

UNIVERSIDADE DE SÃO PAULO
Programa de Pós-Graduação Interunidades em Bioinformática

FERNANDO PACHECO NOBRE ROSSI

Análise comparativa do genoma de bacteriófagos isolados da
compostagem e estudo dos determinantes de sua interação com
Pseudomonas aeruginosa

Versão corrigida da Dissertação defendida em 04/09/2019

São Paulo
Data do Depósito na SPG: 25/07/2019

UNIVERSIDADE DE SÃO PAULO
Programa de Pós-Graduação Interunidades em Bioinformática

FERNANDO PACHECO NOBRE ROSSI

Análise comparativa do genoma de bacteriófagos isolados da
compostagem e estudo dos determinantes de sua interação com
Pseudomonas aeruginosa

Dissertação apresentada ao Instituto de
Matemática e Estatística da Universidade de São Paulo para obtenção
do Título de Mestre em Bioinformática

Orientadora: Profa. Dra. Aline Maria da Silva
Co-orientador: Prof. Dr. João Carlos Setubal

São Paulo
2019

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha Catalográfica

(elaborada pelo autor)

Rossi, Fernando Pacheco Nobre

Análise comparativa do genoma de bacteriófagos isolados da compostagem e estudo dos determinantes de sua interação com *Pseudomonas aeruginosa* / Fernando Pacheco Nobre Rossi. – São Paulo, 2019.

65 p.

Dissertação (mestrado) - Instituto de Matemática e Estatística da Universidade de São Paulo. Programa Interunidades em Bioinformática.

Orientador: da Silva, Aline Maria

Co-Orientador: Setubal, João Carlos

1. Fago. 2. *Pseudomonas aeruginosa*. 3.PA14. 4.pilina. 5. genômica comparativa. 6. interação fago-bactéria. I. T. II. da Silva, Aline Maria, orientador.

Análise comparativa do genoma de bacteriófagos isolados da
compostagem e estudo dos determinantes de sua interação com
Pseudomonas aeruginosa

FERNANDO PACHECO NOBRE ROSSI

Dissertação de Mestrado submetida à Universidade de São Paulo como
parte dos requisitos necessários à obtenção do título de “Mestre em
Ciências”. Programa: Bioinformática

APROVADO POR:

Profa. Dra. Aline Maria da Silva
(Orientadora e Presidente)

Prof. Dra. Regina Lúcia Baldini
IQ-USP

Prof. Dr. José Luiz Caldas Wolff
Universidade Presbiteriana Mackenzie

Prof. Dr. Luciano Antonio Digiampietri
EACH - USP

São Paulo, 04/09/2019

AGRADECIMENTOS

Primeiramente, aos meus pais Antônio e Magali pela educação que me proporcionaram, e por todas as oportunidades que permitiram meu desenvolvimento pessoal e profissional.

A minha irmã Julia pelo apoio e amizade demonstrada em todos esses anos.

A minha namorada Fernanda por todo o suporte, motivação, carinho, amizade e, principalmente, por todo o apoio, paciência e companheirismo durante todos os momentos dessa trajetória.

À Profa. Aline Maria da Silva, minha orientadora, pela incrível orientação, apoio e imprescindíveis dicas de análises e desenvolvimento desse trabalho, assim como seu enorme entusiasmo em relação a ciência e ao desenvolvimento científico.

Ao Prof. João Carlos Setubal pelo apoio e orientação nas metodologias bioinformáticas, além da disponibilidade e atenção ao longo desses dois anos.

Ao meu amigo e companheiro de pesquisa, Deyvid Amgarten, pelas valiosas sugestões, discussões e interpretações de dados computacionais durante a condução desse trabalho, e por sua amizade e bons momentos.

À Dra. Layla Martins pelo imprescindível apoio na interpretação e discussão dos dados, amizade e carinho durante esses dois anos de pesquisa e estudos.

Ao Dr. Paulo Pierry pela realização do sequenciamento genômico das cepas de *Pseudomonas aeruginosa*.

Aos demais colegas de laboratório e de pós-graduação, Guillermo Uceda, Deibs Barbosa, Raquel Riyuzo, Roberta Pereira, Oseias Feitosa, Paulo Burke, Ana Carolina Soares e Suzana Guima, pelas discussões sobre resultados, dicas e suporte em todas as etapas de desenvolvimento desse trabalho e pela amizade.

Aos técnicos de laboratório Alexandre Sanchez e Carlos Morais pelo suporte técnico e profissional constante durante o desenvolvimento desse trabalho.

A minha grande amiga Gabriela Ackel, por compartilhar comigo sua amizade e bons momentos e por fazer parte dessa trajetória.

A minha grande amiga e colega bioinformata Mayla Abraham por compartilhar comigo seus conhecimentos e amizade durante essa trajetória acadêmica.

Ao meu grande amigo Júlio (Índio) pela sua amizade, interesse em inovação científica e por ser uma ótima dupla para escalada.

Aos meus amigos Olivia (Lica), Akauã, Ícaro, Joel e Guilherme (Vergara) por me apoiarem nesse complicado trajeto acadêmico.

À FAPESP (Processo No. 2011/50870-6) e à CAPES (Projeto No. 3385/2013) pelo apoio financeiro aos laboratórios da Profa. Aline M. da Silva e do Prof. João C. Setubal.

À CAPES pela bolsa de Mestrado.

RESUMO

ROSSI, FERNANDO PACHECO NOBRE. Análise comparativa do genoma de bacteriófagos isolados da compostagem e estudo dos determinantes de sua interação com *Pseudomonas aeruginosa*. 2019. 65 páginas. Dissertação de Mestrado - Programa de Pós-Graduação Interunidades em Bioinformática. Universidade de São Paulo, São Paulo.

Bacteriófagos, ou fagos, são vírus que infectam bactérias, muito abundantes na natureza e que exibem extraordinária diversidade genômica. Além de sua importância para o desenvolvimento da biologia molecular, os fagos são considerados potenciais agentes terapêuticos no controle de bactérias multirresistentes a antibióticos. A ampla maioria dos bacteriófagos conhecidos estão classificados morfologicamente nas famílias *Myoviridae*, *Siphoviridae* e *Podoviridae*, que compõem a ordem *Caudovirales*. A adsorção do fago a superfície bacteriana é a primeira etapa da interação e um requerimento absoluto para o sucesso da infecção e posterior replicação do genoma do fago, seguida pela produção e liberação das partículas virais. Os determinantes mais comumente envolvidos na etapa de adsorção são a pilina majoritária (PilA) do *pilus* do tipo 4 (T4P) e o lipolissacarídeo (LPS), ambos presentes na superfície bacteriana, e as estruturas da cauda da partícula viral. A compreensão dos mecanismos moleculares envolvidos na interação dos fagos com cepas bacterianas suscetíveis é um dos requisitos para a avaliação de seu potencial de uso na fagoterapia. Em trabalho anterior de nosso grupo, foram isolados seis fagos (ZC01, ZC03, ZC04, ZC06, ZC07 e ZC08) de amostras de compostagem, utilizando-se *Pseudomonas aeruginosa* PA14 como hospedeira, os quais tiveram seus respectivos genomas sequenciados. Os genomas dos fagos ZC01, ZC03 e ZC08 foram analisados e foi demonstrado experimentalmente que ZC01 (*Siphoviridae*) e ZC03 e ZC08 (*Podoviridae*) têm uma abrangência restrita quanto a suscetibilidade à infecção de isolados clínicos de *P. aeruginosa*. Resultados anteriores também apontaram o T4P como receptor para a adsorção dos fagos ZC01, ZC03 e ZC08. No presente trabalho, tivemos como objetivo aprofundar a caracterização genômica dos outros três fagos isolados da compostagem (ZC04, ZC06 e ZC07), também da família *Podoviridae*, e comparar seus determinantes potencialmente envolvidos na interação específica com *P. aeruginosa*. Além disso, analisamos características de componentes da superfície bacteriana preditos como receptores específicos desses fagos, através da análise de genomas de 18 isolados clínicos que experimentalmente se mostraram suscetíveis ou resistentes à infecção por esses fagos. Os resultados obtidos nos permitem concluir que os fagos de *P. aeruginosa* ZC03, ZC04, ZC06, ZC07 e ZC08, mesmo que muito próximos filogeneticamente e pela alta similaridade de seus genomas, apresentam diferenças potencialmente importantes para suas respectivas capacidades de infecção e propagação. A análise do proteoma predito mostrou importantes diferenças nas fibras das caudas desses fagos que podem ser relevantes para explicar suas distintas capacidades de infecção a alguns isolados clínicos. Por outro lado, não observamos uma relação perfeita entre capacidade de infecção dos fagos ZC01, ZC03 e ZC08 e as diferentes estruturas primárias da pilina majoritária PilA identificada nos genomas dos 18 isolados de *P. aeruginosa*. Também não foi possível identificar relação entre suscetibilidade ou resistência à infecção com os distintos padrões de glicosilação da PilA ou com os sorotipos (tipo de antígeno O do LPS) preditos para estes isolados. Os resultados obtidos sugerem que a relação entre a capacidade de infecção dos fagos e seus determinantes que participam na etapa de adsorção está associada a diferentes interações com estruturas distintas na superfície bacteriana.

Palavras chave: Fago; *Pseudomonas aeruginosa*; PA14; pilina; genômica comparativa; interação fago-bactéria.

ABSTRACT

ROSSI, FERNANDO PACHECO NOBRE. Comparative analysis of genomes from composting-isolated bacteriophages and study of their interaction determinants with *Pseudomonas aeruginosa*. 2019. 65 pages. Masters Dissertation - Graduate Program in Bioinformatics. Universidade de São Paulo, São Paulo.

Bacteriophages, or phages, are virus that infect bacteria. Phages are very abundant in nature and display an extraordinary genomic diversity. Besides their importance to the development of molecular biology, phages are potential agents for therapeutic control of multidrug-resistant bacteria. Based on their morphology, the vast majority of known phages is classified in the families *Myoviridae*, *Siphoviridae* and *Podoviridae*, which comprise the order *Caudovirales*. Phage particle adsorption to bacterial surface is the first step of the interaction and is an absolute requirement for successful infection, followed by phage genome replication, production and release of new viral particles. The most common determinants involved in phage adsorption step are the major pilin (PilA) of the type 4 pilus (T4P) and lipopolysaccharide (LPS) in the host surface, and tail structures of the phage particle. The understanding of the molecular mechanisms involved in the interaction of phages with susceptible host strains is one of the requirements for the evaluation of their therapeutic potential in phage therapy. In a previous work of our group, six phages (ZC01, ZC03, ZC08, ZC04, ZC06 and ZC07) were isolated from composting samples using *Pseudomonas aeruginosa* PA14 as a host, which had their respective genomes sequenced. The genomes of the phages ZC01, ZC03 and ZC08 were analyzed and it was experimentally shown that ZC01 (*Siphoviridae*) and ZC03 and ZC08 (*Podoviridae*) have a narrow host range against *P. aeruginosa* clinical isolates. Moreover, T4P was suggested as the receptor for adsorption of ZC01, ZC03 and ZC08. In the present work we aimed to deepen the characterization of the three remaining phages isolated from composting (ZC04, ZC06 and ZC07), also from the family *Podoviridae*, and compare their determinants potentially involved in the specific interaction with *P. aeruginosa*. In addition, the characteristics of surface components predicted as specific receptors were analyzed using the genomes of 18 *P. aeruginosa* clinical isolates that experimentally showed susceptibility or resistance to infection for these phages. The obtained results lead us to the conclusion that even though the *P. aeruginosa* phages ZC03, ZC04, ZC06, ZC07 and ZC08 were phylogenetically close and shared high genome similarity, they present differences potentially important for their infection and propagation capabilities. The analysis of the predicted proteome showed important differences in the tail fibers from these phages that might be relevant to explain their distinct infection capabilities in some clinical isolates. On the other hand, we have not observed a perfect relation between ZC01, ZC03 and ZC08 infection capabilities and the major pilin (PilA) primary structure identified in the 18 *P. aeruginosa* clinical isolates. We also have not found a relation of the infection susceptibility or resistance to the PilA glycosylation patterns nor the O antigen serotypes predicted for each isolate. Our results suggest that the relation between the phages infection capability and their adsorption determinants is associated with different interactions to distinct bacterial surface structures.

Keywords: Phage; *Pseudomonas aeruginosa*; PA14; pilin; comparative genomics; phage-bacteria interaction.

LISTA DE ABREVIATURAS E SIGLAS

aa:	aminoácido
ADCP:	Proteína com domínio de ATPase
AMR:	<i>Antimicrobial/Antibiotic Resistance</i>
BLAST:	<i>Basic Local Alignment Search Tool</i>
CARD:	<i>Comprehensive Antibiotic Resistance Database</i>
CATG:	Centro Avançado de Tecnologias em Genômica
CDD:	<i>Conserved Domain Database</i>
CDS:	<i>Coding sequence</i> (sequência codificadora de proteína)
Fago:	Bacteriófago
HE:	<i>Homing Endonuclease</i>
HEG:	<i>Homing Endonuclease Gene</i>
HMM:	<i>Hidden Markov Model</i>
ICTV:	<i>International Committee on Taxonomy of Viruses</i>
INDELS:	Inserções ou Deleções
IQ-USP:	Instituto de Química da Universidade de São Paulo
LPS:	Lipopolissacarídeo
MAFFT:	<i>Multiple alignment using fast Fourier transform</i> (Programa de alinhamento múltiplo de sequências de aminoácidos e de nucleotídeos)
MLST:	<i>MultiLocus Sequence Typing</i>
MUMMER:	Sistema para alinhamento ultrarrápido de sequências de DNA
NCBI:	<i>National Center of Biotechnology Information</i>
NJ:	<i>Neighbour Joining</i>
OMP:	<i>Outer Membrane Protein</i>
ORF:	<i>Open Read Frame</i>
PA14:	<i>Pseudomonas aeruginosa</i> cepa PA14
PAO1:	<i>Pseudomonas aeruginosa</i> cepa PAO1
pb:	pares de base
PE:	<i>Paired-end</i>
RaxML:	<i>Randomized Axelerated Maximum Likelihood</i> (Programa de análise filogenética com máxima verossimilhança)
RBP:	<i>Receptor-binding protein</i>
RGI:	<i>Resistance Gene Identifier</i>
RNAP:	RNA polimerase
ROARY:	Programa para análise do pan-genoma de procariotos
SES:	Sistema de exclusão de superinfecção
SNPs:	<i>Single Nucleotide Polymorphisms</i>
ST:	<i>Sequence type</i>
T4P:	Pilus do tipo 4
VOCs:	<i>Viral orthologous clusters</i>
VOG:	<i>Virus orthologous groups</i>
vRNAP:	Víriion-RNAP

SUMÁRIO

1. INTRODUÇÃO	9
1.1. Bacteriófagos	10
1.1.1. <i>Caudovirales</i>	12
1.2. Determinantes de adsorção específica fago-bactéria	13
1.3. Resistência do hospedeiro à adsorção dos fagos	15
1.4. Fagos como agentes antimicrobianos	16
2. MOTIVAÇÃO E OBJETIVOS	17
3. METODOLOGIAS	19
3.1. Sequenciamento, montagem e anotação de genomas	19
3.2. Alinhamento do genoma dos fagos	20
3.3. Análises comparativas de genomas de <i>P. aeruginosa</i>	20
3.4. Análises filogenéticas	20
3.5. Alinhamentos múltiplos de sequências de proteínas	21
3.6. Análise dos genes de MLST nas cepas de <i>P. aeruginosa</i>	21
3.7. Identificação de genes de resistência a antibióticos nos genomas	22
4. RESULTADOS E DISCUSSÃO	22
4.1. Análise dos genomas dos fagos ZC04, ZC06 e ZC07	22
4.1.1. <i>Sequenciamento e comparação dos genomas</i>	22
4.1.2. <i>Anotação dos genomas e comparação das ORFs preditas</i>	24
4.1.3. <i>Análise filogenética</i>	31
4.1.4. <i>Comparação das ORFs preditas como componentes da cauda</i>	32
4.2. Prospecção de determinantes de <i>P. aeruginosa</i> associados à suscetibilidade à infecção pelos fagos ZC01, ZC03 e ZC08	37
4.2.1. <i>Sequenciamento e comparação dos genomas de 18 isolados de P. aeruginosa</i>	40
4.2.2. <i>Análises dos genomas dos 18 isolados e relação com a suscetibilidade por ZC01, ZC03 e ZC08 e a presença de genes de resistência</i>	41
4.2.3. <i>Análise da relação à suscetibilidade por ZC01, ZC03 e ZC08 e variantes de PflE e PflA dos 18 isolados</i>	47
4.2.4. <i>Predição de sítios de glicosilação na PflA e sorotipos dos 18 isolados</i>	52
5. CONCLUSÕES	56
6. REFERÊNCIAS	59
7. LISTA DE APÊNDICES	65

1. Introdução

Bacteriófagos, ou fagos, são vírus que infectam bactérias e são muito abundantes na natureza (Clokier *et al.*, 2011; Rohwer *et al.*, 2014; Hyman & Abedon, 2018).

Desde sua descoberta no início do século 20, os fagos têm sido ferramentas muito importantes no desenvolvimento da biologia molecular e da biotecnologia. Além disso, seu potencial de aplicação como agente antimicrobiano sempre despertou interesse e em alguns poucos casos, os fagos foram utilizados no tratamento de infecções bacterianas (Clark & March, 2006; Haq *et al.*, 2012). Nos últimos anos, o interesse na utilização de fagos como agente terapêutico (*phage therapy*) foi retomado e intensificado, principalmente como alternativa para o controle biológico de bactérias multirresistentes a antibióticos (Nobrega *et al.*, 2015; Cisek *et al.*, 2017; Schooley *et al.*, 2017; Watts, 2017; Dedrick *et al.*, 2019; Saha & Mukherjee, 2019).

Estima-se que 10^{31} partículas virais (vírions), predominantemente bacteriófagos, existam na biosfera (Hendrix, 2003; Hatfull & Hendrix, 2011). Estudos de ecologia microbiana calcularam que ocorrem 10^{23} infecções por bacteriófagos por segundo, indicando que a população viral, além de enorme, é altamente dinâmica (Suttle, 2007; Hatfull & Hendrix, 2011). Esse dinamismo tem o papel ecológico de regular a diversidade e a estrutura de comunidades bacterianas, interferindo em suas funções e na manutenção da homeostase de diversos ecossistemas (Weinbauer, 2004; Clokier *et al.*, 2011; Mirzaei & Maurice, 2017; Ford *et al.*, 2018; Gore, 2018; Morella *et al.*, 2018; Shkorporov & Hill, 2019).

Abordagens metagenômicas têm revelado a extraordinária diversidade de bacteriófagos nos mais variados ambientes, cuja vasta maioria não foi ainda cultivada e caracterizada (Edwards & Rohwer, 2005; Rohwer *et al.*, 2014; Paez-Espino *et al.*, 2016; Hayes *et al.*, 2017; Simmonds *et al.*, 2017). Porém é crescente o número e a diversidade de fagos isolados dos mais variados ambientes, apesar do viés inerente aos métodos de isolamento que dependem da utilização hospedeiro adequado para sua propagação (geralmente uma bactéria modelo bem caracterizada) (Clokier *et al.*, 2011; Lehman, 2018; de Jonge *et al.*, 2019). O isolamento de fagos inéditos possibilita o estudo detalhado de sua morfologia, do seu genoma, dos mecanismos envolvidos na interação com seus hospedeiros específicos e de seu papel ecológico, além de abrir perspectivas para sua aplicação biotecnológica e terapêutica.

Em projeto anterior de nosso grupo, foram isolados seis fagos da compostagem os quais infectam *Pseudomonas aeruginosa* PA14, um clone de relevância clínica por causar infecções hospitalares agressivas e difíceis de tratar com abordagens tradicionais. A caracterização genômica detalhada de três (ZC01, ZC03 e ZC08) desses fagos, juntamente com o teste de eficácia na dissolução do biofilme de *P. aeruginosa* PA14 apontaram seu potencial como agente antimicrobiano em infecções por essa bactéria (Amgarten et al. 2017).

No presente trabalho, tivemos como objetivo aprofundar a caracterização genômica dos outros três fagos isolados da compostagem (ZC04, ZC06 e ZC07) e investigar os determinantes potencialmente envolvidos na interação específica com *P. aeruginosa*. Também analisamos características de componentes da superfície bacteriana preditos como receptores específicos desses fagos no processo de infecção, através da análise de genomas de isolados clínicos que experimentalmente se mostraram suscetíveis ou resistentes à infecção por esses fagos. Este trabalho é parte da pesquisa em andamento em nosso grupo, e que explora o potencial de aplicação destes fagos no tratamento de infecções causadas por *P. aeruginosa*.

1.1. Bacteriófagos

Fagos têm seu material genético, podendo ser DNA ou RNA, fita-dupla ou fita simples, envolvido por capsídeo proteico que frequentemente está acoplado a uma cauda também proteica, constituindo a partícula viral (vírion). Além do genoma viral, o capsídeo pode carrear proteínas envolvidas na internalização e expressão do genoma do fago na célula hospedeira (Kutter & Sulakvelidze, 2005; Ackermann & Prangishvili, 2012; Rohwer *et al.*, 2014; Hyman & Abedon, 2018).

Assim com outros vírus, os fagos dependem de seus hospedeiros para sua replicação e propagação. Os fagos são classificados em líticos (virulentos) ou temperados de acordo com seu ciclo de vida. Os fagos virulentos se multiplicam através do ciclo lítico, isto é, o vírion adsorve a superfície da célula hospedeira, injeta seu genoma e utiliza a maquinaria molecular e metabolismo do hospedeiro para a produção de novas partículas virais que são liberadas pela lise da célula hospedeira. Os fagos temperados, embora capazes de realizar um ciclo lítico, podem estabelecer uma relação mais duradoura com seu hospedeiro, pela integração de seu material genético no genoma bacteriano, e a infecção torna-se latente. Nesse caso, há transmissão vertical do genoma do fago que se replica juntamente com seu

hospedeiro, então dito lisogenizado, podendo permanecer por infinitas gerações como prófago (Kutter & Sulakvelidze, 2005; Howard-Varona *et al.*, 2017). Em alguns casos, fagos líticos podem adotar um estado de pseudolisogenia, em que seu genoma não se integra ao genoma do hospedeiro, mas passa a existir como plasmídeo que é propagado para células filhas na divisão celular (Colavecchio *et al.*, 2017).

A regulação da lisogenização está sob o controle de proteínas repressoras, tal como a proteína *ci* no fago λ , que é expressa pelo prófago e previne a expressão de genes relacionados à ativação do ciclo lítico. Frente a estresses ambientais, principalmente os que causam danos no DNA, o ciclo lítico é disparado (Kutter & Sulakvelidze, 2005). No modelo do fago λ (fago temperado de *Escherichia coli*), a mudança do ciclo lisogênico para lítico deve-se a diminuição da concentração do repressor *ci* com consequente ativação dos genes relacionados ao ciclo lítico (Gandon, 2016).

Assim, os fagos temperados, a cada infecção decidem entre ciclo lítico ou lisogênico. No modelo de fago λ , tal decisão, embora naturalmente probabilística, parece ser influenciada pelo estado nutricional da bactéria infectada e também pelo número de fagos coinfecantes (Oppenheim *et al.*, 2005; Zeng *et al.*, 2010). Já a decisão entre ciclo lítico ou lisogênico para fagos de *Bacillus spp* envolve um sistema de comunicação entre fagos, cuja progênie é capaz de detectar concentrações suficientemente altas de um peptídeo liberado no meio e, nesse caso, decidir pelo ciclo lisogênico (Erez *et al.*, 2017).

Em alguns casos, uma cepa bacteriana lisogenizada se torna resistente à infecção por fagos do mesmo tipo, sendo este tipo de defesa chamada de imunidade à superinfecção ou sistema de exclusão de superinfecção (SES) (Labrie *et al.*, 2010; Seed, 2015). Os mecanismos neste tipo de defesa atrasam a replicação do DNA, impactando tanto a produção de proteínas virais como a sua dinâmica e abundância (Howard-Varona *et al.*, 2018).

Fagos temperados podem transportar genes de um hospedeiro para outro (transdução) e assim conferir características benéficas para a população bacteriana em um dado ambiente ou condição. Dessa forma pode ocorrer a transferência lateral de genes relacionados ao metabolismo de carboidratos e de produção de toxinas e, em casos mais raros, de genes de resistência a antibióticos. Esse mecanismo é consequência da eventual incorporação errônea de trechos do DNA do hospedeiro

durante a excisão do genoma viral quando os prófagos são induzidos ao ciclo lítico. No caso de fagos virulentos, podem ser formados vírions que erroneamente empacotam segmentos do genoma do hospedeiro, e dessa forma também contribuem para transferência horizontal de genes via transdução. (Kutter & Sulakvelidze, 2005; Enault *et al.*, 2017; Howard-Varona *et al.*, 2017; Touchon *et al.*, 2017; Braga *et al.*, 2018).

1.1.1. Caudovirales

A maioria (~95%) dos bacteriófagos conhecidos estão classificados morfologicamente em três famílias: *Myoviridae*, *Siphoviridae* e *Podoviridae*. Essas três famílias compõem a ordem *Caudovirales*, segundo a taxonomia do *International Committee on Taxonomy of Viruses* (ICTV)¹ (Lefkowitz *et al.*, 2018). Os *Caudovirales* (fagos com cauda) têm o genoma de DNA fita-dupla (dsDNA) envolto por um capsídeo icosaédrico, que está conectado com uma cauda que pode conter fibras e espigas, como esquematizado na Figura 1. Os fagos da família *Myoviridae* (myovirus) possuem uma cauda longa e contrátil cujo tubo é recoberto por uma bainha (ex: fago T4) enquanto os *Siphoviridae* (siphovirus) apresentam uma cauda longa não-contrátil (ex: fago λ) e os *Podoviridae* (podovirus) têm uma cauda curta não-contrátil (ex: fago T7). (Veesler & Cambillau, 2011; Ackermann & Prangishvili, 2012; King *et al.*, 2012; Rohwer *et al.*, 2014; Nobrega *et al.*, 2018).

As caudas desses fagos são máquinas moleculares que reconhecem receptores específicos na superfície da célula bacteriana (adsorção), penetram o envelope celular e injetam o material genético viral no citoplasma bacteriano. Uma vez ligado especificamente e irreversivelmente à superfície da célula bacteriana, o fago fica posicionado para injeção do material genético para o interior da célula hospedeira (Kutter & Sulakvelidze, 2005; Silva *et al.*, 2016; Nobrega *et al.*, 2018).

¹<https://talk.ictvonline.org/taxonomy/>

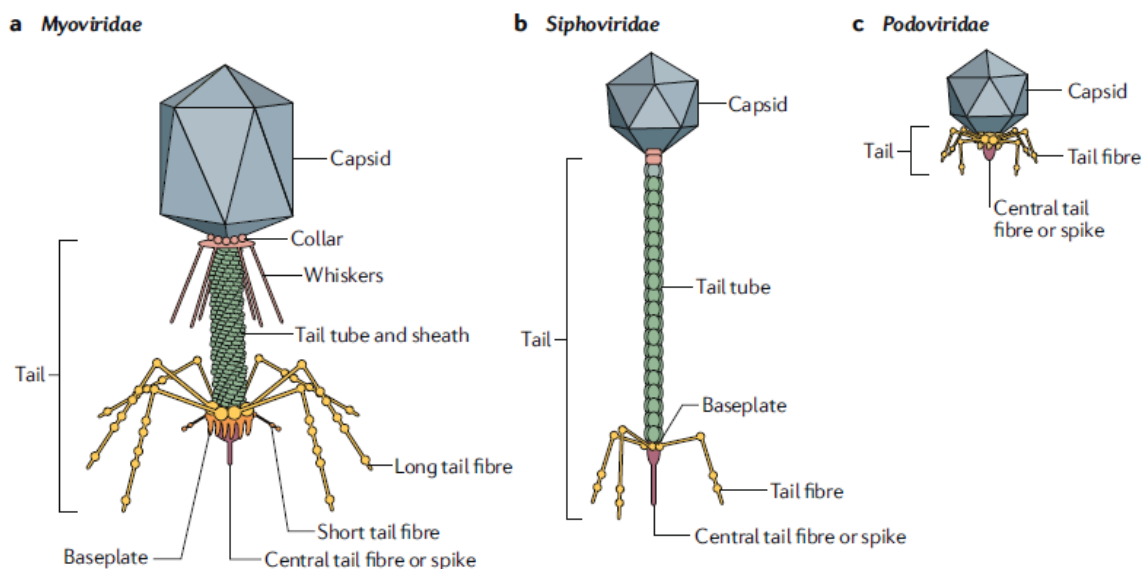


Figura 1: Esquema da estrutura de fagos da ordem *Caudovirales*. Todos os fagos dessa ordem possuem um capsídeo (*capsid*), que contém e protege o genoma, conectado a uma cauda (*tail*). Os membros da família *Myoviridae* possuem uma cauda contrátil cujo tubo da cauda (*tail tube*) é recoberto por uma bainha proteica (*sheath*) e que pode ter fibras do pescoço (*whiskers*). Os fagos da família *Myoviridae* e *Siphoviridae* possuem uma placa basal (*baseplate*) na ponta do tubo da cauda onde estão conectadas as fibras ou espigas da cauda (*tail fibre* ou *tail spike*). Os fagos da família *Podoviridae* não possuem placa basal e suas fibras ou espigas conectam diretamente ao tubo de sua cauda curta. Os fagos da família *Siphoviridae* and *Podoviridae* também possuem uma fibra ou espiga central protuberante conectada à placa basal ou ao tubo da cauda, respectivamente. Figura retirada de (Nobrega *et al.*, 2018).

1.2. Determinantes de adsorção específica fago-bactéria

As interações fago/hospedeiro resultam primeiramente de colisões aleatórias entre vírions e células bacterianas, dependendo, portanto, da proporção fago/hospedeiro e podendo ser influenciada por fatores físico-químicos (pH, temperatura e presença de certos íons) e pelo estado fisiológico do hospedeiro. A etapa de adsorção envolve interações reversíveis e irreversíveis, sendo as estruturas da cauda, consideradas determinantes cruciais na interação com receptores específicos na superfície bacteriana. Nessa etapa, o fago estabelece seu primeiro contato com a bactéria e a identifica, ou não, como um potencial hospedeiro suscetível para infecção. Dessa forma é estabelecida a especificidade e abrangência da interação com o hospedeiro (*host-range specificity*) (Silva *et al.*, 2016; Nobrega *et al.*, 2018). A etapa de adsorção é característica de cada par fago-hospedeiro, e o sucesso

completo da infecção viral dependerá da superação dos mecanismos de defesa do hospedeiro (Rakhuba *et al.*, 2010; Silva *et al.*, 2016; Nobrega *et al.*, 2018).

Nos fagos da ordem *Caudovirales*, as fibras e espigas da cauda funcionam como proteínas ligantes do receptor (RBP, *Receptor-binding protein*) e reconhecem especificamente os receptores na superfície do hospedeiro (Nobrega *et al.*, 2018).

A natureza e localização dos receptores do hospedeiro reconhecidos especificamente pelos fagos varia enormemente dependendo do par fago-hospedeiro, podendo ser proteínas da membrana externa (*Outer Membrane Protein*, OMP), polissacarídeos, lipopolissacarídeos (LPS), fimbrias, *pilus* ou flagelo (Chaturongakul & Ounjai, 2014; Silva *et al.*, 2016). Em alguns casos, além de interagirem com polissacarídeos da superfície bacteriana, os fagos degradam enzimaticamente tais estruturas (Rakhuba *et al.*, 2010).

No caso de *Pseudomonas aeruginosa*, que é o hospedeiro de interesse nesse trabalho, as moléculas de superfície já descritas como receptores de fagos são, principalmente, a pilina PilA do *pilus* do tipo 4 (T4P) e o lipopolissacarídeo (LPS) (Winther-Larsen *et al.*, 2007; Chibeu *et al.*, 2009; Garbe *et al.*, 2011; Bae & Cho, 2013; Kim *et al.*, 2018).

O T4P de *P. aeruginosa* participa da formação do biofilme e adesão às células do hospedeiro eucariótico, além de ser responsável pela motilidade celular do tipo *twitching*. Os T4P têm estrutura alongada, fina e fibrosa, estão divididos em duas classes principais, IVa e IVb, sendo o primeiro um grupo mais homogêneo e mais amplamente distribuído, já o segundo apresenta uma maior heterogeneidade. Ambas as classes de T4P são diferenciadas tanto por suas subunidades como pelo seu maquinário de montagem. Os T4P da classe IVa (T4Pa) são formados por centenas de subunidades da pilina majoritária (PilA) e também por subunidades de pilinas minoritárias (FimU, PilVWX e PilE) que participam da montagem do pilus juntamente com a maquinaria proteica (complexo da secretina PilQ) que forma um canal responsável pela externalização do *pilus*. A formação completa do *pilus*, requer a proteína PilY1, uma adesina na ponta do T4P (Burrows, 2012; Nguyen *et al.*, 2015). Estão descritos cinco alelos da pilina PilA em *P. aeruginosa*, sendo que distintas cepas expressam uma variante desses cinco, dois dos quais são glicosilados. As variantes de PilA diferem quanto ao seu tamanho (150 a 173 aminoácidos), quanto ao número de aminoácidos (12 a 29 aminoácidos) entre as duas cisteínas C-terminal, que formam uma alça com ligação dissulfeto que é crítica para função estrutural e funcional da

proteína, e pela presença de genes acessórios na vizinhança *downstream* do gene *pilA*. A análise de quase 300 cepas revelou que o alelo mais comum é o grupo 1, uma pilina que é O-glicosilada em uma serina C-terminal. A glicosilação de PilA está relacionada a mecanismos de defesa da bactéria à adsorção dos fagos (Kus *et al.*, 2004; Burrows, 2012; Harvey *et al.*, 2018).

Outro determinante importante para a interação de fagos com *P. aeruginosa* é o LPS, que poderia desempenhar a função de receptor reversível, servindo como um primeiro guia para o fago no reconhecimento de seu hospedeiro específico, assim como no caso do fago T4 de *E. coli*. Esse reconhecimento inicial garante ao fago maior probabilidade de sucesso na ligação irreversível para o sucesso da etapa de adsorção (Nobrega *et al.*, 2018)

LPS são complexos formados por três principais partes, o lipídio A, o polissacarídeo central (*core*) e o polissacarídeo-O (ou antígeno-O). O lipídeo A, de modo geral, é composto por ácidos graxos ligados a dissacarídeos de glucosamina fosfatada, o polissacarídeo central se liga ao lipídeo A através de cetodesoxioctanos. O polissacarídeo central e o antígeno-O são formados por muitas unidades de resíduos de açúcares se estendendo além da membrana externa das bactérias Gram negativas. Cepas que possuem todos os três componentes descritos acima são denominadas “lisas”. As cepas que não possuem a região do antígeno-O são denominadas “rugosas”. Os açúcares que compõem o antígeno-O são muito variáveis, enquanto a região central possui maior conservação. Assim, os fagos que são especializados em reconhecer cepas do tipo “lisas” possuem maior especificidade de hospedeiros quando comparados a fagos que reconhecem a região central do LPS (Lam *et al.*, 2011).

1.3. Resistência do hospedeiro à adsorção dos fagos

As bactérias podem apresentar resistência a infecção por determinados fagos, ou adquirir resistência ao longo de processos evolutivos. A resistência adquirida pode ser muito custosa para a célula, uma vez que podem implicar em alterações importantes no seu metabolismo e até perda de fatores de virulência (Ormala & Jalasvuori, 2013; Azam & Tanji, 2019).

Além de alterações diretas nos receptores envolvidos na etapa de adsorção, como a glicosilação da pilina A já mencionada acima (Kim *et al.*, 2018), existem outros

mecanismos de bloqueio da infecção viral na etapa de adsorção, tais como produção de matriz extracelular e produção de inibidores competitivos (Labrie *et al.*, 2010).

A produção de polímeros bacterianos extracelulares acontece em diversos ecossistemas promovendo uma barreira contra condições adversas no ambiente, e em alguns casos, esses polímeros podem promover uma barreira física entre receptores específicos do hospedeiro e as partículas virais presentes no ambiente. Em outros casos, moléculas naturalmente presentes no ambiente bacteriano podem se ligar aos receptores específicos para os fagos tornando-os inacessíveis e bloqueando a infecção. Por exemplo, a proteína FuhA em *E. coli* serve como transportador de ferro além de receptor para fagos, podendo levar a um bloqueio de infecção simplesmente pela necessidade de transporte de ferro para dentro da célula (Labrie *et al.*, 2010).

1.4. Fagos como agentes antimicrobianos

A resistência a antimicrobianos/antibióticos (AMR, *Antimicrobial/Antibiotic Resistance*) em bactérias é cada vez mais preocupante, uma vez que utilização inadequada de antimicrobianos ao longo dos anos levou a seleção e propagação de cepas bacterianas multirresistentes. A resistência a antibióticos ocorre por mutações ou aquisições de genes específicos responsáveis pelo metabolismo de determinada droga, resultando na sobrevivência da bactéria mesmo na presença da molécula antimicrobiana. Os principais mecanismos relacionados a resistência a ação do antibiótico são: modificação do sítio alvo reconhecido pela molécula antibiótica, diminuindo a afinidade do receptor pela molécula e logo resultando em menor aporte dessa molécula para dentro da célula; ativação de mecanismos de efluxo responsáveis por carregar para fora da célula as moléculas tóxicas e; alterações globais na regulação de vias metabólicas importantes associadas a ação/metabolismo do antimicrobiano (Munita & Arias, 2016).

A dispersão de genes associados à AMR bacteriana acontece, na maior parte dos casos, através de plasmídeos via conjugação ou transformação, e de maneira menos frequente por transdução (Michod *et al.*, 2008; Colavecchio *et al.*, 2017). Quando mediada por fagos, a mobilização de genes cromossomais ou plasmidiais relacionados a AMR ocorre durante a infecção. Tais genes podem ser replicados e/ou rearranjados no genoma viral, e posteriormente serem excisados erroneamente, levando a dispersão da resistência (Torres-Barcelo, 2018).

O avanço das infecções por bactérias multirresistentes a antibióticos tem estimulado o interesse na utilização de fagos como agente terapêutico (*phage therapy*) (Nobrega *et al.*, 2015; Cisek *et al.*, 2017; Watts, 2017; Mulani *et al.*, 2019; Saha & Mukherjee, 2019). Devido a elevada especificidade e abrangência limitada de hospedeiros suscetíveis a infecção (*narrow host-range*), características comuns para a maioria dos fagos, se espera menor efeito colateral na microbiota comensal comparativamente aos antibióticos convencionais (Clark & March, 2006; Brussow, 2012; Haq *et al.*, 2012; Nobrega *et al.*, 2015; Cisek *et al.*, 2017; Watts, 2017). Soma-se a essa vantagem, o fato de que o isolamento de fagos é relativamente rápido, simples e de baixo-custo e que os fagos são bastante resistentes a condições ambientais desfavoráveis, se mantendo viáveis (virulentos) e se multiplicando até a densidade da bactéria hospedeira ser drasticamente reduzida. Além disso, postula-se que a resistência ao tratamento com fagos seja adquirida de forma muito mais lenta pelo hospedeiro suscetível à infecção lítica (Doss *et al.*, 2017).

O conceito de bacteriófagos como agentes antibacterianos foi primeiramente proposto por d'Hérelle em 1917, e a primeira tentativa de aplicação terapêutica foi no tratamento de disenteria em humanos no hospital Enfants-Malades de Paris em 1919 (Sulakvelidze *et al.*, 2001). Entretanto, com a descoberta da penicilina os avanços nessa área praticamente cessaram, principalmente no Ocidente. Nos últimos anos, porém, a utilização de fagos como agentes antimicrobianos tem sido gradualmente retomada (Merabishvili *et al.*, 2009; Weber-Dabrowska *et al.*, 2016; Gorski *et al.*, 2018; Moelling *et al.*, 2018; Altamirano & Barr, 2019). Por exemplo, um coquetel de nove bacteriófagos líticos para um isolado clínico de *Acinetobacter baumannii* multirresistente a antibióticos foi eficaz no tratamento de um paciente com pancreatite severa (Schooley *et al.*, 2017). Em outro caso, um coquetel de três fagos foi eficaz no tratamento de infecção generalizada por *Mycobacterium abscessus* em uma paciente com fibrose cística (Dedrick *et al.*, 2019).

2. Motivação e Objetivos

A compreensão dos mecanismos moleculares envolvidos na interação dos bacteriófagos virulentos com cepas bacterianas suscetíveis é um dos requisitos para a avaliação do potencial de uso de fagos em fagoterapia. A adsorção do fago a superfície bacteriana é a primeira etapa da interação e um requerimento absoluto para

infecção e posterior replicação do genoma do fago seguida da produção e liberação de vírions. Os determinantes mais comumente envolvidos na etapa de adsorção são a pilina majoritária (PilA) do *pilus* do tipo 4 (T4P) e o Lipopolissacarídeo (LPS), presentes na superfície bacteriana, e as estruturas da cauda da partícula viral (Silva *et al.*, 2016; Nobrega *et al.*, 2018). Diferenças nessas estruturas são críticas para especificidade da interação fago-célula bacteriana.

Em trabalho anterior de nosso grupo foram isolados seis fagos (ZC01, ZC03, ZC04, ZC06, ZC07 e ZC08) de amostras de compostagem utilizando-se *Pseudomonas aeruginosa* PA14 como hospedeira, os quais tiveram seus respectivos genomas sequenciados. Os genomas dos fagos ZC01, ZC03 e ZC08 foram analisados e foi demonstrado experimentalmente que o fago ZC01 (família *Siphoviridae*) e os fagos ZC03 e ZC08 (família *Podoviridae*) tem uma abrangência restrita de isolados de *P. aeruginosa* quanto a suscetibilidade à infecção (Amgarten *et al.*, 2017). Embora os determinantes envolvidos na interação desses fagos com *P. aeruginosa* não tenham sido identificados, dispomos de evidências experimentais que o T4P é necessário para a adsorção dos fagos ZC01, ZC03 e ZC08 à cepa PA14 (Layla F. Martins & Aline M. da Silva, dados não publicados).

Nesse cenário, o **objetivo geral** do presente trabalho foi prospectar determinantes potencialmente envolvidos na interação dos fagos isolados da compostagem com *P. aeruginosa*, através tanto da análise comparativa dos genomas desses fagos como dos genomas de 18 cepas de *P. aeruginosa* que se mostraram suscetíveis ou resistentes à infecção pelos fagos ZC01, ZC03 e/ou ZC08.

Para tal estabelecemos os seguintes **objetivos específicos**:

1. Montar e anotar o genoma dos fagos ZC04, ZC06 e ZC07 e comparar suas sequências com os genomas dos fagos ZC03 e ZC08, todos eles da família *Podoviridae*;
2. Comparar as ORFs preditas nos genomas dos fagos ZC04, ZC06 e ZC07 e buscar por alterações potencialmente importantes para seu ciclo de vida e potencial de infecção;
3. Montar e anotar os genomas de 18 cepas/isolados de *P. aeruginosa* selecionados após varredura de uma coleção de isolados clínicos e ambientais quanto a suscetibilidade aos fagos ZC01, ZC03 e ZC08;

4. A partir das sequências genômicas dos 18 isolados, inferir características fenotípicas dessas cepas tais como *Sequence Type* (ST) ou grupo clonal (CC) e perfil de resistência a antimicrobianos;
5. Comparar as ORFs potencialmente relacionadas ao processo de adsorção fago-bactéria identificadas nos genomas dos 18 isolados de *P. aeruginosa* e relacionar as similaridades ou dissimilaridades com o fenótipo de resistência ou suscetibilidade à infecção pelos fagos em estudo.

3. Metodologias

3.1. Sequenciamento, montagem e anotação de genomas

O DNA dos bacteriófagos ZC04, ZC06 e ZC07 foi extraído e purificado com a utilização do *MoBio PowerMax Soil DNA kit* (Amgarten *et al.*, 2017). O DNA de isolados de *P. aeruginosa* foi extraído e purificado com *Wizard Genomic DNA Purification kit*. As amostras de DNA foram utilizadas para preparação de bibliotecas de fragmentos com o *Nextera DNA Sample Prep kit* (Illumina, Inc., CA), seguindo-se o protocolo do fabricante. O sequenciamento foi realizado no equipamento Illumina/MiSeq (Illumina), instalado no CATG do IQ-USP, utilizando-se o protocolo *Paired-End* (PE) e o *Illumina/MiSeq Reagent kit v2* de 500 ciclos, que possibilita o sequenciamento de 250 bases de cada extremidade dos fragmentos de DNA. Ao final de cada corrida é gerado um arquivo de saída com as sequências (*read1* e *read2*) acopladas a informações da qualidade de cada nucleotídeo sequenciado (arquivo FastQ). Em média, as corridas no MiSeq tiveram bom desempenho gerando >75% das sequências com um índice de qualidade acima de 30 (Q30). As extrações de DNA, preparação das bibliotecas de fragmentos e sequenciamentos no Illumina/MiSeq foram realizados pela Dra. Layla Farage Martins (fagos) e pelo Dr. Paulo Marques Pierry com auxílio de Joice Cassimiro da Silva (cepas de *P. aeruginosa*).

A montagem dos genomas virais e bacterianos foi realizada utilizando-se o programa MIRA (Chevreux *et al.*, 2004) com os seguintes parâmetros: `hirep_good`, `-NW:cac=warn` e `-NW:cmrnl=no`. Para cada montagem foram utilizados 30 *treads* e 200Gb de memória RAM.

Os genomas dos fagos foram anotados inicialmente com o *pipeline* PROKKA, usando Prodigal como o preditor de genes (Seemann, 2014). A partir da anotação

gerada foi utilizado o banco de dados de ortólogos de proteínas virais (VOG) para curar e avaliar manualmente as anotações. O banco de dados de domínios CDD-NCBI também foi utilizado para anotação manual das proteínas classificadas como hipotéticas.

Para a anotação dos genomas dos isolados de *P. aeruginosa* também foi utilizado o PROKKA (Seemann, 2014) a partir do banco de dados de proteínas anotadas em *P. aeruginosa*.

Em alguns casos foi realizada análise de domínios conservados em proteínas. Para tal, os arquivos do tipo fasta de cada ORF anotada foram recuperados com auxílio da plataforma ARTEMIS (Carver *et al.*, 2012) e analisados contra o banco de dados CDD do NCBI (Marchler-Bauer *et al.*, 2017).

3.2. Alinhamento do genoma dos fagos

O alinhamento dos genomas dos fagos em estudo foi realizado utilizando a plataforma MAFFT (Kato & Standley, 2013) com parâmetros *default*. Para visualização gráfica das diferenças no alinhamento dos genomas foi utilizado o programa Unipro UGENE (Okonechnikov *et al.*, 2012) com geração de uma matriz de similaridade. Ainda utilizando a plataforma MAFFT foi gerada uma árvore pelo método de *Neighbor-Joining* (NJ).

3.3. Análises comparativas de genomas de *P. aeruginosa*

Os genomas de *P. aeruginosa* foram comparados com base na frequência de k-mers (=5) com o pacote “FindMyFriends” disponível no Bioconductor². O mapeamento e análise de SNPs/INDELS foi realizado com MUMMER (função *show-snps*) (Kurtz *et al.*, 2004). Para cálculo do pan-genoma e identificação do *core*-genoma e genoma acessório utilizamos o programa ROARY (Page *et al.*, 2015).

3.4. Análises filogenéticas

Para inferência filogenética dos fagos foram utilizados, além dos 5 fagos em estudo, outros 8 fagos pertencentes à mesma família retirados do repositório VOG's (<http://vogdb.org/>). Os genes do *core*-genoma foram identificados a partir do programa *get_homologues* e em seguida foram alinhados com a plataforma MAFFT (Kato &

² http://www.bioconductor.org/packages/devel/bioc/vignettes/FindMyFriends/inst/doc/FindMyFriends_intro.html

Standley, 2013). A árvore foi gerada a partir de alinhamentos múltiplos de aminoácidos e o método de máxima verossimilhança utilizando-se o programa RAxML(Stamatakis, 2014). O modelo evolutivo utilizado foi o Gamma, a partir da frequência empírica de bases e modelo de substituição LG, com 100 bootstraps para formação da árvore consenso final.

As análises filogenéticas a partir de proteínas dos fagos ou de isolados de *P. aeruginosa* também foram realizadas a partir de alinhamentos múltiplos de aminoácidos de sequências de interesse (ex: PilA, PilE ou *core-genoma*) e com o método de máxima verossimilhança utilizando-se o programa RAxML (Stamatakis, 2014). No caso de PilA e PilE, o modelo evolutivo utilizado foi o Gamma, a partir da frequência empírica de bases e modelo de substituição LG, com 100 *bootstraps* para formação da árvore consenso final. Para inferência utilizando-se o *core-genoma* foi utilizado alinhamento múltiplo de nucleotídeos e modelo evolutivo GTR (*General Time Reversible model*) de substituição de nucleotídeos.

3.5. Alinhamentos múltiplos de sequências de proteínas

As CDS/ORFs de interesse (PilA, PilE e outras) foram identificadas nos genomas anotados utilizando-se BLASTP e sequências de referências anotadas no genoma da cepa PAO1 de *P. aeruginosa*. Os arquivos fasta foram recuperados, os alinhamentos múltiplos foram feitos na plataforma MAFFT (Kato & Standley, 2013) e em seguida visualizados com o programa UGENE (Okonechnikov *et al.*, 2012). Em alguns casos, os alinhamentos foram gerados com o ClustalW2 (Larkin *et al.*, 2007).

3.6. Análise dos genes de MLST nas cepas de *P. aeruginosa*

Para análise de MLST, os genes de referência da cepa PAO1 foram recuperados do *Pseudomonas Genome DB*³ (Winsor *et al.*, 2016) e utilizados para recuperação dos ortólogos nos genomas sequenciados e anotados dos 18 isolados de *P. aeruginosa* em estudo. Os arquivos fasta dessas sequências foram submetidos ao *Pseudomonas aeruginosa MLST Database*⁴ utilizando-se a ferramenta de *sequence query* que faz uma busca por alelos específicos dos genes de MLST para definir o ST (*sequence type*) do isolado de interesse.

³ www.pseudomonas.com

⁴ <https://pubmlst.org/paeruginosa/>

3.7. Identificação de genes de resistência a antibióticos nos genomas

Para identificação de genes de resistência a antibióticos nos genomas sequenciados foi utilizada a ferramenta de linha de comando RGI (*Resistance Genes Identifier*)⁵ da base de dados CARD (*Comprehensive Antibiotic Resistance Database*) (Jia *et al.*, 2017). A partir dos resultados dessa ferramenta foram criadas tabelas que listam quais genes foram encontrados nos genomas analisados.

4. Resultados e Discussão

4.1. Análise dos genomas dos fagos ZC04, ZC06 e ZC07

Os bacteriófagos nomeados ZC04, ZC06 e ZC07 foram isolados a partir de amostras de compostagem utilizando-se a bactéria *P. aeruginosa* PA14 como hospedeira. Esses fagos foram isolados simultaneamente aos fagos ZC01, ZC03 e ZC08 previamente caracterizados em nosso grupo (Amgarten *et al.*, 2017). A partir da base de dados *Viral Orthologs Clusters* (VOCs)⁶, ZC01 foi classificado no *cluster* denominado *Siphoviridae YuA-like* enquanto que ZC03 e ZC08 correspondem a um novo clado no *cluster* denominado *Podoviridae N4-like* (Amgarten *et al.*, 2017).

4.1.1. Sequenciamento e comparação dos genomas

Os genomas dos fagos ZC04, ZC06 e ZC07 foram sequenciados com a plataforma Illumina/MiSeq e montados com o programa MIRA (Chevreux *et al.*, 2004). A cobertura do sequenciamento foi de ~8000 vezes e uniforme para a sequência dos 3 genomas. Tanto o tamanho como conteúdo GC dos genomas de ZC04, ZC06 e ZC07 são mais próximos dos genomas de ZC03 e ZC08 do que do genoma de ZC01 (Tabela 1). Assim, as comparações genômicas de ZC04, ZC06 e ZC07 descritas adiante foram realizadas com ZC03 e ZC08.

⁵ <https://card.mcmaster.ca/analyze/rgi>

⁶ <https://4virology.net/virology-ca-tools/vocs/>

Tabela 1: Resultados da montagem dos genomas dos fagos sequenciados na plataforma Illumina-MiSeq.

Fago	Número de <i>paired-end reads</i>	Tamanho do <i>contig</i> montado	Número de acesso GenBank	Conteúdo GC (%)
ZC04	2.473.936	70.586	-	42,76
ZC06	2.516.901	70.659	-	43,04
ZC07	2.373.662	69.992	-	43,14
ZC03*	2.432.007	69.844	KU356690.1	42,60
ZC08*	2.679.382	70.774	KU356691.1	43,07
ZC01*	2.589.412	57.061	KU356689.1	63,38

* Informações extraídas de (Amgarten *et al.*, 2017).

As sequências de nucleotídeos dos genomas de ZC04, ZC06 e ZC07 foram alinhadas com as sequências de ZC03 e ZC08 utilizando-se o programa MAFFT (Kato & Standley, 2013). O alinhamento mostrou que ZC04 e ZC03 possuem similaridade de 97% e são mais próximos entre si em relação aos outros genomas comparados (Tabela 2). Os fagos ZC06, ZC07 e ZC08 possuem uma similaridade de sequência de 99% e são mais próximos entre si do que com ZC04 e ZC03 (Tabela 2).

Tabela 2: Comparação das identidades das sequências genômicas de ZC04, ZC03, ZC06, ZC07 e ZC08.

	ZC04	ZC03	ZC06	ZC07	ZC08
ZC04	100%	97%	95%	95%	96%
ZC03	97%	100%	94%	93%	94%
ZC06	95%	94%	100%	99%	99%
ZC07	95%	93%	99%	100%	99%
ZC08	96%	94%	99%	99%	100%

Através do alinhamento com MAFFT, foi construída uma árvore para as sequências dos 5 fagos utilizando-se o método de *Neighbor-Joining* (NJ) (Figura 2). O resultado obtido confirma que ZC03 e ZC04 são mais próximos entre si. Também é possível observar que mesmo com a alta similaridade das sequências de ZC07, ZC06 e ZC08 a árvore mostra maior proximidade de ZC07 e ZC06 relativamente a ZC08.

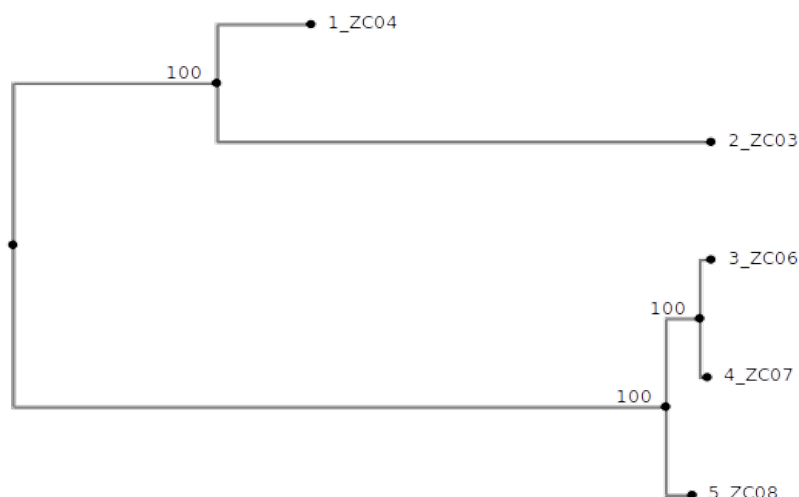


Figura 2: Árvore gerada com o método de *Neighbor-Joining* a partir do alinhamento das sequências genômicas dos fagos com MAFFT. O valor de *bootstrap* está indicado.

4.1.2. Anotação dos genomas e comparação das ORFs preditas

As sequências genômicas dos fagos ZC04, ZC06 e ZC07 foram anotadas utilizando o programa PROKKA (Seemann, 2014). O resultado da anotação mostrou que o número de genes preditos para esses 3 fagos é muito semelhante aos números preditos para ZC03 e ZC08 (Amgarten *et al.*, 2017) (Tabela 3). A semelhança entre os genomas dos cinco fagos é também corroborada pela organização dos genes preditos (Figura 3).

Tabela 3: Anotação dos genomas de ZC04, ZC06 e ZC07.

	ZC04	ZC06	ZC07	ZC03*	ZC08*
Genes preditos (ORFs)	83	83	83	85	83
tRNAs	8	9	9	10	9 + 1 Pseudogene

* Informações extraídas de (Amgarten *et al.*, 2017).

Os genes anotados nos genomas de ZC04, ZC06 e ZC07 estão listados nas Tabelas S1, S2 e S3 (apêndice). Estas tabelas também mostram as identidades de sequências de aminoácidos das ORFs anotadas nesses fagos em relação as suas ortólogas em ZC03 e ZC08. Das 83 ORFs anotadas em ZC04, ZC06 e ZC07, 54 são

compartilhadas entre todos os 5 fagos com identidade de sequência de aminoácidos de 100%. As demais (29 ORFs) tem identidades variáveis (30 a 99%) dependendo dos fagos utilizados na comparação. Assim, apesar da alta identidade global de sequência de nucleotídeos entre os genomas dos cinco fagos (Tabela 2), há diferenças entre as sequências de aminoácidos de algumas ORFs (Tabelas S1, S2 e S3) ou na arquitetura de alguns genes.

Como exemplo dessas diferenças citamos as ORFs ZC06-00089 e ZC06-00090, anotadas como DNA polimerase no fago ZC06. Estas duas ORFs conjuntamente representam a DNA polimerase completa anotada nos demais fagos. A “quebra” da sequência em dois seguimentos é consequência da duplicação de um trecho da sequência desse gene que resulta na mudança do quadro de leitura e surgimento de um códon de parada na posição 1515 de ZC06-00089. A inspeção da sequência do genoma descarta eventuais erros de montagem. A análise de ZC06-00089 (504 aa) e ZC06-00090 (440 aa) contra o banco de dados de domínios conservados (CDD-NCBI) mostrou que a primeira ORF contém o domínio 3'→5' exonuclease, responsável pela atividade revisora da DNA polimerase. A segunda ORF apresenta o domínio da DNA polimerase A, responsável pela replicação do DNA. É provável que este defeito na DNA polimerase do fago ZC06 afete sua replicação, pois assume-se que DNA polimerases de fagos sejam potencialmente importantes para sua propagação no hospedeiro (Liu *et al.*, 2016).

Os membros do gênero *N4-like*, conforme a definição taxonômica do ICTV (King *et al.*, 2012), possuem uma RNA polimerase (Vírus-RNAP ou vRNAP) dentro de seu capsídeo, a qual é injetada na célula hospedeira juntamente com o genoma viral para transcrição dos genes virais precoces (Wittmann *et al.*, 2015). Trata-se de uma RNAP de subunidade única codificada pela ORF #000015 nos genomas de ZC03, ZC04, ZC06, ZC07 e ZC08. A vRNAP de ZC03 e ZC04 tem 3644 aa e a ORF equivalente nos genomas de ZC06, ZC07 e ZC08 tem 3677 aa, devido a inserção de um trecho de 30 aminoácidos na posição 982 de sua sequência.

Além disso, os fagos *N4-like* também têm a característica comum de codificar uma RNAP heterodimérica (subunidades RNAP1 e RNAP2) que é responsável pela transcrição dos genes virais intermediários. As RNAP1 e RNAP2 exibem similaridade a regiões separadas não-superpostas da sequência da RNAP de subunidade única, sugerindo que eventos de fusão ou divisão de genes podem ter sido responsáveis pela evolução das RNAPs (Willis *et al.*, 2002; Amgarten *et al.*, 2017). Assim como

ZC03 e ZC08, os fagos ZC04, ZC06 e ZC07 também codificam RNAP1 e RNAP2 (ex: ZC04_00051 e ZC04_00053, respectivamente). Uma hipótese é que essa divisão tenha sido causada pela inserção de uma HE (*Homing endonuclease*) (Amgarten *et al.*, 2017).

Homing endonucleases são endonucleases de DNA, sítio-específicas, cujos genes (HEG, *Homing endonuclease gene*) são descritos como elementos genéticos egoístas. As HE clivam a sequência alvo em sítios homólogos onde o HEG é inserido como resultado da atuação do sistema de reparo de quebras de fita-dupla. A maioria dos HEGs estão inseridos em genes hospedeiro sem perturbar sua função porque estão frequentemente associados a sequências intervenientes que se auto-processam, como introns do grupo I e inteínas (Burt & Koufopanou, 2004; Gogarten & Hilario, 2006; Stoddard, 2011). HEGs estão presentes em genomas de fungos e plantas e são frequentes em genomas procarióticos e virais (Burt & Koufopanou, 2004), como no caso do fago T4 que possui 15 HEGs (Edgell *et al.*, 2010).

A presença de HEs nos genomas dos fagos ZC03, ZC04, ZC06, ZC07 e ZC08 foi pesquisada com uma varredura de HMM a partir de perfis de *clusters* de HEs obtidos da plataforma Viral Orthologs Clusters (VOCs). Os resultados do HMM *scan* estão sumarizados na Tabela 4.

Essa análise confirmou a inserção de uma HE no gene RNAP2 do fago ZC08 (ZC08_00055) como previamente reportado (Amgarten *et al.*, 2017). A inserção desta HE putativa (HNH endonuclease de 209 aa) também é observada no gene de RNAP2 de ZC06 (ZC06_00055) e ZC07 (ZC07_00054). Por outro lado, o gene da RNAP2 em ZC04 e ZC03 não está interrompido por uma HE (Figura 4). Postula-se que a inserção de uma HE no gene da RNAP2 pode ter alguma relação com história evolutiva das RNAPs de fagos (Amgarten *et al.*, 2017).

Nessa mesma análise também foi predita a presença da HE da família DNA endonuclease I-Hmul (Shen *et al.*, 2004) de 207 aa nos genomas de ZC03 (ZC03_00062), ZC04 (ZC04_00055) e ZC08 (ZC08_00057). Como já reportado (Amgarten *et al.*, 2017), essa HE está inserida em um gene que codifica uma proteína com domínio de ATPase (ADCP), que por causa dessa inserção, é dividido em dois segmentos codificadores de 66 aa e 250 aa (Figura 4). Já em ZC06 e ZC07 o gene da ADCP (342 aa) não está dividido por um HEG. O domínio de ATPase está envolvido em diversas funções que requerem hidrólise de ATP para remodelagens moleculares, como recombinações de DNA, replicação e reparo (Snider *et al.*, 2008).

Uma terceira HE putativa (HNH endonuclease de 143 aa) foi predita nos genomas dos 5 fagos (ZC03_00084; ZC04_00080; ZC06_00080; ZC07_00080 e ZC08_00081).

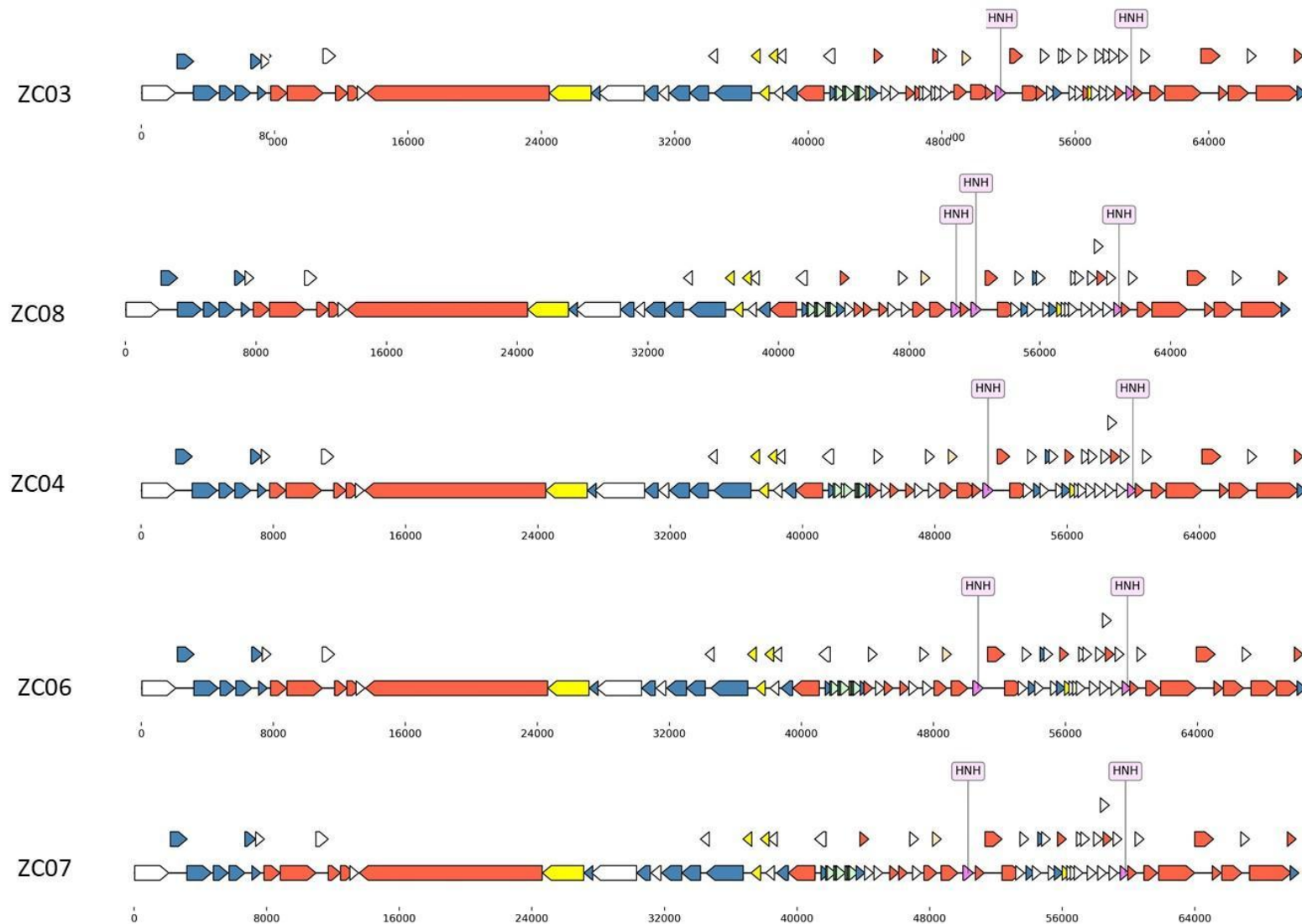


Figura 3: Organização do genoma dos fagos ZC03, ZC04, ZC06, ZC07 e ZC08. As setas indicam os genes preditos nos genomas dos fagos de acordo com sua orientação de transcrição. Em vermelho, ORFs preditas com função de replicação e metabolismo do DNA; em azul, ORFs estruturais; em branco, ORF's com função desconhecida, em amarelo ORFs com função de degradação de estruturas da superfície do hospedeiro. Em rosa claro os elementos moveis preditos e em verde genes dos tRNAs. A descrição dos genes está listada nas Tabelas S1, S2 e S3 (apêndice).

Tabela 4: *Homing Endonucleases* (HE) encontradas pela busca com HMM de VOCs.

Fago/ORF	Posição do domínio predito	Tamanho do domínio predito (aa)	E-value BLASTp	Tamanho predito da ORF (aa)	Anotação
ZC03_00062	50-179	129	1,10E-07	207	DNA endonuclease I-Hmul
ZC03_00084	59-133	74	7,50E-07	143	Putative HNH endonuclease
ZC04_00055	50-179	129	1,10E-07	207	DNA endonuclease I-Hmul
ZC04_00080	59-133	74	7,50E-07	143	Putative HNH endonuclease
ZC06_00055	49-181	132	8,60E-31	209	Putative HNH endonuclease
ZC06_00080	59-133	74	7,50E-07	143	Putative HNH endonuclease
ZC07_00054	49-181	132	8,60E-31	209	HNH endonuclease
ZC07_00080	59-133	74	7,50E-07	143	HNH endonuclease
ZC08_00055	49-181	132	1,30E-30	209	HNH endonuclease
ZC08_00057	49-181	132	1,10E-07	207	DNA endonuclease I-Hmul
ZC08_00081	59-133	74	1,10E-06	143	HNH endonuclease

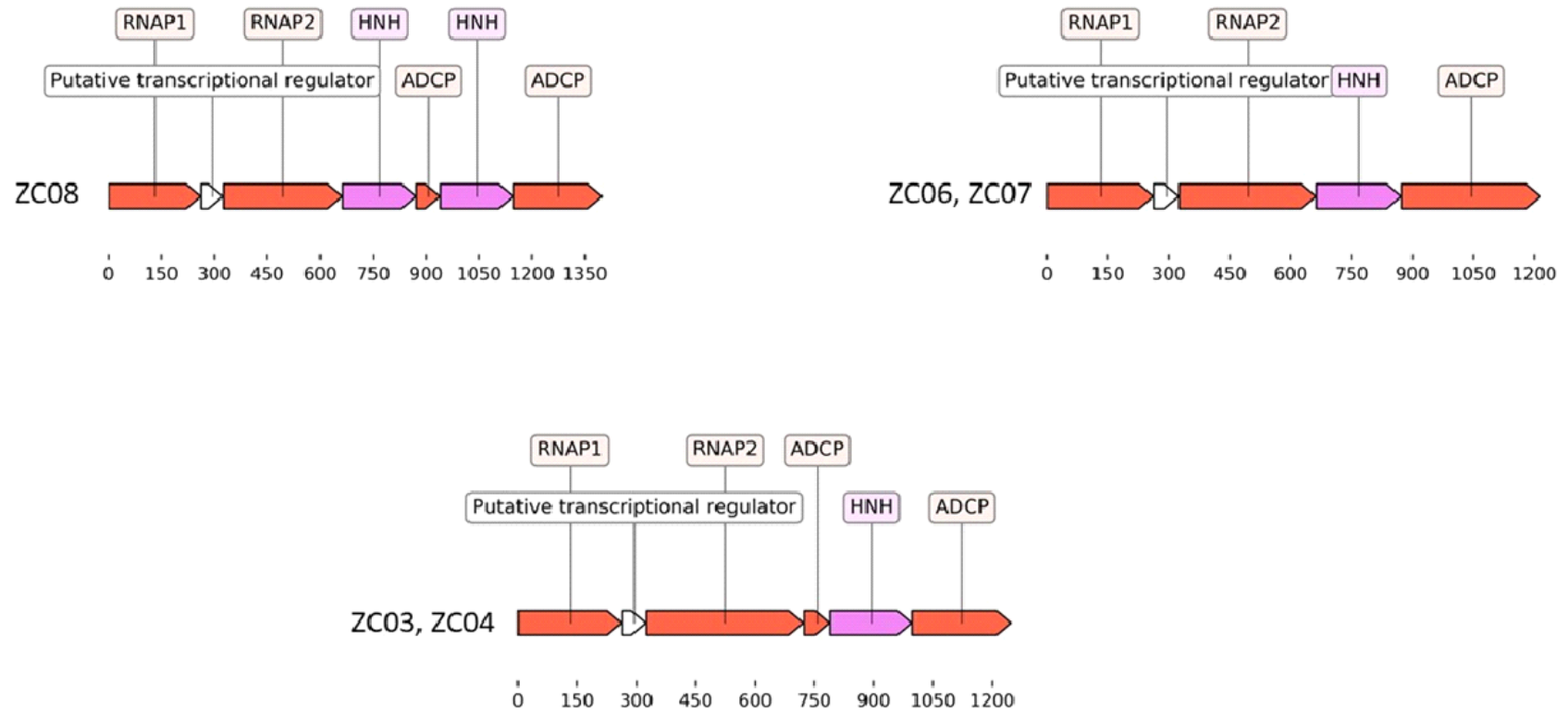


Figura 4: Detalhamento da inserção de genes de *Homing Endonucleases* (HE) nos genomas dos fagos ZC03, ZC04, ZC06, ZC07 e ZC08. As setas indicam os genes preditos nos genomas dos fagos de acordo com sua orientação de transcrição. Os genes de HE da família HNH estão indicados em rosa. Em ZC08, além da HNH que divide a RNAP2, existe uma segunda inserida no gene da ADCP. Em ZC06 e ZC07 apenas a HNH ligada a RNAP2 está presente, e ambos apresentam ADCP intacta. ZC03 e ZC04 apresentam apenas a HNH dividindo ADCP, não apresentando a HNH inserida no gene da RNAP2. Em todos os genomas entre RNAP1 e RNAP2 foi anotada uma ORF predita como regulador transcricional putativo.

4.1.3. Análise filogenética

Como já mencionado, os fagos ZC03 e ZC08 foram previamente classificados no VOCs *cluster* denominado *Podoviridae N4-like* o qual inclui outros 10 fagos (Amgarten *et al.*, 2017). A análise filogenética de ZC04, ZC06 e ZC07 foi realizada com o *core-genoma* do *cluster N4-like*. O *core-genoma* foi definido com a ferramenta *get_homologues* e consiste em 12 genes compartilhados entre os 10 fagos do VOC *N4-like* mais ZC04, ZC06 e ZC07. As sequências proteicas foram concatenadas e a inferência filogenética foi feita com o cálculo de máxima verossimilhança com a ferramenta RAXML (Stamatakis, 2014). A análise filogenética (Figura 5) reforça que os três fagos descritos nesse trabalho, juntamente com ZC03 e ZC08, constituem um novo *clado* do *cluster N4-like* (Amgarten *et al.*, 2017). Ainda que a árvore resultante não mostre um suporte alto para os ramos desse *clado*, ela confirma que ZC06 e ZC07 são mais próximos em relação aos demais, enquanto ZC03 e ZC04 são mais próximos em relação a ZC08.

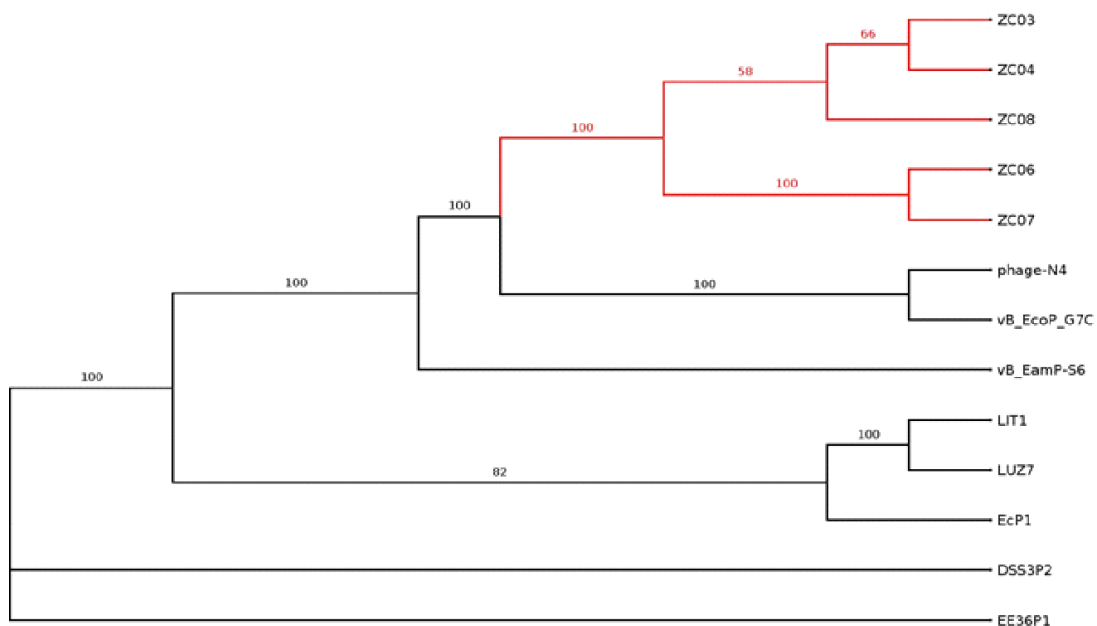


Figura 5: Árvore filogenética não enraizada gerada a partir do *core-genoma* dos fagos ZC04, ZC06 e ZC08 e dos 10 fagos do *cluster N4-like* do VOCs (Amgarten *et al.*, 2017). Foi utilizado o cálculo de máxima verossimilhança com o programa RAXML. Valores de *bootstrap* estão indicados nos ramos da árvore. Além de ZC03 e ZC08, os fagos do *cluster N-4 like* e seus números de acesso no GenBank são: Roseophage DSS3P2 (NC_012697); Erwinia phage vB_EamP-S6 (NC_019514); Enterobacteria phage N4 (NC_008720); Escherichia phage vB_EcoP_G7C (NC_15933); Enterobacter phage EcP1 (NC_019485); Pseudomonas phage LIT1 (NC_013692); Pseudomonas phage (LUZ7 NC_013691) e Roseophage EE36P1 (NC_012696).

4.1.4. Comparação das ORFs preditas como componentes da cauda

As proteínas responsáveis pela interação do fago com seu hospedeiro específico (*receptor-binding proteins*) se localizam na região terminal da cauda, nas fibras (*fibres*) e espigas (*spikes*). Estas estruturas participam do reconhecimento e ligação com diferentes receptores na superfície da bactéria (McPartland & Rothman-Denes, 2009; Leiman *et al.*, 2010; Le *et al.*, 2013; Nobrega *et al.*, 2018).

Já foi demonstrado que os podovirus ZC03 e ZC08, assim como o siphovirus ZC01, têm uma abrangência restrita de hospedeiros, sendo capazes de infectar apenas certas cepas de *P. aeruginosa* (Amgarten *et al.*, 2017). Ademais, parece existir diferenças de suscetibilidade dessas cepas por estes 3 fagos como detalharemos mais adiante nesse trabalho (item 4.2).

Motivados por estas observações, pesquisamos entre as ORFs preditas como estruturais de cauda dos fagos ZC03, ZC04, ZC06, ZC07 e ZC08, aquelas que apresentavam diferenças potencialmente relevantes entre suas sequências de aminoácidos. Identificamos três ORFs distintas, as quais foram comparadas com suas ortólogas nos 5 fagos e com bases de dados para confirmação da predição inicial.

A primeira é uma proteína de 320 aa predita como sendo da fibra da cauda anotada no genoma de todos os cinco fagos. O alinhamento múltiplo (Figura 6) mostra 100% identidade entre as ORFs de ZC06 (ZC06_00005), ZC07 (ZC07_00005) e de ZC08 (ZC08_00005). Por outro lado, a proteína de ZC03 (ZC03_00005) e de ZC04 (ZC04_00005) tem 98% de identidade entre si, e apresentam identidade de 95% com ZC06, ZC07 e ZC08.

A segunda ORF analisada também foi anotada como uma proteína da fibra da cauda (294 aa) e identificada em todos os 5 fagos. O alinhamento múltiplo (Figura 7) de ZC03_00004, ZC04_00004, ZC06_00004, ZC07_00004 e ZC08_00004 revela similaridade de 100% entre as ORFs todos os fagos, com exceção ZC03_00004 que na porção C terminal apresenta diferenças evidentes.

Já foi descrito que o fago (K1-5) apresenta dois tipos distintos de fibras de cauda, e dessa forma pode infectar e replicar tanto as cepas K1 e K5 de *E. coli* (Scholl *et al.*, 2001). Além disso, a alteração de apenas um aminoácido na sequência da proteína da fibra da cauda ampliou o espectro de cepas de *P. aeruginosa* reconhecidas pelo fago PaP1 (Le *et al.*, 2013). Como observado nos testes de infecção realizados para os fagos ZC03 e ZC08 ambos, apesar de muito similares,

apresentaram diferentes resultados de infecção frente a uma mesma cepa, como será discutido adiante.

A terceira ORF analisada (520 aa) foi identificada como *putative central tail spike*. Esta proteína se encontra na região mais terminal da cauda do fago e está relacionada a interação com a membrana externa do hospedeiro (em bactérias Gram-negativas). De maneira geral, a interação acontece entre a *central tail spike* e proteínas de membrana externa (OMP, *Outer Membrane Proteins*). A ligação da partícula viral na membrana externa, é uma etapa crucial que antecede a internalização do DNA viral no citoplasma bacteriano (Nobrega *et al.*, 2018). Essa ORF foi identificada nos cinco fagos (ZC03_00003, ZC04_00003, ZC06_00003, ZC07_00003, ZC08_00003). A partir do alinhamento (Figura 8) é possível observar ZC06, ZC07 e ZC08 apresentam ORFs muito similares entre si e parcialmente similares a ZC04. Por outro lado, a ORF _00003 de ZC03 tem muitas diferenças em relação a de ZC06, ZC07 e ZC08, é parcialmente similar a ZC04_00003. As diferenças verificadas nos permitem especular que a interação dos fagos com a membrana externa da bactéria, através *putative central tail spike*, ocorra a partir de receptores diferentes, ou diferentes mecanismos nos mesmos receptores.

Os resultados apresentados acima indicam que, em relação as ORFs _00003, _00004 e _00005, o fago ZC03 desponta como bastante distinto dos demais, e que ZC08 pode ser considerado um representante ZC06 e ZC07, com similaridade de uma das ORFs com ZC04. Como será discutido adiante, a capacidade de infecção dos fagos ZC03 e ZC08 difere totalmente em alguns isolados de *P. aeruginosa*. Essas observações sugerem que diferenças encontradas nas ORFs preditas como proteínas da fibra da cauda (_00004 e _00005) podem estar relacionadas a efetividade da interação irreversível do fago ao receptor bacteriano. Sabe-se que a adsorção bem sucedida entre fago e a superfície do hospedeiro ocorre em duas etapas: primeiro, uma interação reversível e em seguida uma interação irreversível, as quais podem acontecer com uma mesma estrutura do hospedeiro (Silva *et al.*, 2016). Assim, nossa hipótese é que proteínas da fibra da cauda diferentes interagem de modo distinto (reversível ou irreversível) com um mesmo receptor em uma determinada cepa de *P. aeruginosa*. Além disso, as diferenças observadas nas ORFs referentes a *central tail spike* devem também contribuir para explicar as diferenças observadas nos ensaios de infecção de ZC03 e ZC08 contra certas cepas avaliadas.

```

ZC03_00005    MINLHFGLAGYVDLHKGILDSNGEPVTRHITSFKNLVLDLDTGLQORIGENSDWMNWLHLGT
ZC04_00005    MINLHFGLAGYVDLHKGILDSNGEPVTRHITSFKNLVLDLDTGLQORIGENGDWMNWLHLGT
ZC06_00005    MINLHFGLAGYVDLHKGILDSNGEPVTRHITSFKNLVLDLDTGLQORIGENRDWMSWLHLGT
ZC07_00005    MINLHFGLAGYVDLHKGILDSNGEPVTRHITSFKNLVLDLDTGLQORIGENRDWMSWLHLGT
ZC08_00005    MINLHFGLAGYVDLHKGILDSNGEPVTRHITSFKNLVLDLDTGLQORIGENRDWMSWLHLGT
*****.*****

ZC03_00005    GVQAPHPLQNALQNPTYKGDNLAPSPHTTTGINVSDPLKPYCWIRRTFRVTPRGENRTYA
ZC04_00005    GVQAPHPLQNALQNPTYKGNLAPAPHTTSGINISDPLKPYCWVRRVFRVTPRGENRTYA
ZC06_00005    GVQAPHPLQNSLQNPTYKGNLAPAPHTTTGINISDPLKPYCWIRRIFRVTPRGENRTYA
ZC07_00005    GVQAPHPLQNSLQNPTYKGNLAPAPHTTTGINISDPLKPYCWIRRIFRVTPRGENRTYA
ZC08_00005    GVQAPHPLQNSLQNPTYKGNLAPAPHTTTGINISDPLKPYCWIRRIFRVTPRGENRTYA
*****:*****:****:***:***:*****:*** *****

ZC03_00005    EMGVGWNDDNNLFSRRTLIKDPGPNPTISILGDEYLDVTEVVRMYIPVETAVYSVVPTGDD
ZC04_00005    EMGVGWNDDNNLFSRRTLIKDPGPNPTISILGDEYLDVTEVVRMYIPVETAVYSVVPTGDD
ZC06_00005    EMGVGWNDDNNLFSRRTLIKDPGPNPTISILGDEYLDVTEVVRMYIPVETAVYSVVPTGDD
ZC07_00005    EMGVGWNDDNNLFSRRTLIKDPGPNPTISILGDEYLDVTEVVRMYIPVETAVYSVVPTGDD
ZC08_00005    EMGVGWNDDNNLFSRRTLIKDPGPNPTISILGDEYLDVTEVVRMYIPVETAVYSVVPTGDD
*****

ZC03_00005    VEPRTITVHASRVNTESTSFGWGLAASGSQGTRVPGSILNNWTSNHHNRRFFNGGRGGLYQ
ZC04_00005    VEPRTITVHASRVNTESTSFGWGLAASGSQGTRVPGSILNNWTSNHHNRRFFNGGRGGLYQ
ZC06_00005    VEPRTITVHASRVNTESTSFGWGLAVSGTQGERVPGSILNSYSSTYHNRFFNGGRGGLYQ
ZC07_00005    VEPRTITVHASRVNTESTSFGWGLAVSGTQGERVPGSILNSYSSTYHNRFFNGGRGGLYQ
ZC08_00005    VEPRTITVHASRVNTESTSFGWGLAVSGTQGERVPGSILNSYSSTYHNRFFNGGRGGLYQ
*****.*****.**:** *****.:*:*:*****

ZC03_00005    DPEGSQVGVSTHNRTSMVRTSSTSATITLTRELDPNVGLLRITYQISQEAYCFQMEFDPPFN
ZC04_00005    DPEGSQVGVSTHNRTSMVRTSSTSATITLTRELDPNVGLLRITYQISQEAYCFQMEFDPPFN
ZC06_00005    DPEGSQVGVSVHNRTSMVRTSSTSATITLTRELDPNVGLLRITYQISQEAYCFQMEFDPPFN
ZC07_00005    DPEGSQVGVSVHNRTSMVRTSSTSATITLTRELDPNVGLLRITYQISQEAYCFQMEFDPPFN
ZC08_00005    DPEGSQVGVSVHNRTSMVRTSSTSATITLTRELDPNVGLLRITYQISQEAYCFQMEFDPPFN
*****.*****

ZC03_00005    KSNEDRFSITYSISWGREEE
ZC04_00005    KSNEDRFSITYSISWGREEE
ZC06_00005    KSNEDRFSITYSISWGREEE
ZC07_00005    KSNEDRFSITYSISWGREEE
ZC08_00005    KSNEDRFSITYSISWGREEE
*****

```

Figura 6: Alinhamento da ORF que codifica a proteína de 320 aminoácidos da fibra da cauda dos fagos ZC03, ZC04, ZC06, ZC07 e ZC08. Resíduos conservados e semi-conservados estão indicados com asteriscos, pontos e dois pontos, respectivamente.

```

ZC03_00004    MATVTTLTWQIEGVHDGIGVYRSNQIDSGYEQVAVLAPNTTVFQDEFLLSQLTDKYGTEPE
ZC04_00004    MATVTTLTWQIEGVHDGIGVYRSNQIDSGYEQVAVLAPNTTVFQDEFLLSQLTDKYGTEPE
ZC06_00004    MATVTTLTWQIEGVHDGIGVYRSNQIDSGYEQVAVLAPNTTVFQDEFLLSQLTDKYGTEPE
ZC07_00004    MATVTTLTWQIEGVHDGIGVYRSNQIDSGYEQVAVLAPNTTVFQDEFLLSQLTDKYGTEPE
ZC08_00004    MATVTTLTWQIEGVHDGIGVYRSNQIDSGYEQVAVLAPNTTVFQDEFLLSQLTDKYGTEPE
*****

ZC03_00004    TLFYRLDAFRGLDVKQSDPVPVPLEDTMPEFLDGVGYLEVVPPELKFQDMPATYYNVGVKS
ZC04_00004    TLFYRLDAFRGLDVKQSDPVPVPLEDTMPEFLDGVGYLEVVPPELKFQDMPATYYNVGVKS
ZC06_00004    TLFYRLDAFRGLDVKQSDPVPVPLEDTMPEFLDGVGYLEVVPPELKFQDMPATYYNVGVKS
ZC07_00004    TLFYRLDAFRGLDVKQSDPVPVPLEDTMPEFLDGVGYLEVVPPELKFQDMPATYYNVGVKS
ZC08_00004    TLFYRLDAFRGLDVKQSDPVPVPLEDTMPEFLDGVGYLEVVPPELKFQDMPATYYNVGVKS
*****

ZC03_00004    FNDGNYHVLFGRLRRGPSLFITDPQNGNAPFTEVNQFLRNMPFNASNLEEAIAITMAKQQ
ZC04_00004    FNDGNYHVLFGRLRRGPSLFITDPQNGNAPFTEVNQFLRNMPFNASNLEEAIAITMAKQQ
ZC06_00004    FNDGNYHVLFGRLRRGPSLFITDPQNGNAPFTEVNQFLRNMPFNASNLEEAIAITMAKQQ
ZC07_00004    FNDGNYHVLFGRLRRGPSLFITDPQNGNAPFTEVNQFLRNMPFNASNLEEAIAITMAKQQ
ZC08_00004    FNDGNYHVLFGRLRRGPSLFITDPQNGNAPFTEVNQFLRNMPFNASNLEEAIAITMAKQQ
*****

ZC03_00004    YVLESAADIQMDAMSPFIFREEPILEILPTAEMMQEMQSDILALVAQESLTGYFWTSTMS
ZC04_00004    YVLESAADIQMDAMSPFIFREEPILEILPTAEMMQEMQSDILALVAQESLTGYFWTSTMS
ZC06_00004    YVLESAADIQMDAMSPFIFREEPILEILPTAEMMQEMQSDILALVAQESLTGYFWTSTMS
ZC07_00004    YVLESAADIQMDAMSPFIFREEPILEILPTAEMMQEMQSDILALVAQESLTGYFWTSTMS
ZC08_00004    YVLESAADIQMDAMSPFIFREEPILEILPTAEMMQEMQSDILALVAQESLTGYFWTSTMS
*****

ZC03_00004    PSATNEFGQVVDIDVERFSVAGVSSDSVSI LGDNNYAMTLVFVLIKDI TPLAA
ZC04_00004    PSATNEFGQVVDIDVERFSVAGVSSDSVSI LGDNNYAMTLVFVLIKDI TPLAA
ZC06_00004    PSATNEFGQVVDIDVERFSVAGVSSDSVSI LGDNNYAMTLVFVLIKDI TPLAA
ZC07_00004    PSATNEFGQVVDIDVERFSVAGVSSDSVSI LGDNNYAMTLVFVLIKDI TPLAA
ZC08_00004    PSATNEFGQVVDIDVERFSVAGVSSDSVSI LGDNNYAMTLVFVLIKDI TPLAA
*****.:***** .:*****.: ***** *

```

Figura 7: Alinhamento da ORF que codifica a proteína de 294 aminoácidos da fibra da cauda dos fagos ZC03, ZC04, ZC06, ZC07 e ZC08. Resíduos conservados e semi-conservados estão indicados com asteriscos e pontos, respectivamente.

```

ZC03_00003 MIPMPSQINILIGVVSPGYIP---ESLRTITVNATRKGGQQTNVYKEAVLWSSLRDRVEL
ZC04_00003 MIPMPSQINILIGVVSPGYIP---ESLRTITVNATRKGGQQTNVYKEAVLWSSLRDRVEL
ZC06_00003 -----MDVLTPLYLVFPGSGLAHRITIEVRVTRDGEQTAAQK-LVLMLDGTGQIQQA
ZC07_00003 -----MDVLTPLYLVFPGSGLAHRITIEVRVTRDGEQTAAQK-LVLMLDGTGQIQQA
ZC08_00003 -----MDVLTPLYLVFPGSGLAHRITIEVRVTRDGEQTAAQK-LVLMLDGTGQIQQA
          :.:*: * : : ** *..*.*:** . * ** . .:.:

ZC03_00003 AIFPLKDGITSTRTFAASAVDGKKAIVLYDFDEIPICSVRTLNIITDDTVLYIDLDDGQIP
ZC04_00003 AIFPLKDGITSTRTFAASAVDGKKAIVLYDFDEIPICSVRTLNIITDDTVLYIDLDDGQIP
ZC06_00003 A---FSTDQSTSQVFMASDVDREWLFVARDGDDPPRGAAAVVAVDSDVTIALDLNEHKRP
ZC07_00003 A---FSTDQSTSQVFMASDVDREWLFVARDGDDPPRGAAAVVAVDSDVTIALDLNEHKRP
ZC08_00003 A---FSTDQSTSQVFMASDVDREWLFVARDGDDPPRGAAAVVAVDSDVTIALDLNEHKRP
          * . . :.:.* ** * : : * * * : * . . : .*:.:**.: : *

ZC03_00003 QGGNGLLSGSFPLSLITQGGNLSQANIRILYRPEQGASGDGYVVATTCNADGTWQVSG
ZC04_00003 QGGNGLLSGSFPLSLITQGGNLSQANIRILYRPEQGASGDGYVVATTCNADGTWQVSG
ZC06_00003 EGGNGLLSGSFPLSLITQGGNLSQANIRILYRPEQGASGDGYVVATTCNADGTWQVSG
ZC07_00003 EGGNGLLSGSFPLSLITQGGNLSQANIRILYRPEQGASGDGYVVATTCNADGTWQVSG
ZC08_00003 EGGNGLLSGSFPLSLITQGGNLSQANIRILYRPEQGASGDGYVVATTCNADGTWQVSG
          :*****.*****

ZC03_00003 LLENLKFDIVARI PGFNVDVII SDVQPTGAPLSAHFANLKEEYEGEEVNIQVAMGGRPP
ZC04_00003 LLENLKFDIVARI PGFNVDVII SDVQPTGAPLSAHFANLKEEYEGEEVNIQVAMGGRPP
ZC06_00003 LLENLKFDIVARI PGFNVDVII SDVQPTGAPLSAHFANLKEEYEGEEVNIQVAMGGRPP
ZC07_00003 LLENLKFDIVARI PGFNVDVII SDVQPTGAPLSAHFANLKEEYEGEEVNIQVAMGGRPP
ZC08_00003 LLENLKFDIVARI PGFNVDVII SDVQPTGAPLSAHFANLKEEYEGEEVNIQVAMGGRPP
          *****

ZC03_00003 YTYEAITLPNGLSLSSNGYITGTILDLELTFEVKVTDAEGSVINLVGNSYVAGDPHWK
ZC04_00003 YTYEAITLPNGLSLSSNGYITGTILDLELTFEVKVTDAEGSVINLVGNSYVAGDPHWK
ZC06_00003 YTHEAITLPNGLSLSSNGYITGTILDLELTFEVKVTDAEGSVINLVGNSYVAGDPHWK
ZC07_00003 YTHEAITLPNGLSLSSNGYITGTILDLELTFEVKVTDAEGSVINLVGNSYVAGDPHWK
ZC08_00003 YTHEAITLPNGLSLSSNGYITGTILDLELTFEVKVTDAEGSVINLVGNSYVAGDPHWK
          **:*****.***** :*:*****.*****

ZC03_00003 VVTLHFHDG-DFTDET-----GRSWTPSNGSVIGNIGVFG
ZC04_00003 VVALLHFDGADGSNEAIEEKGILMSVTRGTPSLSSAEFKFGGTSWHTTNGRITSSL----
ZC06_00003 VVALLHFDGADGSNEAIEEKGILMSVTRGTPSLSSAEFKFGGTSWHTTNGRITSSL----
ZC07_00003 VVALLHFDGADGSNEAIEEKGILMSVTRGTPSLSSAEFKFGGTSWHTTNGRITSSL----
ZC08_00003 VVALLHFDGADGSNEAIEEKGILMSVTRGTPSLSSAEFKFGGTSWHTTNGRITSSL----
          **:***** * :.:* * ** .:*. : .:

ZC03_00003 SPGYLHIPSQQAQKIYSTAAPIKRYDHTMEAFIRFGNINQTLRLRPIFSQTDGGSNKQQA
ZC04_00003 -----QSNLSFGAG-----DFTIEAFVMPTNL-----GVRNEQFI
ZC06_00003 -----QSNLSFGAG-----DFTIEAFVMPTNL-----GVRNEQFI
ZC07_00003 -----QSNLSFGAG-----DFTIEAFVMPTNL-----GVRNEQFI
ZC08_00003 -----QSNLSFGAG-----DFTIEAFVMPTNL-----GVRNEQFI
          *.:. * . **:**: * : * **

ZC03_00003 VSI-----NTTTCGLQFHRQVNVGSVVTMNGVTQVPENTWNHVALSFDGSNIRLFLNGN
ZC04_00003 VCIRNQPAAANGQIVSLQNGRLGFSNGTAWLLGSTALSENWVSHVAVSREAGILRFFVNGV
ZC06_00003 VCIRNQPAAANGQIVSLQNGRLGFSNGTAWLLGSTALSENWVSHVAVSREAGILRFFVNGV
ZC07_00003 VCIRNQPAAANGQIVSLQNGRLGFSNGTAWLLGSTALSENWVSHVAVSREAGILRFFVNGV
ZC08_00003 VCIRNQPAAANGQIVSLQNGRLGFSNGTAWLLGSTALSENWVSHVAVSREAGILRFFVNGV
          *.* * ** * . . . : * * :.*.*.*:**: * . . :*:.*

ZC03_00003 LEAIASNTDGWYDLGQMEIGCYTPNPYITYGFLGDIDEIRITKGVARYTENFTPPKPF
ZC04_00003 LDGQVVDNTNFGDRILSIGDAESSH--SSAFVGFIDELRITKGVARYTENFTPPKPF
ZC06_00003 LDGQVVDNTNFGDRILSIGDAESSH--SSAFVGFIDELRITKGVARYTENFTPPKPF
ZC07_00003 LDGQVVDNTNFGDRILSIGDAESSH--SSAFVGFIDELRITKGVARYTENFTPPKPF
ZC08_00003 LDGQVVDNTNFGDRILSIGDAESSH--SSAFVGFIDELRITKGVARYTENFTPPKPF
          *.:. . . :. : . :.* ** .: : .*:.* **:*:***** ** **

ZC03_00003 NY
ZC04_00003 NY
ZC06_00003 NY
ZC07_00003 NY
ZC08_00003 NY
          **

```

Figura 8: Alinhamento da ORF que codifica a *putative central tail spike* dos fagos ZC03, ZC04, ZC06, ZC07 e ZC08. Resíduos conservados e semi-conservados estão indicados com asteriscos e pontos, respectivamente.

4.2. Prospecção de determinantes de *P. aeruginosa* associados à suscetibilidade à infecção pelos fagos ZC01, ZC03 e ZC08

O espectro de ação dos fagos ZC01, ZC03 e ZC08 foi avaliado pela varredura de uma coleção de dezenas de isolados clínicos e ambientais de *P. aeruginosa* quanto a suscetibilidade à infecção por esses fagos ((Amgarten *et al.*, 2017); Layla F. Martins & Aline M. da Silva, dados não publicados). Estas análises revelaram que dentre as cepas resistentes à infecção por esses fagos, está a cepa de referência PAO1. Esta cepa apresenta diferenças genótípicas e fenotípicas em relação a cepa PA14, utilizada no isolamento de fagos a partir de amostras de compostagem e, portanto, suscetível à infecção lítica. PA14 é mais virulenta do que PAO1 e contém duas ilhas de patogenicidade que estão ausentes em PAO1 (He *et al.*, 2004; Lee *et al.*, 2006). Resultados obtidos em nosso laboratório sugerem que T4P parece ser um receptor para os fagos ZC01, ZC03 e ZC08, visto que um mutante da cepa PA14 cujo gene da pilina majoritária (*pilA*) do T4P foi deletado deixa de ser suscetível à infecção. Além disso, resultados iniciais indicam que a expressão de *pilA* da PA14 na cepa PAO1 torna essa cepa suscetível à infecção pelos 3 fagos (Layla F. Martins & Aline M. da Silva, dados não publicados).

Diversos fagos que infectam *P. aeruginosa* utilizam o T4P como receptor (Bradley & Pitt, 1974; James *et al.*, 2012; Bondy-Denomy *et al.*, 2016; Harvey *et al.*, 2018; Kim *et al.*, 2018). Além do T4P, outro receptor frequente é o LPS (Lam *et al.*, 2011), e a interação fago/bactéria pode depender de uma ação combinada desses determinantes (Silva *et al.*, 2016).

Com o propósito de prospectar os determinantes de *P. aeruginosa* associados à suscetibilidade aos fagos ZC01, ZC03 e ZC08, selecionamos, a partir da varredura mencionada acima, isolados suscetíveis e resistentes à infecção para sequenciamento e análise genômica. A Tabela 5 lista os 18 isolados selecionados para sequenciamento incluindo as cepas de referência PA14 e PAO1 em uso em nosso laboratório.

Além da cepa PA14 (P6), os isolados P7, P11 e P17 foram suscetíveis à infecção pelos três fagos. Por outro lado, os isolados P2 (PAO1), P3 e P12 foram resistentes aos três fagos. Três isolados (P5, P13 e P16) se mostraram suscetíveis aos fagos ZC01 e ZC08 porém resistentes a ZC03. Já os isolados P4, P8 e P9 foram

suscetíveis a ZC01 e resistentes a ZC03 e ZC08. Em alguns casos (isolados P1, P10, P14, P15, P18) foi observada a formação de placa de lise túrbida nos ensaios de infecção, o que pode indicar capacidade de adsorção e lise do hospedeiro, mas sem multiplicação do fago (Layla F. Martins & Aline M. da Silva, dados não publicados). Os dados de susceptibilidade aos três fagos (ZC01, ZC03 e ZC08) estão sumarizados na Tabela 5 (colunas 3, 4 e 5).

Tabela 5: Isolados de *P. aeruginosa* selecionados para sequenciamento genômico

ID	Isolado	Suscetibilidade aos fagos			Origem/Descrição	Fonte	Referência
		ZC01	ZC03	ZC08			
P1	SC-61	+	-/+	+	Secreção nasal humana	Dras. Cintya de Oliveira Souza & Karla Lima (Instituto Evandro Chagas)	Não publicado
P2	PAO1	-	-	-	Mutante espontâneo da cepa PAO isolada de ferida (em Melbourne, Austrália, 1954) – cepa de referência	Profa. Regina L. Baldini (IQ-USP)	(Holloway, 1955; Stover <i>et al.</i> , 2000)
P3	SC-116	-	-	-	Secreção traqueal humana	Dras. Cintya de Oliveira Souza & Karla Lima (Instituto Evandro Chagas)	Não publicado
P4	P13.612	+	-	-	Ambiente hospitalar	Profa. Ana Cristina Gales (EPM-UNIFESP)	Não publicado
P5	Fc79a M	+	-	+	Secreção pulmonar de paciente com fibrose cística	Profa. Ana Lúcia da Costa Darini (FCFRP-USP)	Não publicado
P6	PA14	+	+	+	Isolado clínico de paciente – altamente virulento – cepa de referência	Profa. Regina L. Baldini (IQ-USP)	(Rahme <i>et al.</i> , 1995; He <i>et al.</i> , 2004)
P7	Fc7f NM	+	+	+	Secreção pulmonar de paciente com fibrose cística	Profa. Ana Lúcia da Costa Darini (FCFRP-USP)	Não publicado
P8	MT222	+	-	-	Aspirado traqueal de paciente em UTI	Profa. Ana Lúcia da Costa Darini (FCFRP-USP)	(Galetti <i>et al.</i> , 2015)
P9	MT138	+	-	-	Aspirado traqueal de paciente em UTI	Profa. Ana Lúcia da Costa Darini (FCFRP-USP)	(Galetti <i>et al.</i> , 2015)
P10	H6086	+	-/+	-/+	Cultura de sangue de paciente hospitalar	Prof. Nilton Erbet Lincopan Huenuman (ICB-USP)	(Amgarten <i>et al.</i> , 2017)
P11	Fc79a PAB NM	+	+	+	Secreção pulmonar de paciente com fibrose cística	Profa. Ana Lúcia da Costa Darini (FCFRP-USP)	Não publicado
P12	SC-84	-	-	-	Secreção ocular humana	Dras. Cintya de Oliveira Souza & Karla Lima (Instituto Evandro Chagas)	Não publicado
P13	3845/GSP-3	+	-	+	Fezes de paciente com HIV	Dras. Cintya de Oliveira Souza & Karla Lima (Instituto Evandro Chagas)	Não publicado
P14	ALERTA 275/VIM-7	-/+	-/+	-/+	Ambiente hospitalar	Profa. Ana Cristina Gales (EPM-UNIFESP)	Não publicado
P15	ALERTA 226/GES-5	-/+	-/+	-/+	Ambiente hospitalar	Profa. Ana Cristina Gales (EPM-UNIFESP)	Não publicado
P16	ALERTA 395/IMP-18	+	-	+	Ambiente hospitalar	Profa. Ana Cristina Gales (EPM-UNIFESP)	Não publicado
P17	H6044	+	+	+	Cultura de sangue de paciente hospitalar	Prof. Nilton Erbet Lincopan Huenuman (ICB-USP)	(Amgarten <i>et al.</i> , 2017)
P18	5757	-/+	-/+	-	Cultura de sangue/urina de paciente hospitalar	Prof. Nilton Erbet Lincopan Huenuman (ICB-USP)	(Amgarten <i>et al.</i> , 2017)

O ensaio de infecção foi realizado pelo gotejamento de 3 diluições da suspensão de fagos ZC01, ZC03 e ZC08 sobre o tapete de bactérias. A formação placa de lise em pelo menos uma diluição foi interpretada com suscetível a lise (+). Quando nenhuma lise foi observada, o isolado foi considerado resistente (-). Em alguns casos foi observada a formação de placa túrbida (-/+) na diluição mais concentrada, o que pode indicar capacidade de adsorção e lise de fora para dentro (sem multiplicação do fago)

4.2.1. Sequenciamento e comparação dos genomas de 18 isolados de *P. aeruginosa*

Os genomas dos 18 isolados listados na Tabela 5 foram sequenciados com a plataforma Illumina/MiSeq, montados com o programa MIRA (Chevreux *et al.*, 2004) e anotados com o programa PROKKA (Seemann, 2014). As métricas do sequenciamento, montagem e anotação estão sumarizadas na Tabela 6.

Os genomas foram primeiramente comparados quanto as suas respectivas similaridades com base na frequência de k-mers (Figura 9). Essa análise mostrou que, de modo geral, as cepas possuem similaridade muito alta (maior do que 75%), exceto a cepa SC-116 que apresentou similaridade entre 50% e 75% em relação as demais cepas.

Os genomas das cepas PAO1 (P2) e PA14 (P6) que sequenciamos foram comparados aos respectivos genomas de referência publicamente disponíveis. O alinhamento em ambos os casos foi de 99,9% com 186 e 283 SNPs/INDELS encontrados, respectivamente, entre os dois genomas de PAO1 e os dois genomas de PA14 (Tabela S4, apêndice). Além disso, na montagem do genoma de P2 e do genoma de P6 foram encontrados *scaffolds* distintos, com respectivamente 1800 pb e 1250 pb que não mapearam contra os genomas de referência. Tais diferenças podem refletir erros de sequenciamento ou variações que essas cepas acumularam ao longo do cultivo em laboratório. De todo modo, as variações observadas não parecem ser relevantes no contexto desse trabalho, e serão melhor investigadas futuramente.

Também determinamos o pan-genoma (Figura 10) e analisamos a presença e ausência de genes (Figura 11) pela comparação dos 18 genomas sequenciados e anotados. Na análise do pan-genoma é possível observar que do total de genes anotados nos 18 genomas, em média 5927 ± 205 (Tabela 6), 5088 foram considerados como *core-genoma* (compartilhados dentre todas as 18 cepas com identidade de sequência de nucleotídeos $\geq 95\%$). Utilizando esse parâmetro, observamos que aproximadamente 1000 genes são diferentes entre as 18 cepas. A matriz de a presença e ausência de genes ilustra essa observação (Figura 11).

4.2.2. Análises dos genomas dos 18 isolados e relação com a suscetibilidade por ZC01, ZC03 e ZC08 e a presença de genes de resistência

Com o objetivo de ampliar as informações sobre os 18 isolados em estudo, analisamos seus genomas quanto a presença de genes de resistência a antibióticos. Para tal, foram feitas buscas de genes de resistência com a versão de linha de comando do programa CARD RGI (*Resistance Gene Identifier*), que utiliza o banco de dados CARD (*Comprehensive Antibiotic Resistance Database*) para predição de genes de resistência em arquivos de sequências de nucleotídeos. Essa análise identificou 70 genes de resistência dentre as 18 cepas em estudo (Tabela S5, apêndice). Dentre todos os genes encontrados, 47 foram identificados em pelo menos 17 das 18 cepas, sugerindo que estas cepas sejam resistentes a múltiplas-drogas, exibindo resistência a classes específicas de antibióticos (aminoglicosídeos, beta-lactâmicos, biciclomicina, cloranfenicol, fosfomicina e polimixina). Este resultado era esperado, pois sabe-se que *P. aeruginosa* possui resistência intrínseca a muitas classes de antibióticos, além de possuir grande capacidade de adquirir resistência a antibióticos através de mutações (Livermore, 2002; Jeukens *et al.*, 2019).

Também determinamos o ST (*Sequence Type*) de cada isolado (Tabela 7), pela análise dos genes utilizados para MLST (*MultiLocus Sequence Typing*). As duas cepas de referência (P2 e P6), como esperado, foram identificadas como grupo clonal (CC), respectivamente PAO1 e PA14. Os isolados apresentaram grande variação quanto ao ST, sendo poucas as sobreposições de STs. Os isolados P8 e P15 não foram classificados com STs conhecidos, sendo que P15 não mostrou alelos similares aos registrados nessa base de dados. Será necessária uma análise mais detalhada para confirmar se estes dois isolados correspondem de fato a novos STs.

Não foi possível identificar qualquer relação entre os STs dos 18 isolados e suas respectivas suscetibilidades à infecção pelos 3 fagos (Tabela 7). Tão pouco identificamos relação evidente entre as relações filogenéticas dos 18 isolados e suas respectivas suscetibilidades à infecção pelos 3 fagos (Figura 12).

Tabela 6: Resultados da montagem dos genomas dos isolados de *P. aeruginosa* sequenciados com a plataforma Illumina-MiSeq.

ID	Isolado	Número de <i>paired-end reads</i>	Tamanho total do genoma (pb)	Número de contigs montados	Número de <i>scaffolds</i>	Total de genes	Conteúdo GC (%)
P1	SC-61	1.649.536	6,4M	53	14	6068	66,24
P2	Pao1	3.409.900	6,1M	29	5	5687	66,56
P3	SC116	1.136.598	6,1M	266	6	5742	66,61
P4	P13.612	1.533.026	6,2M	33	4	5838	66,45
P5	Fc79aM	1.252.852	6,2M	80	22	5744	66,49
P6	PA14	2.958.229	6,2M	59	11	5920	66,30
P7	Fc7fNM	3.799.678	6,3M	34	5	5852	66,41
P8	MT222	2.711.010	6,1M	17	4	5699	66,53
P9	MT138	2.980.152	6,6M	62	22	6122	66,07
P10	H6086	2.461.616	6,5M	32	12	6128	66,10
P11	FC79aNM	2.516.462	6,2M	29	9	5748	66,48
P12	Sc-84	2.275.350	6,2M	36	11	5798	66,39
P13	3845/GSP-3	2.517.594	6,3M	22	3	5917	66,32
P14	ALERTA 275/VIM-7	3.061.074	6,7M	72	31	6279	66,02
P15	ALERTA 226/GES-5	2.461.504	6,9M	61	15	6413	66,81
P16	IMP-18	1.662.608	6,5M	44	10	6112	66,27
P17	H6044	1.847.650	6,3M	31	7	5791	66,49
P18	5757	2.608.210	6,4M	34	8	5831	66,38

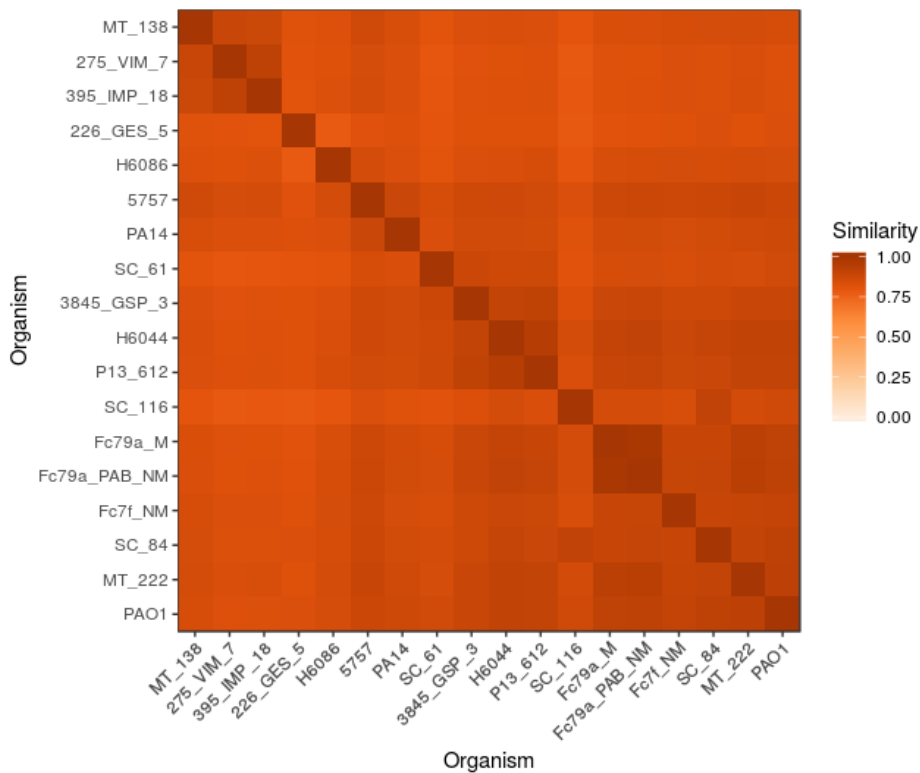


Figura 9: Comparação dos genomas dos 18 isolados com base na frequência de k-mers (k=5). A escala de similaridade está indicada na barra ao lado.

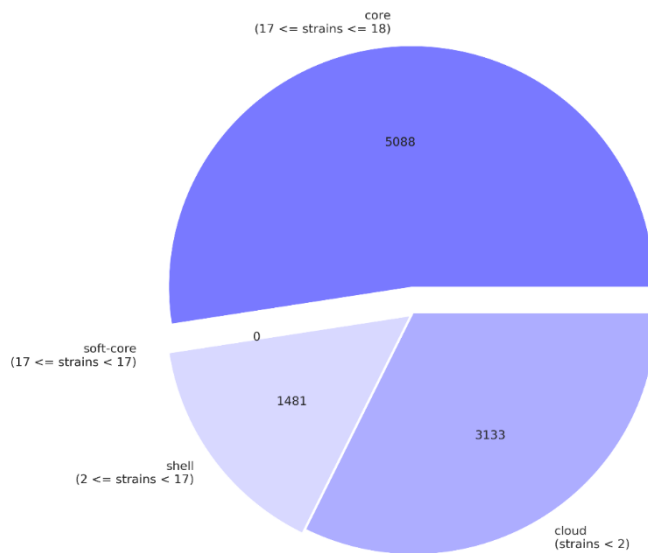


Figura 10: Análise do pan-genoma dos 18 isolados sequenciados. Estão indicados o número de genes do *core*-genoma e do genoma acessório (*cloud* e *shell*) como resultado da análise com o programa ROARY.

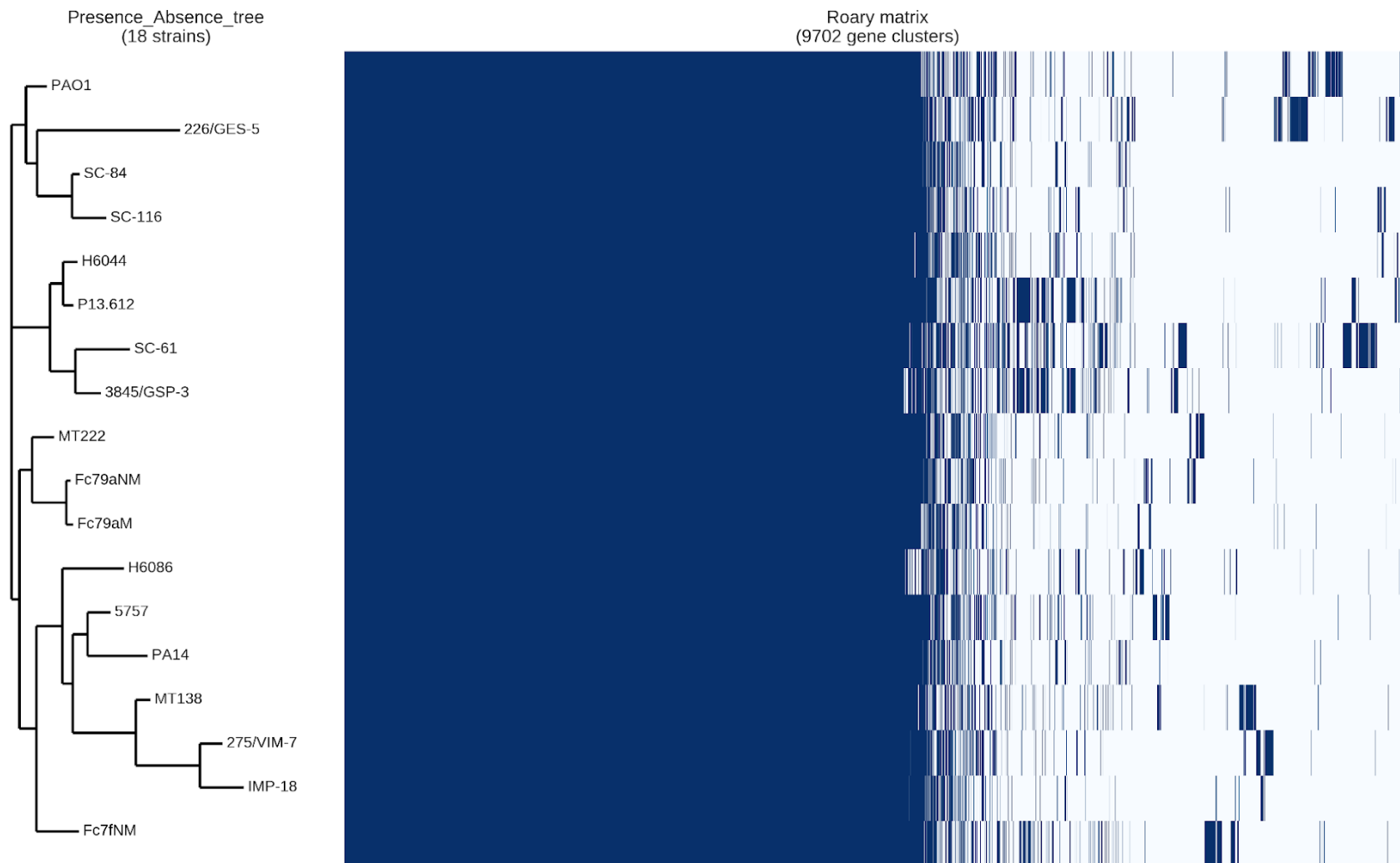


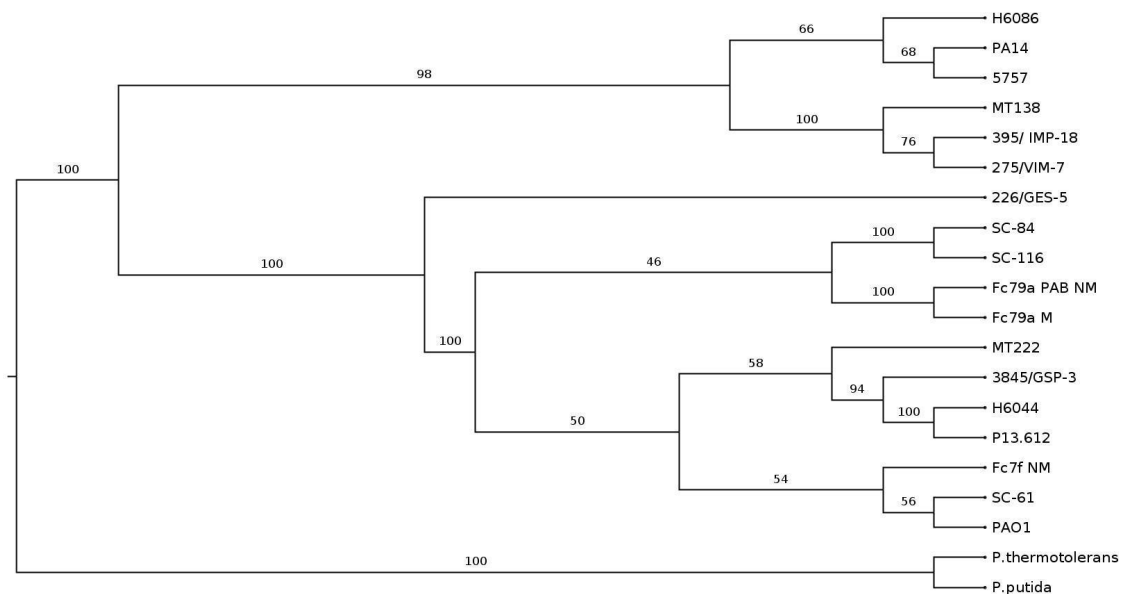
Figura 11: Matriz de presença e ausência de genes nos 18 isolados sequenciados. Os genes com identidade $\geq 95\%$ estão representados em azul. As linhas indicam os 18 genomas sequenciados, conforme indicado na árvore que é gerada pelo programa ROARY.

Tabela 7: Classificação dos isolados de acordo com o ST e descrição dos alelos de MLST.

Isolado	ST	CC	acs	aro	gua	mut	nuo	pps	trp
P1	170		36	5	29	7	4	10	7
P2	549	PAO1	7	5	12	3	4	1	7
P3	2524		17	5	11	72	3	4	3
P4	274		23	5	11	7	1	12	7
P5	2734		6	5	5	3	3	6	19
P6	253	PA14	4	4	16	12	1	6	3
P7	3038		16	5	36	11	4	28	254
P8	?		125	3	12	3	1	135	9
P9	235,2613		38, 172	11	3	13	1	2	4
P10	309		13	8	9	3	1	17	15
P11	2734		6	5	5	3	3	6	19
P12	2524		17	5	11	72	3	4	3
P13	803		15	5	11	3	4	4	7
P14	235,2613		38, 172	11	3	13	1	2	4
P15	?		?	?	?	?	?	?	?
P16	235,2613		172, 38	11	3	13	1	2	4
P17	274		23	5	11	7	1	12	7
P18	316		13	8	9	3	1	6	9

Isolado	Suscetibilidade aos fagos		
	ZC01	ZC03	ZC08
SC-61	+	-/+	+
PAO1	-	-	-
SC-116	-	-	-
<i>P13.612</i>	+	-	-
<i>Fc79a M</i>	+	-	+
PA14	+	+	+
<i>Fc7f NM</i>	+	+	+
<i>MT222</i>	+	-	-
<i>MT138</i>	+	-	-
<i>H6086</i>	+	-/+	-/+
<i>Fc79a PAB NM</i>	+	+	+
SC-84	-	-	-
3845/GSP-3	+	-	+
<i>ALERTA 275/VIM-7</i>	-/+	-/+	-/+
<i>ALERTA 226/GES-5</i>	-/+	-/+	-/+
<i>ALERTA 395/IMP-18</i>	+	-	+
<i>H6044</i>	+	+	+
<i>5757</i>	-/+	-/+	-

Nota: A suscetibilidade dos isolados aos fagos ZC01, ZC03 e ZC08 está mostrada na tabela a direita (dados reproduzidos da Tabela 5).



Isolado	Suscetibilidade aos fagos		
	ZC01	ZC03	ZC08
<i>H6086</i>	+	-/+	-/+
PA14	+	+	+
5757	-/+	-/+	-
MT138	+	-	-
ALERTA 395/IMP-18	+	-	+
ALERTA 275/VIM-7	-/+	-/+	-/+
ALERTA 226/GES-5	-/+	-/+	-/+
SC-84	-	-	-
SC-116	-	-	-
<i>Fc79a PAB NM</i>	+	+	+
<i>Fc79a M</i>	+	-	+
MT222	+	-	-
3845/GSP-3	+	-	+
P13.612	+	-	-
H6044	+	+	+
<i>Fc7f NM</i>	+	+	+
SC-61	+	-/+	+
PAO1	-	-	-

Figura 12: Árvore filogenética gerada a partir do core-genoma dos 18 isolados sequenciados mais as espécies *P. putida* e *P. thermotolerans*. Para o cálculo da árvore foi utilizada o método de máxima verossimilhança do programa RAxML. A suscetibilidade dos isolados aos fagos ZC01, ZC03 e ZC08 está mostrada na tabela a direita (dados reproduzidos da Tabela 5).

4.2.3. Análise da relação à suscetibilidade por ZC01, ZC03 e ZC08 e variantes de PilE e PilA dos 18 isolados

Como já mencionado, o *pilus* do tipo 4 (T4P) parece ser um receptor para os fagos ZC01, ZC03 e ZC08. Dessa forma, investigamos diferenças nas sequências das pilinas e se tais diferenças se relacionam ou não com a suscetibilidade aos fagos. A pilina PilA é o componente majoritário da fibra do T4P enquanto que a pilina minoritária PilE participa da montagem do T4P (Burrows 2012; Nguyen et al. 2015; Harvey et al. 2018).

O alinhamento múltiplo da sequência de PilE dos 18 isolados mostra que esta pilina é bastante similar (entre 99% e 100%) entre Fc79aM, 395/IMP-18, MT138, H6086, Fc79a PAB NM, 275/VIM-7, MT222, PAO1, SC-116, SC84 e 226/GES-5 mas difere em SC-61, 3845/GSP3, H6044, PA14 e P13.612 que, por sua vez, compartilham sequência de PilE entre 92% e 100% similar. Já o isolado P13.612 apresentou PilE com sequência muito similar às citadas acima mas com uma porção da região N-terminal ausente (92% de similaridade), como pode ser observado no alinhamento apresentado na Figura 13. Essas sequências de PilE foram utilizadas para gerar uma árvore filogenética visando observar se há relação entre a sequência de PilE e a capacidade de infecção pelos fagos (Figura 14). Novamente, não encontramos relação direta entre suscetibilidade aos fagos e tipo de sequência primária de PilE.

A Figura 15 mostra o alinhamento múltiplo de PilA indicando que essa proteína é muito variável no genoma dos isolados de *P. aeruginosa* sequenciados. Estas observações estão de acordo com a literatura que reporta que *P. aeruginosa* tem diferentes classes de PilA (Kus et al. 2004; Burrows 2012; Harvey et al. 2018).

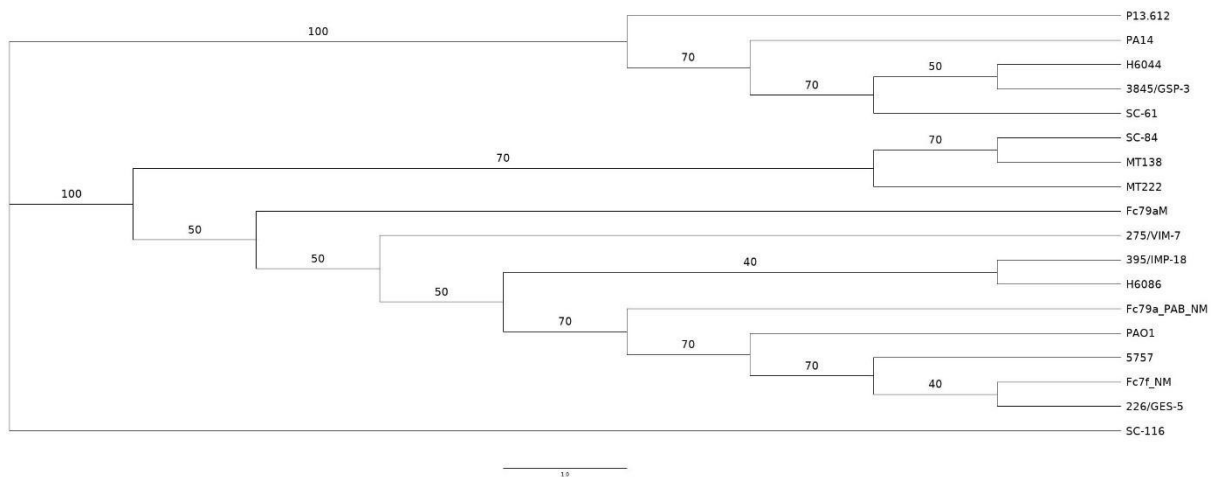
As sequências de PilA foram utilizadas para construção de uma árvore filogenética (Figura 16) e a partir desses resultados buscamos uma possível relação entre estrutura primária da proteína e suscetibilidade aos fagos ZC01, ZC03 e ZC08. Porém a relação não é perfeita, existem agrupamentos de PilA onde uma estrutura similar correlaciona com suscetibilidade em um isolado e resistência em outro. Mas de qualquer modo, esta análise se mostrou mais promissora do que as anteriores, indicando que determinada sequência de PilA pode indicar suscetibilidade aos fagos.


```

Fc79aM      MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
395/_IMP-18 MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
MT138      MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
H6086     MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
Fc79a_PAB_NM MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
5757     MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
SC-84     MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
Fc7f_NM   MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
275/_VIM-7 MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
MT222     MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
PA01     MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
226/GES-5 MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
SC-116    MRTRQKGFLLLEMVVVAVIGILLGIAIPSYQNYVIRSNRTEGQALLSDAAARQERYYSQ 60
SC-61     MKS-NRGFTLIELMIVVVI IAILAGIAYPSYDEYVVRGNRTEGQALLSEAAATQERYYSQ 60
3845/GSP-3 MKS-NRGFTLIELMIVVVI IAILAGIAYPSYDEYVVRGNRTEGQALLSEAAATQERYYSQ 60
H6044     MKS-NRGFTLIELMIVVVI IAILAGIAYPSYDEYVVRGNRTEGQALLSEAAATQERYYSQ 60
PA14     MKS-NRGFTLIELMIVVVI IAILAGIAYPSYDEYVVRGNRTEGQALLSEAAATQERYYSQ 60
P13.612   -----MIVVVI IAILAGIAYPSYDEYVVRGNRTEGQALLSEAAATQERYYSQ 60
          :*:*. ** * ** *::* *.*****:*** ***:**
Fc79aM      NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
395/_IMP-18 NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
MT138      NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
H6086     NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
Fc79a_PAB_NM NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
5757     NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
SC-84     NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
Fc7f_NM   NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
275/_VIM-7 NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
MT222     NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
PA01     NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
226/GES-5 NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
SC-116    NPGVGYTK-DVAKLGM-----SSANSPNNLYNLT IAT-PTSTTYTLTATPINSQTRDKTC 120
SC-61     NNTYITTQADIGKLMHRNTSGTTVKSSSTGKYSLTVDTVANDGGYRLIA---NQAFNDLDC 120
3845/GSP-3 NNTYITTQADIGKLMHRNTSGTTVKSSSTGKYSLTVDTVANDGGYRLIA---NQAFNDLDC 120
H6044     NNTYITTQADIGKLMHRNTSGTTVKSSSTGKYSLTVDTVANDGGYRLIA---NQAFNDLDC 120
PA14     NNTYITTQADIGKLMHRNTSGTTVKSSSTGKYSLTVDTVANDGGYRLIA---NQAFNDLDC 120
P13.612   NNTYITTQADIGKLMHRNTSGTTVKSSSTGKYSLTVDTVANDGGYRLIA---NQAFNDLDC 120
          *   * :*. ** *   :*:*. ** * :*: * ... * * * * . . * *
Fc79aM      GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
395/_IMP-18 GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
MT138      GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
H6086     GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
Fc79a_PAB_NM GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
5757     GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
SC-84     GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
Fc7f_NM   GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
275/_VIM-7 GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
MT222     GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
PA01     GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
226/GES-5 GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
SC-116    GKLTNLQLGERGAAGKTGNNSTVNDCWR 148
SC-61     GNLTLTANGEK---GRTGSKKSVAECWR 148
3845/GSP-3 GNLTLTANGEK---GRTGSKKSVAECWR 148
H6044     GNLTLTANGEK---GRTGSKKSVAECWR 148
PA14     GNLTLTANGEK---GRTGSKKSVAECWR 148
P13.612   GNLTLTANGEK---GRTGSKKSVAECWR 148
          *:***. ** : * :*:*. ** * :*:***

```

Figura 13: Alinhamento múltiplo das sequências de Pile dos 18 isolados. As sequências de Pile foram recuperadas do genoma de cada isolado listado nas Tabelas 5 e 6.



Isolado	Suscetibilidade aos fagos		
	ZC01	ZC03	ZC08
<i>P13.612</i>	+	-	-
PA14	+	+	+
H6044	+	+	+
3845/GSP-3	+	-	+
SC-61	+	-/+	+
SC-84	-	-	-
MT138	+	-	-
MT222	+	-	-
<i>Fc79a M</i>	+	-	+
ALERTA 275/VIM-7	-/+	-/+	-/+
ALERTA 395/IMP-18	+	-	+
H6086	+	-/+	-/+
<i>Fc79a PAB NM</i>	+	+	+
PAO1	-	-	-
5757	-/+	-/+	-
<i>Fc7f NM</i>	+	+	+
ALERTA 226/GES-5	-/+	-/+	-/+
SC-116	-	-	-

Figura 14: Árvore filogenética gerada a partir das P_{ilE} das 18 cepas. Para o cálculo da árvore foi utilizada a técnica de máxima verossimilhança com o programa RAXML. A suscetibilidade dos isolados aos fagos ZC01, ZC03 e ZC08 está mostrada na tabela (dados reproduzidos da Tabela 5).

4.2.4. Predição de sítios de glicosilação na PilA e sorotipos dos 18 isolados

O mecanismo de interação do fago com o T4P (ou com a pilina majoritária PilA) ainda é desconhecido, porém sabe-se que a glicosilação de PilA é um mecanismo de defesa da bactéria à infecção por fagos. Entretanto, nem todos grupos de PilA são alvo de glicosilação, como é o caso da PilA do grupo 2 (Harvey *et al.*, 2018). Dessa forma, fizemos uma predição dos sítios de glicosilação de PilA e buscamos alguma relação com os fenótipos de resistência à infecção exibido por alguns dos isolados sequenciados. Como mostrado na Tabela 8 não há relação clara entre o tipo de PilA, sítios de glicosilação e resistência à infecção pelos fagos, visto que um mesmo grupo de PilA está associado tanto a isolados suscetíveis como resistentes.

A interação dos fagos a uma determinada bactéria depende, em parte dos casos, de mais de um determinante de adsorção (Silva *et al.*, 2016). Uma mesma cepa pode possuir receptores diferentes para interação dos fagos, sendo eles reversíveis (ou intermediários) e irreversíveis, onde a ligação no determinante irreversível depende de uma ligação com receptor intermediário/reversível. Segundo a base de dados Phage Receptor Database (PhReD)⁷, o principal receptor irreversível observado em *P. aeruginosa* é o T4P e que LPS atua como um receptor intermediário. Como já mencionado, LPS é um receptor frequente na interação de fagos, além do T4P (Lam *et al.*, 2011).

O LPS possui três principais domínios que compõem sua estrutura final: lipídio A, oligossacarídeo central (*core*) e antígeno O (Lam *et al.*, 2011). O antígeno O se estende para fora da membrana externa de bactérias Gram negativas e é altamente variável, resultando em 20 sorotipos descritos para *P. aeruginosa*. Cada sorotipo é dependente de um *cluster* de genes flanqueado pelos genes *himD* e *wbpM*, para ser formado. Esses clusters são altamente variáveis no genoma de *P. aeruginosa* e podem ser divididos em 11 padrões de *clusters* gênicos (Raymond *et al.*, 2002; Lam *et al.*, 2011).

Para comparar essas regiões nos genomas sequenciados fizemos buscas em bases de dados a fim de encontrar os dois genes (*himD* e *wbpM*), os quais

⁷ <https://phred.herokuapp.com/>

são muito conservados. Os *clusters* gênicos dos 18 genomas que sequenciamos foram identificados e foi realizado um alinhamento múltiplo com os *clusters* dos 20 sorotipos referência (O1 a O20) recuperados do NCBI (Raymond *et al.*, 2002). As similaridades observadas para os *clusters* das 18 cepas com as sequências de referências estão listadas na Tabela 9, indicando o sorotipo predito.

A cepa PAO1 que é resistente a infecção apresenta um sorotipo bastante distinto das demais, e como já mencionado é resistente aos 3 fagos. Nos demais casos não foi possível identificar uma relação precisa dos diferentes sorotipos isoladamente com a suscetibilidade aos fagos.

Em resumo, a análise das características gerais dos genomas dos 18 isolados e de seus prováveis determinantes de interação não evidenciou individualmente relações diretas com suscetibilidade ou resistência aos fagos em estudo, apontadas nos ensaios de infecção. Os dados apresentados sugerem que um único fator não é capaz de relacionar satisfatoriamente a capacidade de infecção dos fagos com o genoma dos isolados. Estes resultados serão melhor compreendidos quando realizarmos ensaios que avaliem com mais precisão a eficiência de adsorção dos fagos nos diferentes isolados.

Além disso, outros determinantes da interação, como proteínas da membrana externa (OMPs) que atuam como receptores irreversíveis da *central tail spike* de alguns fagos (Nobrega *et al.*, 2018), serão futuramente investigados.

Tabela 8: Predição dos sítios de glicosilação na PilA dos 18 isolados de *P. aeruginosa*.

ID	Isolado	Grupo da Pilina PilA	Sítios de Glicosilação Preditos								Suscetibilidade aos fagos			
											ZC01	ZC03	ZC08	
J5	Fc79aM	2										+	-	+
J11	Fc79a PAB	2										+	+	+
J16	395 IMP 18	2										+	-	+
J9	MT 138	2										+	-	-
J7	Fc7f NM	2										+	+	+
J14	275/VIM7	2										-/+	-/+	-/+
J8	MT222	2										+	-	-
J12	SC-84	2										-	-	-
J10	H6086	2										+	-/+	-/+
J3	SC-116	2										-	-	-
J2	PAO1	2a	32	47	117	125	149					-	-	-
J18	5757	2a	32	47	65	85	117					-/+	-/+	-
J1	SC-61	3	71	76	101	104	107	111	122	151		+	-/+	+
J13	3845 GSP3	3	72	76	77	94	114	123	126	152		+	-	+
J17	H6044	3	72	76	77	94	114	123	126	152		+	+	+
J4	P13.612	3	72	76	77	94	114	123	126	152		+	-	-
J6	PA14	3	71	76	101	104	107	122	151			+	+	+
J15	226/GES-5	1b	66	71	141	145	150					-/+	-/+	-/+

Tabela 9: Predição dos tipos de LPS nos 18 isolados de *P. aeruginosa*.

ID	Isolado	Grupo da Pilina PilA	% similaridade / Sorotipo predito	Suscetibilidade aos fagos		
				ZC01	ZC03	ZC08
J15	226/GES-5	1b	88% O15	-/+	-/+	-/+
J5	Fc79aM	2	97% O1	+	-	+
J11	Fc79a PAB	2	97% O1	+	+	+
J8	MT222	2	97% O1	+	-	-
J6	PA14	3	97% O10, 97% O19	+	+	+
J16	395 IMP 18	2	94% O11, 89% O17	+	-	+
J9	MT 138	2	96% O11, 91% O17	+	-	-
J7	Fc7f NM	2	97% O11, 91% O17	+	+	+
J14	275/VIM7	2	96% O11, 91% O17	-/+	-/+	-/+
J12	SC-84	2	97% O11, 91% O17	-	-	-
J10	H6086	2	96% O11, 91% O17	+	-/+	-/+
J3	SC-116	2	77% O11, 76% O17	-	-	-
J18	5757	2a	96% O11, 91% O17	-/+	-/+	-
J1	SC-61	3	85% O15, 94% O3	+	-/+	+
J13	3845 GSP3	3	88% O15, 96% O3	+	-	+
J17	H6044	3	88% O15, 97% O3	+	+	+
J4	P13.612	3	88% O15, 97% O3	+	-	-
J2	PAO1	2a	96% O16, 97% O18, 97% O2, 93% O20, 97% O5	-	-	-

5. Conclusões

A partir das análises apresentadas nesse trabalho é possível concluir que os fagos de *Pseudomonas aeruginosa* estudados nesse trabalho (ZC03, ZC04, ZC06, ZC07 e ZC08 da família *Podoviridae*), mesmo que muito próximos filogeneticamente e em nível de similaridade de genoma, apresentam diferenças importantes que podem levar a alterações em seus ciclos de vida. A análise do proteoma predito também mostrou importantes diferenças presentes nesses fagos que justificam a realização de futuros estudos para compreendermos em maior detalhe sua estrutura, tais como investigar se os vírions de ZC03 e de ZC08, de fato, expressam fibras da cauda distintas, isto é, com ambas as proteínas codificadas pelas ORFs _00003 e _00004.

Os resultados obtidos indicam que capacidade de infecção dos fagos ZC03 e ZC08 é oposta para alguns isolados de *P. aeruginosa*, como sumarizado na Figura 17. Postulamos que as diferenças encontradas nas ORFs preditas como proteínas da fibra da cauda (_00004 e _00005) podem estar relacionadas a efetividade da interação irreversível do fago ao receptor bacteriano. Como já mencionado, a adsorção bem-sucedida entre fago e a superfície do hospedeiro ocorre em duas etapas, primeira uma interação reversível e em seguida uma interação irreversível, as quais podem acontecer com um mesmo receptor do hospedeiro (Silva *et al.*, 2016). Assim, nossa hipótese é que proteínas da fibra da cauda diferentes interagem de modo distinto (reversível ou irreversível) com um mesmo receptor em uma determinada cepa de *P. aeruginosa*. Além disso, as diferenças observadas nas ORFs referentes a *central tail spike* devem também contribuir para explicar as diferenças observadas nos ensaios de infecção de ZC03 e ZC08 contra certas cepas avaliadas (Figura 17).

Por outro lado, não observamos uma relação perfeita entre capacidade de infecção tanto pelos fagos ZC03 e ZC08 (família *Podoviridae*) como pelo fago ZC01 (família *Siphoviridae*) e as diferentes estruturas primárias da pilina majoritária (PilA) do T4P identificada nos genomas sequenciados dos 18 isolados de *P. aeruginosa* (Figuras 17). Já foi mostrado que diferenças na sequência primária de pilinas de um mesmo grupo interferem na interação com estruturas do fago (Kim *et al.*, 2018). Tão pouco identificamos relações entre a suscetibilidade ou resistência à infecção com os distintos padrões de glicosilação

da PilA ou distintos sorotipos (tipo de antígeno O) preditos para os 18 isolados de *P. aeruginosa*.

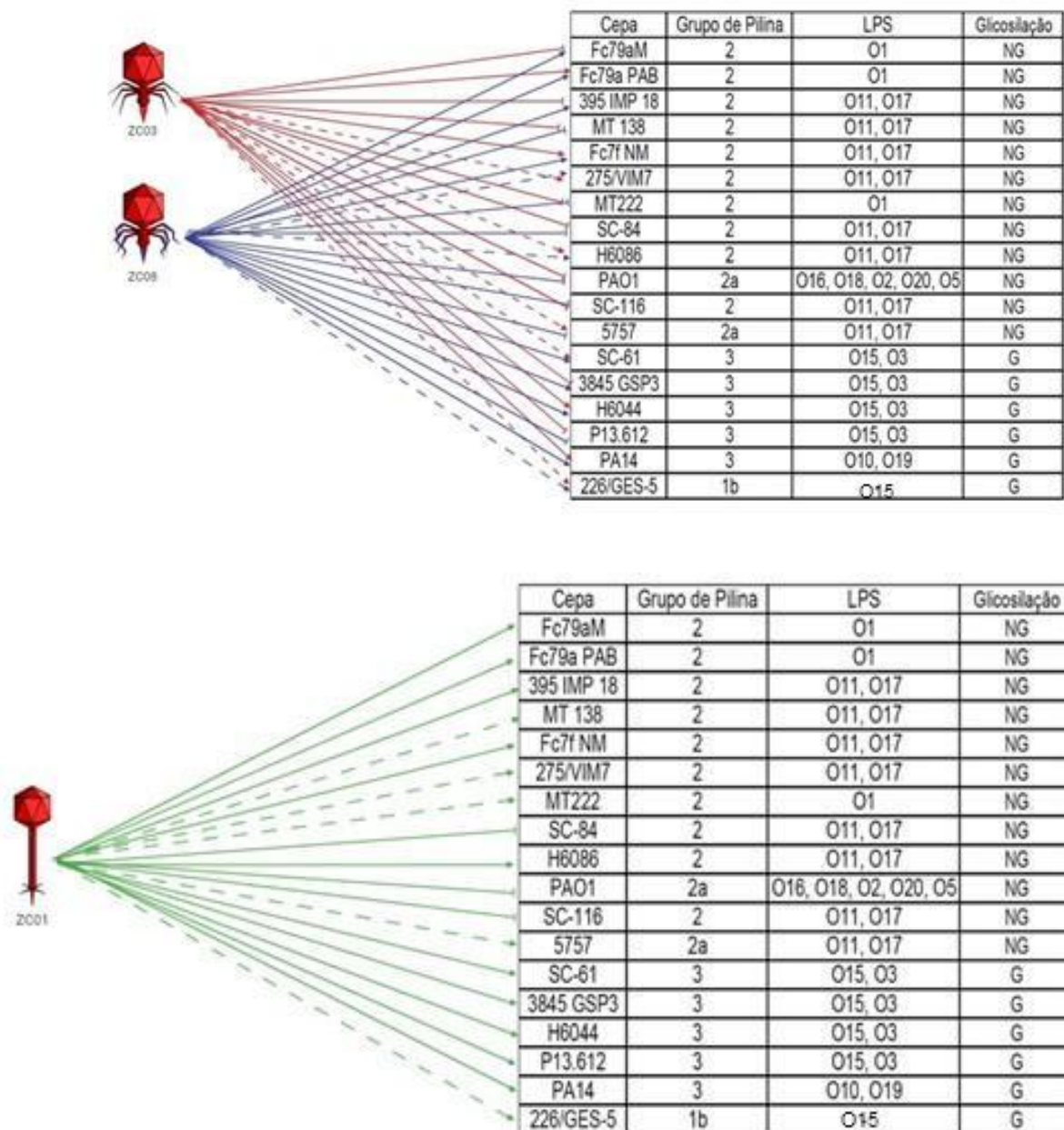


Figura 17: Sumário das interações dos fagos ZC01 (verde), ZC03 (vermelho) e ZC08 (azul) com os 18 isolados clínicos de *P. aeruginosa*. Suscetibilidade à infecção está indicada por linhas sólidas. Resistência à infecção está indicada por linhas sólidas com uma barra na sua extremidade. Linhas tracejadas indicam capacidade de adsorção e/ou lise de fora para dentro (sem multiplicação do fago), ou seja, placa de lise túrbida. As tabelas indicam as predições do grupo da PilA, tipo do antígeno do O do LPS e ausência (NG) ou presença (G) de glicosilação de PilA. Para a confecção dos esquemas foram utilizadas as informações das Tabelas 5, 8 e 9.

Nossas observações indicam, como esperado, que a relação entre capacidade de infecção e estruturas ligantes dos fagos está associada a diferentes interações com estruturas distintas na superfície bacteriana. Dessa forma é plausível supor que os isolados de *P. aeruginosa* estudados apresentam duas etapas na interação, talvez uma etapa que pode envolver a interação com LPS (antígeno O) em um contato reversível, e a segunda dependente de uma ligação irreversível das proteínas de cauda dos fagos com o T4P. A etapa da interação com LPS parece, entretanto, não ser um requerimento absoluto, ao contrário da interação com PilA, visto que a expressão de PilA da cepa PA14 confere suscetibilidade aos fagos para cepa PAO1, a qual é resistente aos três fagos. O conjunto de nossos resultados não apontaram para um padrão único de interação, e dessa forma, é muito provável que diversas estruturas tanto dos fagos como das bactérias apresentam papel importante na interação.

Ensaio de infecção em cepas mutantes de PilA e/ou que não produzem antígeno O, além de ensaios para avaliar com maior precisão a eficácia da adsorção dos fagos a alguns dos isolados clínicos cujos genomas foram sequenciados, são necessários para verificar as hipóteses levantadas neste trabalho.

6. Referências

1. **Ackermann, H. W. & Prangishvili, D.** 2012. Prokaryote viruses studied by electron microscopy. *Archives of Virology* 157:1843-1849.
2. **Altamirano, F. L. G. & Barr, J. J.** 2019. Phage Therapy in the Postantibiotic Era. *Clinical Microbiology Reviews* 32.
3. **Amgarten, D., Martins, L. F., Lombardi, K. C., Antunes, L. P., de Souza, A. P. S., Nicastro, G. G., Kitajima, E. W., Quaggio, R. B., Upton, C., Setubal, J. C. & da Silva, A. M.** 2017. Three novel *Pseudomonas* phages isolated from composting provide insights into the evolution and diversity of tailed phages. *BMC Genomics* 18:ARTN 346.
4. **Azam, A. H. & Tanji, Y.** 2019. Bacteriophage-host arm race: an update on the mechanism of phage resistance in bacteria and revenge of the phage with the perspective for phage therapy. *Applied Microbiology and Biotechnology* 103:2121-2131.
5. **Bae, H. W. & Cho, Y. H.** 2013. Complete Genome Sequence of *Pseudomonas aeruginosa* Podophage MPK7, Which Requires Type IV Pili for Infection. *Genome Announc* 1.
6. **Bondy-Denomy, J., Qian, J., Westra, E. R., Buckling, A., Guttman, D. S., Davidson, A. R. & Maxwell, K. L.** 2016. Prophages mediate defense against phage infection through diverse mechanisms. *Isme Journal* 10:2854-2866.
7. **Bradley, D. E. & Pitt, T. L.** 1974. Pilus-Dependence of 4 *Pseudomonas-Aeruginosa* Bacteriophages with Non-Contractile Tails. *Journal of General Virology* 24:1-15.
8. **Braga, L. P. P., Soucy, S. M., Amgarten, D. E., da Silva, A. M. & Setubal, J. C.** 2018. Bacterial Diversification in the Light of the Interactions with Phages: The Genetic Symbionts and Their Role in Ecological Speciation. *Frontiers in Ecology and Evolution* 6.
9. **Brussow, H.** 2012. What is needed for phage therapy to become a reality in Western medicine? *Virology* 434:138-142.
10. **Burrows, L. L.** 2012. *Pseudomonas aeruginosa* Twitching Motility: Type IV Pili in Action. *Annual Review of Microbiology*, Vol 66 66:493-520.
11. **Burt, A. & Koufopanou, V.** 2004. Homing endonuclease genes: the rise and fall and rise again of a selfish element. *Current Opinion in Genetics & Development* 14:609-615.
12. **Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A.** 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28:464-9.
13. **Chaturongakul, S. & Ounjai, P.** 2014. Phage-host interplay: examples from tailed phages and Gram-negative bacterial pathogens. *Frontiers in Microbiology* 5:ARTN 442.
14. **Chevreur, B., Pfisterer, T., Drescher, B., Driesel, A. J., Muller, W. E. G., Wetter, T. & Suhai, S.** 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Research* 14:1147-1159.
15. **Chibeu, A., Ceysens, P. J., Hertveldt, K., Volckaert, G., Cornelis, P., Matthijs, S. & Lavigne, R.** 2009. The adsorption of *Pseudomonas aeruginosa* bacteriophage phiKMV is dependent on expression regulation of type IV pili genes. *FEMS Microbiol Lett* 296:210-8.
16. **Cisek, A. A., Dabrowska, I., Gregorczyk, K. P. & Wyzewski, Z.** 2017. Phage Therapy in Bacterial Infections Treatment: One Hundred Years After the Discovery of Bacteriophages. *Current Microbiology* 74:277-283.
17. **Clark, J. R. & March, J. B.** 2006. Bacteriophages and biotechnology: vaccines, gene therapy and antibacterials. *Trends in Biotechnology* 24:212-218.

18. **Clokic, M. R. J., Millard, A. D., Letarov, A. V. & Heaphy, S.** 2011. Phages in nature. *Bacteriophage* 1:31-45.
19. **Colavecchio, A., Cadieux, B., Lo, A. & Goodridge, L. D.** 2017. Bacteriophages Contribute to the Spread of Antibiotic Resistance Genes among Foodborne Pathogens of the Enterobacteriaceae Family - A Review. *Frontiers in Microbiology* 8.
20. **de Jonge, P. A., Nobrega, F. L., Brouns, S. J. J. & Dutilh, B. E.** 2019. Molecular and Evolutionary Determinants of Bacteriophage Host Range. *Trends in Microbiology* 27:51-63.
21. **Dedrick, R. M., Guerrero-Bustamante, C. A., Garlena, R. A., Russell, D. A., Ford, K., Harris, K., Gilmour, K. C., Soothill, J., Jacobs-Sera, D., Schooley, R. T., Hatfull, G. F. & Spencer, H.** 2019. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nature Medicine* 25:730-+.
22. **Doss, J., Culbertson, K., Hahn, D., Camacho, J. & Berekzi, N.** 2017. A Review of Phage Therapy against Bacterial Pathogens of Aquatic and Terrestrial Organisms. *Viruses-Basel* 9.
23. **Edgell, D. R., Gibb, E. A. & Belfort, M.** 2010. Mobile DNA elements in T4 and related phages. *Virol J* 7:290.
24. **Edwards, R. A. & Rohwer, F.** 2005. Viral metagenomics. *Nature Reviews Microbiology* 3:504-510.
25. **Enault, F., Briet, A., Bouteille, L., Roux, S., Sullivan, M. B. & Petit, M. A.** 2017. Phages rarely encode antibiotic resistance genes: a cautionary tale for virome analyses. *Isme Journal* 11:237-247.
26. **Erez, Z., Steinberger-Levy, I., Shamir, M., Doron, S., Stokar-Avihail, A., Peleg, Y., Melamed, S., Leavitt, A., Savidor, A., Albeck, S., Amitai, G. & Sorek, R.** 2017. Communication between viruses guides lysis-lysogeny decisions. *Nature* 541:488.
27. **Ford, B. E., Baloh, M. & Dennehy, J. J.** 2018. Evolutionary Ecology of the Viruses of Microorganisms, p. 53-76. *In* P. Hyman & S. T. Abedon (ed.), *Viruses of Microorganisms*. Caister Academic Press, Norfolk, UK.
28. **Galetti, R., Andrade, L. N. & Darini, A. L. D.** 2015. *Pseudomonas aeruginosa* carrying blaCTX-M-2 Brazil: The occurrence of 'high-risk clones'? *Journal of Global Antimicrobial Resistance* 3:153-154.
29. **Gandon, S.** 2016. Why Be Temperate: Lessons from Bacteriophage lambda. *Trends in Microbiology* 24:356-365.
30. **Garbe, J., Bunk, B., Rohde, M. & Schobert, M.** 2011. Sequencing and Characterization of *Pseudomonas aeruginosa* phage JG004. *Bmc Microbiology* 11:Artn 102.
31. **Gogarten, J. P. & Hilario, E.** 2006. Inteins, introns, and homing endonucleases: recent revelations about the life cycle of parasitic genetic elements. *BMC evolutionary biology* 6:94-94.
32. **Gore, J.** 2018. Simple organizing principles in microbial communities. *Current Opinion in Microbiology* 45:195-202.
33. **Gorski, A., Miedzybrodzki, R., Lobočka, M., Glowacka-Rutkowska, A., Bednarek, A., Borysowski, J., Jonczyk-Matysiak, E., Lusiak-Szelachowska, M., Weber-Dabrowska, B., Baginska, N., Letkiewicz, S., Dabrowska, K. & Scheres, J.** 2018. Phage Therapy: What Have We Learned? *Viruses-Basel* 10.
34. **Haq, I. U., Chaudhry, W. N., Akhtar, M. N., Andleeb, S. & Qadri, I.** 2012. Bacteriophages and their implications on future biotechnology: a review. *Virology Journal* 9:Artn 9.
35. **Harvey, H., Bondy-Denomy, J., Marquis, H., Sztanko, K. M., Davidson, A. R. & Burrows, L. L.** 2018. *Pseudomonas aeruginosa* defends against phages through type IV pilus glycosylation. *Nature Microbiology* 3.

36. **Hatfull, G. F. & Hendrix, R. W.** 2011. Bacteriophages and their genomes. *Current Opinion in Virology* 1:298-303.
37. **Hayes, S., Mahony, J., Nauta, A. & van Sinderen, D.** 2017. Metagenomic Approaches to Assess Bacteriophages in Various Environmental Niches. *Viruses-Basel* 9.
38. **He, J. X., Baldini, R. L., Deziel, E., Saucier, M., Zhang, Q. H., Liberati, N. T., Lee, D., Urbach, J., Goodman, H. M. & Rahme, L. G.** 2004. The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proceedings of the National Academy of Sciences of the United States of America* 101:2530-2535.
39. **Hendrix, R. W.** 2003. Bacteriophage genomics. *Current Opinion in Microbiology* 6:506-511.
40. **Holloway, B. W.** 1955. Genetic recombination in *Pseudomonas aeruginosa*. *J Gen Microbiol* 13:572-81.
41. **Howard-Varona, C., Hargreaves, K. R., Abedon, S. T. & Sullivan, M. B.** 2017. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *Isme Journal* 11:1511-1520.
42. **Howard-Varona, C., Hargreaves, K. R., Solonenko, N. E., Markillie, L. M., White, R. A., Brewer, H. M., Ansong, C., Orr, G., Adkins, J. N. & Sullivan, M. B.** 2018. Multiple mechanisms drive phage infection efficiency in nearly identical hosts. *Isme Journal* 12:1605-1618.
43. **Hyman, P. & Abedon, S. T.** 2018. *Viruses of Microorganisms*. Caister Academic Press, Norfolk, UK.
44. **James, C. E., Fothergill, J. L., Kalwij, H., Hall, A. J., Cottell, J., Brockhurst, M. A. & Winstanley, C.** 2012. Differential infection properties of three inducible prophages from an epidemic strain of *Pseudomonas aeruginosa*. *Bmc Microbiology* 12.
45. **Jeukens, J., Freschi, L., Kukavica-Ibrulj, I., Emond-Rheault, J. G., Tucker, N. P. & Levesque, R. C.** 2019. Genomics of antibiotic-resistance prediction in *Pseudomonas aeruginosa*. *Annals of the New York Academy of Sciences* 1435:5-17.
46. **Jia, B. F., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P. Y. et al.** 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research* 45:D566-D573.
47. **Katoh, K. & Standley, D. M.** 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30:772-780.
48. **Kim, E. S., Bae, H. W. & Cho, Y. H.** 2018. A Pilin Region Affecting Host Range of the *Pseudomonas aeruginosa* RNA Phage, PP7. *Frontiers in Microbiology* 9.
49. **King, A. M. Q., Adams, M. J., Carstens, E. B. & Lefkowitz, E. J.** 2012. Order - Caudovirales, p. 39-45. *In* A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz (ed.), *Virus Taxonomy*. Elsevier, San Diego.
50. **Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C. & Salzberg, S. L.** 2004. Versatile and open software for comparing large genomes. *Genome Biol* 5:R12.
51. **Kus, J. V., Tullis, E., Cvitkovitch, D. G. & Burrows, L. L.** 2004. Significant differences in type IV pilin allele distribution among *Pseudomonas aeruginosa* isolates from cystic fibrosis (CF) versus non-CIF patients. *Microbiology-Sgm* 150:1315-1326.
52. **Kutter, E. & Sulakvelidze, A.** 2005. *Bacteriophages: Biology and Applications*. CRC Press, Boca Raton, FL.
53. **Labrie, S. J., Samson, J. E. & Moineau, S.** 2010. Bacteriophage resistance mechanisms. *Nature Reviews Microbiology* 8:317-327.

54. **Lam, J. S., Taylor, V. L., Islam, S. T., Hao, Y. A. & Kocincova, D.** 2011. Genetic and functional diversity of *Pseudomonas aeruginosa* lipopolysaccharide. *Frontiers in Microbiology* 2.
55. **Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G.** 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947-8.
56. **Le, S., He, X., Tan, Y., Huang, G., Zhang, L., Lux, R., Shi, W. & Hu, F.** 2013. Mapping the tail fiber as the receptor binding protein responsible for differential host specificity of *Pseudomonas aeruginosa* bacteriophages PaP1 and JG004. *PLoS One* 8:e68562.
57. **Lee, D. G., Urbach, J. M., Wu, G., Liberati, N. T., Feinbaum, R. L., Miyata, S., Diggins, L. T., He, J. X., Saucier, M., Deziel, E., Friedman, L., Li, L., Grills, G., Montgomery, K., Kucherlapati, R., Rahme, L. G. & Ausubel, F. M.** 2006. Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biology* 7.
58. **Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G. & Smith, D. B.** 2018. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research* 46:D708-D717.
59. **Lehman, S. M.** 2018. Bacteriophage Diversity, p. 145-165. *In* P. Hyman & S. T. Abedon (ed.), *Viruses of Microorganisms*. Caister Academic Press, Norfolk, UK.
60. **Leiman, P. G., Arisaka, F., van Raaij, M. J., Kostyuchenko, V. A., Aksyuk, A. A., Kanamaru, S. & Rossmann, M. G.** 2010. Morphogenesis of the T4 tail and tail fibers. *Virology* 7:355.
61. **Liu, B. Y., Gu, S. L., Liang, N. S., Xiong, M., Xue, Q. Z., Lu, S. G., Hu, F. Q. & Zhang, H. D.** 2016. *Pseudomonas aeruginosa* phage PaP1 DNA polymerase is an A-family DNA polymerase demonstrating ssDNA and dsDNA 3'→5' exonuclease activity. *Virus Genes* 52:538-551.
62. **Livermore, D. M.** 2002. Multiple mechanisms of antimicrobial resistance in *Pseudomonas aeruginosa*: Our worst nightmare? *Clinical Infectious Diseases* 34:634-640.
63. **Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J. et al.** 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200-D203.
64. **McPartland, J. & Rothman-Denes, L. B.** 2009. The tail sheath of bacteriophage N4 interacts with the *Escherichia coli* receptor. *J Bacteriol* 191:525-32.
65. **Merabishvili, M., Pirnay, J. P., Verbeken, G., Chanishvili, N., Tediashvili, M. et al.** 2009. Quality-Controlled Small-Scale Production of a Well-Defined Bacteriophage Cocktail for Use in Human Clinical Trials. *Plos One* 4.
66. **Michod, R. E., Bernstein, H. & Nedelcu, A. M.** 2008. Adaptive value of sex in microbial pathogens. *Infection Genetics and Evolution* 8:267-285.
67. **Mirzaei, M. K. & Maurice, C. F.** 2017. Menage a trois in the human gut: interactions between host, bacteria and phages. *Nature Reviews Microbiology* 15:397-408.
68. **Moelling, K., Broecker, F. & Willy, C.** 2018. A Wake-Up Call: We Need Phage Therapy Now. *Viruses-Basel* 10.
69. **Morella, N. M., Gomez, A. L., Wang, G., Leung, M. S. & Koskella, B.** 2018. The impact of bacteriophages on phyllosphere bacterial abundance and composition. *Molecular Ecology* 27:2025-2038.
70. **Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S. & Pardesi, K. R.** 2019. Emerging Strategies to Combat ESKAPE Pathogens in the Era of Antimicrobial Resistance: A Review. *Frontiers in Microbiology* 10.
71. **Munita, J. M. & Arias, C. A.** 2016. Mechanisms of Antibiotic Resistance. *Microbiol Spectr* 4.

72. **Nguyen, Y., Harvey, H., Sugiman-Marangos, S., Bell, S. D., Buensuceso, R. N. C., Junop, M. S. & Burrows, L. L.** 2015. Structural and Functional Studies of the *Pseudomonas aeruginosa* Minor Pilin, PilE. *Journal of Biological Chemistry* 290:26856-26865.
73. **Nobrega, F. L., Costa, A. R., Kluskens, L. D. & Azeredo, J.** 2015. Revisiting phage therapy: new applications for old resources. *Trends in Microbiology* 23:185-191.
74. **Nobrega, F. L., Vlot, M., de Jonge, P. A., Dreesens, L. L., Beaumont, H. J. E., Lavigne, R., Dutilh, B. E. & Brouns, S. J. J.** 2018. Targeting mechanisms of tailed bacteriophages. *Nature Reviews Microbiology* 16:760-773.
75. **Okonechnikov, K., Golosova, O., Fursov, M. & Team, U.** 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28:1166-1167.
76. **Oppenheim, A. B., Kobilier, O., Stavans, J., Court, D. L. & Adhya, S.** 2005. Switches in bacteriophage lambda development. *Annu Rev Genet* 39:409-29.
77. **Ormal, A. M. & Jalasvuori, M.** 2013. Phage therapy: Should bacterial resistance to phages be a concern, even in the long run? *Bacteriophage* 3:e24219.
78. **Paez-Espino, D., Eloie-Fadrosh, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N. N. & Kyrpides, N. C.** 2016. Uncovering Earth's virome. *Nature* 536:425-430.
79. **Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., Fookes, M., Falush, D., Keane, J. A. & Parkhill, J.** 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691-3.
80. **Rahme, L. G., Stevens, E. J., Wolfort, S. F., Shao, J., Tompkins, R. G. & Ausubel, F. M.** 1995. Common Virulence Factors for Bacterial Pathogenicity in Plants and Animals. *Science* 268:1899-1902.
81. **Rakhuba, D. V., Kolomiets, E. I., Dey, E. S. & Novik, G. I.** 2010. Bacteriophage Receptors, Mechanisms of Phage Adsorption and Penetration into Host Cell. *Polish Journal of Microbiology* 59:145-155.
82. **Raymond, C. K., Sims, E. H., Kas, A., Spencer, D. H., Kutyavin, T. V., Ivey, R. G., Zhou, Y., Kaul, R., Clendenning, J. B. & Olson, M. V.** 2002. Genetic variation at the O-antigen biosynthetic locus in *Pseudomonas aeruginosa*. *Journal of Bacteriology* 184:3614-3622.
83. **Rohwer, F., Youle, M., Maughan, H. & Hisakawa, N.** 2014. Life in our phage world: a centennial field guide to the Earth's most diverse inhabitants, 1st ed ed, vol. 1. Wholon, San Diego, CA.
84. **Saha, D. & Mukherjee, R.** 2019. Ameliorating the antimicrobial resistance crisis: phage therapy. *Iubmb Life* 71:781-790.
85. **Scholl, D., Rogers, S., Adhya, S. & Merrill, C. R.** 2001. Bacteriophage K1-5 encodes two different tail fiber proteins, allowing it to infect and replicate on both K1 and K5 strains of *Escherichia coli*. *J Virol* 75:2509-15.
86. **Schooley, R. T., Biswas, B., Gill, J. J., Hernandez-Morales, A., Lancaster, J. et al.** 2017. Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails To Treat a Patient with a Disseminated Resistant *Acinetobacter baumannii* Infection. *Antimicrobial Agents and Chemotherapy* 61.
87. **Seed, K. D.** 2015. Battling Phages: How Bacteria Defend against Viral Attack. *PLoS Pathog* 11:e1004847.
88. **Seemann, T.** 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068-2069.
89. **Shen, B. W., Landthaler, M., Shub, D. A. & Stoddard, B. L.** 2004. DNA binding and cleavage by the HNH homing endonuclease I-Hmul. *Journal of Molecular Biology* 342:43-56.
90. **Shkoporov, A. N. & Hill, C.** 2019. Bacteriophages of the Human Gut: The "Known Unknown" of the Microbiome. *Cell Host & Microbe* 25:195-209.

91. **Silva, J. B., Storms, Z. & Sauvageau, D.** 2016. Host receptors for bacteriophage adsorption. *Fems Microbiology Letters* 363:ARTN fnw002.
92. **Simmonds, P., Adams, M. J., Benko, M., Breitbart, M., Brister, J. R. et al.** 2017. Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology* 15:161-168.
93. **Snider, J., Thibault, G. & Houry, W. A.** 2008. The AAA plus superfamily of functionally diverse proteins. *Genome Biology* 9.
94. **Stamatakis, A.** 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312-3.
95. **Stoddard, B. L.** 2011. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure* 19:7-15.
96. **Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrener, P. et al.** 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* 406:959-964.
97. **Sulakvelidze, A., Alavidze, Z. & Morris, J. G.** 2001. Bacteriophage therapy. *Antimicrobial Agents and Chemotherapy* 45:649-659.
98. **Suttle, C. A.** 2007. Marine viruses - major players in the global ecosystem. *Nature Reviews Microbiology* 5:801-812.
99. **Torres-Barcelo, C.** 2018. The disparate effects of bacteriophages on antibiotic-resistant bacteria. *Emerging Microbes & Infections* 7.
100. **Touchon, M., de Sousa, J. A. M. & Rocha, E. P. C.** 2017. Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Current Opinion in Microbiology* 38:66-73.
101. **Veesler, D. & Cambillau, C.** 2011. A Common Evolutionary Origin for Tailed-Bacteriophage Functional Modules and Bacterial Machineries. *Microbiology and Molecular Biology Reviews* 75:423-433.
102. **Watts, G.** 2017. Phage therapy: revival of the bygone antimicrobial. *Lancet* 390:2539-2540.
103. **Weber-Dabrowska, B., Jonczyk-Matysiak, E., Zaczek, M., Lobočka, M., Lusiak-Szelachowska, M. & Gorski, A.** 2016. Bacteriophage Procurement for Therapeutic Purposes. *Frontiers in Microbiology* 7:Article 1177.
104. **Weinbauer, M. G.** 2004. Ecology of prokaryotic viruses. *Fems Microbiology Reviews* 28:127-181.
105. **Willis, S. H., Kazmierczak, K. M., Carter, R. H. & Rothman-Denes, L. B.** 2002. N4 RNA polymerase II, a heterodimeric RNA polymerase with homology to the single-subunit family of RNA polymerases. *Journal of Bacteriology* 184:4952-4961.
106. **Winsor, G.L., Griffiths, E.J., Lo R., Dhillon, B.K., Shay, J.A., Brinkman, F.S.** 2016. Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Research* 44: D646–D653.
107. **Winther-Larsen, H. C., Wolfgang, M. C., van Putten, J. P. M., Roos, N., Aas, F. E., Egge-Jacobsen, W. M., Maier, B. & Koomey, M.** 2007. *Pseudomonas aeruginosa* type IV pilus expression in *Neisseria gonorrhoeae*: Effects of pilin subunit composition on function and organelle dynamics. *Journal of Bacteriology* 189:6676-6685.
108. **Wittmann, J., Klumpp, J., Switt, A. I. M., Yagubi, A., Ackermann, H. W., Wiedmann, M., Svircev, A., Nash, J. H. E. & Kropinski, A. M.** 2015. Taxonomic reassessment of N4-like viruses using comparative genomics and proteomics suggests a new subfamily - "Enquartavirinae". *Archives of Virology* 160:3053-3062.
109. **Zeng, L., Skinner, S. O., Zong, C., Sippy, J., Feiss, M. & Golding, I.** 2010. Decision making at a subcellular level determines the outcome of bacteriophage infection. *Cell* 141:682-91.

7. Lista de apêndices

1. Tabela S1: Anotação do genoma do fago ZC04.
2. Tabela S2: Anotação do genoma do fago ZC06.
3. Tabela S3: Anotação do genoma do fago ZC07.
4. Tabela S4: SNPs/INDELS entre os genomas de PAO1 e de PA14 que sequenciamos (draft) e os respectivos genomas de referência.
5. Tabela S5: Anotação do genoma do fago ZC04.