transcription factor		95
A0A060BMZ5	Endonuclease II	JS09_0218	136
A0A067ZK98	Endonuclease subunit	PhAPEC2_58	339
A0A060BMT9	Exonuclease	JS09_0167	225
O64299	Exonuclease subunit 2 (EC 3.1.11.-) (Protein Gp46)	46	562
A0A060BN42	Fibrin	JS09_011	482
Q7Y267	Frd (Frd dihydrofolate reductase)	Frd	195
Q7Y557	Gp49 EndoVII packaging and recombination endonuclease VII	49	157
Q7Y4P7	Gp52 topoisomerase II, medium subunit	52	441
Q7Y5F6	Gp59	59	217
Q7Y5A8	Gp61 DNA primase subunit	58 ^a	340
A0A060BHL8	Head completion protein	JS09_023	149
A0A088FRC5	Head scaffolding protein		270
A0A097J4K1	Homing endonuclease	segG RB9_237	210
A0A067ZJ23	Inhibitor of host transcription	PhAPEC2_216	187
A0A067ZJ00	Inhibitor of prohead protease	PhAPEC2_176	222
A0A060BDN2	Internal head protein	JS09_029	95
Q06EE8	Large head outer capsid protein	hoc RB32ORF178c	376
X2KQP8	Large terminase protein	JS09_06	611
A0A067ZKJ5	Late promoter transcription accessory protein	PhAPEC2_228	112
A0A088FQU5	Late sigma transcription factor		185
A0A060BN58	Long tail fiber distal subunit	JS09_0201	1080
A0A067ZKF7	Lysozyme (EC 3.2.1.17)	PhAPEC2_148	577
A0A060BDQ1	Lysozyme (EC 3.2.1.17)	JS09_054	162
A0A088FR50	Major capsid protein		522
A0A067ZIU2	Modifier of transcription	PhAPEC2_11	151
A0A088FV80	mRNA metabolism modulator		136
A0A088FRB6	Neck protein		308
A0A088FVK3	Neck protein		254
Q0ILG8	NrdG	nrdG	156
A0A060BHC7	Nucleoid disruption protein	JS09_0186	148
Q7Y516	NudE nudix hydrolase	nudE	152
I7LF52	Phage holin	g256 BN81_256	219
I7K3Z1	Phage tail fibers	g255 BN81_255	183
Q7Y555	Pin protease inhibitor	Pin	137
A0A088FR45	Portal protein		523
X2KLC5	Prohead core protein	JS09_02	78
A0A088FS90	Prohead core protein		141
A0A067ZK06	Prohead core protein protease	PhAPEC2_167	213

Q7Y5C8	RB69ORF010c similar to motB	RB69ORF010c	139
A0A060BDY0	RecA-like recombination protein	JS09_0139	390
A0A060BHI9	Recombination, repair and ssDNA binding protein	JS09_0258	156
Q7Y582	RegA translational repressor protein	regA	125
A0A076G8C3	RI lysis inhibition regulator	Av05_00199	100
A0A088FVR8	Ribonuclease H		305
A0A060BE47	Ribonucleoside-diphosphate reductase (EC 1.17.4.1)	JS09_0216	751
A0A088FV76	RIIA protein		737
A0A060BE15	RIIB protector from prophage-induced early lysis	JS09_0180	315
A0A060BE70	RIII lysis inhibition accessory protein	JS09_0236	82
A0A067ZK08	RNA ligase	PhAPEC2_172	332
A0A060BNQ3	RNA ligase A	JS09_0219	374
A0A060BNK0	RNA polymerase ADP-ribosylase	JS09_0163	202
A0A088FVX9	RNA polymerase ADP-ribosylase		695
A0A088FQT7	RNA polymerase-binding protein		138
G1FHG5	RNaseH ribonuclease		45
A0A088FR40	Sheath completion		273
A0A088FRC1	Sheath protein		660
A0A067ZJ29	Single-stranded DNA binding protein	PhAPEC2_226	300
A0A088FR72	Site-specific RNA endonuclease		152
A0A088FVC1	Sliding clamp		228
A0A088FVK4	Spackle		97
A0A067ZHE8	Tail completion and sheath stabilizer protein	PhAPEC2_144	194
A0A088FVK7	Tail tube		163
A0A088FVV8	Terminase small subunit		164
A0A060BDT5	Thioredoxin	JS09_094	87
A0A060BHR2	Thymidine kinase (EC 2.7.1.21)	JS09_073	193
A0A067ZK35	Thymidylate synthase	PhAPEC2_222	286
A0A067ZIU6	Transcription modulator under heat shock	PhAPEC2_21	154
Q7Y4V3	UvsW RNA-DNA and DNA-DNA helicase, ATPase	uvsW	504
Q7Y529	Vs valyl-tRNA synthetase modifier	VS	115

FONTE: elaborado pelo autor.

TABELA 11 – Resultado da anotação dos genes do bacteriófago F9.

Identificador	Proteína	Gene	Tamanho
A0A075E0Y1	Base plate family protein	DA66_0026	268
U5PUI9	DNA binding protein	Marshall_98	92
U5PV36	DNA helicase	Marshall_99	572
W8JKH4	DNA ligase		483
K4I3T2	DNA polymerase (EC 2.7.7.7)		999

Q0MUT0	DNA primase-helicase subunit	gp41	243
K4I4X1	DNA topoisomerase		637
S4TV77	DNA topoisomerase II (Fragment)	SP029_01052	310
A0A059XCJ3	DnaB-like helicase C terminal domain protein		133
C8XUS3	Gp14 neck protein	orf00200	216
C8XUS2	Gp15 proximal tail sheath stabilization protein	orf00199	231
C8XUS1	Gp16 terminase DNA packaging enzyme small subunit	orf00198	233
C8XUS0	Gp17 terminase DNA packaging enzyme large subunit	orf00196	736
C8XUR7	Gp20 portal vertex protein of head	orf00191	563
C8XUR3	Gp22 prohead core protein	orf00186	288
A8WEB1	Gp23 major capsid protein (Fragment)		214
C8XUP9	Gp3 tail completion and sheath stabilizer protein	orf00171	166
C8XUH0	Gp32 T4-like ssDNA binding protein	orf00074	343
C8XUH2	Gp33 T4-like late promoter transcription accessory protein	orf00076	81
C8XUD4	Gp4 head completion protein	orf00036	204
C8XUP3	Gp44 clamp loader subunit DNA polymerase accessory protein	orf00164	329
C8XUP4	Gp45 sliding clamp DNA polymerase accessory protein	orf00165	222
C8XUD5	Gp48 putative tail tube associated base plate protein	orf00037	322
C8XUI0	Gp5 baseplate hub subunit and tail lysozyme	orf00084	536
C8XUB8	Gp52 topoisomerase II medium subunit	orf00017	446
C8XUD6	Gp53 baseplate wedge subunit	orf00038	184
C8XUT5	Gp6 baseplate wedge subunit	orf00217	593
C8XUJ5	Gp61 DNA primase subunit	orf00102	352
C8XUP2	Gp62 clamp loader subunit DNA polymerase accessory protein	orf00163	140
K4I6U3	Helicase loader		219
X4YDU4	Hemolysin-type calcium binding protein		1013
A0A1K9	Initiator RepB protein (Replication protein)	repE ERS085366_04868 UC41_06395	233
M4MEQ6	Lyase (Fragment)	gp17	713
E3UVU0	Major capsid protein g23 (Fragment)		269
C8XUJ7	MobD.6 conserved hypothetical phage protein	orf00104	117
K4I3N1	Molecular chaperone, HSP90 family		303
P11111	Neck protein gp14 (Gene product 14) (gp14)	14	256
K4I4P8	Nicotinamide phosphoribosyltransferase (EC 2.4.2.12)		554
U5PV19	Nucleoside triphosphate pyrophosphohydrolase	Marshall_79	290
U5PYF5	O-spanin	Marshall_95	108
E1XTC6	Phage associated protein	Vi01_117c	149

I7AT86	Phage lysis inhibitor		497
M4MCT7	Recombination and repair protein UvsX	VPDG_00004	349
C8XUP1	RegA translational repressor protein	orf00162	154
K4I3L2	Ribonucleoside-diphosphate reductase (EC 1.17.4.1)		760
K4I6R6	Ribonucleotide reductase of class Ia (Aerobic) beta subunit (EC 1.17.4.1)		367
K4I2Z8	Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)		284
C8XUP6	RNA-DNA and DNA-DNA helicase UvsW	orf00168	502
U5PYN2	RNaseH	Marshall_101	175
R9TEF4	SegD homing endonuclease	segD VPFG_00129	167
K4I3R1	Tail fiber protein		269
S6CFE9	Tail tube protein		177
X4YGA2	Tailspike protein		185
C8XUF9	UvsX RecA-like recombination protein	orf00061	361

FONTE: elaborado pelo autor.

TABELA 12 – Resultado da anotação dos genes do bacteriófago F10.

Identificador	Proteína	Gene	Tamanho
X2KTT1	6-carboxy-5,6,7,8-tetrahydropterin synthase	9g_00005	177
X2KPA0	7-cyano-7-deazaguanine synthase	9g_00006	471
X2KU72	ATPase	9g_00014	256
X2KTT7	ATP-dependent DNA ligase	9g_00015	291
X2L019	Deoxyuridine 5'-triphosphatenucleotidohydrolase	9g_00018	169
X2KSN7	DNA polymerase beta subunit	9g_00002	330
X2KP98	DNA polymerase elongation subunit 1	9g_00001	657
X2KTU7	DNA primase	9g_00025	755
X2KU64	GTP cyclohydrolase 1	9g_00004	204
X2L058	Host specificity protein	9g_00068	1029
X2KTW4	Lysis protein	9g_00045	166
X2L047	Main coat protein	9g_00053	367
X2L013	Organic radical activating enzyme	9g_00008	254
I7LF52	Phage holin	g256 BN81_256	219
X2L008	Queuosine tRNA-ribosyltransferase	9g_00003	317
X2KPA7	RNase H	9g_00016	196
X2KTY7	Side tail fiber	9g_00070	1211
X2KPA5	Superfamily II DNA/RNA helicase	9g_00011	667
X2KPE1	Tail fiber assembly protein	9g_00071	192
X2KUA6	Tail protein	9g_00059	139
X2KPD4	Tail subunit	9g_00061	393
X2KUA9	Tape measure protein	9g_00064	1086
X2KTW9	Terminase large subunit	9g_00050	532

FONTE: elaborado pelo autor.

TABELA 13 – Resultado da anotação dos genes do bacteriófago F12.

Identificador	Proteína	Gene	Tamanho
A4SH38	Colicin E1 immunity protein	imm SSON_PB02	113
A4SH39	Colicin E1 lysis protein	lys SSON_PB03	45
D8BRD0	Colicin pore forming domain protein	HMPREF9551_00172	460
C8XUC4	DexA exonuclease	orf00026	210
U5PV36	DNA helicase	Marshall_99	572
L7TVA8	DNA polymerase (Fragment)	g43	137
I6YMU9	DNA primase-helicase subunit		137
S4TT58	DNA topoisomerase 2 (Fragment)	SP029_00380	332
A0A075LAW4	Fimbrial usher protein	L960_3881c	249
C8XUS3	Gp14 neck protein	orf00200	216
C8XUS2	Gp15 proximal tail sheath stabilization protein	orf00199	231
A8WEB1	Gp23 major capsid protein (Fragment)		214
C8XUP9	Gp3 tail completion and sheath stabilizer protein	orf00171	166
C8XUH2	Gp33 T4-like late promoter transcription accessory protein	orf00076	81
H8ZNF2	Gp43 DNA polymerase		104
H8ZNF3	Gp43 DNA polymerase		255
C8XUP4	Gp45 sliding clamp DNA polymerase accessory protein	orf00165	222
C9DFX9	Gp48 baseplate tail tube cap	7	195
C8XUI0	Gp5 baseplate hub subunit and tail lysozyme	orf00084	536
C8XUB8	Gp52 topoisomerase II medium subunit	orf00017	446
C8XUD6	Gp53 baseplate wedge subunit	orf00038	184
C8XUJ5	Gp61 DNA primase subunit	orf00102	352
X4YDU4	Hemolysin-type calcium binding protein		1013
A0A1K9	Initiator RepB protein (Replication protein)	repE ERS085366_04868 UC41_06395	233
G8GDQ5	Large terminase subunit		736
M4MEQ6	Lyase (Fragment)	gp17	713
K4I3N1	Molecular chaperone, HSP90 family		303
U5PYF5	O-spanin	Marshall_95	108
E1XTC6	Phage associated protein	Vi01_117c	149
C8XUI8	PhoH-like protein	orf00095	280
A0A067XQW4	Recombination endonuclease subunit		557
N9V9U1	Ribonucleoside-diphosphate reductase (EC 1.17.4.1) (Fragment)	G114_09720	292
I2NM60	Ribonucleoside-diphosphate reductase, beta subunit-like protein (Fragment)	HMPREF1051_0404	102
E3SMV7	RNA-DNA + DNA-DNA helicase	uvrW PHM1_126	279
U5PYN2	RNAseH	Marshall_101	175
C8XUV3	SegD-like protein	orf00242	228
K4I3R1	Tail fiber protein		269
S6CFE9	Tail tube protein		177

X4YGA2	Tailspike protein		185
A0A0A0YRA5	Terminase DNA packaging enzyme large subunit	NW77_114	580
D9IE34	Topoisomerase II large subunit (Topoisomerase II large subunit C-terminal region)	60 T4147_003 T4GT7_003 T4Tp003 T4wild_003	160
U3M077	Viral capsid assembly protein (Fragment)	g20	149

FONTE: elaborado pelo autor.

5 DISCUSSÃO

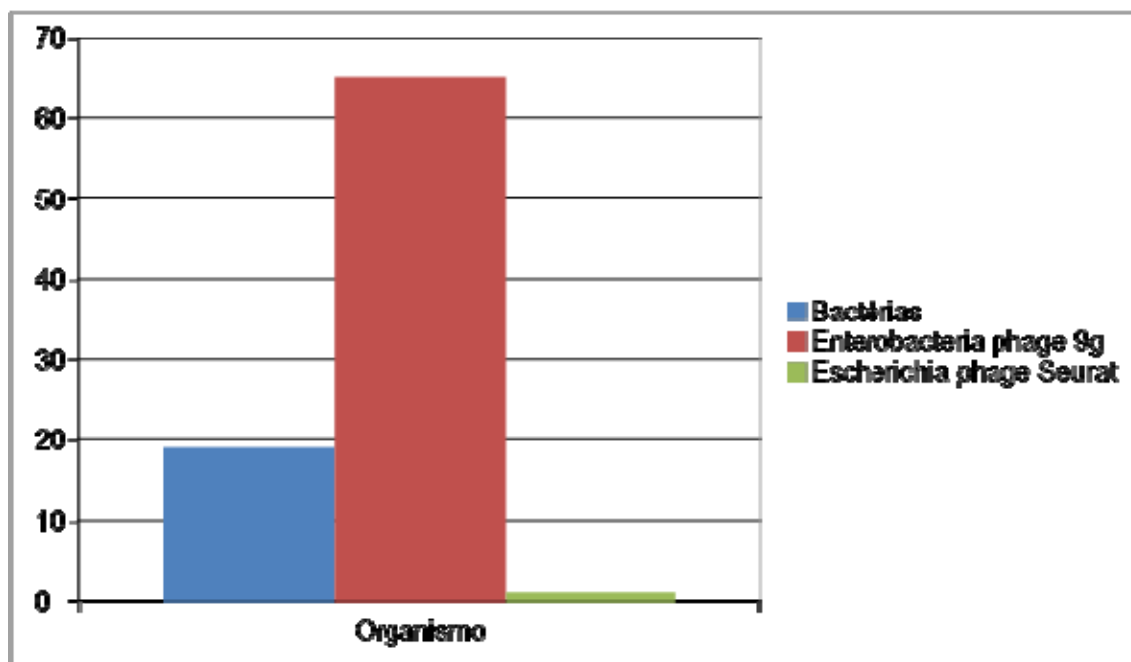
Os fagos F1 e F10 possuem genes codificadores de 7-ciano-7-deazaguanina sintase, GTP ciclodrolase, enzima ativadora de radicais orgânicos e 6-carboxi-5,6,7,8-tetrahidropterina sintase, responsáveis pela síntese de queuosina. A queuosina é um nucleosídeo modificado, capaz de influenciar a estrutura e função do tRNA dos hospedeiros e danifica a transcrição. Além disso, ela também altera processos metabólicos e fisiológicos da bactéria e parece ter relação com a eficiência da síntese de proteínas estruturais de fagos, porém esses processos não são muito bem compreendidos (SABRI et al., 2011). A presença desses genes é uma característica rara encontrada no genoma do bacteriófago 9g, do fago *Dp-1* e de alguns outros bacteriófagos. Este pode ser um dos mecanismos através dos quais os bacteriófagos F1 e F10 causam a lise da célula hospedeira (KULIKOV et al., 2014).

Os fagos F1 e F10 também apresentaram o gene *radD*, que produz a proteína RadD, responsável pelo reparo do DNA. Essa proteína é de grande importância para a adaptação do fago a situações adversas às quais é submetido no interior do seu hospedeiro, como o estresse oxidativo (CHEN et al., 2015).

Outro gene de F1 é o *sopA*, que produz uma protease capaz de degradar a proteína *IcsA*, importante componente da parede celular de enterobactérias como *Shigella* (EGILE et al., 1997). Considerando-se as semelhanças fisiopatológicas existentes entre *Shigella spp.* e *Escherichia coli* enteroinvasiva (EIEC), a presença desse gene sugere o potencial do bacteriófago F1 de romper a célula hospedeira (UD-DIN; WAHID, 2014).

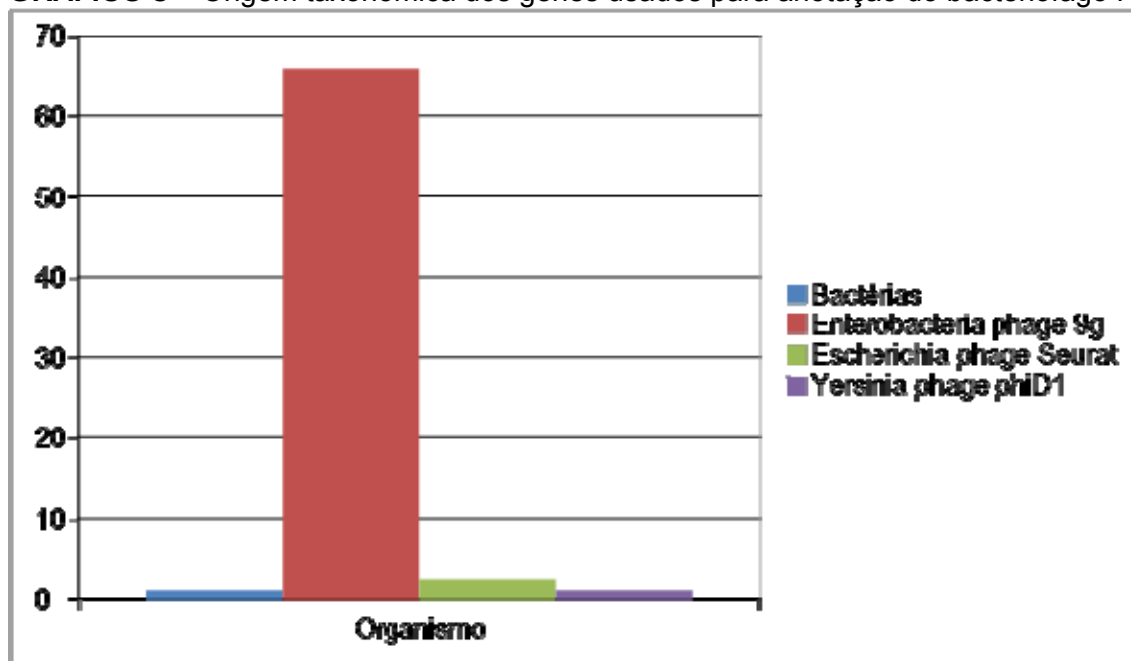
A análise da origem dos genes utilizados para a anotação de F1 e F10 demonstrou que possivelmente são bacteriófagos semelhantes ao 9g (NC_024146.1), descrito como capaz de infectar *E. coli*. O genoma de F1 apresentou 65 de seus genes anotados como pertencentes a esse fago e um gene de *Escherichia phage Seurat*, também capaz de invadir enterobactérias (**Gráfico 2**). O bacteriófago F10 é classificado na ordem *Caudovirales*, família *Siphoviridae*. F10 teve seus genes classificados de forma semelhante aos de F1 (**Gráfico 3**).

GRÁFICO 2 – Origem taxonômica dos genes usados para anotação do bacteriófago F1.



FONTE: elaborado pelo autor.

GRÁFICO 3 – Origem taxonômica dos genes usados para anotação do bacteriófago F10.

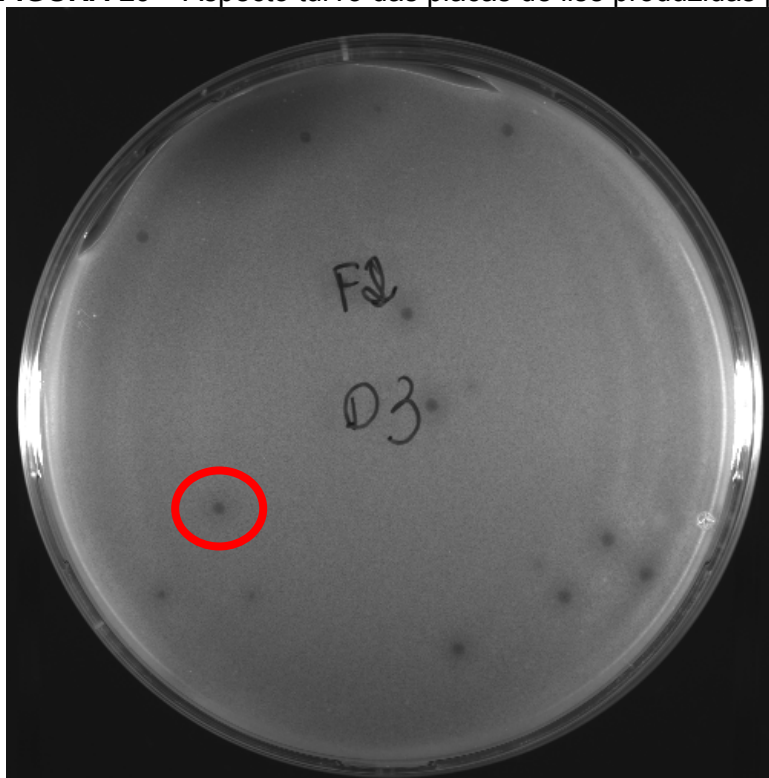


FONTE: elaborado pelo autor.

A partir da análise do repertório gênico dos bacteriófagos descritos nesse trabalho é possível sugerir que os vírus F3, F6, F8, F9, F10 e F12 são bacteriófagos líticos, pois seu genoma não apresentou genes como os que

sintetizam as proteínas integrase/excisionase. Além disso, esses fagos possuem diversos genes responsáveis pela síntese, empacotamento e reparo de DNA e todos possuem genes codificadores de proteínas líticas. A exceção foi o fago F2, que apresentou o gene responsável pela produção da integrase (GROTH; CARLOS, 2004). Esse fato pode ser confirmado pela análise das placas de lise geradas pelo F2, que têm aspecto turvo, característico de fagos temperados (Figura 10).

FIGURA 10 – Aspecto turvo das placas de lise produzidas pelo bacteriófago F2.



O aspecto turvo das placas de lise e a presença de halos (circulado de vermelho) ao redor das placas permite a classificação do fago F2 como lisogênico.

FONTE: elaborado pelo autor.

Os genomas de F3, F6 e F8 contém diversos genes capazes de produzir enzimas líticas como: nudix hidrolase, holina e lisozima, o que indica que esses bacteriófagos apresentam grande potencial de lisar o seu hospedeiro bacteriano após a infecção, sendo essa uma importante característica de bacteriófagos líticos (WENLIAN et al., 2002). Os fagos F2, F3, F6 e F8, F9 e F12 apresentam genes que sintetizam proteínas como a proteína finalizadora da cauda e a proteína estabilizadora da bainha, capazes de sintetizar a bainha da cauda, sendo

classificados na família *Myoviridae*, já que a presença de bainha envolvendo a cauda viral é uma característica única dessa família. A bainha ajuda no momento da infecção, contraindo-se e facilitando a injeção do DNA no interior da célula bacteriana (YAP; ROSSMANN, 2014).

6 CONCLUSÃO

- Foi possível concluir, ao final do desenvolvimento desse trabalho, que a água de rejeitos provinda da estação de tratamento de esgoto Arrudas (ETE-Arrudas) é uma fonte que possibilitou o isolamento de bacteriófagos capazes de infectar uma linhagem enteroinvasiva ATCC 43893 de *E. coli*, usada neste trabalho como modelo de enterobactéria patogênica.
- A metodologia utilizada nessa pesquisa possibilitou o isolamento de diferentes vírus, assim como a extração e purificação de seu material genético com bom rendimento e quantidade reduzida de DNA bacteriano contaminante.
- 8 desses bacteriófagos tiveram seu material genético extraído, sequenciado e seu genoma montado e caracterizado e o repertório gênico desses vírus indicou que são vírus da ordem *Caudovirales*, sendo 2 da família *Siphoviridae* e 6 da família *Myoviridae*. Nenhum dos bacteriófagos foi classificado na família *Podoviridae*.
- Apenas um bacteriófago apresentou potencial de ser temperado.
- 7 dos bacteriófagos isolados não tiveram nenhum gene relacionado à lisogenia.
- Apesar de 7 dos bacteriófagos isolados não terem apresentado genes relacionados ao ciclo lisogênico, análises mais aprofundadas devem ser realizadas para comprovar que são realmente exclusivamente líticos, já que muitos não apresentam seu genoma completo e mais de 50% dos genes anotados não têm função definida.
- É necessário também que sejam feitos estudos complementares para que o potencial de uso dos bacteriófagos isolados como bioterápicos seja comprovado.
- Testes *in vivo* devem ser realizados para garantir a eficiência e segurança desses agentes na terapia antimicrobiana.

7 PERSPECTIVAS

- Realizar a coleta de bacteriófagos em ecossistemas sadios onde o controle populacional de bactérias pela infecção bacteriofágica seja mais eficiente.
- Fazer a PCR utilizando diferentes *primers* para analisar se há contaminação das amostras por DNA plasmidial.
- Realizar testes de resistência dos bacteriófagos isolados à radiação ultravioleta, à ação de hipoclorito de sódio, assim como a diferentes temperaturas.
- Fazer a confirmação da classificação dos vírus nas famílias *Siphoviridae*, *Myoviridae* e *Podoviridae* utilizando imagens de microscopia eletrônica de transmissão.

REFERÊNCIAS

- ACKERMANN, H. W. **5500 Phages examined in the electron microscope**. Arch Virol, v. 152, p. 227-43. 2007.
- ACKERMANN, H. W.; EISENSTARK, A. **The present state of phage taxonomy**. Intervirology, v. 3, p. 201-19. 1974.
- ADAMS, M. H. **Bacteriophages**. New York: Interscience Publishers, Inc.; 1959.
- ADRIAENSSENS, E. M.; COWAN, D. A. **Using Signature Genes as Tools To Assess Environmental Viral Ecology and Diversity**. Applied and Environmental Microbiology. v. 80, n. 15, p. 4470-4480. 2014.
- AKSYUK, A. A.; ROSSMANN, M. G. **Bacteriophage assembly**. Viruses, v. 3, p. 172-203. 2011.
- ALEKSHUN, M. N.; LEVY, S. B. **Molecular mechanisms of antibacterial multidrug resistance**. Cell, v. 128, p. 1037-50. 2007.
- ALIKHAN, N. F. et al. **BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons**. BMC Genomics. 12: 402. Ago. 2011.
- ANDREWS-PFANNKUCH, C. et al. **Hydroxyapatite-Mediated Separation of Double-Stranded DNA, Single-Stranded DNA, and RNA Genomes from Natural Viral Assemblages**. Appl Environ Microbiol 76: 5039–5045. 2010.
- ANVISA. **Principais Síndromes Infeciosas**. 2004. Disponível em: <http://www.anvisa.gov.br/servicosaude/microbiologia/mod_1_2004.pdf>. Acesso em: 22 dez. 2014.
- ARANDA, K. R.; FAGUNDES-NETO, U.; SCALETSKY I. C. **Evaluation of multiplex PCRs for diagnosis of infection with diarrheagenic Escherichia coli and Shigella spp.** J Clin Microbiol, v. 42, p. 5849-53. 2004.
- BANKEVICH, A. et al. **SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing**. J Comput Biol. 19(5): 455–477. Mai. 2012.
- BATEMAN, A. et al. **UniProt: a hub for protein information**. Nucleic Acids Res. 43(Database issue):D204-12. Jan. 2015.
- BEAUDOIN, R. N. et al. **Isolation of a Bacteriophage from Sewage Sludge and Characterization of Its Bacterial Host Cell**. Rivier Academic Journal. 3(1):1-8. 2007.

BIKARD, D.; MARRAFFINI, L. A. **Innate and adaptive immunity in bacteria: mechanisms of programmed genetic variation to fight bacteriophages.** *Curr Opin Immunol.* 24(1):15-20. 2012.

BLAIR, J. M. A. et al. **Molecular mechanisms of antibiotic resistance.** *Nat Rev Micro.* 13:42-51. 2015.

BOLGER, A. M. et al. **Trimmomatic: a flexible trimmer for Illumina sequence data.** *Bioinformatics.* 30:2114–2120. Ago. 2014.

BONDY-DENOMY, J. et al. **Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system.** *Nature*, v. 493, p. 429-32. 2013.

BRÜSSOW, H.; HENDRIX, R. W. **Phage genomics: small is beautiful.** *Cell.* 11;108(1):13-6. Jan. 2012.

CABRAL, J. P. S. **Water Microbiology. Bacterial Pathogens and Water.** *Int. J. Environ. Res. Public Health.* v. 7, p. 3657-3703. 2010.

CALENDAR, R. **The Bacteriophages.** 2. ed. New York: Oxford University Press, 2006.

CASJENS, S. R. **Comparative genomics and evolution of the tailed-bacteriophages.** *Curr Opin Microbiol*, v. 8, p. 451-8. 2005.

CHATURONGAKUL, S.; OUNJAI, P. **Phage-host interplay: examples from tailed phages and Gram-negative bacterial pathogens.** *Front Microbiol*, v. 5, p. 442. 2014.

CHEN, S. H. et al. ***Escherichia coli radD (yejH)* gene: a novel function involved in radiation resistance and double-strand break repair.** *Mol Microbiol.* 95(5): 754–768. Mar. 2015.

CHO, E. N. et al. **Interactions between Integrase and Excisionase in the Phage Lambda Excisive Nucleoprotein Complex.** *J Bacteriol.* 184(18): 5200–5203. Set. 2002.

CLEARY, J. et al. **Enteropathogenic *Escherichia coli* (EPEC) adhesion to intestinal epithelial cells: role of bundle-forming pili (BFP), EspA filaments and intimin.** *Microbiology.* 150:527-538. 2004.

CLOKIE, M. R. J.; KROPINSKI, A. M. **Bacteriophages: Methods and Protocols. Isolation, Characterization, and Interactions.** v. 501. 2009.

DAVIES, J.; DAVIES, D. **Origins and evolution of antibiotic resistance.** *Microbiol Mol Biol Rev*, v. 74, p. 417-33. 2010.

D'HÉRELLE, F. **Sur un microbe invisible antagoniste des bacilles dysentérique.** Acad Sci Paris. 165:373–5. 1917.

DELCHER, A. L et al. **Fast algorithms for large-scale genome alignment and comparison.** Nucleic Acids Res. 1;30(11):2478-83. Jun 2002.

DRULIS-KAWA, Z. et al. **Learning from bacteriophages - advantages and limitations of phage and phage-encoded protein applications.** Curr Protein Pept Sci, v. 13, p. 699-722. 2012.

DRULIS-KAWA, Z. et al. **Bacteriophages and Phage-Derived Proteins – Application Approaches.** Curr Med Chem. 22(14): 1757–1773. Mai. 2015.

EGILE, C. et al. **SopA, the outer membrane protease responsible for polar localization of IcsA in Shigella flexneri.** Mol Microbiol. 23(5):1063-73. Mar. 1997.

ENVIRONMENTAL PROTECTION AGENCY. **Escherichia coli K-12 Derivatives Final Risk Assessment.** 1997. Disponível em: <http://epa.gov/biotech_rule/pubs/fra/fra004.htm>. Acesso em: 19 dez. 2014.

FDA. **Bad Bug Book: Foodborne Pathogenic Microorganisms and Natural Toxins Handbook Enteroinvasive Escherichia coli.** 2012. Disponível em: <<http://www.fda.gov/Food/FoodbornenessContaminants/CausesOfIllnessBadBugBook/ucm071298.htm>>. Acesso em: 3 jan. 2015.

FISCHER, S. et al. **Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups.** Curr Protoc Bioinformatics. Cap. 6: Unit 6.12.1-19. Set. 2011.

FRUCIANO, D. E.; BOURNE, S. **Phage as an antimicrobial agent: d'Herelle's heretical theories and their role in the decline of phage prophylaxis in the West.** Can J Infect Dis Med Microbiol, v. 18, p. 19-26. 2007.

GRAINGE, I.; JAYARAM, M. **The integrase family of recombinase: organization and function of the active site.** Mol Microbiol. 33(3):449-56. Ago. 1999. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/10577069>>. Acesso em: 17 abr. 2016.

GREGORACCI, G. B. **Levantamento de bacteriófagos líticos : isolamento e caracterização de vírus provenientes de esgoto comum com potencial aplicação antimicrobiana.** Campinas, SP. 2006.

GROTH, A. C.; CALOS, M. P. **Phage Integrases: Biology and Applications.** Journal of Molecular Biology. v. 335. Issue 3. p. 667-678. Jan. 2004. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022283603013561>>. Acesso em 23 fev. 2016.

HERRERO, J. et al. **Ensembl comparative genomics resources**. The Journal of Biological Databases and Curation. 2016. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4761110/>>. Acesso em: 10 jan. 2015.

HUDSON, J. A. et al. **Bacteriophages as biocontrol agents in food**. J Food Prot. 68:426-437. 2005.

IGNACIO-ESPINOZA, J. C. et al. **The global virome: not as big as we thought?** Curr Opin Virol. 3(5): 566-571. 2013.

KANEHISA, M. et al. **KEGG as a reference resource for gene and protein annotation**. Nucleic Acids Res. 4; 44(Database issue): D457–D462. Jan. 2016.

KESIK-SZELOCH, A. et al. **Characterising the biology of novel lytic bacteriophages infecting multidrug resistant *Klebsiella pneumoniae***. Virol J. 10(100):1-12. 2013.

KLUMPP, J. et al. **Next generation sequencing technologies and the changing landscape of phage genomics**. Bacteriophage. 2(3):190-199. 2012.

KLUMPP, J. et al. **Bacteriophage functional genomics and its role in bacterial pathogen detection**. Brief Funct Genomics. 12(4):354-65. Jul. 2013.

KRÜGER, D. H.; BICKLE, T. A. **Bacteriophage Survival: Multiple Mechanisms for Avoiding the Deoxyribonucleic Acid Restriction Systems of Their Hosts**. Microbiological Reviews. p. 345-360. Set. 1983.

KULIKOV, E. E. et al. **Genomic Sequencing and Biological Characteristics of a Novel Escherichia Coli Bacteriophage 9g, a Putative Representative of a New Siphoviridae Genus**. Viruses. 6(12): 5077–5092. Dez, 2014.

LABRIE, S. J. et al. **Bacteriophage resistance mechanisms**. Nat Rev Microbiol. 8(5):317-27. Mai. 2010.

LANGMEAD, B.; SALZBERG, S. L. **Fast gapped-read alignment with Bowtie 2**. Nat Methods. 4;9(4):357-9. Mar. 2012.

LOPES, A. et al. **Automated classification of tailed bacteriophages according to their neck organization**. BMC Genomics, v. 15, p. 1027. 2014.

LOSEE, L. L. et al. **A new antibiotic kills pathogens without detectable resistance**. *Nature* 517:455–459. 2015.

LU, T. K.; COLLINS, J. J. **Dispersing biofilms with engineered enzymatic bacteriophage**. Proc. Natl. Acad. Sci. 104, 11197-11202. USA. 2007.

LUKASHIN, A. V.; BORODOVSKY, M. **GeneMark.hmm: new solutions for gene finding**. Nucleic Acids Research. v. 26, n. 4, p. 1107-1115. 1998.

LUKJANCENKO, O.; WASSENAAR, T. M.; USSERY, D. W. **Comparison of 61 sequenced *Escherichia coli* genomes**. Microb Ecol, v. 60, p. 708-20. 2010.

LWOFF, A. et al. **A system of viruses**. In **Basic Mechanisms in Animal Virology**. Cold Spring Harb. Symp. Quant. Biol. 27, 51–62. 1962.

MARKINE-GORIAYNOFF, N. et al. **Glycosyltransferases encoded by viruses**. J Gen Virol, v. 85, p. 2741-54. 2004.

MAURICE, C. F. et al. **Linking the lytic and lysogenic bacteriophage cycles to environmental conditions, host physiology and their variability in coastal lagoons**. Environ Microbiol, v. 15, p. 2463-75. 2013.

MCCALLIN, S. et al. **Safety analysis of a Russian phage cocktail: From MetaGenomic analysis to oral application in healthy human subjects**. Virology. 443(2):187-196. 2013.

MCDANIEL, T. K. et al. **A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens**. Proc Natl Acad Sci U S A, v. 92, p. 1664-8. 1995.

MILLER, E. S. et al. **Bacteriophage T4 genome**. Microbiol Mol Biol Rev, v. 67, p. 86-156, table of contents. 2003.

MUNIESA, M. et al. **Shiga toxin-producing *Escherichia coli* O104:H4: a new challenge for microbiology**. Appl Environ Microbiol. 78(12):4065-73. Jun. 2012.

MURPHY, F. A. et al. **Virus taxonomy: Classification and nomenclature of viruses**. Sixth report of the International Taxonomy of Viruses. Arch Virol. 10:1-585. 1995.

NATARO, J. P.; KAPER, J. B. **Diarrheagenic *Escherichia coli***. Clin Microbiol. Rev. 11(1):142-201.1998.

NELSON, D. **Phage taxonomy: we agree to disagree**. J Bacteriol, v. 186, p. 7029-31. 2004.

OKEKE, I. N. **Diarrheagenic *Escherichia coli* in sub-Saharan Africa: status, uncertainties and necessities**. J Infect Dev Ctries. 3(11):817-842. 2009.

PAISANO, A. F.; BOMBONA, A. C. **Fagoterapia como alternativa no combate às infecções endodônticas**. Rev Gaúch Odontol. 58(2):243-252. 2010.

PAUL, J. H. **Microbial gene transfer: an ecological perspective**. J Mol Microbiol Biotechnol, v. 1, p. 45-50. 1999.

RAKHUBA, D.V. et al. **Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell**. Polish Journal of Microbiology. v. 59, n. 3, p. 145-155. 2010.

REARDON, S. **Phage therapy gets revitalized**. Nature international weekly journal of science. v. 510, n. 7503, p. 15-16. Jun. 2014. Disponível em: <<http://www.nature.com/news/phage-therapy-gets-revitalized-1.15348#auth-1>> Acesso em: 13 abr. 2016.

REYES, A. et al. **Going viral: next-generation sequencing applied to phage populations in the human gut**. Nat Rev Microbiol. 10(9):607-17. Set. 2012.

ROHWER, F. **Global phage diversity**. Cell, v. 113, p. 141. 2003.

ROHWER, F.; EDWARDS, R. **The Phage Proteomic Tree: a genome-based taxonomy for phage**. J Bacteriol, v. 184, p. 4529-35. 2002.

ROHWER, F. et al. **Roles of viruses in the environment**. Environ Microbiol 11:2771–2774. 2009

SABRI, M. et al. Genome annotation and intraviral interactome for the *Streptococcus pneumoniae* virulent phage Dp-1. J Bacteriol. v. 193. p. 551–562. Jan. 2011.

SAMSON, J. E. et al. **Revenge of the phages: defeating bacterial defences**. Nat Rev Microbiol, v. 11, p. 675-87. 2013.

SHARMA M. **Lytic bacteriophages: Potential interventions against enteric bacterial pathogens on produce**. Bacteriophage. v. 3. Issue 2. Jun. 2013.

SHIN, H. et al. **Receptor Diversity and Host Interaction of Bacteriophages Infecting *Salmonella enterica* Serovar Typhimurium**. PLoS ONE 7(8): e43392. Ago. 2012.

STURINO, J. M.; KLAENHAMMER, T. R. **Engineered bacteriophage-defence systems in bioprocessing**. Nat Rev Microbiol, v. 4, p. 395-404. 2006.

SUMMERS, W. C. **Bacteriophage therapy**. Annu Rev Microbiol, v. 55, p. 437-51. 2001.

TURNER, M. **Phage on the rampage**. Nature international weekly journal science. v. 360. Jun. 2011. Disponível em: <http://www.nature.com/news/2011/110609/full/news.2011.360.html>. Acesso em 12 jan. 2016.

UD-DIN, A.; WAHID, S. **Relationship among Shigella spp. and enteroinvasive Escherichia coli (EIEC) and their differentiation.** *Braz J Microbiol.* 4;45(4):1131-8. Mar. 2014.

VIERTEL, T. M. et al. **Viruses versus bacteria - novel approaches to phage therapy as a tool against multidrug-resistant pathogens.** *J Antimicrob Chemother.* 69(9):2326-2336. 2014.

WEINBAUER, M. G. **Ecology of prokaryotic viruses.** *FEMS Microbiol Rev*, v. 28, p. 127-81. 2004.

WENLIAN, X. et al. **The Gene, e.1(nudE.1), of T4 Bacteriophage Designates a New Member of the Nudix Hydrolase Superfamily Active on Flavin Adenine Dinucleotide , Adenosine (5')triphospho(5')adenosine, and ADP-ribose.** *The journal of biological chemistry.* Abr. 2002.

WHO. **Meeting the MDG drinking water and sanitation target: the urban and rural challenge of the decade.** WHO technical report series, Geneva, 2006. Disponível em: http://www.who.int/water_sanitation_health/monitoring/jmpfinal.pdf. Acesso em: 15 jan. 2015.

_____. **Burden of disease and cost-effectiveness estimates.** WHO technical report series, Geneva, 2014. Disponível em: http://www.who.int/water_sanitation_health/diseases/burden/en/. Acesso em: 19 jan. 2015.

_____. **Diarrhoea: Why children are still dying and what can be done.** WHO technical report series, Geneva, 2009. Disponível em: http://whqlibdoc.who.int/publications/2009/9789241598415_eng.pdf. Acesso em: 15 dez. 2014.

_____. **Diarrhoeal Disease: Fact Sheet N°330.** Net, Geneva, abr. 2013. Disponível em: <http://www.who.int/mediacentre/factsheets/fs330/en/index.html>. Acesso em: 7 jan. 2015.

_____. **Antimicrobial resistance: global report on surveillance.** WHO technical report series, Geneva, 2014. Disponível em: <http://www.who.int/drugresistance/documents/surveillancereport/en/>. Acesso em: 14 jan. 2015.

_____. **WHO's first global report on antibiotic resistance reveals serious, worldwide threat to public health.** WHO technical report series, Geneva, 2014. Disponível em: <http://www.who.int/mediacentre/news/releases/2014/amr-report/en/>. Acesso em: 27 dez. 2014.

_____. **Lack of sanitation for 2.4 billion people is undermining health improvements.** WHO technical report series, Geneva, 2015. Disponível em: <<http://who.int/mediacentre/news/releases/2015/jmp-report/en/>>. Acesso em: 14 abr. 2016.

YAP, M. L.; ROSSMANN, M. G. **Structure and function of bacteriophage T4.** *Future Microbiol.* 9: 1319–1327. Out. 2014.

YE, J. et al. **Control of Salmonella on sprouting mung bean and alfalfa seeds by using a biocontrol preparation based on antagonistic bacteria and lytic bacteriophages.** *J Food Prot*, 73:9-17. 2010.

YOUNG, R. BLÄSI, U. **Holins: form and function in bacteriophage lysis.** FEMS Microbiol Rev. 17(1-2):191-205. Ago. 1995.