

JOSICELLI SOUZA CRISPIM

**CARACTERIZAÇÃO GENÔMICA DE PROFAGOS E DE  
BACTÉRIAS DO GÊNERO *Desulfovibrio***

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós Graduação em Microbiologia Agrícola, para obtenção do título de *Doctor Scientiae*.

VIÇOSA  
MINAS GERAIS - BRASIL  
2018

**Ficha catalográfica preparada pela Biblioteca Central da Universidade  
Federal de Viçosa - Câmpus Viçosa**

T

C932c  
2018  
Crispim, Josicelli Souza, 1989-  
Caracterização genômica de profagos e de bactérias do  
gênero *Desulfovibrio* / Josicelli Souza Crispim. – Viçosa, MG,  
2018.  
x, 86 f. : il. (algumas color.) ; 29 cm.

Texto em português e inglês.

Orientador: Sérgio Oliveira de Paula.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. *Desulfovibrio*. 2. Bacteriófagos. 3. Genoma.  
I. Universidade Federal de Viçosa. Departamento de  
Microbiologia. Programa de Pós-Graduação em Microbiologia  
Agrícola. II. Título.

CDD 22. ed. 579.33

JOSICELLI SOUZA CRISPIM

## CARACTERIZAÇÃO GENÔMICA DE PROFAGOS E DE BACTÉRIAS DO GÊNERO *Desulfovibrio*


Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós Graduação em Microbiologia Agrícola, para obtenção do título de *Doctor Scientiae*.


APROVADA: 16 de julho de 2018.

  
Pedro Marcus Pereira Vidigal

  
Mateus Ferreira Santana

  
Michelle Dias de Oliveira

  
Cynthia Canêdo da Silva  
(Coorientadora)

  
Roberto Sousa Dias  
(Coorientador)

  
Sérgio Oliveira de Paula  
(Orientador)

“Não troque o que mais quer na vida pelo que  
mais quer no momento, pois o momento  
passa e a vida continua”  
**Renato Russo**

Dedico este trabalho aos meus pais, Milton e Célia,  
aos meus irmãos, Micelli, Jacyelli e Ewerton, à  
minha Família e aos eternos Amigos.

## AGRADECIMENTOS

À Universidade Federal de Viçosa, ao Departamento de Microbiologia, ao Programa de Pós-graduação em Microbiologia Agrícola e ao CNPq pela oportunidade de realização do Doutorado.

Ao corpo técnico-administrativo, em especial ao pessoal da secretaria (Sandra, Letícia, Nilcéa e Gabriel) pela imensa paciência e ajuda em todos os momentos. Ao Esquilo, Raquel e Aline, por toda ajuda sempre.

Ao meu orientador professor Sérgio Oliveira de Paula, pela oportunidade de realização do meu Doutorado em Microbiologia Agrícola.

À professora Cynthia e ao Roberto pela orientação e ensinamentos.

Aos integrantes da banca: Pedro Vidigal, um profissional que admiro muito, obrigada por acreditar no trabalho e por me dar forças para seguir em frente; professor Mateus, obrigada pelo apoio em todos os momentos e por acreditar junto comigo no sucesso dos profagos; professora Michelle, obrigada por me acolher no LIVM, pelos ensinamentos e por toda ajuda sempre.

A todos os professores do Departamento de Microbiologia, em especial aos professores Hilário e Tótola, que abriram as portas dos seus laboratórios. À professora Miriam pela ajuda nos eventos voltados aos alunos do COLUNI e ao Ciência na praça.

Aos professores do Departamento de Biologia Geral, em especial ao Leandro, Lino e Serrão por todo ensinamento e ajuda nos experimentos.

Ao Núcleo de Microscopia e Microanálise e ao Núcleo de Análise de Biomoléculas, pela estrutura e pela ajuda de todos os funcionários, especialmente ao Tns. Edvaldo e Gilmar pela atenção e aprendizado.

Aos colegas do Laboratório de Imunovirologia Molecular, em especial à Fran e à Jéssica pela amizade, pelos momentos de desabafo, pelos cafés, pelas melhores companhias de eventos científicos e pelos conselhos. Às minhas ICs, Clara e Larissa, por toda ajuda sempre, não sei o que seria de mim sem vocês!!!. Às ICs que não se encontram mais em nosso laboratório, Kassiana e Phedra, obrigada pela ajuda.

Aos grandes amigos que conheci na pós-graduação, em especial: Thyara, Aneli, Rita, Kaliane, Tomás, Vanessa, Lucas, Alan, Paulo e Ju, obrigada pela amizade, pelos conselhos, pelas horas de desabafo, pelos almoços e encontros da

turma. Ao Newton, por toda ajuda em todos os momentos que se possa imaginar, obrigada por ser nosso mestre da pós e da vida. Às companheiras de laboratório agregado: Elsa, Sofia, Cleriane e Bruna, obrigada por me ouvirem nos momentos difíceis e pelas risadas nos momentos felizes. Aos amigos que fiz nas disciplinas, Fernando e André, obrigada pelos estudos em grupo e amizade. Aos companheiros da Biologia Celular, Del, Neila, Luis, Micaela e Sabrina por demonstrar que amizade vai além da disciplina da pós-graduação.

À república Aloha pelos grandes momentos... só tenho a agradecer por ter conhecido vocês, pelos momentos de descontração e por terem entendido meus muitos momentos de ausência. Obrigada Amanda, Karina, Thais, Letícia, Lorena, Lis e Kamilla.

Aos amigos da eterna PYGS: Rafael Tchatcha, Mari, Ethel, Fernanda e Pedro pelo apoio sempre e conversas de conforto.

À minha segunda casa durante a pós-graduação: COLUNI. Agradeço aos técnicos administrativos Flávio, Tiago, Romildo, Fydiol, Ronaldo, Karlla, Luiza, Barbara, Zé Mauro e Adalberto, por toda ajuda. Em especial à Vanessa, minha companheira de laboratório, obrigada por tudo. Aos professores Isnard, Daniel, Adenilson, Duarte, Edson, Eunice, Odilaine, Hélio, Leomar, Zé Marcelo, Mário Alino, Renata Pires, Glauco, Renan e Vânia, por tornarem o ambiente de trabalho tão especial. Às professoras de Biologia Átima e Jildete, por me acolherem na área e pelos ensinamentos nesses 4,5 anos. Ao professor Luciano (Lep) por me acolher como pai, como amigo, como colega de trabalho e como um grande mestre, serei sempre grata por tudo que me ensinou, pelos momentos de desabafo e conselhos.

Aos meus pais, Milton e Célia, pelo amor incondicional, por todo apoio e pela compreensão dos meus momentos de ausência. Aos meus irmãos, pela amizade e por sempre me darem força para seguir em frente. À minha sobrinha Manuelli, pelo imenso amor e carinho que transmite à todos nós. Aos meus avós, pelo carinho e por todo apoio.

Ao Ácacio, pelas idas intermináveis à noite no laboratório e nos finais de semana comigo, por estar sempre ao meu lado comemorando as vitórias, por me dar força nos momentos difíceis e por me mostrar que o amor fortalece. Obrigada por tudo!!

A todos aqueles que acreditaram em mim e me deram todo apoio para chegar até aqui. À vocês, o meu muito obrigada!!!

## ÍNDICE

RESUMO .....	vii
ABSTRACT .....	ix
INTRODUÇÃO GERAL .....	1
Referência bibliográfica.....	10
CAPÍTULO 1 .....	16
<b>Screening and characterization of prophages in <i>Desulfovibrio</i> genomes</b>	
Abstract.....	17
Introduction .....	17
Results .....	19
Discussion.....	26
Conclusions .....	30
Methods .....	30
References.....	32
Supplementary information .....	36
CAPÍTULO 2 .....	44
<b>Profagos de <i>Desulfovibrio alaskensis</i> e seu possível envolvimento na     transferência horizontal de genes por vesículas de membrana externa</b>	
Referências.....	62
CAPÍTULO 3 .....	78
<b>Sequência genômica de <i>Desulfovibrio indonesiensis</i> P37SLT</b>	
Referências.....	81
CAPÍTULO 4 .....	82
<b>Sequência genômica de <i>Desulfovibrio marinus</i> P48SEP, o primeiro genoma     da espécie</b>	
Referências.....	85
CONCLUSÕES GERAIS E CONSIDERAÇÕES FINAIS .....	86

## RESUMO

CRISPIM, Josicelli Souza, D.Sc., Universidade Federal de Viçosa, julho de 2018. **Caracterização genômica de profagos e de bactérias do gênero *Desulfovibrio***. Orientador: Sérgio Oliveira de Paula. Coorientadores: Cynthia Canêdo da Silva e Roberto Sousa Dias.

Bactérias do gênero *Desulfovibrio* pertencem ao grupo de Bactérias Redutoras de Sulfato (BRS). As BRS causam significativos prejuízos à indústria de petróleo, principalmente devido à capacidade de causar corrosão microbiologicamente induzida, formação de biofilme e produção de H<sub>2</sub>S. Bacteriófagos são uma forma alternativa para o controle biológico das BRS, cuja informação para este grupo de bactérias, entretanto, é escassa. Assim, este trabalho teve como objetivos: 1) identificar e caracterizar profagos em genomas de *Desulfovibrio*; 2) verificar a expressão de profagos de *Desulfovibrio alaskensis*; 3) Isolar, sequenciar e montar genomas de BRS. Para identificação de profagos de *Desulfovibrio* foram utilizados 46 genomas depositados em banco de dados. 53 profagos completos foram identificados em 46 dos genomas analisados. Estes foram classificados dentro da ordem *Caudovirales*, com 69,82% pertencentes à família *Myoviridae*. 18 profagos identificados possuem genes que codificam para lisozima e holina. Quatro dos genomas analisados apresentam profagos com identidade acima de 50% na mesma linhagem, cuja análise comparativa demonstrou a existência de colinearidade entre as sequências. Dos 17 genomas bacterianos fechados analisados, 6 possuem sistema CRISPR-Cas classificado como inativo. A identificação da poli-lisogenia, a proximidade entre os profagos completos e a possível inatividade do CRISPR-Cas nos genomas fechados analisados, permitiu a escolha de linhagens poli-lisogênicas com profagos pertencentes à família *Myoviridae* para isolamento de profagos e testes em linhagens relacionadas para estudos posteriores. A linhagem *D. alaskensis* DSM16109 possui três profagos, cuja a expressão e a relação com vesículas de membrana externa foram verificados neste trabalho. A linhagem DSM16109 apresentou menor crescimento após a adição de mitomicina C em relação à cultura controle. Esta redução foi acompanhada da presença de partículas *virus-like* (VLPs), indicando indução de profagos por mitomicina C. Foi verificado o aumento do número de cópias dos genes *cap* dos três profagos e transcrição dos mesmos na amostra controle, entretanto, sem a formação de VLPs. Genes de profagos foram identificados em vesículas de membrana externa de culturas tratadas

e não tratadas com mitomicina C. Os profagos de DSM16109 são induzidos de forma espontânea, mas, apenas a presença de mitomicina C levou à formação de VLPs. O material genético dos profagos detectado dentro de vesículas de membrana externa podem estar relacionados com a transferência horizontal de genes virais. Nas análises genômicas de isolados de amostras de plataformas de extração de petróleo da Bacia de Campos (RJ) foi verificado que a linhagem P37SLT pertence à espécie *Desulfovibrio indonesiensis* com 4,2Mb e a linhagem P48SEP pertence à espécie *Desulfovibrio marinus* com 5,1Mb. Ambas as linhagens foram sequenciadas pelo sistema HiSeq (2x300) e apresentam sequências relacionadas à profagos, uma característica comum do gênero *Desulfovibrio* encontrada nos genomas anteriormente analisados.

## ABSTRACT

CRISPIM, Josicelli Souza, D.Sc., Universidade Federal de Viçosa, July, 2018. **Genomic characterization of prophages and bacteria of the genus *Desulfovibrio***. Adviser: Sérgio Oliveira de Paula. Co-advisers: Cynthia Canêdo Silva and Roberto Sousa Dias.

Bacteria of the genus *Desulfovibrio* belong to the group of Sulphate Reducing Bacteria (SRB). SRB generate significant liabilities in the petroleum industry, mainly due to their ability to microbiologically induce corrosion, biofilm formation and H<sub>2</sub>S production. Bacteriophages are an alternative control method for SRB, whose information for this group of bacteria however, is scarce. Thus, this work had as objectives: 1) identify and characterize prophages in *Desulfovibrio* genomes; 2) verify the expression of *Desulfovibrio alaskensis* prophages; 3) Isolate, sequencing and assembling BRS genomes. For the identification of *Desulfovibrio* prophages, 47 genomes deposited in a database were used. In the 46 genomes analyzed 53 complete prophages were identified. These were classified within the order *Caudovirales*, with 69.82% belonging to the *Myoviridae* family. 18 prophages identified have genes coding for lysozyme or holin. Four of the analyzed bacterial genomes present prophages with identity above 50% in the same strain, whose comparative analysis demonstrated the existence of colinearity between the sequences. Of the 17 closed bacterial genomes analyzed, 6 have the CRISPR-Cas system classified as inactive. The identification of bacterial poly-lysogeny, the proximity between the complete prophages and the possible inactivity of the CRISPR-Cas in closed bacterial genomes analyzed allowed the choice of polylysogenic strains with prophages belonging to the *Myoviridae* family for the isolation of prophages and testing of related strains for subsequent studies. The *D. alaskensis* DSM16109 has three prophages, whose expression and relation with outer membrane vesicles were verified in this work. The DSM16109 strain presented lower growth after the addition of mitomycin C in relation to the control culture. This reduction was accompanied by the presence of virus-like particles (VLPs), indicating induction of mitomycin C prophages. The increase in the number of copies of the *cap* genes of the three prophages and their transcription in the control sample was verified, however, without the formation of VLPs. Prophage genes were identified in outer membrane vesicles from cultures treated and not treated with mitomycin C. The prophages of DSM16109 are spontaneously induced, but only the presence of mitomycin C led to the formation

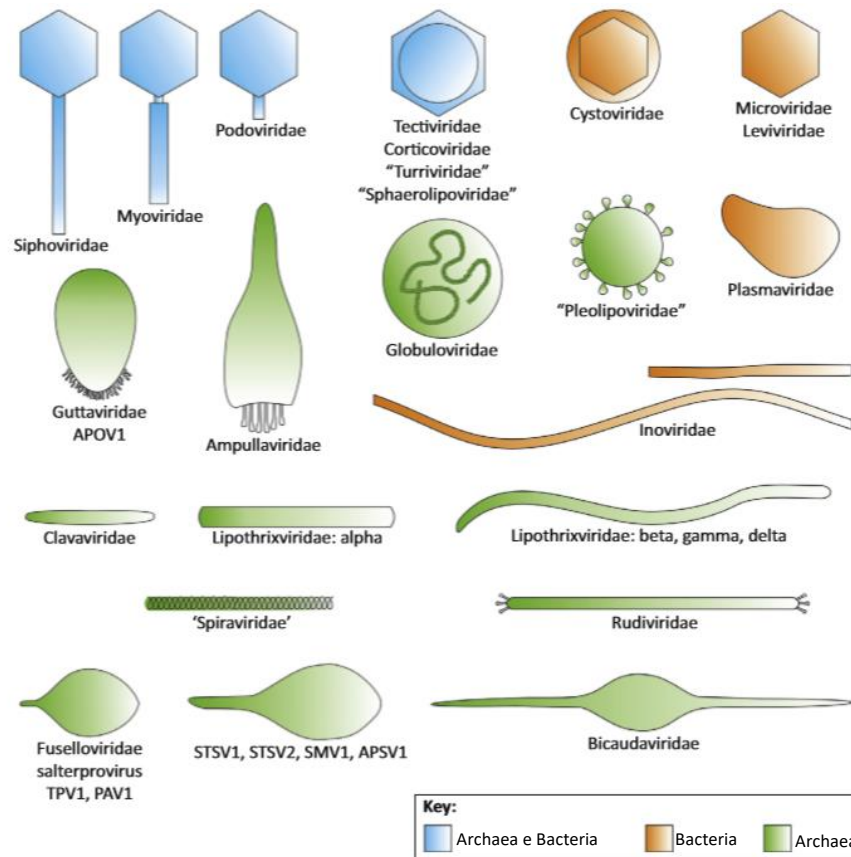
of VLPs. The genetic material of the prophages detected within outer membrane vesicles may be related to the horizontal transfer of viral genes. In the genomic analyzes of isolates from samples of oil extraction platforms of the Bacia de Campos (RJ) samples, the P37SLT strain was classified as *Desulfovibrio indonesiensis* with 4.2Mb and the P48SEP strain belongs to the species *Desulfovibrio marinus* with 5.1Mb. Both strains were sequenced by the HiSeq system (2x300) and presented sequences related to the prophages, a common characteristic of the genus *Desulfovibrio* found in the genomes previously analyzed.

## INTRODUÇÃO GERAL

Vírus que infectam Bacteria ou Archaea são chamados de bacteriófagos e arqueovírus, respectivamente. Esses vírus podem carregar material genético de DNA ou RNA, de fita simples ou dupla, e podem apresentar formas diversas (poliédrica, filamentosa, pleomórfica ou caudada) (Ackermann e Prangishvili, 2012; Pietila et al., 2014). A forma poliédrica é caracterizada pela estrutura icosaédrica sem a presença de cauda, como *Leviviridae*, *Corticoviridae*, *Microviridae*, *Tectiviridae* SH-1 e STIV. Já a forma filamentosa é caracterizada pela estrutura alongada, como em *Lipothrixviridae*, *Clavaviridae*, *Rudiviridae* e *Inoviridae*. A pleomorfia compõe um grupo de vírus com morfologia distinta, como *Ampullaviridae* e *Bicaudaviridae*. A forma caudada é caracterizada pela presença de uma cabeça icosaédrica, onde se localiza o material genético, e uma cauda proteica, que é responsável pelo reconhecimento do receptor celular e posterior injeção do material genético na célula hospedeira. Esse grupo caracteriza a ordem *Caudovirales*, que abrange 96% dos vírus de procariotos e é composta por quatro famílias: *Myoviridae* (cauda curta e contrátil), *Siphoviridae* (cauda longa e não contrátil), *Podoviridae* (cauda curta e não contrátil) e *Ackermannviridae* (Ackermann e Prangishvili, 2012; Adriaenssens et al., 2018). A diversidade morfológica desses vírus pode ser visualizada na Figura 1.

No processo de infecção da célula hospedeira, os vírus reconhecem receptores celulares (processo de adsorção) e liberam seu material genético no interior do hospedeiro, onde irão seguir o ciclo lítico ou o ciclo lisogênico, extensivamente estudados no bacteriófago lambda (Oppenheim et al., 2005). No ciclo lítico ocorre propagação do fago, cujas novas partículas virais são liberadas da célula por meio da lise celular. A lise celular é gerada pela ação de enzimas holinas que danificam a membrana plasmática, permitindo o acesso de endolisinas que atacam o peptidoglicano (Ackermann, 1998). No ciclo lisogênico ocorre a integração do material genético do fago no genoma bacteriano, cuja célula hospedeira continua seu crescimento, gerando novas partículas virais de forma passiva juntamente com sua própria replicação. Além desses dois ciclos de replicação, os bacteriófagos também apresentam infecções crônicas e pseudolisogenia. Nas infecções crônicas ocorre liberação constante de partículas virais da célula hospedeira por brotamento ou extrusão, e na infecção

pseudolisogênica, os fagos estão presentes no interior celular na forma semelhante a um plasmídeo (epissomal) (Weinbauer, 2006).



**Figura 1.** Morfotipos de vírus de infectam Bacteria e Archaea. Figura adaptada de Pietila e colaboradores (2014).

Acredita-se que a lisogenia é favorecida quando a célula hospedeira encontra-se em um ambiente desfavorável, pobre em nutrientes, sendo o ciclo lítico favorecido em um ambiente oposto, onde a quantidade de células hospedeiras é alta para a propagação viral (Kourilsky, 1973). O controle entre esses ciclos é realizada principalmente pelas proteínas CI e CII no bacteriófago lambda (Wegrzyn et al., 2012). Quando CII se encontra em altas concentrações, a lisogenia é estabilizada, pois ativa os promotores  $P_{RE}$  e  $P_I$ , que levam a transcrição de CI e do gene da integrase. CI liga-se aos promotores  $P_L$  e  $P_R$ , prevenindo o ciclo lítico e permitindo a integração do genoma do fago no genoma da célula hospedeira. Quando CII não está em concentrações suficientes, devido a regulação da sua produção ou por degradação proteica, o ciclo lítico é ativado, ocorrendo a transcrição de genes relacionados à replicação e à estrutura viral. Esse controle de ciclos também pode

ser governado por meio de peptídeos sinalizadores, semelhante a um *quorum sensing* bacteriano, como demonstrado para o grupo de bacteriófagos SpBeta de *Bacillus*, onde a alta concentração de peptídeos sinalizadores induz a lisogenia (Erez et al., 2017). Entretanto, os ciclos lítico e lisogênico podem coexistir quando ocorre a infecção da célula hospedeira por mais de um fago, devido ao controle de ciclos diferencial de cada um (Zeng et al., 2010). Um outro estudo mostra que o ciclo lítico e o ciclo lisogênico também podem coexistir quando ocorre a infecção de apenas um fago na célula hospedeira e sua integração durante o ciclo lítico (Shao et al., 2017).

Os bacteriófagos lisogênicos, chamados de profagos, estão relacionados com a diversidade bacteriana, podendo compor até 20% do genoma da célula hospedeira (Casjens, 2003). Como os bacteriófagos são capazes de carregar genes bacterianos, atuando como elemento de transferência horizontal de genes, contribuem na evolução, *fitness* e virulência bacteriana (Brueggemann et al., 2017; Walker et al., 2009). Quando colocados frente à estresses químicos ou físicos, os profagos podem seguir o ciclo lítico (Gandon, 2016). Isso ocorre porque o sistema SOS é ativado durante tratamentos que danificam o DNA ou inibem a replicação (Radman, 1975). A proteína RecA, quando ativada por meio de danos no DNA, cliva o repressor LexA do sistema SOS, permitindo a transcrição de genes relacionados ao reparo do DNA. A ativação de RecA também leva a clivagem de repressores de bacteriófagos, semelhantes ao repressor CI do bacteriófago lambda (Pawlowski e Koudelka, 2004), retirando assim, o bloqueio da transcrição de genes que caracterizam o ciclo lítico.

Entre os mecanismos utilizados para a indução de profagos em laboratório encontra-se o antibiótico mitomicina C, que demonstra ser a técnica mais eficiente (Dziewit e Radlinska, 2016; Hammerl et al., 2016; Lorenz et al., 2016). A indução de profagos é empregada com intuito de isolamento dos mesmos para posterior realização de modificações genéticas para torná-los infectivos para outras células hospedeiras ou retirar a capacidade de integração para torná-los apenas líticos (Chen et al., 2017; Hudson et al., 2016). Com o isolamento de profagos é possível realizar também o sequenciamento do seu genoma e assim identificar genes de interesse para a prospecção de enzimas virais (Song et al., 2016). Além disso, com profagos isolados pode-se realizar testes de infecção em linhagens relacionadas e distintas (Hamdi et al., 2017; Mercanti et al., 2011; Walker et al., 2006).

Nessa interação entre vírus e célula hospedeira, as bactérias desenvolveram mecanismos de escape da infecção viral, assim como os vírus desenvolveram outras formas para burlar estes mecanismos (Golais et al., 2013; Shabbir et al., 2016). Dois tipos de sistema de defesa bacteriano são encontrados: um que impede a entrada do material genético viral e outro que degrada o material genético após a inoculação do mesmo na célula. No primeiro caso, modificações no receptor de adsorção viral podem ocorrer, impedindo o reconhecimento pelos vírus. A entrada do material genético viral na célula também pode ser evitada por meio do bloqueio de canais transportadores. Caso ocorra a entrada do material genético viral, enzimas de restrição bacteriana podem clivar o material genético invasor. Mas, os vírus são capazes de sobressair às modificações da célula bacteriana: alterações na estrutura viral podem ocorrer e serem capazes de reconhecer o novo receptor celular e liberar o material genético no interior da célula por outro transportador, além de carregarem sua própria DNA metilase, o que impede que sejam reconhecidos pelas enzimas de restrição bacteriana (Golais et al., 2013). Outra forma de defesa é o sistema CRISPR (*Clustered Regularly Interspace Short Palindromic Repeats*), que é capaz de criar uma memória do material genético invasor, impedindo assim, uma super infecção (Makarova et al., 2015). Por outro lado, modificações podem ocorrer no genoma viral, impedindo o reconhecimento pela memória criada pelo sistema CRISPR (Golais et al., 2013).

Apesar das células bacterianas apresentarem formas de escape à infecção viral, sendo esta infecção uma forma de controle biológico por meio do ciclo lítico, os bacteriófagos apresentam enzimas que podem ser utilizadas no controle do biofilme bacteriano sem a necessidade de infecção da célula hospedeira (Chan e Abedon, 2015). Essas enzimas são depolimerases de polissacarídeos, hidrolases codificadas por bacteriófagos que desestruturam a matriz do biofilme bacteriano por meio da degradação das substâncias poliméricas extracelulares (Yan et al., 2014). A destruição do biofilme faz com que as células se tornem planctônicas, ficando mais susceptíveis a ação de antibióticos e de outros agentes antimicrobianos, como os biocidas, que não conseguem penetrar no biofilme (Flemming e Wingender, 2010). Outra ação enzimática é caracterizada pela desestabilização celular por meio do alto número de partículas virais no processo de adsorção (Abedon, 2011). Este processo é conhecido como *lysis from without*, no qual os vírus não injetam o material genético na célula hospedeira, mas a presença de enzimas virais desestabiliza a parede

celular, levando à morte da bactéria. Essas enzimas são hidrolases de peptidoglicano, que agem sobre a parede celular bacteriana criando um pequeno túnel, no qual o bacteriófago injeta o material genético para dentro da célula durante o processo de infecção (Rodríguez-Rubio et al., 2016). Além das hidrolases de peptidoglicano, outras proteínas líticas são as holinas, enzimas que desestabilizam a membrana celular durante o processo de liberação de novas partículas no ciclo lítico (Song et al., 2016). Esta variedade de enzimas codificadas por bacteriófagos são potenciais candidatos à agentes antibacterianos, sendo já testadas no tratamento de doenças infecciosas em animais e desinfecção de superfícies (Rodríguez-Rubio et al., 2016). Assim, muitas formas de controle bacteriano podem ser realizadas por meio de bacteriófagos, mas pouco é conhecido sobre o uso dos mesmos no controle de Bactérias Redutoras de Sulfato (BRS).

As BRS são procariotos do Domínio Bacteria que utilizam o sulfato como aceptor final de elétrons durante o processo de respiração anaeróbia (Postgate, 1965), sendo ubíquas em ambientes anóxicos (Muyzer e Stams, 2008). As BRS representam um grupo diversificado de procariotos, abrangendo bactérias pertencentes à classe *Deltaproteobacteria* (ordens: *Desulfovibrionales*, *Desulfobacterales*, *Desulfarculales* e *Syntrophobacterales*); organismos formadores de esporo Gram-positivos das classes *Negativicutes* e *Clostridia* (famílias: *Peptococcaceae*, *Thermoanaerobacteraceae* e *Thermodesulfobiaceae*); classe *Thermodesulfobacteria* (gêneros *Thermodesulfobacterium* e *Thermodesulfatator*) e classe *Nitrospira* (gênero *Thermodesulfovibrio*) (Rabus et al., 2015).

Apesar de serem caracterizadas pela capacidade de redução de sulfato, as BRS também podem utilizar outros aceptores finais de elétrons, como Mn (IV) (Myers e Nealson, 1988) e Fe (III) (Lovley et al., 2004). As BRS tem um importante papel no ciclo do carbono e do enxofre em ambientes anaeróbios (Muyzer e Stams, 2008), mas também podem ser encontradas em ambientes aeróbios, sendo tolerantes à algumas variações de temperatura, pH e pressão (Heggenhorn et al., 2013; Robador et al., 2016). Algumas BRS são capazes de realizar respiração aeróbia, mas este processo gera energia apenas para o microrganismo se manter (Fareleira et al., 2003; Lemos et al., 2001).

Durante o processo de oxidação de compostos orgânicos ocorre a formação de H<sub>2</sub>, que cria um ambiente intracelular muito reduzido (Pereira et al., 2008; Plugge et al., 2011). A via de redução de sulfato é uma alternativa para mudar esse ambiente

intracelular, pois o  $H_2$  gerado é utilizado como doador de elétrons para a cadeia transportadora de elétrons que tem como produto final o  $H_2S$ . A formação e liberação para o ambiente de  $H_2S$  por BRS é caracterizada pelo processo dissimilativo de sulfato, no qual o sulfato é primeiramente ativado em uma reação com gasto de energia e posteriormente transformado em  $H_2S$ , tendo  $HSO_3^-$  como intermediário. A energia necessária para ativação do sulfato vem da fosforilação a nível de substrato e da oxidação a partir de uma fonte de carbono, como o lactato ou da fosforilação oxidativa pela ATPsintetase.

As BRS são encontradas principalmente em ambientes anóxicos ricos em sulfato, como o ambiente marinho (Hoehler e Jorgensen, 2013; Leloup et al., 2009), e estão relacionadas com o fenômeno de *souring* na indústria de petróleo (Gieg et al., 2011). Esse fenômeno é caracterizado pela produção de  $H_2S$  pelas BRS durante a oxidação anaeróbia da matéria orgânica presente no petróleo. O *souring* do petróleo ocorre dentro dos reservatórios (Agrawal et al., 2010) e pode ser induzido pela introdução de água do mar nos mesmos, utilizada para aumentar a pressão local, permitindo a contínua retirada de petróleo (Gieg et al., 2011). Esta água recebe o nome de água de injeção, sendo fonte de BRS e sulfato. Além dos reservatórios, a redução de sulfato por BRS pode ocorrer em todos os pontos durante a exploração e processamento de petróleo, como separadores, na água de injeção reutilizada e nos tanques de armazenamento e transporte (Gieg et al., 2011).

O  $H_2S$  representa um impacto negativo para a indústria de petróleo, pois é altamente tóxico, podendo levar à morte dos trabalhadores (Beauchamp et al., 1984). Além disso, o  $H_2S$  diminui o pH do petróleo e aumenta a corrosão de tubulações e tanques, que diminuem o valor do produto e aumentam os custos com infraestrutura (Enning e Garrelfs, 2014). A corrosão ocorre porque o  $H_2S$  produzido pelas BRS interage com o  $Fe^0$ , que age como doador de elétrons, levando a formação de  $FeS$  e  $H_2$ , sendo o  $H_2$  utilizado pelas BRS na redução de sulfato novamente. Em um meio com matéria orgânica limitada e excesso de sulfato, as BRS podem utilizar indiretamente o  $Fe^0$  como doador de elétrons para formação de  $H_2$ , que será utilizado como doador de elétrons para a redução de sulfato. Mas essa biocorrosão também pode ocorrer de forma direta, na qual algumas BRS utilizam diretamente o  $Fe^0$  como doador de elétrons para a redução de sulfato. Assim, a biocorrosão de ferro por BRS pode ocorrer pelas três formas citadas acima (Enning e Garrelfs, 2014).

Além da biocorrosão, as BRS possuem capacidade de formação de biofilme, caracterizado por um conjunto de células aderidas em uma superfície e envoltas por uma camada de matriz polimérica extracelular (Flemming e Wingender, 2010; Rosa et al., 2016). A formação de biofilme favorece a corrosão microbiologicamente induzida (CMI) e dificultam as formas de controle bacteriano (Flemming e Wingender, 2010; Rajasekar et al., 2010). Para o controle das BRS são empregados biossurfactantes (Zhao et al., 2016), biocidas (Wen et al., 2012; Xu et al., 2012) e nitrato (Bodtker et al., 2008), sendo o nitrato um aceptor final de elétrons que atua como competidor do sulfato, privilegiando bactérias redutoras de nitrato em detrimento das redutoras de sulfato no uso da fonte de carbono. Além dessas formas de controle, bacteriófagos líticos de BRS também vem sendo isolados (Eydal et al., 2009; Kamimura e Araki, 1989) e patentes tem sido desenvolvidas (*Use of Prokaryote Viruses to Remediate Bio-Fouling* - patente de nº 8252576). A indução de profagos também tem se mostrado promissora no controle de BRS, dado verificado na patente de nº WO2002040642A1 (*Bacterial control by bacteriophages*).

De acordo com Rabus e colaboradores (2015), existem 37 gêneros dentro do grupo de BRS, sendo que apenas um gênero apresenta trabalhos publicados relacionados à bacteriófagos, o gênero *Desulfovibrio*. *Desulfovibrio* é o segundo gênero bacteriano que abrange o maior número de espécies dentro do grupo de BRS (Rabus et al., 2015), e o que apresenta o maior número de genomas depositados NCBI (*National Center for Biotechnology Information*) no presente ano.

O primeiro bacteriófago de *Desulfovibrio* foi isolado em 1973, sendo um profago de *Desulfovibrio vulgaris* induzido utilizando mitomicina C (Handley et al., 1973). Este estudo abriu caminhos para posteriores investigações da presença de fagos em BRS. Em 1987, um profago capaz de mediar o processo de transdução foi identificado em *Desulfovibrio desulfuricans* ATCC 27774 (Rapp e Wall, 1987), demonstrando a possibilidade de ferramentas de biologia molecular para a manipulação de BRS. Já em 1989, o primeiro fago lítico foi isolado, tendo *Desulfovibrio salexigens* como hospedeira (Kamimura e Araki, 1989). Este bacteriófago foi isolado de amostra de sedimentos marinho, sendo classificado dentro da ordem *Caudovirales* pela presença de uma cabeça icosaédrica e uma cauda longa, característica da família *Siphoviridae*. Em 1991, Seyedirashti e colaboradores retornaram às análises com profagos. Este autores induziram os profagos de *D. vulgaris* Hildenborough, a mesma linhagem utilizada em 1973, e

*Desulfovibrio desulfuricans* ATCC 13541 por luz ultravioleta (UV) (Seyedirashti et al., 1991). Os profagos encontrados para as duas linhagens bacterianas não apresentaram similaridade por meio de sondas, mesmo tendo bactérias de mesmo gênero como hospedeiras. Seyedirashti e colaboradores empregaram novamente essa mesma metodologia em *D. vulgaris* Hildenborough, onde verificaram a presença de dois tipos de partículas virais no sobrenadante da cultura da linhagem por meio de microscopia eletrônica e de gradiente por cloreto de cério (Seyedirashti et al., 1992). Após 14 anos, outros pesquisadores realizaram novos estudos com *D. vulgaris* Hildenborough, mas agora utilizando uma linhagem altamente relacionada (DePue) como hospedeira dos profagos induzidos de Hildenborough (Walker et al., 2006). Por meio dessa linhagem hospedeira, estes pesquisadores conseguiram isolar dois profagos da linhagem Hildenborough. O posterior sequenciamento dessas duas linhagens demonstrou a relação de regiões de profagos com a variabilidade encontrada entre elas (Walker et al., 2009), o que demonstra a importância dos profagos para a variabilidade bacteriana. Além disso, uma das características que permitiu o uso da linhagem DePue como hospedeira de profagos da Hildenborough foi a ausência de espaçadores no sistema CRISPR similares aos profagos da linhagem Hildenborough, dados identificados após o sequenciamento (Walker et al., 2009).

Como visto anteriormente, o sistema CRISPR pode impedir a super infecção viral (Makarova et al., 2015). Outro mecanismo apresentado por bactérias que pode impedir a infecção por bacteriófagos é a produção de vesículas de membrana externa (*Outer Membrane Vesicles* - OMVs). OMVs são produzidas por bactérias Gram-negativas por meio da vesiculação da parede celular, podendo apresentar de 10-300 nm (Jan, 2017). Essas vesículas protegem da infecção por bacteriófagos por apresentarem receptores virais, o que leva ao reconhecimento viral das vesículas, impedindo que o fago chegue até a célula hospedeira (Manning and Kuehn, 2011; Reyes-Robles et al., 2018). As vesículas podem carregar proteínas, material genético e RNAs regulatórios, estando envolvidas na virulência e formação de biofilme bacteriano (Jan, 2017). Além de material genético da própria célula hospedeira, já foi verificado a presença de material genético viral no interior dessas vesículas (Gaudin et al., 2014). Esse transporte intercelular de material genético viral caracteriza um mecanismo de transferência horizontal de genes (THG) mediado por OMVs (Domingues and Nielsen, 2017). A THG está diretamente relacionada com a

evolução e diversidade bacteriana (Paquola *et al.*, 2018). Apenas um trabalho sobre OMVs é encontrado para o gênero *Desulfovibrio* (Haynes et al., 1995), o qual não explora o envolvimento das OMVs com o transporte intercelular de material genético viral.

Como demonstrado, poucos estudos foram realizados com bacteriófagos e com vesículas de membrana externa no grupo de BRS. Assim, levando em consideração também a grande quantidade de genomas bacterianos do gênero *Desulfovibrio* disponibilizados no banco de dados, o estudo de profagos presentes em genomas deste gênero e a verificação da expressão de profagos e de vesículas de membrana externa, demonstram serem áreas atrativas. Além disso, a obtenção de culturas puras de BRS é necessária para o isolamento de bacteriófagos líticos em estudos posteriores e o sequenciamento dos genomas desses isolados importante na busca de profagos. Portanto, o objetivo deste trabalho foi identificar e caracterizar profagos em genomas de *Desulfovibrio*, verificar a expressão de profagos e de vesículas de membrana externa em *Desulfovibrio alaskensis*, assim como, isolar, sequenciar e analisar o genoma de dois isolados de BRS em busca de sequências de profagos.

## Referência bibliográfica

- Abedon ST. (2011). Lysis from without. *Bacteriophage* 1: 46–49.
- Ackermann H-W. (1998). Tailed bacteriophages: the order *Caudovirales*. In: Maramorosch K, Murphy FA, Shatkin AJBT-A in VR (eds) Vol. 51. Academic Press, pp 135–201.
- Ackermann H-W, Prangishvili D. (2012). Prokaryote viruses studied by electron microscopy. *Arch Virol* 157: 1843–1849.
- Adriaenssens EM, Wittmann J, Kuhn JH, Turner D, Sullivan MB, Dutilh BE, et al. (2018). Taxonomy of prokaryotic viruses: 2017 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Arch Virol* 163: 1125–1129.
- Agrawal A, Vanbroekhoven K, Lal B. (2010). Diversity of culturable sulfidogenic bacteria in two oil-water separation tanks in the north-eastern oil fields of India. *Anaerobe* 16: 12–18.
- Beauchamp ROJ, Bus JS, Popp JA, Boreiko CJ, Andjelkovich DA. (1984). A critical review of the literature on hydrogen sulfide toxicity. *Crit Rev Toxicol* 13: 25–97.
- Bodtker G, Thorstenson T, Lillebo B-LP, Thorbjornsen BE, Ulvoen RH, Sunde E, et al. (2008). The effect of long-term nitrate treatment on SRB activity, corrosion rate and bacterial community composition in offshore water injection systems. *J Ind Microbiol Biotechnol* 35: 1625–1636.
- Brueggemann AB, Harrold CL, Rezaei Javan R, Van Tonder AJ, McDonnell AJ, Edwards BA. (2017). Pneumococcal prophages are diverse, but not without structure or history. *Sci Rep* 7: 1–13.
- Casjens S. (2003). Prophages and bacterial genomics: What have we learned so far? *Mol Microbiol* 49: 277–300.
- Chan BK, Abedon ST. (2015). Bacteriophages and their enzymes in biofilm control. *Curr Pharm Des* 21: 85–99.
- Chen M, Zhang L, Xin S, Yao H, Lu C, Zhang W. (2017). Inducible prophage mutant of *Escherichia coli* can lyse new host and the key sites of receptor recognition identification. *Front Microbiol* 8: 1–13.
- Domingues S, Nielsen KM. (2017). Membrane vesicles and horizontal gene transfer in prokaryotes. *Curr Opin Microbiol* 38: 16–21.
- Dziewit L, Radlinska M. (2016). Two inducible prophages of an antarctic *Pseudomonas* sp. ANT\_H14 use the same capsid for packaging their genomes

- characterization of a novel phage helper-satellite system. PLoS One 11: e0158889.
- Enning D, Garrelfs J. (2014). Corrosion of iron by sulfate-reducing bacteria: new views of an old problem. Appl Environ Microbiol 80: 1226–1236.
- Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, Peleg Y, et al. (2017). Communication between viruses guides lysis–lysogeny decisions. Nature 541: 488.
- Eydal HSC, Jagevall S, Hermansson M, Pedersen K. (2009). Bacteriophage lytic to *Desulfovibrio aespoeensis* isolated from deep groundwater. ISME J 3: 1139–1147.
- Fareleira P, Santos BS, Antonio C, Moradas-Ferreira P, LeGall J, Xavier A V, et al. (2003). Response of a strict anaerobe to oxygen: survival strategies in *Desulfovibrio gigas*. Microbiology 149: 1513–1522.
- Flemming H-C, Wingender J. (2010). The biofilm matrix. Nat Rev Microbiol 8: 623–633.
- Gandon S. (2016). Why be temperate: lessons from bacteriophage λ. Trends Microbiol 24: 356–365.
- Gaudin M, Krupovic M, Marguet E, Gaudiard E, Cvirkaite-Krupovic V, Le Cam E, et al. (2014). Extracellular membrane vesicles harbouring viral genomes. Environ Microbiol 16: 1167–1175.
- Gieg LM, Jack TR, Foght JM. (2011). Biological souring and mitigation in oil reservoirs. Appl Microbiol Biotechnol 92: 263–282.
- Golais F, Hollý J, Vítková J. (2013). Coevolution of bacteria and their viruses. Folia Microbiol (Praha) 58: 177–186.
- Hamdi S, Rousseau GM, Labrie SJ, Tremblay DM, Kourda RS, Slama K Ben, et al. (2017). Characterization of two polyvalent phages infecting Enterobacteriaceae. Sci Rep 7. e-pub ahead of print, doi: 10.1038/srep40349.
- Hammerl JA, Göllner C, Al Dahouk S, Nöckler K, Reetz J, Hertwig S. (2016). Analysis of the first temperate broad host range Brucellaphage (BiPBO1) isolated from *B. inopinata*. Front Microbiol 7: 24.
- Handley J, Adams V, Akagi JM. (1973). Morphology of bacteriophage-like particles from *Desulfovibrio vulgaris*. J Bacteriol 115: 1205–7.

- Haynes TS, Klemm DJ, Ruocco JJ, Barton LL. (1995). Formate dehydrogenase activity in cells and outer membrane blebs of *Desulfovibrio gigas*. *Anaerobe* 1: 175–182.
- Heggendorn FL, Goncalves LS, Dias EP, Silva Junior A, Galvao MM, Lutterbach MTS. (2013). Detection of sulphate-reducing bacteria in human saliva. *Acta Odontol Scand* 71: 1458–1463.
- Hoehler TM, Jørgensen BB. (2013). Microbial life under extreme energy limitation. *Nat Rev Microbiol* 11: 83.
- Hudson J, Billington C, Premaratne A, On S. (2016). Inactivation of *Escherichia coli* O157:H7 using ultraviolet light-treated bacteriophages. *Rev Agroquimica y Tecnol Aliment* 22: 3–9.
- Jan AT. (2017). Outer Membrane Vesicles (OMVs) of Gram-negative bacteria: a perspective update. *Front Microbiol* 8: 1053.
- Kamimura K, Araki M. (1989). Isolation and characterization of a bacteriophage lytic for *Desulfovibrio salexigens*, a salt-requiring, sulfate-reducing bacterium. *Appl Environ Microbiol* 55: 645–648.
- Kourilsky P. (1973). Lysogenization by bacteriophage lambda. *Mol Gen Genet* 122: 183–195.
- Leloup J, Fossing H, Kohls K, Holmkvist L, Borowski C, Jørgensen BB. (2009). Sulfate-reducing bacteria in marine sediment (Aarhus Bay, Denmark): abundance and diversity related to geochemical zonation. *Environ Microbiol* 11: 1278–1291.
- Lemos RS, Gomes CM, Santana M, LeGall J, Xavier A V, Teixeira M. (2001). The 'strict' anaerobe *Desulfovibrio gigas* contains a membrane-bound oxygen-reducing respiratory chain. *FEBS Lett* 496: 40–43.
- Lorenz N, Reiger M, Toro-Nahuelpan M, Brachmann A, Poettinger L, Plener L, et al. (2016). Identification and initial characterization of prophages in *Vibrio campbellii*. *PLoS One* 11: 1–14.
- Lovley DR, Holmes DE, Nevin KP. (2004). Dissimilatory Fe(III) and Mn(IV) reduction. *Adv Microb Physiol* 49: 219–286.
- Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* 13: 722–736.

- Manning AJ, Kuehn MJ. (2011). Contribution of bacterial outer membrane vesicles to innate bacterial defense. *BMC Microbiol* 11: 258.
- Mercanti DJ, Carminati D, Reinheimer JA, Quiberoni A. (2011). Widely distributed lysogeny in probiotic lactobacilli represents a potentially high risk for the fermentative dairy industry. *Int J Food Microbiol* 144: 503–510.
- Muyzer G, Stams AJM. (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nat Rev Microbiol* 6: 441–454.
- Myers CR, Nealson KH. (1988). Bacterial manganese reduction and growth with manganese oxide as the sole electron acceptor. *Science* 240: 1319–1321.
- Oppenheim AB, Kobilier O, Stavans J, Court DL, Adhya S. (2005). Switches in bacteriophage lambda development. *Annu Rev Genet* 39: 409–429.
- Paquola ACM, Asif H, Pereira CA de B, Feltes BC, Bonatto D, Lima WC, et al. (2018). Horizontal gene transfer building prokaryote genomes: genes related to exchange between cell and environment are frequently transferred. *J Mol Evol* 86: 190–203.
- Pawlowski DR, Koudelka GB. (2004). The preferred substrate for RecA-mediated cleavage of bacteriophage 434 repressor is the DNA-bound dimer. *J Bacteriol* 186: 1–7.
- Pereira PM, He Q, Valente FMA, Xavier A V, Zhou J, Pereira IAC, et al. (2008). Energy metabolism in *Desulfovibrio vulgaris* Hildenborough: insights from transcriptome analysis. *Antonie Van Leeuwenhoek* 93: 347–362.
- Pietilä MK, Demina TA, Atanasova NS, Oksanen HM, Bamford DH. (2014). Archaeal viruses and bacteriophages: comparisons and contrasts. *Trends Microbiol* 22: 334–344.
- Plugge CM, Zhang W, Scholten JCM, Stams AJM. (2011). Metabolic flexibility of sulfate-reducing bacteria. *Front Microbiol* 2: 81.
- Postgate JR. (1965). Recent advances in the study of the sulfate-reducing bacteria. *Bacteriol Rev* 29: 425–441.
- Rabus R, Venceslau SS, Wöhlbrand L, Voordouw G, Wall JD, Pereira IAC. (2015). A post-genomic view of the ecophysiology, catabolism and biotechnological relevance of sulphate-reducing prokaryotes. *Adv Microb Physiol* 66: 55–321.
- Radman M. (1975). SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sci* 5A: 355–367.

- Rajasekar A, Anandkumar B, Maruthamuthu S, Ting Y-P, Rahman PKSM. (2010). Characterization of corrosive bacterial consortia isolated from petroleum-product-transporting pipelines. *Appl Microbiol Biotechnol* 85: 1175–1188.
- Rapp BJ, Wall JD. (1987). Genetic transfer in *Desulfovibrio desulfuricans*. *Proc Natl Acad Sci U S A* 84: 9128–9130.
- Reyes-Robles T, Dillard RS, Cairns LS, Silva-Valenzuela CA, Housman M, Ali A, et al. (2018). *Vibrio cholerae* outer membrane vesicles inhibit bacteriophage infection. *J Bacteriol.* e-pub ahead of print, doi: 10.1128/JB.00792-17.
- Robador A, Muller AL, Sawicka JE, Berry D, Hubert CRJ, Loy A, et al. (2016). Activity and community structures of sulfate-reducing microorganisms in polar, temperate and tropical marine sediments. *ISME J* 10: 796–809.
- Rodriguez-Rubio L, Gutierrez D, Donovan DM, Martinez B, Rodriguez A, Garcia P. (2016). Phage lytic proteins: biotechnological applications beyond clinical antimicrobials. *Crit Rev Biotechnol* 36: 542–552.
- Rosa JP da, Tiburcio SRG, Marques JM, Seldin L, Coelho RRR. (2016). *Streptomyces lunalinharesii* 235 prevents the formation of a sulfate-reducing bacterial biofilm. *Braz J Microbiol* 47: 603–609.
- Seyedirashti S, Wood C, Akagi JM. (1991). Induction and partial purification of bacteriophages from *Desulfovibrio vulgaris* (Hildenborough) and *Desulfovibrio desulfuricans* ATCC 13541. *Microbiology* 137: 1545–1549.
- Seyedirashti S, Wood C, Akagi JM. (1992). Molecular characterization of two bacteriophages isolated from *Desulfovibrio vulgaris* NCIMB 8303 (Hildenborough). *Microbiology* 138: 1393–1397.
- Shabbir MAB, Hao H, Shabbir MZ, Wu Q, Sattar A, Yuan Z. (2016). Bacteria vs. bacteriophages: parallel evolution of immune arsenals. *Front Microbiol* 7: 1292.
- Shao Q, Trinh JT, McIntosh CS, Christenson B, Balázsi G, Zeng L. (2017). Lysis-lysogeny coexistence: prophage integration during lytic development. *Microbiologyopen* 6: e00395.
- Song J, Xia F, Jiang H, Li X, Hu L, Gong P, et al. (2016). Identification and characterization of HolGH15: the holin of *Staphylococcus aureus* bacteriophage GH15. *J Gen Virol* 97: 1272–1281.
- Walker CB, Stolyar S, Chivian D, Pinel N, Gabster JA, Dehal PS, et al. (2009). Contribution of mobile genetic elements to *Desulfovibrio vulgaris* genome plasticity. *Environ Microbiol* 11: 2244–2252.

- Walker CB, Stolyar SS, Pintel N, Yen HCB, He Z, Zhou J, et al. (2006). Recovery of temperate *Desulfovibrio vulgaris* bacteriophage using a novel host strain. *Environ Microbiol* 8: 1950–1959.
- Węgrzyn G, Licznarska K, Węgrzyn A. (2012). Phage  $\lambda$ —new insights into regulatory circuits. In: Łobocka M, Szybalski WTBT-A in VR (eds) Vol. 82. *Bacteriophages, Part A*. Academic Press, pp 155–178.
- Weinbauer MG. (2006). Ecology of prokaryotic viruses. *FEMS Microbiol Rev* 28: 127–181.
- Wen J, Xu D, Gu T, Raad I. (2012). A green triple biocide cocktail consisting of a biocide, EDDS and methanol for the mitigation of planktonic and sessile sulfate-reducing bacteria. *World J Microbiol Biotechnol* 28: 431–435.
- Xu D, Li Y, Gu T. (2012). A synergistic D-tyrosine and tetrakis hydroxymethyl phosphonium sulfate biocide combination for the mitigation of an SRB biofilm. *World J Microbiol Biotechnol* 28: 3067–3074.
- Yan J, Mao J, Xie J. (2014). Bacteriophage polysaccharide depolymerases and biomedical applications. *BioDrugs* 28: 265–274.
- Zeng L, Skinner SO, Sippy J, Feiss M, Golding I. (2010). Decision making at a subcellular level determines the outcome of bacteriophage infection. *Cell* 141: 682–691.
- Zhao F, Zhou J-D, Ma F, Shi R-J, Han S-Q, Zhang J, et al. (2016). Simultaneous inhibition of sulfate-reducing bacteria, removal of H<sub>2</sub>S and production of rhamnolipid by recombinant *Pseudomonas stutzeri* Rhl: Applications for microbial enhanced oil recovery. *Bioresour Technol* 207: 24–30.

## CAPÍTULO 1

### **Screening and characterization of prophages in *Desulfovibrio* genomes**

Published: Scientific Reports (doi.org/10.1038/s41598-018-27423-z)

Josicelli Souza Crispim<sup>1</sup>, Roberto Sousa Dias<sup>1</sup>, Pedro Marcus Pereira Vidigal<sup>2</sup>,  
Maíra Paula de Sousa<sup>3</sup>, Cynthia Canêdo da Silva<sup>1</sup>, Mateus Ferreira Santana<sup>1</sup>,  
Sérgio Oliveira de Paula<sup>4\*</sup>

<sup>1</sup> Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa 36570-900, Brazil.

<sup>2</sup> Núcleo de Análise de Biomoléculas, Universidade Federal de Viçosa, Viçosa, Brazil.

<sup>3</sup> Centro de Pesquisas e Desenvolvimento Leopoldo Américo Miguez de Mello, CENPES, Rio de Janeiro, Brazil.

<sup>4</sup> Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, Brazil.

Running head: *Desulfovibrio* prophages

Keywords: Bacteriophage, CRISPR-Cas, Sulphate Reducing Bacteria, Workflow

\*Corresponding author: Sérgio Oliveira de Paula (depaula@ufv.br); phone: +55 (31) 3899-2559

## **Abstract**

Bacteria of the genus *Desulfovibrio* belong to the group of Sulphate Reducing Bacteria (SRB). SRB generate significant liabilities in the petroleum industry, mainly due to their ability to microbiologically induce corrosion, biofilm formation and H<sub>2</sub>S production. Bacteriophages are an alternative control method for SRB, whose information for this group of bacteria however, is scarce. The present study developed a workflow for the identification of complete prophages in *Desulfovibrio*. Poly-lysogenesis was shown to be common in *Desulfovibrio*. In the 47 genomes analyzed 53 complete prophages were identified. These were classified within the order *Caudovirales*, with 69.82% belonging to the *Myoviridae* family. More than half the prophages identified have genes coding for lysozyme or holin. Four of the analyzed bacterial genomes present prophages with identity above 50% in the same strain, whose comparative analysis demonstrated the existence of colinearity between the sequences. Of the 17 closed bacterial genomes analyzed, 6 have the CRISPR-Cas system classified as inactive. The identification of bacterial poly-lysogeny, the proximity between the complete prophages and the possible inactivity of the CRISPR-Cas in closed bacterial genomes analyzed allowed the choice of poly-lysogenic strains with prophages belonging to the *Myoviridae* family for the isolation of prophages and testing of related strains for subsequent studies.

## **Introduction**

Viruses have a diversity of hosts and are considered obligate intracellular parasites. Viruses that infect bacteria are referred to as bacteriophages<sup>1</sup>. These viruses have great abundance in the biosphere<sup>2</sup>, with the marine floor being the richest in bacteriophages<sup>3</sup>. When in contact with their host, bacteriophages release their genetic material into the cell, where they follow the lytic or lysogenic cycle<sup>4</sup>. The lytic cycle is characterized by viral replication and subsequent viral particle release with host cell lysis. On the other hand, the integration of the viral genome into the host genome, named the prophage, occurs in the lysogenic cycle. Prophages remain in the lysogenic state until the lytic cycle is activated through chemical and physical stresses or by spontaneous induction<sup>5,6</sup>. While they are present in the bacterial genome, these prophages are directly related to the genome diversity of the host cell, contributing positively or not, to bacterial fitness<sup>7,8</sup>. In addition to the better understanding of the involvement of bacteriophages in bacterial development, the

identification and isolation of prophages allows, (i) the development of forms of bacterial control based on genetic modifications<sup>9,10</sup>, (ii) identification of genes of interest<sup>11</sup> and (iii) infection tests in related and distinct strains<sup>12–14</sup>.

The replication and integration of the bacteriophage genetic material into the host strain can be avoided through the CRISPR-Cas system<sup>15</sup>. This system is characterized by a genomic region composed by spacers alternated with small palindromic sequences and associated with Cas proteins. Spacers are sequences related to the invader genetic material, such as bacteriophages and plasmids, inserted into this region by proteins that constitute the system. The CRISPR-Cas acts as a defense system in prokaryotes because the spacers store memory sequences against the invading target, leading to breakdown after recognition by the system<sup>16</sup>. Due to the small number of studies on prophages and CRISPR-Cas systems present in the Sulphate Reducing Bacteria (SRB) group, this area is of great interest to understand the involvement of bacteriophages in the diversity of the group and for future development of SRB control methods.

SRB are composed of anaerobic bacteria capable of reducing sulfate during metabolism, with hydrogen sulphide (H<sub>2</sub>S) as the final product<sup>17</sup>. These bacteria are ubiquitous in anaerobic environments, but have a negative impact on the oil industry. Additionally, H<sub>2</sub>S gas is toxic to workers, causes metal bio-corrosion and a loss of oil quality<sup>18,19,20</sup>. According to Rabus et al.<sup>17</sup>, there are approximately 37 bacterial genera in the SRB group. *Desulfovibrio* is the second genus that encompasses the largest number of species within this bacteria group, being characterized by Gram-negative bacteria. The first SRB was isolated in 1895 in the Netherlands by W. M. Beijerinck and named *Spirillum desulfuricans*<sup>21</sup>. Later, this strain was reclassified as the genus *Desulfovibrio*, based on cellular vibrio morphology and the ability to reduce sulfate. *Desulfovibrio vulgaris* is among the most well-known *Desulfovibrio* species and was isolated from clay soil, near Hildenborough (Kent, UK). This location gave its name to the strain, being the first SRB species sequenced and characterized as a model strain<sup>22</sup>. Another model species is *Desulfovibrio gigas*, which was isolated from a water sample from a pond in France (Etang de Berre, near Marseilles) by Jean Legall in 1963 and named for its unusual size of 11 µm<sup>23</sup>. Some strains are related to the petroleum industry, like as *Desulfovibrio alaskensis*, which was isolated in 1991 by E. van der Vende in an oil reservoir in Alaska, and once again named after its place of discovery<sup>24</sup>. The species *D. vulgaris* and *D. gigas*, although not isolated from

petroleum extraction or processing environments, are used as model strains in SRB control analyses for these environments<sup>25</sup>.

SRB genome studies are increasing, with approximately 47 *Desulfovibrio* genomes deposited in NCBI (National Center for Biotechnology Information) to date. Thus, the objective of the present study was to verify the presence of prophages and CRISPR-Cas systems in the *Desulfovibrio* genomes. Poly-lysogeny has been shown to be common in *Desulfovibrio* strains, with complete prophages having genes encoding holins and lysozyme proteins. Prophages of the *Myoviridae* family have been shown to be the most abundant and, together with the inactivity of the CRISPR-Cas system found in some bacterial strains, the use of prophages and their genes proves to be an interesting tool for SRB control. This is the first analysis of prophages present in this important SRB genus.

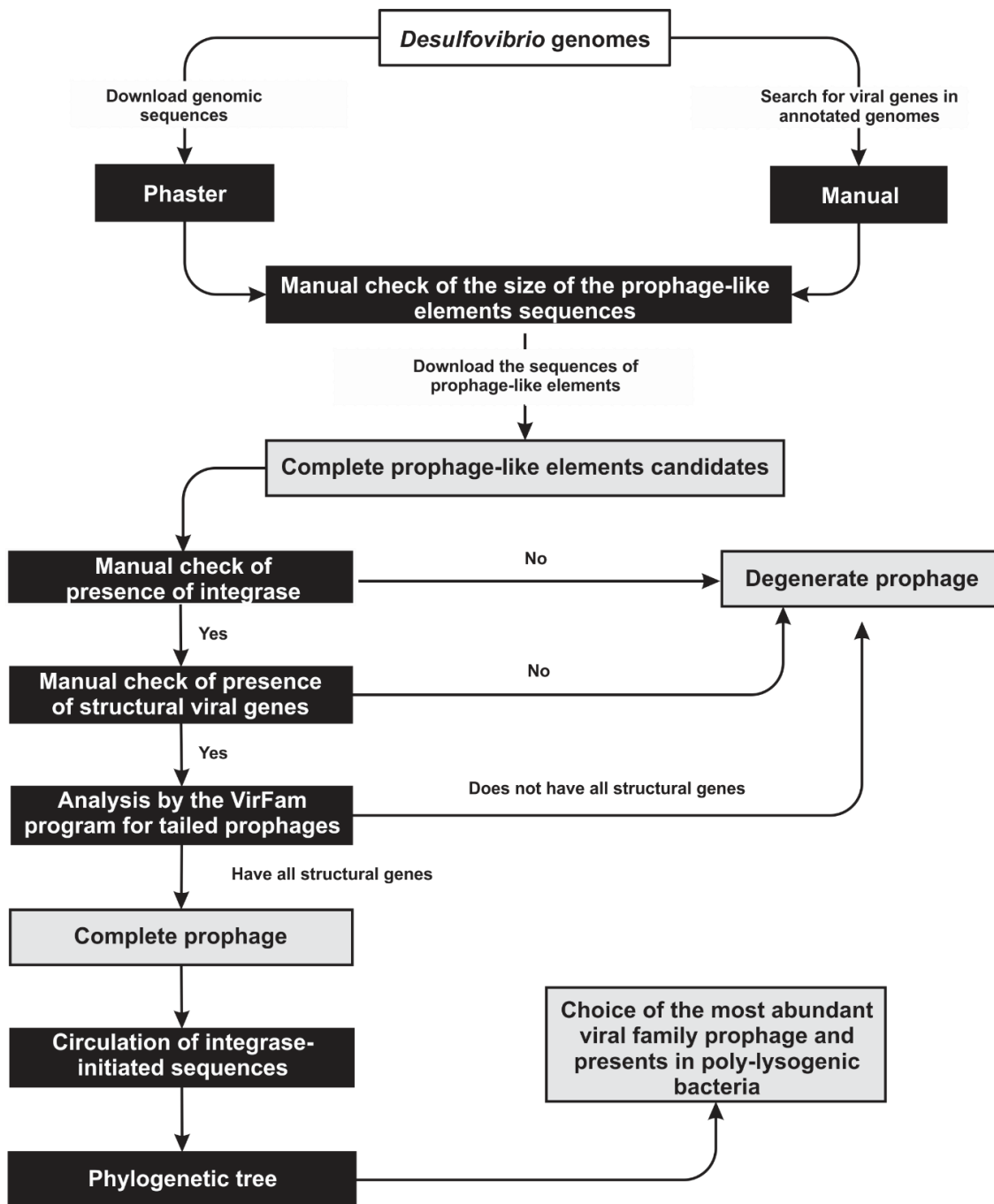
## Results

### Computational identification of *Desulfovibrio* prophages

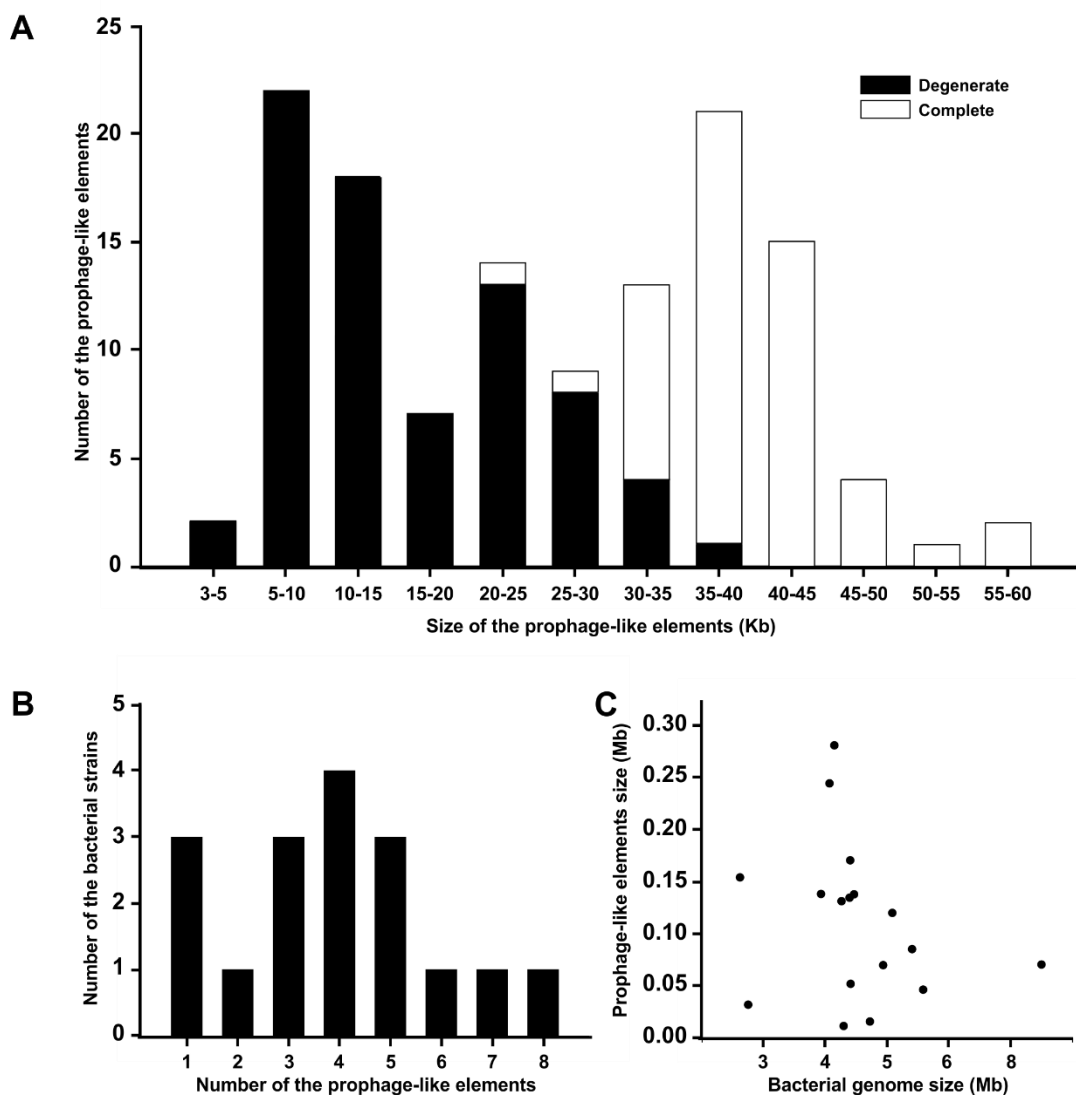
Two prophage identification strategies were used: identification through the PHASTER program and manual search for phage related sequences in annotated genomes (Fig. 1). The PHASTER program was able to identify 109 sequences related to the prophages and the manual search led to the addition of 19 sequences, totaling 128 prophage-like elements. According to the established method, the 128 elements are present in 46 of the 47 analyzed genomes (Supplementary Table S1).

The identified prophage-like sequences have an average size of 25.5 kb, the smallest (3.6 kb) found in *D. vulgaris* DP4 and the largest (57.1 kb) in *D. fairfieldensis* CCUG45958 (Supplementary Table S1). 41.40% (53) of the prophage-like elements were classified as complete prophages and 58.60% (75) as degenerate prophages (Fig. 2A). Within this classification, the prophages have a variable size, with complete prophages ranging from 20 to 60 kb (mean 39.0 kb) and the degenerate prophages from 3 to 40 kb (mean 16.09 kb), with complete and degenerate prophages with the same size being observed. All the complete prophages presented structural components related to the order *Caudovirales*, which allowed the classification of 69.82% (37) of them into the *Myoviridae* family, 22.64% (12) into *Siphoviridae* and 7.54% (4) into the *Podoviridae* family (Supplementary Table S2). Of the 53 complete prophages, 38 showed genes encoding lysozyme and 29 for the holin enzyme. Both are found in 18 complete prophages (Supplementary Table S2). Complete prophages

are present in 23 strains, with 4 being from human samples and 19 from environmental samples (Supplementary Table S2). Of these 23 strains, 14 are polylysogenic, with 5 strains presenting prophages characterized in more than one viral family (Supplementary Table S2).



**Figure 1.** Identification and characterization strategies for *Desulfovibrio* prophages. The black rectangles represent the tools used for prophage identification and characterization. The white rectangles represent the input data and the gray rectangles represent the results obtained at each step represented by arrows.



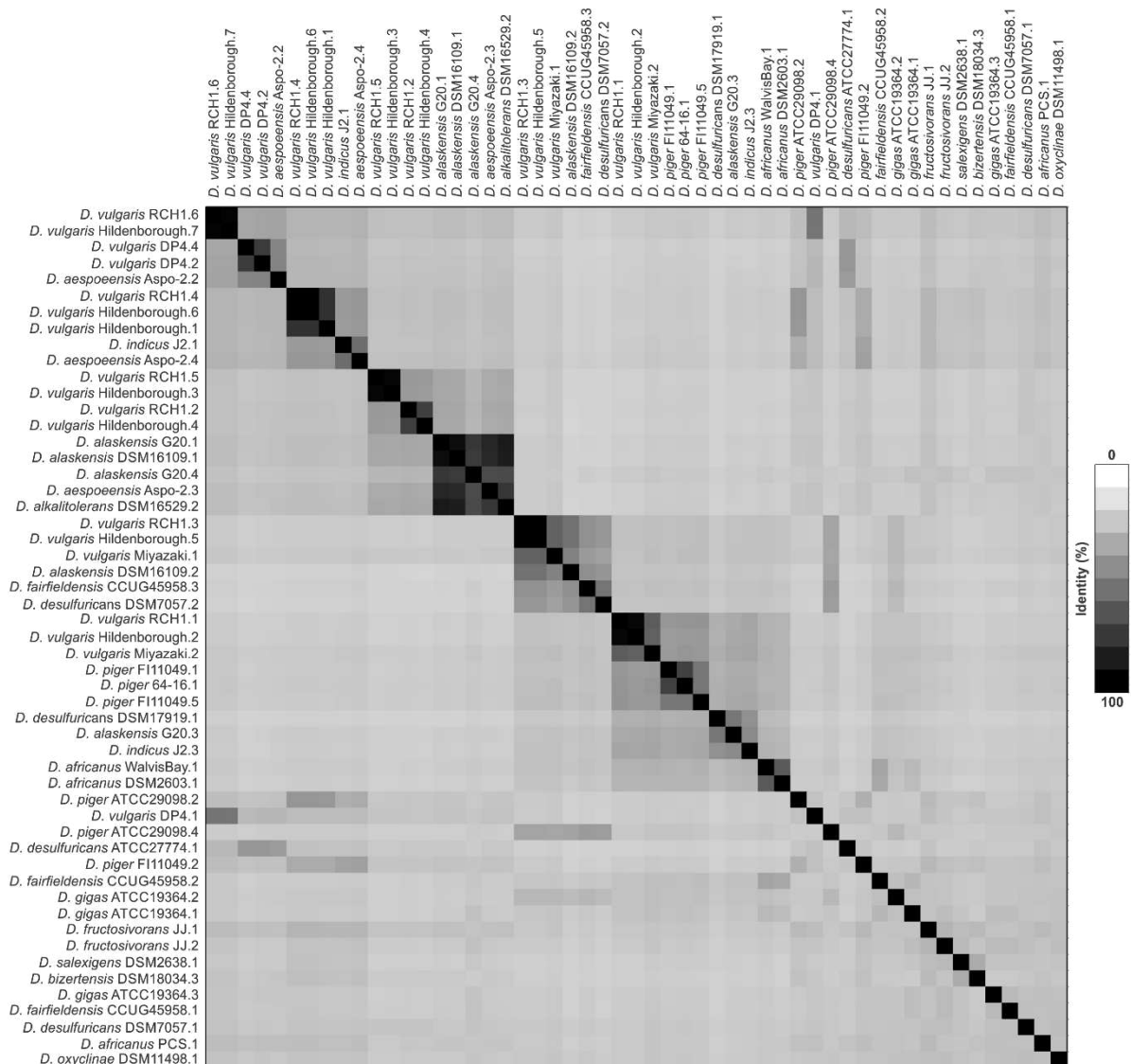
**Figure 2.** Characteristics of prophage-like elements in *Desulfovibrio*. (A) Distribution of prophage-like elements in 46 *Desulfovibrio* genomes in each category: degenerate and complete. The bars represent the number of elements corresponding to the size range of the sequence found. (B) Frequency of integration of prophage-like elements in 17 closed bacterial genomes. (C) Correlation between the size of closed bacterial genomes and prophage-like elements.

From 17 strains with closed genomes (Supplementary Table S1), 81.25% (13) presented poly-lysogenicity with strains ranging from 2 to 8 prophage-like elements (Fig. 2B). The *D. vulgaris* Hildenborough strain presented the highest number of elements, with 7 of the 8 elements classified as complete, and the highest percentage of prophage-like elements in the genome (7.81%) (Supplementary Table S1). Of the three strains that did not present poly-lysogeny, the *D. piezophilus* strain CITLV30

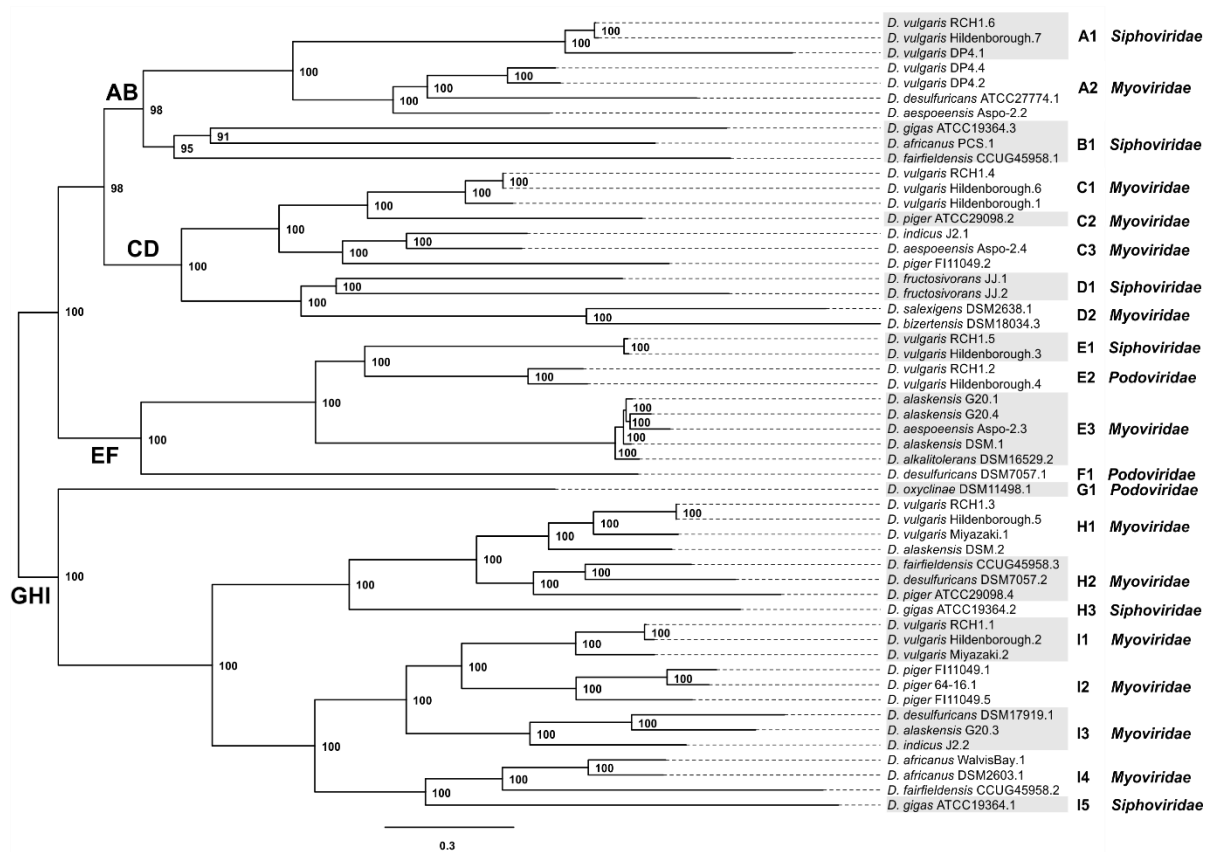
had the lowest content of prophage-like elements (0.31%). Despite this large difference in prophage content in the closed genome strains, no correlation between bacterial genome size and prophage content was observed ( $R^2 = 0.05$ ) (Fig. 2C).

### Phylogenetic and structural relationships

All 53 complete prophage sequences were aligned and a mean identity of 25.82% was found. Despite the low identity, matrix and phylogenetic tree data allowed the distinction of 9 monophyletic groups (A-I) consisting of 21 subgroups (A1-2, B1, C1-3, D1-2, E1-3, F1, G1, H1 -3 and I1-5) (Fig. 3 and 4).



**Figure 3.** Heatmap among 53 complete *Desulfovibrio* prophages. The map describes the identity between two pairs of sequences.



**Figure 4.** Phylogenetic relationships between complete *Desulfovibrio* prophage sequences. In the phylogenetic tree we can distinguish 9 phylogenetic groups consisting of 21 subgroups. The putative viral families of the subgroups are indicated to the right of the tree.

The phylogenetic analysis showed that complete prophages present a great diversity with prophages belonging to the same viral family forming different phylogenetic clusters, with 12 subgroups for *Myoviridae*, 6 for *Siphoviridae* and 3 for the *Podoviridae* family being found (Fig. 4). The phylogenetic tree also allowed the identification of subgroups with prophages belonging to the same bacterial species (A1, E1-2 and I1-2), which may indicate an acquisition of related prophages. In the present study, poly-lysogenesis has been shown to be common in the genus *Desulfovibrio*. Analyzing the prophages belonging to the same strain, we found that they are distributed among the phylogenetic subgroups, such as the *Desulfovibrio vulgaris* Hildenborough strain, which has elements in subgroups A1, C1, E1, E2, H1 and I1. This indicates the presence of various prophages found in a single bacterial genome.

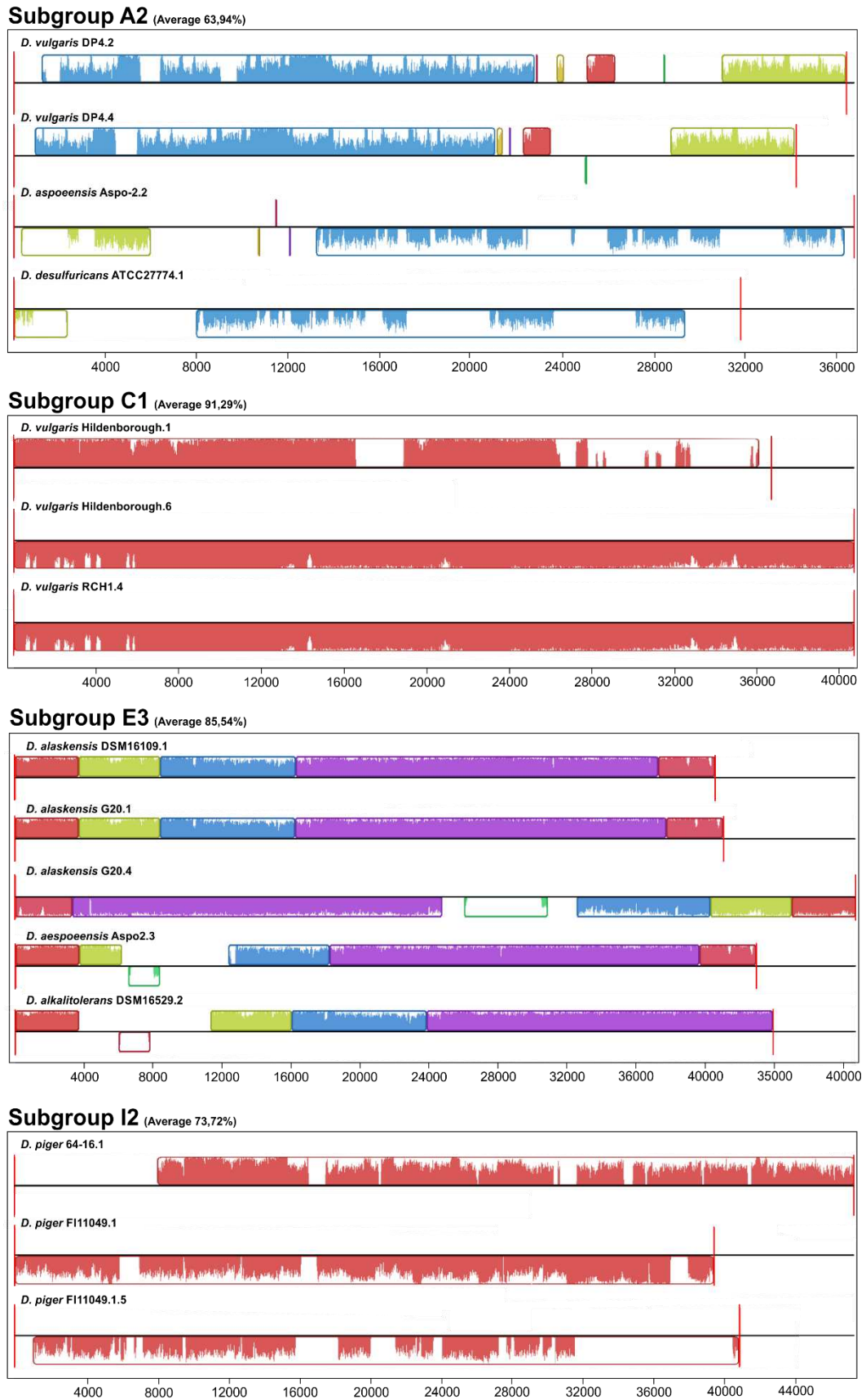
The 9 groups were organized into 4 larger clusters (AB, CD, EF and GHI) according to the phylogenetic relationship (Fig. 4). The AB, CD, EF and GHI clusters

have 479, 576, 417 and 1023 proteins distributed in 125, 161, 139 and 122 groups of orthologs, respectively (Supplementary Fig. S1). A Venn Diagram shows that only 3 groups of proteins are shared between all phylogenetic groupings, with most protein groups being specific to each cluster (Supplementary Fig. S1). The groups of orthologs shared between all clusters (AB, CD, EF and GHI) were related to transcriptional regulators (CI and Cro), structural tail protein and integrase.

Despite the diversity of prophages in poly-lysogenic strains, closely related prophages were also found in 4 strains (Fig. 4). The *D. vulgaris* DP4 strain (subgroup A2), *D. vulgaris* Hildenborough (subgroup C1), *D. alaskensis* G20 (subgroup E3) and *D. piger* FI11049 (subgroup I2) present two prophages with an identity above 50% (Fig. 3). In the D1 subgroup, prophages belonging to the same strain, *D. fructosivorans*, were also found, but these prophages have low similarity (29%) between nucleotide sequences (Fig. 3). The synteny analysis of the A2, C1, E3 and I2 subgroups showed that all sequences of complete prophages present collinearity, which is characterized by an absence of rearrangement of gene blocks (Fig. 5). The difference between prophages occurs through insertions and deletions. In the E3, the main difference between these two prophages of *D. alaskensis* G20 is in a region present in prophage 4 and absent in prophage 1. This region of prophage 4 is related to bacterial genes, demonstrating a relationship of the prophages with horizontal transference of genes.

### **Analysis of CRISPR-Cas systems in *Desulfovibrio***

Seeking to understand the acquisition of prophages, the CRISPR-Cas system of *Desulfovibrio* strains with closed genomes was analyzed. Of the 17 strains analyzed, 64.7% (11) have complete CRISPR arrangements, and 1 to 5 arrangements with a total variation of 4 to 98 spacers were found (Table 1). In relation to the locus *cas*, 14 loci belonging to 5 subtypes (I-C, I-E, I-F, I-U and III-B) were found (Table 1 and Supplementary Fig. S2). These loci were found in 11 strains, with three strains presenting more than one locus, at least one of them being complete (Table 1). The *D. hydrothermalis* DSM14728 and *D. magneticus* RS-1 strains have 2 complete *cas* loci, with the subtype I-F being common to both strains, and subtype III-B and I-C present in each strain, respectively. The *D. gigas* ATCC19364 strain has two *cas* loci, the subtype I-F being classified as complete and the subtype U-I as incomplete due to its not presenting Cas8, Cas4/Cas1 and Cas2 proteins, according to the classification given by Makarova et al.<sup>16</sup>.



**Figure 5.** Comparison of prophage sequences of subgroups A2, C1, E3 and I2. The collinearity between the sequences is represented by the conservation of the location of the blocks in all subgroups.

The type I-C *cas* locus of the DP4, Hildenborough and RCH1 strains of *D. vulgaris* have 97% identity. This result differs from the comparison with the *cas* locus of the *D. vulgaris* Miyazaki strain, which presents coverage of only 1% in relation to the loci of the other strains of the species. Unlike *D. vulgaris*, *D. desulfuricans* strains present different types of *cas* loci.

We are unable to classify the activity of the system using only *in silico* analysis, however the complete CRISPR-Cas systems were classified as possibly active, being found in 64.70% of the closed genome strains (Table 1). No positive correlation was found between the number of spacers in the strains with possibly active CRISPR-Cas and the extent of the prophage content. The CRISPR-Cas system characteristics of the closed genomes were re-evaluated by the CRISPRone software (Supplementary Table S3), confirming previous findings.

**Table 1** CRISPR-Cas system in closed genomes of *Desulfovibrio*

N°	Strain	N° of prophage-like elements	CRISPR arrays (n° of spacers)	CRISPR-Cas systems type	Putative activity
1	<i>D. aespoeensis</i> Aspo-2	4	ND	ND	Inactive
2	<i>D. africanus</i> Walvis Bay	4	ND	ND	Inactive
3	<i>D. alaskensis</i> G20	4	1 (19)	I-E	Active
4	<i>D. desulfuricans</i> ND132	3	1 (98)	I-C	Active
5	<i>D. desulfuricans</i> ATCC 27774	1	2 (29)	I-E	Active
6	<i>D. fairfieldensis</i> CCUG45958	5	Questionable	ND	Inactive
7	<i>D. gigas</i> ATCC19364	3	1 (4)	I-F e I-U Incomplete	Active
8	<i>D. hydrothermalis</i> DSM14728	3	5 (35)	III-B e I-F	Active
9	<i>D. indicus</i> J2	2	1 (19)	I-C	Active
10	<i>D. magneticus</i> RS-1	5	1 (30)	I-C e I-F	Active
11	<i>D. piezophilus</i> C1TLV30	1	Questionable	ND	Inactive
12	<i>D. piger</i> F111049	5	Questionable	ND	Inactive
13	<i>D. salexigens</i> DSM2638	1	Questionable	ND	Inactive
14	<i>D. vulgaris</i> DP4	6	1 (44)	I-C	Active
15	<i>D. vulgaris</i> Hildenborough	8	1 (27)	I-C	Active
16	<i>D. vulgaris</i> Miyazaki	4	2 (54)	I-C	Active
17	<i>D. vulgaris</i> RCH1	7	1 (27)	I-C	Active

## Discussion

The *in silico* workflow developed in the present study allowed the identification of 128 prophage-like elements in 46 *Desulfovibrio* genomes. However, other prophage-like elements can be identified in the future because most of the genomes used in the analysis (56.60%) were deposited in contigs form and without gene annotation, as with some strains identified in this study that did not present prophage-like elements (*D. gracilis* DSM 16080). Of the 17 closed *Desulfovibrio* genomes, 82.35% (14) present poly-lysogeny. No relationship was found between the genomic size of poly-lysogenic strains and the content of prophage-like elements that they

possess. On the other hand, this data may be modified due to the low number of closed genomes available in the database. In contrast, a positive correlation was found in other bacterial species such as *Cronobacter sakazakii*, where the increase of bacterial genomes is found to contain a greater number of prophages<sup>26,27</sup>. The maximum value of the genomic viral content was 7.8% in *D. vulgaris* Hildenborough. A maximum value of 20% viral content has already been detected in *Borrelia burgdorferi*<sup>28</sup>. Some of the prophages of *D. vulgaris* Hildenborough have already been described in the literature<sup>12,29</sup>. Handley et al.<sup>29</sup> described 50 nm icosahedral and contractile tail structures present in the supernatant of the strain induced by mitomycin C. Walker et al.<sup>12</sup> used a closely related *D. vulgaris* DePue strain as a recipient for prophages induced by mitomycin C of the *D. vulgaris* Hildenborough, finding structures with an icosahedral head of 50 and 100 nm in diameter. This work is an indication that the prophage-like elements of *Desulfovibrio* are induced by mitomycin C, with its isolation from related strains being possible. Additionally, the sequence analysis of these two related strains showed that the DePue shares only one of the seven phage-related elements present in the Hildenborough, demonstrating that the genetic diversity of these strains is mainly related to the regions of the prophages<sup>7</sup>.

53 elements were classified as complete, characterizing possible *Desulfovibrio* prophages. The 75 elements classified as degenerate may be related to gene transfer agents<sup>30</sup>. The complete prophages showed an average size of 39 kb, which is close to that of caudal prophages present in the bacterial Enterobacteriaceae family, with approximately 40-50 Kb<sup>31</sup>. In addition, 71.7% of the prophages have genes encoding lysozymes and 54.7% have genes encoding holins. Lysozymes are peptidoglycan hydrolases and holins cause plasmatic membrane damage, the action of these enzymes leading to cell disruption<sup>11,32</sup>. The presence of these enzymes characterizes these phages as infective, highlighting an attractive feature for their isolation by induction and infection assays with different strains, as seen by Walker et al.<sup>12</sup>. Additionally, these genes can be used for the development of bacterial control tools, such as protein expression in bacterial systems for subsequent environmental application<sup>11</sup>.

Complete prophages found have genes related to the tail, being classified into three families in the order *Caudovirales*, with 69.82% of the prophages belonging to the *Myoviridae* family. Poly-lysogeny was found in different phage families in the

same bacterial genome. The presence of prophages belonging to the different viral families is also seen in *Vibrio campbellii*, which present four complete prophages, two related to the *Myoviridae* and two to the *Inoviridae* family<sup>33</sup>. The classification by viral family performed in this study was according to the proteins of head and tail connectivity, a parameter used by the VirFam program. Sequence alignment was performed using the total prophage sequences. The sequences can share other genes, such as integrases, transcriptional regulators and tail proteins, as found in OrthoVenn results (Supplementary Fig. S1). The sharing of these other genes may explain the diversification of viral families in the phylogenetic tree (Fig. 4). Although most of the complete prophages have genes encoding holins and lysozymes, no shared orthologous protein was found among the orthologous protein in intersectional groups, indicating that the *Desulfovibrio* prophages exhibit significant protein diversity (Supplementary Fig. S1). This result indicates the need for several holin and lysozyme conjunctions for the control of *Desulfovibrio* strains. In addition, many hypothetical genes are also found among the complete prophages, as described for other phages<sup>34,35</sup>. The analysis of these genes may promote a better understanding of the prophage-host interactions and reveal other interesting genes.

Although a great diversity of complete prophages were found, the presence of similar elements within the same strain was also seen, which was analyzed by synteny analysis (Fig. 5). Related prophages were found in *D. vulgaris* DP4 (subgroup A2), *D. vulgaris* Hildenborough (subgroup C1), *D. alaskensis* G20 (subgroup E3) and *D. piger* FI111049 (subgroup I2) (Fig. 5). Although the *D. alaskensis* G20 and DSM16109 strains show prophages with 95.16% identity, the DSM16109 does not present related prophages in its genome (Fig. 5). The *D. alaskensis* DSM16109 is deposited in contig form, which may mask the presence of another similar prophage. The same is true for *D. piger* 64-16, which is deposited in contig form. Differently, the *D. vulgaris* RCH1 has a closed genome and did not present a related prophage, such as the strain of the same species *D. vulgaris* Hildenborough (Fig. 5). The results showed that some prophages found in the same species remain unchanged between the strains, while others have some changes at the nucleotide level (Fig. 5). Selective pressures maintain these prophages in the genome, therefore having some beneficial characteristics for the host<sup>31</sup>. The same was found for *Helicobacter pylori* strains, whose prophages showed low variation between strains from the same species<sup>35</sup>. On the other hand, the acquisition of related

prophages is a characteristic of the strain and not of the species in *Desulfovibrio*. One factor that may be involved in this acquisition is the CRISPR-Cas system, whose spacers have a wide variety in the strains, as found for *Salmonella*<sup>36</sup>.

*Desulfovibrio* strains mostly have type I *cas* loci, but type III was also found in one of the 17 strains analyzed with closed genomes. The presence of more than one *cas* locus allows the system to have different targets, such as type-I and type-II that have DNA strands and type III that have RNA strands as a target<sup>37</sup>. Three strains presented more than one *cas* locus. This allows the absence of a given gene at one locus to be compensated for by the presence of a related gene at another locus<sup>16</sup>, as with *D. gigas* ATCC19364, which presents one of these incomplete *cas* loci. The *cas* locus of the subtype I-C has been shown to be the most common among the *Desulfovibrio* strains, representing 50% (7) of the total *cas* loci observed (Table 1 and Supplementary Fig. S2). Although it is the most common, this subtype showed variability between strains from the same species, as found for *D. vulgaris* (Table 1). In addition, the *D. desulfuricans* species presents *cas* loci of different subtypes between the two analyzed strains (Table 1). This demonstrates that variability of *cas* loci is common within the same species.

The absence of correlation between the number of spacers in possibly active CRISPR-Cas and the size of the prophage content found in *Desulfovibrio* strains was demonstrated in other bacterial strains<sup>26</sup>. On the other hand, *Cronobacter sakazakii* strains presented lower prophage content with the greater presence of spacers<sup>27</sup>. This relationship found in *C. sakazakii* demonstrates the involvement of the CRISPR system in the protection against bacteriophages, which was not found in *Desulfovibrio* strains, according to the closed genomes analysed in this work. The spacer composition of the CRISPR system was also analyzed by Walker et al.<sup>7</sup> in *D. vulgaris* DePue, the strain used as receptor for induced prophages of *D. vulgaris* Hildenborough. The authors found no similarity between the spacers and the prophages, a feature that allowed the infection of the recipient bacterium by prophages of the Hildenborough strain, since they were not recognized by the CRISPR-Cas system. The presence of prophages in bacteria with important role in the food industry has a negative aspect since that can be induced and follow the lytic cycle<sup>13</sup>. In the present study, the presence of these prophages in the strains analyzed is of great importance for the discovery of phages that can be used in SRB control.

## Conclusions

Given these results, the attempt to isolate prophages using nearby receptor host cells is feasible for this group, since part of the strains with closed genomes present the possibly inactive CRISPR-Cas system. There is also an indication that the spacers of the *Desulfovibrio* CRISPR-Cas system do not protect against bacteriophages. Additionally, the presence of related prophages in different strains demonstrated to be common, indicates that a bacteriophage could have the ability to infect strains of the same bacterial genus. All these characteristics demonstrate that bacteriophages have great potential for SRB control, because a single bacteriophage could infect different strains without being recognized by the CRISPR-Cas system. Some bacteriophages can be isolated by inducing prophages inserted in bacterial genomes, using related bacterial species as new host. The strains that show the greatest potential for prophage isolation are those that present the highest number of complete prophages belong to the *Myoviridae* family, which is the most common family in the bacterial genus. Among the strains with this potential are *D. vulgaris* (Myiazaki, Hildenborough, RCH1 and DP4), *D. piger* FI11049, *D. fairfieldenses*, *D. alaskensis* G20 and *D. aespoeensis*. Although the strains *D. piger* and *D. fairfieldenses* are of human origin, they present prophages that can be used for control of bacteria of environmental origin. In addition, many genus related to holins and lysozymes have been identified, which can be used for the development of control tools. On the other hand, the function of many genes of the prophages characterized in this study are not known, representing a vast reservoir of new information to be determined.

## Methods

### Identification of prophages and CRISPR-Cas system

Forty-seven genomes belonging to 33 species of the genus *Desulfovibrio* were downloaded from NCBI Genome database (<https://www.ncbi.nlm.nih.gov/genome/>), with 17 complete genome sequences. The presence of prophages in the genomes was initially verified using the PHASTER (PHAge Search Tool Enhanced Release) (<http://phaster.ca/>) web server<sup>38</sup>. The prophage sequences identified by PHASTER were manually inspected using the integrase position and the last phage-related gene as the criterion to determine the genome boundaries.

The annotation of complete bacterial genomes was also manually inspected by text-mining to screen for the presence of other phage-related regions containing proteins such as integrase, tail and capsid. Prophage sequences were screened for the presence of genes encoding integrase and structural genes, with the hypothetical proteins being further analyzed by PSI-BLAST ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE=Proteins&PROGRAM=blastp&RUN\\_PSIBLAST=on](https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE=Proteins&PROGRAM=blastp&RUN_PSIBLAST=on))<sup>39</sup>. The presence of structural genes in tailed prophage sequences was verified by the VirFam (<http://biodev.extra.cea.fr/virfam/>) web server<sup>40</sup>. The prophages, which in addition to integrase, have all structural genes for classification within the families *Myoviridae*, *Siphoviridae* and *Podoviridae*, were classified as complete prophages. The prophages that do not have integrase or that show integrase also in addition to an absence of genes related to the viral structure were classified as degenerate prophages.

CRISPR-Cas systems were detected by the CRISPRFinder (<http://crispr.i2bc.paris-saclay.fr/Server/>) program<sup>41</sup> and *cas* genes by manual analysis of closed genomes. The results obtained by these analyzes were confirmed by the CRISPRone program (<http://omics.informatics.indiana.edu/CRISPRone/>).

### **Genomic analysis of complete prophages**

The complete prophage sequences were manually analyzed using Geneious version 9.1 (Biomatters Lts), adjusting the direction of nucleotide sequences and opening the genomes for the integrase gene. Multiple sequence alignment was performed on Geneious using the ClustaW algorithm with standard parameters<sup>42</sup>. A pairwise identity matrix was also calculated by Geneious from the alignment of genome sequences.

The GTR+G model was selected by jModeltest version 2.1.10<sup>43</sup> as the best fit nucleotide substitution model for phylogenetic analysis. The phylogenetic tree was obtained by Bayesian Inference using the MrBayes version 3.2.6<sup>44</sup>. A Bayesian Markov Chain Monte Carlo (MCMC) analysis was performed in two runs with 5000000 generations. The parameter convergence was analyzed and 10% of the trees generated were burnt to produce the consensus tree.

For comparative purposes, the proteins shared between the phage groups that were separated in the phylogenetic tree, were identified using OrthoVenn (<http://www.bioinfo-genome.net/OrthoVenn/>) web server<sup>45</sup>. In addition, a comparative

genomic analysis between phages from each group was performed using Mauve version 2.3.1<sup>46</sup>.

### **Acknowledgements**

We thank CNPq, FAPEMIG, Capes and Petrobras for their support.

### **Author Contributions**

J.S.C., R.S.D., P.M.P.V and M.F.S. developed the prophage screening workflow. J.S.C. and P.M.P. processed data. R.S.D., M.P.S., C.C.S., M.F.S. and S.O.P assisted in data analysis. J.S.C. prepared manuscript figures and wrote the manuscript. All authors analysed data and reviewed the manuscript before submission.

### **Additional Information**

**Competing Interests:** The authors declare that they have no competing interests.

### **References**

1. Sharma, S. et al. Bacteriophages and its applications: an overview. *Folia Microbiol. (Praha)*. 62, 17–55 (2017).
2. Ackermann, H.-W. 5500 Phages examined in the electron microscope. *Arch. Virol.* 152, 227–243 (2007).
3. Mizuno, C. M., Ghai, R., Saghai, A., López-García, P. & Rodriguez-Valera, F. Genomes of abundant and widespread viruses from the deep ocean. *MBio* 7, 1–9 (2016).
4. Erez, Z., Steinberger-levy, I., Shamir, M., Doron, S. & Stokar, A. Communication between viruses guides lysis-lysogeny decisions. 541, 488–493 (2017).
5. Gandon, S. Why be temperate: lessons from bacteriophage  $\lambda$ . *Trends Microbiol.* 24, 356–365 (2016).
6. Owen, S. V. et al. Characterization of the prophage repertoire of African *Salmonella* Typhimurium ST313 reveals high levels of spontaneous induction of novel phage BTP1. *Front. Microbiol.* 8, 1–17 (2017).
7. Walker, C. B. et al. Contribution of mobile genetic elements to *Desulfovibrio vulgaris* genome plasticity. *Environ. Microbiol.* 11, 2244–2252 (2009).
8. Brueggemann, A. B. et al. Pneumococcal prophages are diverse, but not without structure or history. *Sci. Rep.* 7, 1–13 (2017).

9. Hudson, J., Billington, C., Premaratne, A. & On, S. Inactivation of *Escherichia coli* O157:H7 using ultraviolet light-treated bacteriophages. *Rev. Agroquimica y Tecnol. Aliment.* 22, 3–9 (2016).
10. Chen, M. et al. Inducible prophage mutant of *Escherichia coli* can lyse new host and the key sites of receptor recognition identification. *Front. Microbiol.* 8, 1–13 (2017).
11. Song, J. et al. Identification and characterization of HolGH15: the holin of *Staphylococcus aureus* bacteriophage GH15. *J. Gen. Virol.* 97, 1272–1281 (2016).
12. Walker, C. B. et al. Recovery of temperate *Desulfovibrio vulgaris* bacteriophage using a novel host strain. *Environ. Microbiol.* 8, 1950–1959 (2006).
13. Mercanti, D. J., Carminati, D., Reinheimer, J. A. & Quiberoni, A. Widely distributed lysogeny in probiotic lactobacilli represents a potentially high risk for the fermentative dairy industry. *Int. J. Food Microbiol.* 144, 503–510 (2011).
14. Hamdi, S. et al. Characterization of two polyvalent phages infecting Enterobacteriaceae. *Sci. Rep.* 7, (2017).
15. Edgar, R. & Qimron, U. The *Escherichia coli* CRISPR system protects from  $\lambda$  lysogenization, lysogens, and prophage induction. *J. Bacteriol.* 192, 6291–6294 (2010).
16. Makarova, K. S. et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736 (2015).
17. Rabus, R. et al. A post-genomic view of the ecophysiology, catabolism and biotechnological relevance of sulphate-reducing prokaryotes. *Adv. Microb. Physiol.* 66, 55–321 (2015).
18. Gieg, L. M., Jack, T. R. & Foght, J. M. Biological souring and mitigation in oil reservoirs. *Appl. Microbiol. Biotechnol.* 92, 263–282 (2011).
19. Vigneron, A. et al. Complementary microorganisms in highly corrosive biofilms from an offshore oil production facility. *Appl. Environ. Microbiol.* 82, 2545–2554 (2016).
20. Tian, H. et al. Compositions and abundances of sulfate-reducing and sulfur-oxidizing microorganisms in water-flooded petroleum reservoirs with different temperatures in China. *Front. Microbiol.* 8, 1–14 (2017).
21. Voordouw, G. The genus *Desulfovibrio*: the centennial. *Appl. Environ. Microbiol.* 61, 2813–2819 (1995).

22. Heidelberg, J. F. et al. The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *Nat. Biotechnol.* 22, 554–559 (2004).
23. Legall, J. a New Species of *Desulfovibrio*. *J. Bacteriol.* 86, 1120 (1963).
24. Feio, M. J. et al. *Desulfovibrio alaskensis* sp. nov., a sulphate-reducing bacterium from a soured oil reservoir. *Int. J. Syst. Evol. Microbiol.* 54, 1747–1752 (2004).
25. Lavania, M., Sarma, P. M., Mandal, A. K., Cheema, S. & Lai, B. Efficacy of natural biocide on control of microbial induced corrosion in oil pipelines mediated by *Desulfovibrio vulgaris* and *Desulfovibrio gigas*. *J. Environ. Sci.* 23, 1394–1402 (2011).
26. Touchon, M., Bernheim, A. & Rocha, E. P. C. Genetic and life-history traits associated with the distribution of prophages in bacteria. *ISME J.* 10, 2744–2754 (2016).
27. Zeng, H. et al. The driving force of prophages and CRISPR-Cas system in the evolution of *Cronobacter sakazakii*. *Sci. Rep.* 7, (2017).
28. Casjens, S. Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* 49, 277–300 (2003).
29. Handley, J., Adams, V. & Akagi, J. M. Morphology of bacteriophage-like particles from *Desulfovibrio vulgaris*. *J. Bacteriol.* 115, 1205–7 (1973).
30. Lang, A. S., Zhaxybayeva, O. & Beatty, J. T. Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* 10, 472–482 (2012).
31. Bobay, L.-M., Touchon, M. & Rocha, E. P. C. Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci.* 111, 12127–12132 (2014).
32. Joshi, H. et al. Dissecting the structure–function relationship in lysozyme domain of mycobacteriophage D29-encoded peptidoglycan hydrolase. *FEBS Lett.* 591, 3276–3287 (2017).
33. Lorenz, N. et al. Identification and initial characterization of prophages in *Vibrio campbellii*. *PLoS One* 11, 1–14 (2016).
34. Pope, W. H. et al. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife* 4, e06416 (2015).
35. Vale, F. F. et al. Genomic structure and insertion sites of *Helicobacter pylori* prophages from various geographical origins. *Sci. Rep.* 7, 1–12 (2017).

36. Shariat, N., Timme, R. E., Pettengill, J. B., Barrangou, R. & Dudley, E. G. Characterization and evolution of *Salmonella* CRISPR-Cas systems. *Microbiology* 161, 374–386 (2015).
37. Van Der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 12, 479–492 (2014).
38. Arndt, D. et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21 (2016).
39. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402 (1997).
40. Lopes, A., Tavares, P., Petit, M.-A., Guérois, R. & Zinn-Justin, S. Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics* 15, 1027 (2014).
41. Grissa, I., Vergnaud, G. & Pourcel, C. CRISPRcompar: a website to compare clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 36, 52–57 (2008).
42. Thompson, J. D., Gibson, T. J. & Higgins, D. G. in *Current protocols in bioinformatics* (John Wiley & Sons, Inc., 2002).
43. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772 (2012).
44. Ronquist, F. et al. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542 (2012).
45. Wang, Y., Coleman-Derr, D., Chen, G. & Gu, Y. Q. OrthoVenn: A web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 43, W78–W84 (2015).
46. Darling, A. E., Mau, B. & Perna, N. T. Progressivemauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5, (2010).

## Supplementary information

### Supplementary Table S1 Prophage-like elements of *Desulfovibrio*

N°	Strain	Access number	GC%	Prophage-like elements	% Genome	Contig (location)	Size (pb)	Classification	GC %	CDS
1	<i>D. vulgaris</i> Hildenborough	NC_002937.3	63.1	1	1,02	233752-270425	36674	Complete	63.7	48
2				2	0,99	1205411-1241031	35621	Complete	64.8	47
3				3	1,10	1556492-1595902	39411	Complete	58.2	52
4				4	1,02	1770879-1807430	36552	Complete	59.9	45
5				5	0,98	2799302-2834336	35035	Complete	64.9	48
6				6	1,13	2936472-2977150	40679	Complete	62.6	54
7				7	1,24	2250828-2295194	44367	Complete	60.8	51
8				8	0,33	2714953-2726897	11945	Degenerate	67.6	17
9	<i>D. vulgaris</i> RCH1	NC_017310.1	63.1	1	1,14	1207755-1248036	40282	Complete	63.1	52
10				2	1,05	1760491-1797698	37208	Complete	60.1	49
11				3	0,99	2802434-2837468	35035	Complete	65.9	49
12				4	1,15	2939604-2980282	40679	Complete	62.6	53
13				5	0,98	1987316-2022273	34958	Complete	59.0	47
14				6	1,51	2254349-2297214	42866	Complete	61.0	45
15				7	0,36	2718253-2731058	12806	Degenerate	67.1	19
16	<i>D. piger</i> FI11049	NZ_LT630450.1	64.2	1	1,40	25995-65412	39418	Complete	60.2	50
17				2	1,51	148736-191335	42600	Complete	46.8	56
18				3	0,25	777899-785196	7298	Degenerate	48.1	7
19				4	0,83	786568-810100	23533	Degenerate	63.1	31
20				5	1,45	430305-471180	40876	Complete	62.0	50
21	<i>D. piger</i> ATCC29098	NZ_ABXU00000000.1	63.0	1	1,33	3.3:858-37806	37806	Degenerate	63.2	50
22				2	1,34	4.1:16868-54731	37864	Complete	62.1	52
23				3	0,22	4.3:1-6352	6352	Degenerate	69.3	8
24				4	1,24	4.4:1-35308	35309	Complete	59.2	51
25				5	1,03	4.5:1-29333	29333	Degenerate	58.5	48
26	<i>D. fairfieldensis</i> CCUG45958	NZ_CP014229.1	60.9	1	1,54	1076345-1133484	57140	Complete	59.6	63
27				2	1,04	1555216-1593958	38473	Complete	62.5	53
28				3	0,82	2149353-2186193	30416	Complete	61.9	36
29				4	0,59	263020-284989	21970	Degenerate	58.3	18
30				5	0,59	484160-506251	22092	Degenerate	64.3	25
31	<i>D. vulgaris</i> DP4	NC_008751.1	63.0	1	1,23	1272991-1315732	42742	Complete	61.6	50
32				2	1,05	1745393-1781803	36411	Complete	61.6	48
33				3	0,10	1779304-1782906	3603	Degenerate	58.1	7
34				4	0,98	3198595-3232809	34215	Complete	62.3	42

35				5	0,34		1068859-1080736	11878	Degenerate	57.7	14
36				6	0,26		1271645-1280709	9065	Degenerate	59.3	13
37				1	1,10		939853-980902	41050	Complete	63.1	42
38	<i>D. alaskensis</i> G20	NC_007519.1	57.8	2	0,20	3,67	2787989-2795585	7597	Degenerate	47.8	9
39				3	1,07		3331749-3371978	40230	Complete	55.2	52
40				4	1,30		1918124-1966812	48689	Complete	60.7	50
41				1	1,21		990478-1035433	44956	Complete	65.1	56
42	<i>D. gigas</i> ATCC19364	NC_022444.1	63.8	2	0,89	3,62	2474033-2507188	33156	Complete	63.1	42
43				3	1,52		2127703-2183892	56190	Complete	54.9	52
44				1	0,36		352121-365295	13175	Degenerate	64.9	8
45	<i>D. aespoeensis</i> Aspo-2	NC_014844.1	62.6	2	1,01	3,59	474227-510996	36770	Complete	61.8	51
46				3	1,18		1775493-1818421	42929	Complete	42.9	44
47				4	1,04		2595084-2633147	38064	Complete	38.0	58
48				1	0,91		864144-901131	36988	Complete	65.2	43
49	<i>D. vulgaris</i> Miyazaki	NC_011769.1	67.1	2	0,97	2,94	1558929-1598331	39403	Complete	65.2	46
50				3	0,56		2016554-2039449	22896	Degenerate	64.4	30
51				4	0,50		1302496-1322867	20372	Degenerate	70.2	22
52				1	1,35		5:2749-46467	43719	Complete	55.2	54
53	<i>D. desulfuricans</i> DSM7057	NZ_FPIW00000000.1	57.3	2	0,88	2,89	8:58080-86702	28623	Complete	58.2	39
54				3	0,39		50:8532-21340	12809	Degenerate	60.6	12
55				4	0,27		80:3-9082	9082	Degenerate	58.1	13
56				1	0,79		49:33508-70620	37113	Complete	62.4	48
57	<i>D. fructosivorans</i> JJ	NZ_AECZ00000000.1	63.9	2	1,08	2,59	49:138295-188970	50676	Complete	61.3	64
58				3	0,72		55:63769-97635	33867	Degenerate	64.2	48
59	<i>D. alkalitolerans</i> DSM16529	NZ_ATHI00000000.1	64.5	1	1,11	2,48	12:30050-65856	35807	Degenerate	62.6	47
60				2	1,37		02:19980-63856	43877	Complete	62.1	45
61				1	0,88		2:441837-473030	31194	Complete	49.5	47
62	<i>D. desulfuricans</i> DSM 17919	NZ_FQZR00000000.1	45.1	2	0,60	2,27	2:1139634-1160999	21366	Degenerate	47.1	29
63				3	0,79		11:4474-32796	28323	Degenerate	48.0	37
64				1	0,74		5:86038-118974	32937	Complete	65.4	47
65	<i>D. africanus</i> DSM2603	NZ_AULZ00000000.1	61.1	2	0,59	2,14	7:81990-108084	26095	Degenerate	64.0	33
66				3	0,54		10:45240-69368	24129	Degenerate	63.2	29
67				4	0,27		12:47622-59836	12215	Degenerate	65.4	19
68				1	0,84		727103-762542	35440	Complete	64.2	52
69	<i>D. africanus</i> Walvis Bay	NC_016629.1	61.4	2	0,63	2,00	2133499-2160143	26645	Degenerate	65.0	32
70				3	0,29		3893257-3905761	12505	Degenerate	61.5	10
71				4	0,24		2790428-2800848	10421	Degenerate	55.7	10
72				1	0,45		1:283510-298460	14951	Degenerate	55.9	21
73	<i>D. bizertensis</i> DSM 18034	NZ_FUYA00000000.1	52.2	2	0,40	1,95	4:1-13153	13153	Degenerate	55.9	19
74				3	1,10		11:60957-97045	36089	Complete	49.9	36

75	<i>D. piger</i> 64-16	MNTW00000000.1	63.7	1	1,84	1,84	1:22449-69767	47319	Complete	58.3	57
76	<i>D. alaskensis</i> DSM16109	NZ_AXWQ00000000.1	58.0	1	1,14	1,82	1:145720-186292	40573	Complete	63.2	42
77				2	0,68		6:317764-342216	24453	Complete	58.5	30
78	<i>D. indicus</i> J2	NZ_CP014206.1	63.5	1	0,92	1,74	19149-56014	36866	Complete	64.7	55
79				2	0,82		1479948-1512740	32793	Complete	61.9	50
80	<i>D. bastinii</i> DSM16055	NZ_AUCX00000000.1	43.1	1	0,31		8:228603-240880	12278	Degenerate	45.9	19
81				2	0,79	1,62	22:50-31249	31200	Degenerate	45.0	17
82				3	0,52		31:734-21301	20568	Degenerate	41.4	26
83	<i>D. litoralis</i> DSM11393	NZ_FRDI00000000.1	36.9	1	1,27	1,55	8:23822-58699	34878	Degenerate	38.0	38
84				2	0,28		13:1-7877	7877	Degenerate	35.5	14
85	<i>D. hydrothermalis</i> DSM14728	NC_020055.1	45.1	1	0,71		220039-246518	26480	Degenerate	46.0	38
86				2	0,44	1,38	264063-280712	16650	Degenerate	41.0	25
87				3	0,23		7509-16126	8618	Degenerate	38.8	8
88	<i>D. magneticus</i> RS-1	NC_012796.1	62.8	1	0,54		2787492-2816248	28757	Degenerate	64.1	34
89				2	0,14		1598917-1606756	7840	Degenerate	60.8	10
90				3	0,25	1,31	1651912-1665043	13132	Degenerate	44.3	8
91				4	0,19		3812347-3822640	10294	Degenerate	69.0	9
92				5	0,19		4745792- 4755972	10181	Degenerate	49.3	8
93	<i>D. oxyclinae</i> DSM11498	NZ_AQXE00000000.1	59.1	1	1,29	1,29	1:255253-298505	43253	Complete	55.0	44
94	<i>D. cuneatus</i> DSM11391	NZ_AUCY00000000.1	53.5	1	0,51	1,21	6:21787-39205	17419	Degenerate	54.6	24
95				2	0,70		9:76936-100771	23836	Degenerate	54.8	28
96	<i>D. desulfuricans</i> DSM 642	NZ_ATUZ00000000.1	57.4	1	0,66	1,10	11:301623-324411	22789	Degenerate	59.8	29
97				2	0,44		11:716541-731554	15014	Degenerate	51.2	10
98	<i>D. desulfuricans</i> ATCC 27774	NC_011883.1	58.1	1	1,10	1,10	277508-309276	31769	Complete	55.7	42
99	<i>D. salexigens</i> DSM2638	NC_012881.1	47.1	1	1,07	1,07	2315338-2361506	46169	Complete	45.5	53
100	<i>D. zosteræ</i> DSM11974	NZ_AUDC00000000.1	41.8	1	0,34	0,99	10:126911-140903	13993	Degenerate	41.9	14
101				2	0,65		11:407097-434019	26923	Degenerate	41.9	29
102	<i>D. magneticus</i> IFRC170	NZ_JAGC00000000.1	64.7	1	0,58	0,91	5:10256-38573	28318	Degenerate	63.9	32
103				2	0,33		5:48224-64398	16175	Degenerate	66.4	20
104	<i>D. aminophilus</i> DSM12254	NZ_AUMA00000000.1	66.2	1	0,23	0,80	4:20345-28622	8278	Degenerate	68.3	8
105				2	0,57		19:13012-32703	19692	Degenerate	68.5	30
106	<i>D. africanus</i> PCS	NZ_AOSV00000000.1	61.2	1	0,79	0,79	63:67267-98453	31187	Complete	60.2	40
107	<i>D. termitidis</i> HI1	NZ_AZAO00000000.1	66.1	1	0,17		2:6809-14189	7381	Degenerate	66.7	10
108				2	0,19	0,69	6:863784-872136	8353	Degenerate	67.3	15
109				3	0,14		7:67044-73337	6294	Degenerate	64.7	7
110				4	0,19		9:222471-230896	8426	Degenerate	65.3	8
111	<i>D. brasiliensis</i> JCM12178	NZ_BBCB00000000.1	59.7	1	0,61	0,61	28:4203-26287	22085	Degenerate	61.1	24
112	<i>D. frigidus</i> DSM17176	NZ_JONL00000000.1	42.8	1	0,54	0,54	12:1639-24526	22888	Degenerate	45.6	26
113	<i>D. putealis</i> DSM16056	NZ_AUBQ00000000.1	62.8	1	0,49	0,49	4:295731-319243	23513	Degenerate	65.8	34
114	<i>D. dechloracetivorans</i> BerOc1	NZ_LKAQ00000000.1	63.8	1	0,48	0,48	04:2266245-2285871	19627	Degenerate	65.2	26

115	<i>D. inopinatus</i> DSM10711	NZ_AUBP00000000.1	49.2	1	0,28	0,45	12:77918-94286	16369	Degenerate	49.9	18
116				2	0,17		17:89525-99426	9902	Degenerate	51.0	16
117	<i>D. desulfuricans</i> ND132	NC_016803.1	65.2	1	0,16	0,43	1652158-1657915	5758	Degenerate	53.3	6
118				2	0,12		2638934-2643452	4519	Degenerate	66.1	6
119				3	0,15		2438616-2443963	5348	Degenerate	55.9	6
120	<i>D. longus</i> DSM6739	NZ_ATVA00000000.1	63.7	1	0,39	0,39	13:177164-191691	14528	Degenerate	63.4	17
121	<i>D. ferrireducens</i> DSM16995	FNGA00000000.1	42.8	1	0,19	0,37	1:98115-105581	7467	Degenerate	46.3	8
122				2	0,18		1:680280-687550	7271	Degenerate	42.6	8
123	<i>D. magneticus</i> MBC34	ALAO00000000.1	65.7	1	0,17	0,35	303.133:16744-24517	7774	Degenerate	64.3	10
124				2	0,18		435.212:995-8978	7984	Degenerate	63.1	7
125	<i>D. alcoholivorans</i> DSM5433	NZ_JNJA00000000.1	64.7	1	0,13	0,33	3:79563-86728	7166	Degenerate	63.3	9
126				2	0,20		7:63594-74297	10704	Degenerate	48.2	14
127	<i>D. piezophilus</i> C1TLV30	NC_020409.1	50.0	1	0,31	0,31	1174880-1186253	11374	Degenerate	54.5	13
128	<i>D. legallii</i> KHC7	FNBX00000000.1	64.8	1	0,29	0,29	18: 46636-54571	7936	Degenerate	66.4	10

**Supplementary Table S2** Complete prophages in *Desulfovibrio*

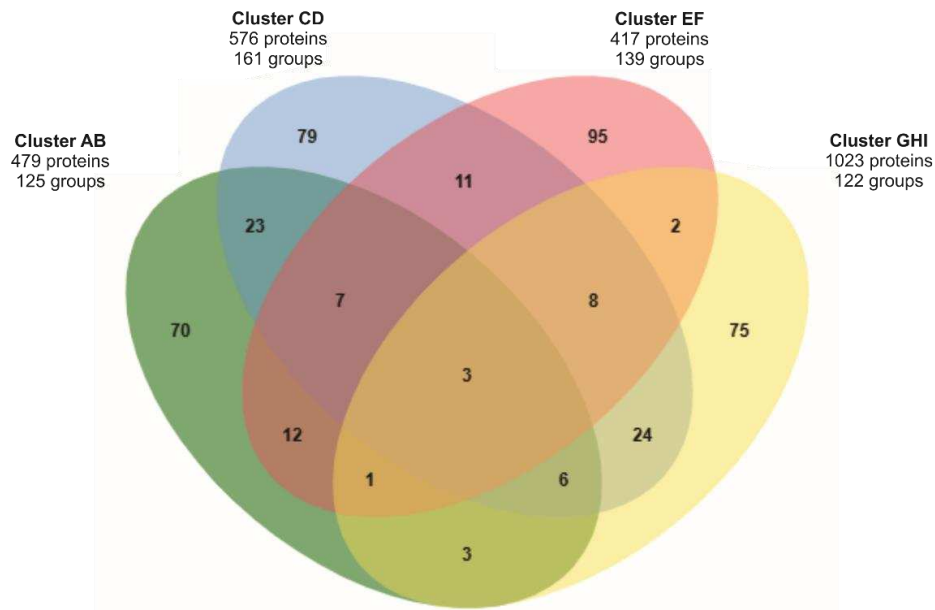
N°	Strain	Origin <sup>a</sup>	Prophage	Size (pb)	Family <sup>b</sup>	Proteins <sup>c</sup>
1	<i>D. aespoeensis</i> Aspo-2	Env	2	36770	<i>Myo</i>	Lys hyp (WP_083808685.1)
2			3	42929	<i>Myo</i>	Lys hyp (WP_013514570.1)
3			4	38064	<i>Myo</i>	Lys (WP_013515329.1)
4	<i>D. africanus</i> Walvis Bay	Env	1	35440	<i>Myo</i>	Hol hyp (WP_014258869.1) e Lys hyp (WP_014258893.1)
5	<i>D. africanus</i> DSM2603	Env	1	32937	<i>Myo</i>	-
6	<i>D. africanus</i> PCS	Env	1	31187	<i>Sip</i>	-
7	<i>D. alaskensis</i> G20	Env	1	41050	<i>Myo</i>	Lys hyp (WP_011366992.1)
8			3	40230	<i>Myo</i>	-
9			4	48689	<i>Myo</i>	Lys (WP_011367798.1)
10	<i>D. alaskensis</i> DSM16109	Env	1	40573	<i>Myo</i>	-
11			2	24453	<i>Myo</i>	-
12	<i>D. alkalitolerans</i> DSM16529	Env	2	43877	<i>Myo</i>	-
13	<i>D. bizertensis</i> DSM 18034	Env	3	36089	<i>Myo</i>	Lys hyp (SKA81922.1)
14	<i>D. desulfuricans</i> ATCC 27774	Env	1	31769	<i>Myo</i>	-
15	<i>D. desulfuricans</i> DSM7057	Env	1	43719	<i>Pod</i>	Hol (SFW23071.1) e Lys (SFW23095.1)
16			2	28623	<i>Myo</i>	Lys (SFW30341.1)
17	<i>D. desulfuricans</i> DSM17919	Env	1	31194	<i>Myo</i>	Hol hyp (SHI60426.1)
18	<i>D. fairfieldensis</i> CCUG45958	Hum	1	57140	<i>Sip</i>	Hol hyp (WP_083521954.1) e Lys hyp (WP_062251836.1)
19			2	38473	<i>Myo</i>	Lys hyp (WP_062252149.1)
20			3	30416	<i>Myo</i>	Hol hyp (WP_062252788.1) e Lys hyp (WP_062252832.1)
21	<i>D. fructosivorans</i> JJ	Env	1	37113	<i>Sip</i>	Hol hyp (EFL53014.1) e Lys (EFL53013.1)
22			2	50676	<i>Sip</i>	Hol hyp (EFL53098.1) e Lys (EFL53099.1)
23	<i>D. gigas</i> ATCC19364	Env	1	44956	<i>Sip</i>	Lys hyp (WP_021759525.1)
24			2	33156	<i>Sip</i>	Hol hyp (WP_021761111.1)
25			3	56190	<i>Sip</i>	Lys hyp (WP_051286217.1)
26	<i>D. indicus</i> J2	Env	1	36866	<i>Myo</i>	Hol hyp (WP_066798947.1) e Lys (WP_066798949.1)
27			2	32793	<i>Myo</i>	-
28	<i>D. oxyclinae</i> DSM11498	Env	1	43253	<i>Pod</i>	Lys hyp (292456-292824 complementar)
29	<i>D. piger</i> Fl11049	Hum	1	39418	<i>Myo</i>	Hol hyp (WP_072331516.1) e Lys (WP_083575220.1)
30			2	42600	<i>Myo</i>	-
31			5	40876	<i>Myo</i>	Hol hyp (WP_072332501.1) e Lys (WP_083575254.1)
32	<i>D. piger</i> ATCC29098	Hum	2	37864	<i>Myo</i>	-
33			4	35309	<i>Myo</i>	-
34	<i>D. piger</i> 64-16	Hum	1	47319	<i>Myo</i>	Hol hyp (OLA87393.1) e Lys (OLA87415.1)
35	<i>D. salexigens</i> DSM2638	Env	1	46169	<i>Myo</i>	Hol hyp (WP_081434591.1)
36	<i>D. vulgaris</i> Miyazaki	Env	1	36988	<i>Myo</i>	Lys (WP_012611872.1) e Lys hyp (WP_012611897.1)
37			2	39403	<i>Myo</i>	Hol (WP_012612411.1) e Lys (WP_012612412.1)

38	<i>D. vulgaris</i> Hildenborough	Env	1	36674	<i>Myo</i>	Hol (YP_009427.1) e Lys hyp (YP_009428.1)
39			2	35621	<i>Myo</i>	Hol hyp (YP_010350.1) e Lys (YP_010349.1)
40			3	39411	<i>Sip</i>	Hol (YP_010724.1)
41			4	36552	<i>Pod</i>	Hol hyp (YP_010945.1)
42			5	35035	<i>Myo</i>	Lys hyp (YP_011939.1)
43			6	40679	<i>Myo</i>	Hol (YP_012079.1) e Lys hyp (YP_012078.1)
44			7	44367	<i>Sip</i>	Hol (YP_011379.1) e Lys (YP_011378.1)
45	<i>D. vulgaris</i> DP4	Env	1	42742	<i>Sip</i>	Hol (WP_081428991.1) e Lys (WP_011792031.1)
46			2	36411	<i>Myo</i>	Hol hyp (WP_011792275.1)
47			4	34215	<i>Myo</i>	Hol hyp (WP_011793063.1)
48	<i>D. vulgaris</i> RCH1	Env	1	40282	<i>Myo</i>	Lys (WP_010938427.1)
49			2	37208	<i>Pod</i>	Hol hyp (WP_010939015.1)
50			3	35035	<i>Myo</i>	Lys (WP_010939994.1)
51			4	40679	<i>Myo</i>	Hol (WP_010940129.1) e Lys hyp (WP_010940128.1)
52			5	34958	<i>Sip</i>	Hol hyp (WP_010938799.1)
53			6	42866	<i>Sip</i>	Hol (WP_010939443.1) e Lys (WP_010939442.1)

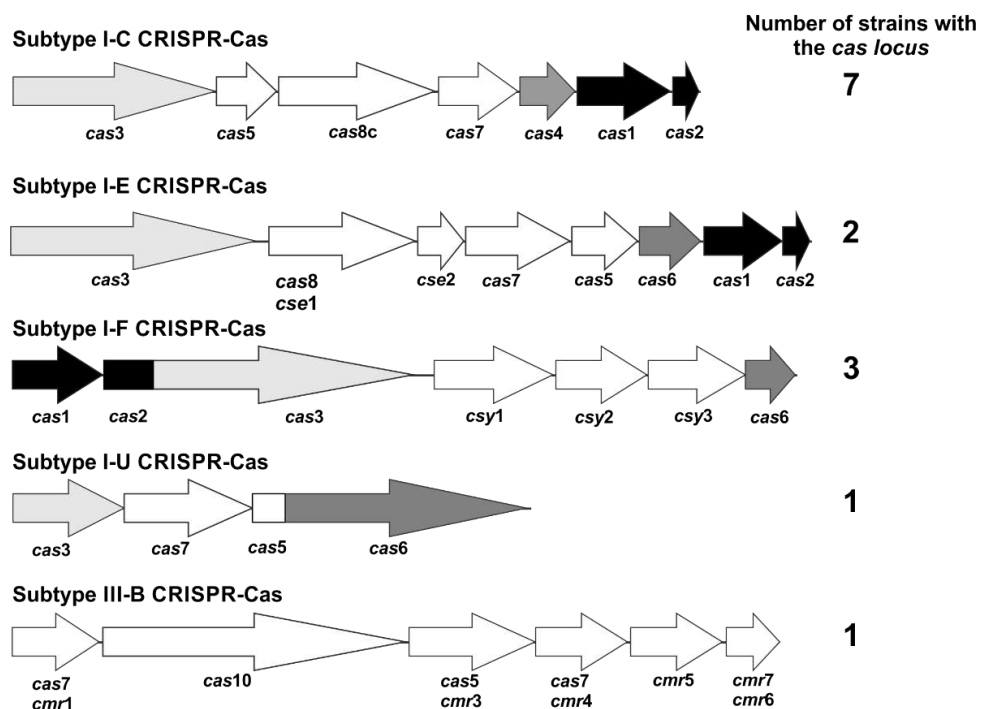
<sup>a</sup> Strains were characterized in environmental (Env) and human (Hum).

<sup>b</sup> Phages families were characterized in *Myoviridae* (Myo), *Siphoviridae* (Sip) and *Podoviridae* (Pod).

<sup>c</sup> Proteins were characterized in lysozyme (Lys), holina (Hol), hypothetical (hyp).



**Supplementary Fig. S1.** Sharing of proteins between *Desulfovibrio* complete prophages. The Venn Diagram was drafted with the total proteins. The numbers indicate protein groups shared between the four phylogenetic clusters (AB, CD, EF and GHI) of 53 *Desulfovibrio* prophages.



**Supplementary Fig. S2.** Architecture of the gene region of the CRISPR-Cas system in closed *Desulfovibrio* genomes. Each *cas* locus is represented according to a *Desulfovibrio* strain: *D. desulfuricans* ND132 (I-C), *D. alaskensis* G20 (I-E), *D. gigas* (I-F and I-U) and *D. hydrothermalis* (III-B). On the right is the number of genomes that have the *cas* locus type highlighted. The genes are colored according to their function: adaptation (black), pre-crRNA processing (dark gray), interference modulus (white) and target cleavage (light gray).

**Supplementary Table S3** CRISPR-Cas system in closed genomes of *Desulfovibrio* by CRISPRone

N°	Strain	CRISPR arrays (n° of spacers)	CRISPR-Cas systems type	Putative activity
1	<i>D. aespoeensis</i> Aspo-2	0	I-incomplete	Inactive
2	<i>D. africanus</i> Walvis Bay	ND	ND	Inactive
3	<i>D. alaskensis</i> G20	1(20)	I-E	Active
4	<i>D. desulfuricans</i> ND132	1(100)	I-C	Active
5	<i>D. desulfuricans</i> ATCC 27774	1 (30)	I-E	Active
6	<i>D. fairfieldensis</i> CCUG45958	ND	ND	Inactive
7	<i>D. gigas</i> ATCC19364	1(6)	I-F e I-U incomplete	Active
8	<i>D. hydrothermalis</i> DSM14728	6 (45)	III-B e I-F	Active
9	<i>D. indicus</i> J2	2(29)	I-C	Active
10	<i>D. magneticus</i> RS-1	2(52)	I-C e I-F	Active
11	<i>D. piezophilus</i> C1TLV30	0	ND	Inactive
12	<i>D. piger</i> F111049	1(3)	ND	Inactive
13	<i>D. salexigens</i> DSM2638	0	I-incomplete	Inactive
14	<i>D. vulgaris</i> DP4	1(46)	I-C	Active
15	<i>D. vulgaris</i> Hildenborough	1(28)	I-C	Active
16	<i>D. vulgaris</i> Miyazaki	2(60)	I-C	Active
17	<i>D. vulgaris</i> RCH1	1(28)	I-C	Active

ND: Not detected.

## CAPÍTULO 2

### **Profagos de *Desulfovibrio alaskensis* e seu possível envolvimento na transferência horizontal de genes por vesículas de membrana externa**

*De acordo com as normas da Journal of genetics and genomics*

Josicelli Souza Crispim<sup>a</sup>, Roberto Sousa Dias<sup>b</sup>, Clara Nogueira Laguardia<sup>b</sup>, Larissa Cristina Araújo<sup>b</sup>, Jéssica Duarte da Silva<sup>a</sup>, Pedro Marcus Pereira Vidigal<sup>c</sup>, Maíra Paula de Sousa<sup>d</sup>, Cynthia Canêdo da Silva<sup>a</sup>, Mateus Ferreira Santana<sup>a</sup>, Sérgio Oliveira de Paula<sup>b\*</sup>

<sup>a</sup>Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa 36570-900, Brazil.

<sup>b</sup>Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, Brazil.

<sup>c</sup>Núcleo de Análise de Biomoléculas, Universidade Federal de Viçosa, Viçosa, Brazil.

<sup>d</sup>Centro de Pesquisas e Desenvolvimento Leopoldo Américo Miguez de Mello, CENPES, Rio de Janeiro, Brazil.

Running head: Profagos de *Desulfovibrio alaskensis*

\*Corresponding author: Sérgio Oliveira de Paula (depaula@ufv.br); phone: +55 (31) 3899-2559

## Resumo

*Desulfovibrio alaskensis* é uma espécie pertencente ao grupo de Bactérias Redutoras de Sulfato (BRS) e que apresenta profagos no seu genoma, uma característica comum do gênero *Desulfovibrio*. Material genético pode ser transportado por vesículas de membrana externa. Porém, nenhum dado sobre a produção dessas vesículas foi descrito para *D. alaskensis*. Para a verificação da expressão de profagos de *D. alaskensis* e seu envolvimento com vesículas de membrana externa foi utilizada a linhagem DSM16109. A linhagem DSM16109 possui três profagos e apresentou menor crescimento após a adição de mitomicina C em relação à cultura controle. Esta redução foi acompanhada da presença de partículas *virus-like* (VLPs), indicando indução de profagos por mitomicina C. Foi verificado o aumento do número de cópias dos genes *cap* dos três profagos e transcrição dos mesmos na amostra controle, entretanto, sem a formação de VLPs. Genes de profagos foram identificados em vesículas de membrana externa de culturas tratadas e não tratadas com mitomicina C. Os profagos de DSM16109 são induzidos de forma espontânea, mas, apenas a presença de mitomicina C levou à formação de VLPs. O material genético dos profagos detectado dentro de vesículas de membrana externa podem estar relacionados com a transferência horizontal de genes virais.

**Palavras chaves:** Bacteriophage, mitomycin C, sulphate reducing bacteria, prophage

## Introdução

*Desulfovibrio alaskensis* pertence ao grupo de Bactérias Redutoras de Sulfato (BRS), que possuem a capacidade de reduzir sulfato durante seu metabolismo, levando à produção de sulfeto de hidrogênio (H<sub>2</sub>S) (Postgate, 1965). O gás H<sub>2</sub>S é altamente tóxico, sendo um dos principais problemas para a indústria petrolífera, tanto para a saúde dos trabalhadores quanto para a qualidade do petróleo (Gieg et al., 2011). Além da produção de H<sub>2</sub>S, as BRS são capazes de formar biofilme, o que as torna resistentes à tratamentos químicos, e favorece a corrosão microbiologicamente induzida (MIC) de tubulações, gerando grandes prejuízos na indústria de petróleo (Rajasekar et al., 2010). De acordo com a literatura, duas

linhagens de *Desulfovibrio alaskensis* são atualmente estudadas e possuem o genoma depositado no banco de dados: DSM16109 e G20. A linhagem *D. alaskensis* DSM16109 foi isolada de material coletado em reservatório de óleo no Alasca (EUA) por E. van der Vende em 1991, mas sua caracterização ocorreu apenas em 2004, quando recebeu o nome do local onde foi encontrada (Feio et al., 2004). *D. alaskensis* G20 é uma linhagem derivada de *Desulfovibrio desulfuricans* G100A, que também foi isolada de reservatório de óleo (Califórnia, EUA), sendo reclassificada como *D. alaskensis* G20 em 2011 (Hauser et al., 2011). Essas duas linhagens possuem o genoma sequenciado, nos quais foram identificados a presença de profagos (Crispim et al., 2018).

Profagos, genomas de bacteriófagos que estão presentes no genoma bacteriano, são encontrados em diversos grupos bacterianos (Bobay et al., 2013; Crispim et al., 2018; Pratama and van Elsas, 2017). Em *D. alaskensis* G20, sequências relacionadas à profagos ocupam 3,67% do genoma (Crispim et al., 2018). Os profagos podem entrar no ciclo lítico espontaneamente ou quando induzidos por algum estresse (Gandon, 2016; Nanda et al., 2015; Tang et al., 2017), o que permite seu isolamento e posterior caracterização. Muitos profagos são defectivos, não sendo indúzíveis ou se indúzíveis, formam partículas virais incompletas (Bobay et al., 2014; Durmaz et al., 2008; Yuan et al., 2014). Entretanto, a presença desses profagos nos genomas bacterianos representam um significativo papel na evolução e diversidade bacteriana (Canchaya et al., 2004; Wang et al., 2010). Nesse contexto, vesículas de membrana externa podem atuar como veículos de transporte intercelular de genomas de fagos, facilitando a recombinação entre fagos e/ou genomas bacterianos (Gaudin et al., 2014).

Vesículas de membrana externa (*Outer Membrane Vesicles* – OMVs) são estruturas membranosas liberadas por brotamento da membrana externa de bactérias Gram-negativas (Kulp and Kuehn, 2010). As OMVs apresentam forma esférica de 10-300nm e podem carregar sRNAs regulatórios, proteínas e DNA, sendo já caracterizado o seu envolvimento com aderência, formação de biofilme bacteriano e transferência horizontal de genes (Bonnington and Kuehn, 2014; Domingues and Nielsen, 2017). Além disso, as OMVs podem agir como uma defesa inata bacteriana contra bacteriófagos ao adsorver partículas virais (Manning and Kuehn, 2011; Reyes-Robles et al., 2018).

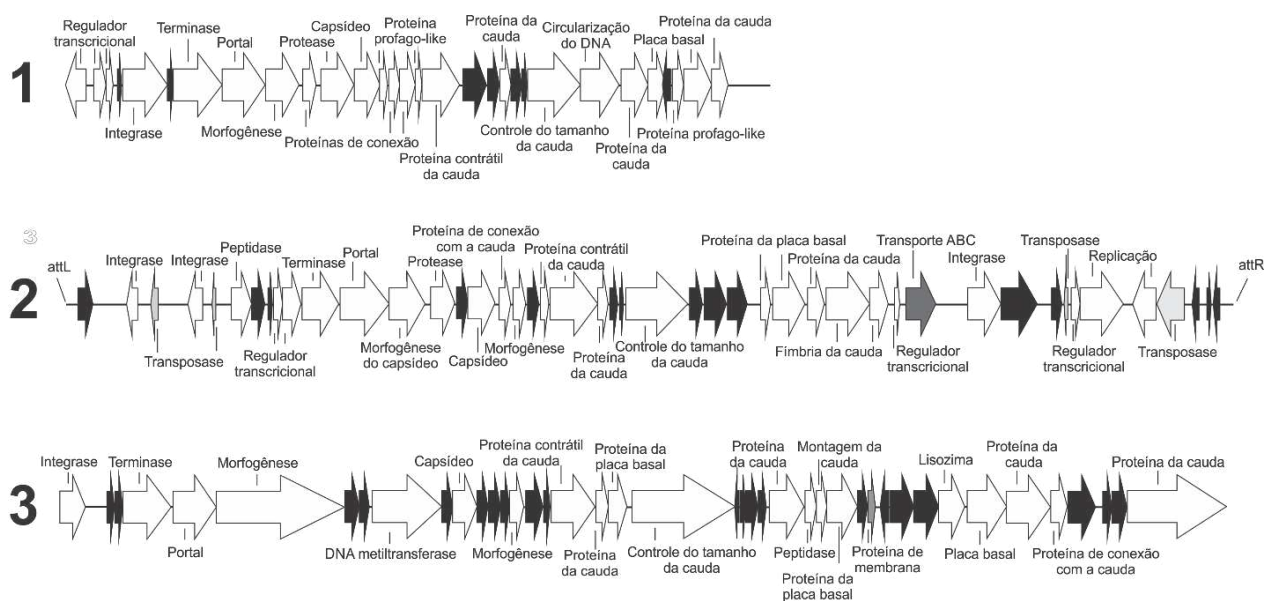
Apenas um trabalho foi realizado sobre bacteriófagos em *D. alaskensis* (Crispim et al., 2018) e nenhum estudo sobre OMVs foi encontrado para esta espécie, o que torna essa área de grande interesse para o estudo de formas de controle das BRS. Este trabalho realiza a verificação da expressão de profagos em *D. alaskensis*, anteriormente estudados em análises *in silico* (Crispim et al., 2018), e seu envolvimento com vesículas de membrana externa.

## **Resultados**

### **Profagos em *D. alaskensis***

Em um trabalho anterior foi identificado três profagos completos no genoma fechado de *D. alaskensis* G20 (elementos profago-*like* 1, 3 e 4) e dois profagos no genoma depositado na forma de *contigs* de *D. alaskensis* DSM16109 (elementos profago-*like* 1 e 2) (Crispim et al., 2018). Desses profagos, sabe-se que os elementos profago-*like* 1 e 4 da linhagem G20 e o elemento profago-*like* 1 da linhagem DSM16109 possuem 94 % de identidade. No presente estudo, verificamos por meio da ferramenta Blast que o *contig* 15 da linhagem DSM16109 possui 48 % de cobertura e 99 % de identidade com o elemento profago-*like* 3 da linhagem G20 e que o elemento profago-*like* 2 da linhagem DSM16109 é exclusivo dela. Em vista da identidade entre os profagos completos das linhagens de *D. alaskensis*, foi determinado o total de 3 profagos na espécie para a realização de análises posteriores. A representação dos três profagos encontrados em *D. alaskensis* pode ser visualizada na Fig. 1 e a caracterização completa das regiões codificadoras nas Tabelas S1-3.

O profago 1 de *D. alaskensis* (Fig. 1), nomeado como proDAD2, possui reguladores transcricionais *upstream* ao gene da integrase, que possui a mesma direção dos genes virais. Genes tardios foram encontrados a jusante ao gene da integrase, com posterior localização de genes estruturais da cauda. Esse profago, apresenta ausência de genes que codificam para fibras da cauda e relacionados à replicação viral.



**Fig. 1.** Organização genômica de profagos de *D. alaskensis*. Os profagos foram representados de acordo com o observado nas linhagens G20 e DSM16109: 1 (proDAD2) - elemento profago-like 2 da DSM16109, 2 (proDAG3) - elemento profago-like 3 da G20 e 3 (proDAG1) - elemento profago-like 1 da G20, de acordo com Crispim e colaboradores (Crispim et al., 2018). As setas indicam a orientação transcricional dos genes e estão coloridas de acordo com a função: transposase (cinza claro), hipotética (preto), virais (branco) e bacteriana (cinza escuro).

O profago 2 (Fig. 1), nomeado como proDAG3, possui duas integrases próximas e com direção contrária aos genes virais. Uma terceira integrase é encontrada nesse profago apresentando a mesma direção que os genes virais. Esse profago possui módulo tardio, que codifica para proteínas do portal, terminase e capsídeo, a jusante aos dois genes da integrase de localização próxima. Posteriormente, são encontrados os genes estruturais da cauda, incluindo genes que codificam as fibras da cauda e proteínas relacionadas à replicação viral. Além disso, esse profago possui três transposases e um gene que codifica uma proteína bacteriana da família SunT, um transportador do tipo ABC.

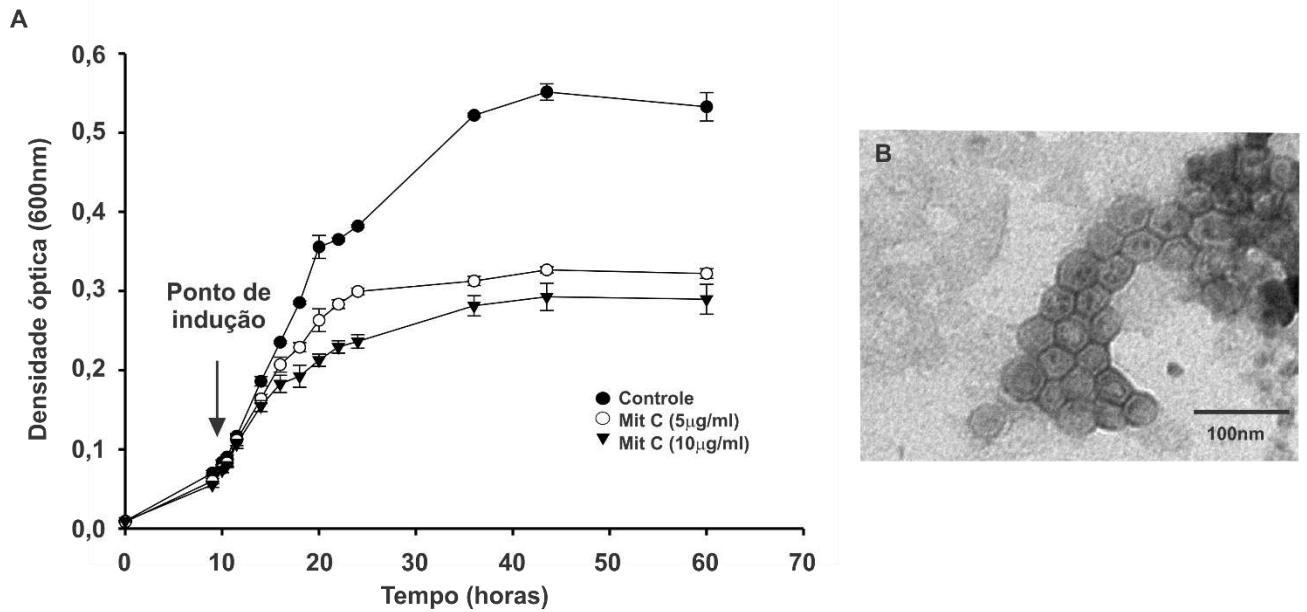
O profago 3 (Fig. 1), nomeado como proDAG1, apresenta o gene que codifica para integrase disposto no mesmo sentido que os genes virais. Entre os genes virais encontrados destacam-se os que codificam a terminase, a proteína do portal e o capsídeo, localizados a jusante ao gene da integrase, caracterizando o módulo de transcrição tardio. Posteriormente, encontram-se genes que codificam proteínas da cauda, proteína de conexão capsídeo/cauda e proteínas da placa basal. Nesse profago não foi encontrada proteínas da fibra da cauda e proteínas relacionadas à replicação e regulação viral. Por outro lado, foram identificados os genes que codificam a enzima DNA metiltransferase e a proteína lisozima.

Ao comparar esses profagos com genomas disponíveis no banco de dados foi encontrado que o proDAG1 possui aproximadamente de 99 % de cobertura e 96 % de identidade com profagos presentes nos genomas das linhagens *Syntrophobacter fumaroxidans* MPOB e *Geobacter* sp. M18. Os outros profagos apresentaram baixa identidade com as sequências presentes no banco de dados, sendo 7% de cobertura e 72% de identidade do proDAD2 com profago presente no genoma de *Desulfovibrio vulgaris* Miyazaki F e de 4% de cobertura e 90% de identidade do proDAG3 com profago presente no genoma de *Pseudodesulfovibrio indicus* J2.

### **Profagos induzíveis de *D. alaskensis* DSM16109**

Na análise de verificação de profagos *in vivo* foi observada a redução do crescimento bacteriano após a inoculação de mitomicina C nas concentrações de 5 e 10 µg/mL na cultura de DSM16109 (Fig. 2A). Para análises posteriores, a concentração de 5 µg/ml foi usada.

Partículas compostas apenas por capsídeo e com ausência de material eletrodense no seu interior foram visualizadas nas amostras induzidas com mitomicina C por 6 horas (Fig. 2B). Nenhuma partícula foi visualizada na amostra controle no mesmo ponto de coleta da amostra induzida. A PCR com os genes *cap1*, *cap2* e *cap3* confirmou a ausência de material genético no interior das partículas pela não amplificação dos genes virais (dados não mostrados), caracterizando essas partículas como *virus-like particles* (VLPs). Os genes que codificam para a proteína do capsídeo dos três profagos (*cap1*, 2 e 3) apresentaram transcrição no tempo de 6 horas, ponto da visualização das VLPs na amostra induzida (Fig. S1). O mesmo resultado foi encontrado para a amostra não induzida (controle).

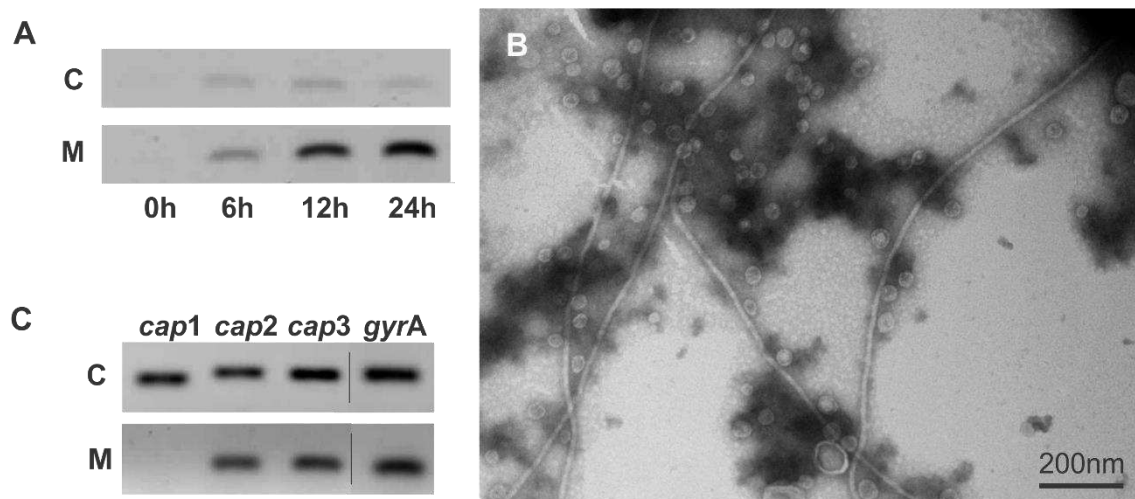


**Fig. 2.** Verificação de VLPs em *D. alaskensis* DSM16109 induzida com mitomicina C. A: Curva de crescimento de *D. alaskensis* DSM16109 sem e com mitomicina C (5µg/ml e 10µg/ml) inoculada na DO 0,1. B: Microscopia eletrônica de transmissão de partículas *virus-like*. Imagem obtida de amostra induzida com mitomicina C. Bars: 100 nm.

### Material genético em vesículas de membrana externa

Foi verificada a presença de material genético na matriz extracelular nos tempos de 6, 12 e 24 horas após a inoculação de mitomicina C, cujo material genético acumula-se no tempo de 24 horas (Fig. 3A). A amostra controle apresentou menor quantidade de material extracelular nos mesmos tempos verificados na amostra induzida. Este material genético pode estar dentro das vesículas ou livre no meio extracelular.

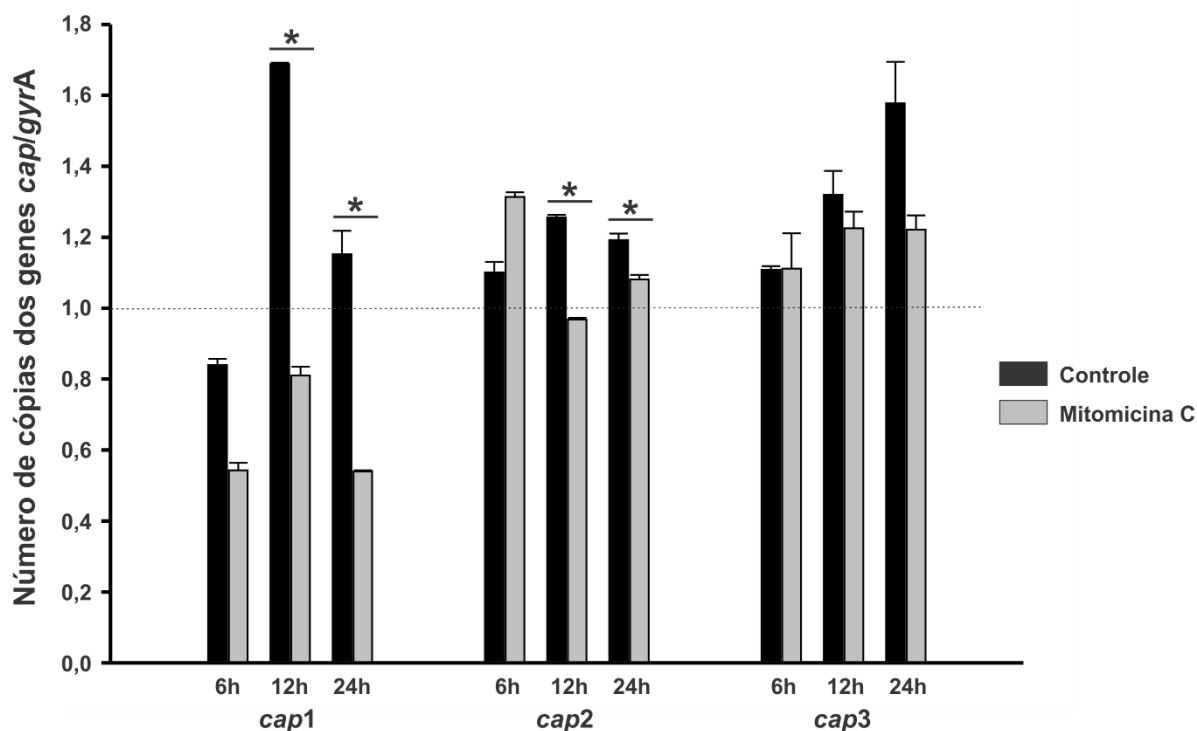
Vesículas de membrana externa foram visualizadas em amostras controle e induzida por mitomicina C após 24 horas de crescimento, Fig. 3A. As vesículas apresentaram tamanho variando de 34-125 nm com grande contrastação do interior, indicando presença de material intravesicular. A reação de PCR detectou a presença de genes virais (*cap1*, 2 e 3) e bacteriano (*gyrA*) em ambas as amostras (controle e induzida com mitomicina C), com exceção do gene *cap1* na amostra induzida, Fig. 3B.



**Fig. 3.** Vesículas de membrana externa de *D. alaskensis* DSM16109 carregam seqüências relacionadas à profagos. A: PCR da matriz extracelular da linhagem DSM16109 com (M) e sem (C) a inoculação de mitomicina C nos tempos de 6, 12 e 24 horas. Para verificação de material genético bacteriano no sobrenadante da cultura foi usado o primer GyrA, os mesmos resultados foram encontrados com os primers Cap1, 2 e 3. B: Microscopia eletrônica de transmissão de vesículas de membrana externa. A amostra representa vesículas de membrana externa da amostra controle, que também foram visualizadas na amostra induzida com mitomicina C. Bars: 200 nm. C: Presença de material genético em vesículas de membrana externa por PCR. *cap*: gene que codifica para a proteína do capsídeo dos profagos 1, 2 e 3. GyrA: gene que codifica para a proteína DNA girase. C: controle; M: mitomicina C.

### Número de cópias dos genes *cap* na matriz extracelular

A qPCR permitiu a verificação do número de cópias dos genes dos capsídeos virais em relação ao gene da DNA girase nas amostras analisadas anteriormente por PCR, Fig. 3B. O gene *cap1* apresenta diminuição do número de cópias no tempo de 6 horas de crescimento, tanto na amostra controle quanto na amostra com mitomicina C, sem diferença estatística entre as duas amostras. Em 12 e 24 horas de crescimento, observa-se diminuição significativa do número de cópias de *cap1* ( $p < 0,05$ ) na amostra induzida em relação a amostra controle. Já o gene *cap2* apresentou pequeno aumento do número de cópias na amostra induzida em relação à amostra controle no tempo de 6 horas, mas sem diferença significativa entre as duas amostras. Entretanto, nos tempos de 12 e 24 foi verificada a diminuição do número de cópias desse gene na amostra induzida em relação a amostra controle ( $p < 0,05$ ). O gene *cap3* não apresentou diferença no número de cópias em 6 horas de crescimento da cultura. Em 12 e 24 horas foi observado redução do número de cópias desse gene na amostra induzida em relação a amostra controle, porém, sem diferença significativa.



**Fig. 4.** Variação do material genético de profagos na matriz extracelular de *D. alaskensis* DSM16109. Número de cópias dos genes *cap1*, 2 e 3 em 6, 12 e 24 horas após inoculação de mitomicina C. O material extracelular da cultura do ponto 0, quando foi realizada a inoculação de mitomicina C, foi utilizado como ponto de referência para a análise relativa, representado pela linha tracejada. *cap*: gene que codifica para a proteína do capsídeo dos profagos 1, 2 e 3. O gene *gyrA* foi utilizado como controle endógeno. O asterístico (\*) representa os pontos com diferença estatística pelo teste-*t*,  $p < 0,05$ , entre amostra sem (controle) e com mitomicina C.

## Discussão

A análise comparativa de profagos entre as linhagens DSM16109 e G20 de *D. alaskensis*, inicialmente caracterizados por Crispim e colaboradores (Crispim et al., 2018), permitiu identificar a presença de três profagos na espécie. Outros profagos podem ser identificados, uma vez que o genoma da linhagem DSM16109 encontra-se depositado na forma de *contig*. O proDAD2, exclusivo da linhagem DSM16109, é o menor profago da linhagem e não possui todas as características para ser profago induzível. Porém, este profago está localizado na extremidade de um *contig*, o que pode mascarar a continuidade do mesmo. O proDAG3, diferentemente do proDAD2, possui a presença de genes que codificam as fibras da cauda, que tem a função de reconhecimento da célula hospedeira (Chen et al., 2017). Este profago possui todas as características de um profago induzível, mas proteínas podem ser defectivas, impedindo a formação da partícula viral infectiva. O proDAG3 possui também transposases, as quais são responsáveis pela

relocalização de sequências dentro e entre genomas (Frost et al., 2005). A presença de transposase e de genes exógenos neste elemento pode ser um indício de aquisição de genes de outras bactérias, como o gene que codifica a proteína da família SunT. Proteínas dessa família estão relacionadas com transportadores do tipo ABC envolvidos no transporte de lantibióticos, que são peptídeos conhecidos como bacteriocinas (McAuliffe et al., 2001). As bacteriocinas são produzidas por bactérias Gram-positivas e Gram-negativas e são capazes de inibir o crescimento de outras bactérias devido a interação com a parede celular e formação de poros (Chikindas et al., 2018). A presença desse gene neste profago demonstra o envolvimento de bacteriófagos na transferência horizontal de genes. No proDAG1 não foram encontradas proteínas da fibra da cauda e proteínas relacionadas à replicação e regulação viral, mas, genes hipotéticos estão presentes e podem estar relacionados com estas funções. Este profago também apresentou os genes que codificam para DNA metiltransferase e proteína lisozima. A DNA metiltransferase é encontrada em todas as bactérias, mas não em todos os bacteriófagos, estando envolvida em regulação de genes bacterianos e na metilação dos genomas virais replicados, conferindo proteção contra enzimas de restrição do hospedeiro (Smith and Jeddloh, 2005). A proteína lisozima tem a função de hidrolase de peptideoglicano, sendo capaz de agir na parede celular bacteriana como endolisinas, seja para a liberação das partículas virais após a replicação ou para a formação de poros na parede celular para injeção do material genético na célula hospedeira (Latka et al., 2017; Rodriguez-Rubio et al., 2013). A presença desse gene em profagos de *D. alaskensis* mostra ser um interessante alvo para o desenvolvimento de ferramentas para controle bacteriano por meio da sua expressão em sistema heterólogo (Guo et al., 2017).

No teste de indução de profagos com mitomicina C houve a redução de crescimento da linhagem DSM16109 com ambas as concentrações de mitomicina C (5 e 10 µg/ml). Apesar da redução de crescimento, não foi observada uma queda exponencial do crescimento como observado em alguns trabalhos (Hammerl et al., 2016; Pratama and van Elsas, 2017), o que demonstraria lise celular pela liberação de partículas virais e não um efeito inibitório da mitomicina C. Essa característica da redução da curva de crescimento sem queda exponencial também foi comum na linhagem *Lactobacillus casei* BL23, na qual foi verificada a presença de partículas virais no material extracelular da cultura (Dieterle et al., 2016). Neste trabalho, a lise

celular foi caracterizada como incompleta, na qual nem todas as células sofreram lise celular, mantendo assim uma fase estacionária de crescimento. No presente estudo, também foram verificadas partículas virais na amostra induzida com mitomicina C, porém, as partículas apresentaram-se formadas apenas por capsídeo, sem a presença de cauda e sem material genético, caracterizando a formação de partículas *virus-like* (VLPs). Profagos defectivos são muito comuns e em alguns casos não são induzíveis por mitomicina C (Chen et al., 2006; Mobberley et al., 2010), devido à, por exemplo, mutações nos genes que codificam para a proteína do capsídeo viral, como no profago proCM3 de *Bacillus thuringiensis* (Yuan et al., 2014). Já o profago Lrm1 de *Lactobacillus rhamnosus* M1 é induzível por mitomicina C, mas poucas partículas virais completas são formadas, encontrando capsídeo e cauda separados no material extracelular da cultura bacteriana induzida (Durmaz et al., 2008).

A ausência de material genético relacionado aos profagos nas VLPs encontradas pode ser devido a ausência ou a não expressão de proteínas da interface cabeça-cauda, que são importantes para a incorporação do DNA no interior do capsídeo (Tavares, 2018). Entretanto, já foi demonstrado que profagos defectivos são capazes de empacotar regiões aleatórias do genoma bacteriano de até 8 kb (Jin et al., 2014). A análise de transcrição reversa demonstrou que os três profagos apresentam transcrição do gene que codifica para o capsídeo viral. O processo de liberação de partículas virais é resultante da ação de holinas e lisozimas presentes em bacteriófagos (Shi et al., 2012), cujo gene que codifica a proteína lisozima foi detectado no proDAG1. Assim, a liberação das VLPs encontradas pode estar relacionada com a ativação transcricional deste gene ocasionada pela indução por mitomicina C, sendo já demonstrado estar relacionada com a ativação transcricional de genes virais do ciclo lítico (Chou et al., 2017). Por outro lado, a transcrição de genes que codificam as proteínas dos capsídeos virais também foi encontrada na amostra controle, o que indica indução espontânea dos mesmos, porém, sem a liberação de VLPs. A indução espontânea já foi demonstrada em outras linhagens (Nanda et al., 2014; Pratama and van Elsas, 2017), podendo representar um impacto para o *fitness* bacteriano, bem como estar envolvido na formação de biofilme (Carrolo et al., 2010; Nanda et al., 2015).

Apesar da ausência de material genético relacionados à profagos nas VLPs, foi identificado um aumento de material genético na matriz extracelular da amostra

induzida. Este material genético pode estar livre na matriz extracelular, liberado junto com as VLPs formadas, ou dentro de vesículas de membrana externa, uma vez que mitomicina C induz a vesiculação (Dutta et al., 2004). A análise de microscopia eletrônica de transmissão confirmou a formação de vesículas de membrana externa por *D. alaskensis* DSM16109 na fase exponencial de crescimento nas amostras controle e induzida com mitomicina C. Estas vesículas podem carregar componentes citoplasmáticos, como proteínas, polissacarídeos e material genético de até 370 kb (Domingues and Nielsen, 2017). O interior eletrodens dessas vesículas indicou a possível presença de material intravesicular. Na análise do conteúdo das vesículas de fase estacionária de *D. alaskensis*, foi verificada a presença de material genético dos três profagos na amostra controle e dos profagos proDAG3 e proDAG1 na amostra induzida com mitomicina C. Ho e colaboradores (Ho et al., 2015) demonstraram que vesículas de membrana externa de *Porphyromonas gingivalis* são capazes de carregar grande diversidade de suas regiões genômicas. A presença de sequências genômicas relacionadas à vírus em vesículas de membrana externa já foi demonstrada (Gaudin et al., 2014) e representa um mecanismo de transferência horizontal de genes (THG). A THG mediada por OMVs foi confirmada em *P. gingivalis* (Ho et al., 2015) e em outras espécies bacterianas (Domingues and Nielsen, 2017; Klieve et al., 2005).

No presente estudo, uma possível THG pode ser justificada pela análise comparativa das sequências de profagos de *D. alaskensis* com o banco de dados. Foi encontrada alta identidade do proDAG1 de *D. alaskensis* com sequências relacionadas à profagos de *Syntrophobacter fumaroxidans* e *Geobacter* sp. A bactéria *S. fumaroxidans* também é uma redutora de sulfato e pertence à família *Syntrophobacteraceae*, que está dentro da mesma classe que *D. alaskensis*, Deltaproteobacteria (Plugge et al., 2012). Já *Geobacter* sp. é uma bactéria redutora de Fe, encontrada em diversos ambientes anaeróbios (Lovley et al., 2011). Ambas espécies são Gram-negativas e possíveis produtoras de OMVs. Além disso, estas espécies já foram descritas e identificadas em ambientes relacionados à exploração e beneficiamento de petróleo (Nazina et al., 2017), indicando uma possível relação de profagos interespecífica que pode ter ocorrido por THG mediada por OMVs. Apesar da baixa identidade do proDAG3 ter sido encontrada com outras linhagens presentes no banco de dados, suas características como a presença de transposase e gene que codifica proteína bacteriana, demonstram possível envolvimento com a

THG. O proDAD2 também apresentou baixa identidade com sequências presentes no banco de dados, o que indica que *D. alaskensis* possui dois profagos caracterizados pela primeira vez neste trabalho.

A reação de qPCR permitiu verificar variação do material genético na matriz extracelular em diferentes tempos da curva de crescimento. Observa-se que a indução de mitomicina diminui o número de cópias dos genes *cap1*, 2 e 3 em 12 e 24 horas em relação ao controle, sendo apenas significativa nos genes *cap1* e 2 ( $p < 0,05$ ). Esses dados condizem com os resultados de PCR do conteúdo fágico nas vesículas de membrana externa, que não transportam o gene *cap1* na fase estacionária da amostra induzida, representada pelo ponto de 24 horas na curva de crescimento. Essa diminuição do número de cópias pode representar uma possível retroinibição, uma vez que foi demonstrado que a transcrição desses genes ocorre de forma espontânea e a mitomicina C tem a capacidade de ativar genes do ciclo lítico (Chou et al., 2017). O gene *cap2* foi o único que apresentou aumento do número de cópias na presença de mitomicina C, porém sem diferença estatística em relação a amostra controle. Mas, este resultado indicar que as VLPs encontradas após 6 horas de indução podem pertencer ao proDAG3, o profago que possui todas as características para ser um profago induzível. Essa variação do material genético na matriz extracelular ocorre devido a presença de OMVs, que podem carregar de forma aleatória componentes intracelulares.

As OMVs também podem agir como defesa inata bacteriana contra bacteriófagos (Manning and Kuehn, 2011), na qual os bacteriófagos injetam seu material genético nas vesículas, diminuindo a chance de infecção das células hospedeiras. A infecção de uma célula hospedeira por bacteriófagos também pode ser impedida pela presença de profagos no genoma (Pratama and van Elsas, 2017) e pelo sistema CRISPR-cas (Edgar and Qimron, 2010). Estas características podem impedir o isolamento de bacteriófagos líticos para a célula hospedeira alvo. A dificuldade de isolamento de bacteriófagos líticos foi encontrada por Pratama e colaboradores em *Paraburkholderia terrae* (Pratama and van Elsas, 2017), que identificaram a presença de profagos no genoma da bactéria hospedeira. Devido a estes fatos, justifica-se a procura de novas alternativas para o controle bacteriano e o melhor entendimento da função de OMVs e indução espontânea de profagos em *D. alaskensis*.

## Conclusão

Neste estudo foi verificado que profagos de *D. alaskensis* DSM16109 apresentam indução espontânea com aumento do número de cópias dos genes *cap* sem a formação de partículas virais, as quais foram observadas na amostra induzida com mitomicina C. As partículas encontradas são defectivas e sem material genético, caracterizando VLPs. Material genético tanto fágico quanto bacteriano foi encontrado compondo vesículas de membrana externa. Este trabalho mostra pela primeira vez que vesículas de membrana externa podem estar relacionadas com a THG nesta espécie de BRS e também abre portas para posteriores estudos de caracterização de OMVs de *D. alaskensis*.

## Métodos

### Caracterização de profagos em genomas de *D. alaskensis*

Profagos em *Desulfovibrio alaskensis* (G20 e DSM16109) foram anteriormente identificados (Crispim et al., 2018), Tabela 1. Para verificar a relação de identidade dos profagos entre as duas linhagens foi utilizada a ferramenta Blast (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). As sequências dos profagos foram avaliadas manualmente por meio dos seguintes programas: FGENESV0 (*Gene Finding in Viral Genomes* – <http://softberry.com>) e ViruSITE (*Integrated database for viral genomics*) (Stano et al., 2016), para encontrar e identificar a função de genes virais, respectivamente. O Phyre2 (Kelley et al., 2015) e PSI-BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) foram utilizados para identificação de proteínas com função desconhecida.

**Tabela 1: Elementos profago-like em *Desulfovibrio alaskensis* identificado por Crispim e colaboradores (2018).**

Linhagem (Nº de acesso)	Elemento profago-like	Contig (localização)	Tamanho (pb)	Classificação	GC (%)	Genoma (%)	Família
G20	1	939853-980902	41049	Completo	63.1	1,10	<i>Myoviridae</i>
	2	2787989-2795585	7596	Hipotético	47.8	0,20	-
	3	3331749-3371978	40229	Completo	55.2	1,07	<i>Myoviridae</i>
	4	1918124-1966812	48688	Completo	60.7	1,30	<i>Myoviridae</i>
DSM16109	1	1:145720-186292	40573	Completo	63.2	1,14	<i>Myoviridae</i>
	2	6:317764-342216	24452	Completo	58.5	0,68	<i>Myoviridae</i>

### Cultivo bacteriano e curva de crescimento

A verificação da expressão de profagos de *D. alaskensis* foi realizada na linhagem DSM16109, que foi rotineiramente cultivada em meio Postgate C (Postgate, 1965) sem FeSO<sub>4</sub>, em condição anaeróbica à 30°C. O crescimento

bacteriano foi realizado em tubos do tipo Hungate, com a leitura da densidade óptica em comprimento de onda de 600 nm ( $DO_{600}$ ) realizada em espectrofotômetro SPECTRONIC 20D (Thermo-Fisher Scientific, Waltham, MA USA). Para a indução de profagos foi inoculado mitomicina C (Sigma-Aldrich, St. Louis, MO, USA) na concentração final de 5 e 10  $\mu\text{g}/\text{mL}$  quando o crescimento bacteriano atingiu a  $DO_{600}$  de 0,1. Posteriormente, a amostra foi cultivada por 60 horas para acompanhamento do crescimento, seguindo a metodologia descrita por Dieterle e colaboradores (Dieterle et al., 2016). Os ensaios posteriores foram realizados com 5  $\mu\text{g}/\text{ml}$  de mitomicina C.

### **Concentração de partícula viral**

A cultura bacteriana de *D. alaskensis* DSM16109 foi induzida com mitomicina C por 6 horas, tratada com clorofórmio 4 % e centrifugada à 5.000 g por 10 minutos. O sobrenadante foi filtrado em membrana 0,22  $\mu\text{m}$  e tratado com NaCl 1,0 M e polietilenoglicol 8 KDa (PEG 8000) na concentração final de 10 %. As partículas viras foram concentradas por centrifugação à 15.000 g por 30 minutos, à 4 °C e suspendidas em 1 mL de tampão SM [10mM de NaCl, 50mM de Tris (pH 7,5), 10mM de  $\text{MgSO}_4$  e 0,01% de gelatina]. Posteriormente, foi adicionado 1 mL de clorofórmio para a retirada do PEG. Após uma nova centrifugação de 15.000 g por 30 minutos, a parte superior da amostra foi armazenada à 4°C para posteriores análises. O mesmo procedimento foi realizado com a cultura não induzida com mitomicina C.

### **Microscopia eletrônica**

Aproximadamente 10  $\mu\text{L}$  das amostras reservadas foram adicionados à grids de 200 mesh (Electron Microscopy Sciences, Hatfield, PA, USA) cobertas com Formvar® e incubadas por 5 minutos. A contrastação da amostra foi realizada com acetato de uranila 5 % por 10 segundos. As amostras foram analisadas em microscópio eletrônico de transmissão Zeiss Libra 120 (Zeiss, Oberkochen Germany).

### **Verificação de material genético nas partículas virais**

Este teste foi realizado por PCR, de acordo com Pratama e colaboradores (Pratama and van Elsas, 2017). A amostra induzida por 6 horas foi tratada com

DNase para degradar qualquer DNA presente fora do capsídeo viral. Para isso, 12 µL de cada amostra foram tratados com 4 u de DNase RQ1 (Promega, Madison, USA) e 2 µL de tampão (10X), à 37°C por 30 minutos. Posteriormente, a amostra foi submetida à 95°C por 30 minutos para inativação da DNase e abertura do capsídeo viral. A reação de PCR foi realizada com 100 ng de DNA nas seguintes condições: desnaturação à 95°C por 30 segundos, anelamento à 60°C por 30 segundos e alongação à 72°C por 15 segundos, em um total de 30 ciclos. Foram usados os primers que codificam a proteína do capsídeo de cada um dos três profagos: Cap1 (capsídeo do profago 1), Cap2 (capsídeo do profago 2) e Cap3 (capsídeo do profago 3), Tabela 2. Os amplicons foram visualizados em gel de agarose 1,5 %.

**Tabela 2: Primers usados para amplificação específica dos genes dos capsídeos virais dos três profagos de *Desulfovibrio alaskensis*.**

Nº	Região	Nome do primer	Sequência (5' → 3')	Tamanho do amplicon (pb)
1	Capsídeo 1 - profago 1	Cap1F	CGGTGCCGTTTCTCTGGTCA	91
2		Cap1R	CTTGAGCGGGCGGGATACAT	
3	Capsídeo 2 - profago 2	Cap2F	TGAGACTCGCATCGGCATGG	107
4		Cap2R	GCTGCAGTCCAACGGCTCTT	
5	Capsídeo 3 - profago 3	Cap3F	GACGCCATCCGCAAGGAGAT	107
5		Cap3R	GCCTCTTCGGTGAGCAGGTC	
6	DNA girase	GyrAF	GGTGCCCAACCTGCTGCTTA	99
7		GyrAR	GAAGTGGTTGACGGCCTGCT	

### **Transcrição reversa do RNAm que codificam os capsídeos virais**

A transcrição de genes dos profagos identificados em *D. alaskensis* por Crispim e colaboradores (Crispim et al., 2018) foi verificada por transcrição reversa. Para isso, células da linhagem DSM16109, sem e com tratamento com mitomicina C após 6 horas de crescimento, foram submetidas ao processo de extração de RNA utilizando o TRI Reagente (Sigma-Aldrich, St. Louis, MO, USA), de acordo com as recomendações do fabricante. As amostras foram tratadas com DNase RQ1 (Promega, Madison, USA) e a síntese do DNA complementar (cDNA) realizada por meio do Kit GoScript (Promega, Madison, USA). A reação de PCR foi realizada com 100 ng de DNA nas condições descritas anteriormente. O DNA genômico da linhagem DSM16109 foi usado como controle positivo, as amostras de RNA (controle e tratamento com mitomicina C) sem transcrição reversa foram usadas como controle negativo e utilizados os primers Cap1, Cap2 e Cap3, Tabela 2.

## **Análise de OMVs**

A linhagem bacteriana foi induzida com mitomicina C por 24 horas, centrifugada à 6000 g por 10 minutos e o sobrenadante filtrado em membrana de 0,22 µm. A concentração de vesículas de membrana externa foi realizada de acordo com Persson e colaboradores (Persson et al., 2018). Posteriormente, foi realizado ultracentrifugação (Sorvall™ WX+, rotor T-8100, Thermo-Fisher Scientific, Waltham, MA USA) à 49.900 rpm (aproximadamente 200.000 g) por 3 horas à 4°C. O precipitado formado foi suspenso em PBS estéril e a presença de vesículas verificada por microscopia eletrônica de transmissão, como descrito anteriormente. O mesmo procedimento foi realizado com a amostra controle.

A presença de material genético nas vesículas foi verificada por PCR, como realizado por Persson e colaboradores (Ho et al., 2015). Para isso, 13 µL da amostra foram tratadas com 5 u de DNase RQ1 (Promega, Madison, USA) e 2 µL de tampão (10X) à 37°C por 30 minutos, para eliminar qualquer material genético extravesicular. A reação de PCR foi realizada com os primers Cap1, Cap2, Cap3 e GyrA, como descrito anteriormente.

## **Verificação de material genético extracelular por PCR**

Para verificação de material genético extracelular em amostras controles e induzidas por mitomicina C foi realizada uma PCR com quatro pontos da curva de crescimento da linhagem DSM16109: ponto 1 – ponto a partir do qual foi inoculado mitomicina C (DO<sub>600</sub> 0,1), ponto 2 – 6 horas após indução, ponto 3 – 12 horas após indução e ponto 4 – 24 após a indução. Essas amostras foram centrifugadas à 6000 g por 10 minutos, o sobrenadante filtrado em membrana 0,22 µm e tratado com clorofórmio 10%. Após quantificadas, alíquotas de 100 ng de DNA de cada amostra foram usadas na reação de PCR semi-quantitativo nas seguintes condições: desnaturação à 95°C por 30 segundos, anelamento à 60°C por 30 segundos e alongação à 72°C por 15 segundos, em um total de 20 ciclos, usando os primers para o gene da DNA girase (*gyrA*) e para os genes das proteínas dos capsídeos (*cap1*, *cap2* e *cap3*), Tabela 2.

## **PCR quantitativo (qPCR)**

Para verificação do número de cópias dos profagos no material extracelular nos 4 pontos da curva de crescimento foi realizada uma qPCR. Os primers para os

genes das proteínas dos capsídeos (*cap1*, *cap2* e *cap3*) foram usados na reação, como realizado por Pratama e colaboradores (Pratama and van Elsas, 2017). O primer para o gene *gyrA* foi usado como controle endógeno do genoma bacteriano, Tabela 2. A qPCR foi realizada com 5 µL de SYBR Green PCR Master Mix (Bio-Rad, USA), 0,5 µL de cada primer (5 µmol/l), 3 µL de água ultrapura e 1 µL da amostra (100 ng), para um total de 10 µL de reação. A amplificação e detecção foram realizadas por meio do equipamento Eco™ Real-Time PCR System (Illumina, CA, USA) com as seguintes condições da reação de qPCR: desnaturação 95°C por 30 segundos, anelamento à 60°C por 60 segundos e elongação à 72°C por 60 segundos. O ponto 0 foi utilizado como ponto de referência e o experimento foi realizado em triplicata biológica.

### **Análise estatística**

A análise estatística do número de cópias dos genes que codificam as proteínas dos capsídeos dos profagos em relação ao gene da DNA girase foi realizada por meio do teste-*t*, sendo considerado estatisticamente diferente os valores de  $p < 0,05$ .

### **Agradecimentos**

Agradecemos à CNPq, FAPEMIG, CAPES e Petrobrás pelo suporte financeiro. Ao Núcleo de Microscopia e Microanálise da Universidade Federal de Viçosa e aos técnicos deste Núcleo pelo apoio e auxílio na microscopia eletrônica. Ao Núcleo de Análise de Biomoléculas pelo suporte na análise dos dados. À Jamile Fernanda S. Cossolin e ao professor José Eduardo Serrão do Departamento de Biologia Geral da UFV pela ajuda nos experimentos.

## Referências

- Bobay, L.-M., Rocha, E.P.C., Touchon, M., 2013. The adaptation of temperate bacteriophages to their host genomes. *Mol. Biol. Evol.* 30, 737–751.
- Bobay, L.-M., Touchon, M., Rocha, E.P.C., 2014. Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci.* 111, 12127–12132.
- Bonnington, K.E., Kuehn, M.J., 2014. Protein selection and export via outer membrane vesicles. *Biochim. Biophys. Acta* 1843, 1612–1619.
- Canchaya, C., Fournous, G., Brussow, H., 2004. The impact of prophages on bacterial chromosomes. *Mol. Microbiol.* 53, 9–18.
- Carrolo, M., Frias, M.J., Pinto, F.R., Melo-Cristino, J., Ramirez, M., 2010. Prophage spontaneous activation promotes DNA release enhancing biofilm formation in *Streptococcus pneumoniae*. *PLoS One* 5, e15678.
- Chen, F., Wang, K., Stewart, J., Belas, R., 2006. Induction of multiple prophages from a marine bacterium: a genomic approach. *Appl. Environ. Microbiol.* 72, 4995–5001.
- Chen, M., Zhang, L., Xin, S., Yao, H., Lu, C., Zhang, W., 2017. Inducible prophage mutant of *Escherichia coli* can lyse new host and the key sites of receptor recognition identification. *Front. Microbiol.* 8, 1–13.
- Chikindas, M.L., Weeks, R., Drider, D., Chistyakov, V.A., Dicks, L.M., 2018. Functions and emerging applications of bacteriocins. *Curr. Opin. Biotechnol.* 49, 23–28.
- Chou, W.-C., Huang, S.-C., Chiu, C.-H., Chen, Y.-Y.M., 2017. YMC-2011, a temperate phage of *Streptococcus salivarius* 57.I. *Appl. Environ. Microbiol.* 83.
- Crispim, J.S., Dias, R.S., Vidigal, P.M.P., de Sousa, M.P., da Silva, C.C., Santana, M.F., de Paula, S.O., 2018. Screening and characterization of prophages in *Desulfovibrio* genomes. *Sci. Rep.* 8, 9273.
- Dieterle, M.E., Fina Martin, J., Durán, R., Nemirovsky, S.I., Sanchez Rivas, C., Bowman, C., Russell, D., Hatfull, G.F., Cambillau, C., Piuri, M., 2016. Characterization of prophages containing “evolved” Dit/Tal modules in the genome of *Lactobacillus casei* BL23. *Appl. Microbiol. Biotechnol.* 100, 9201–9215.
- Domingues, S., Nielsen, K.M., 2017. Membrane vesicles and horizontal gene transfer in prokaryotes. *Curr. Opin. Microbiol.* 38, 16–21.
- Durmaz, E., Miller, M.J., Azcarate-Peril, M.A., Toon, S.P., Klaenhammer, T.R., 2008. Genome sequence and characteristics of Lrm1, a prophage from industrial

- Lactobacillus rhamnosus* strain M1. *Appl. Environ. Microbiol.* 74, 4601–4609.
- Dutta, S., Iida, K., Takade, A., Meno, Y., Nair, G.B., Yoshida, S., 2004. Release of Shiga toxin by membrane vesicles in *Shigella dysenteriae* serotype 1 strains and in vitro effects of antimicrobials on toxin production and release. *Microbiol. Immunol.* 48, 965–969.
- Edgar, R., Qimron, U., 2010. The *Escherichia coli* CRISPR system protects from  $\lambda$  lysogenization, lysogens, and prophage induction. *J. Bacteriol.* 192, 6291–6294.
- Feio, M.J., Zinkevich, V., Beech, I.B., Llobet-Brossa, E., Eaton, P., Schmitt, J., Guezennec, J., 2004. *Desulfovibrio alaskensis* sp. nov., a sulphate-reducing bacterium from a soured oil reservoir. *Int. J. Syst. Evol. Microbiol.* 54, 1747–1752.
- Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* 3, 722–732.
- Gandon, S., 2016. Why be temperate: lessons from bacteriophage  $\lambda$ . *Trends Microbiol.* 24, 356–365. <https://doi.org/10.1016/j.tim.2016.02.008>
- Gaudin, M., Krupovic, M., Marguet, E., Gauthier, E., Cvirkaite-Krupovic, V., Le Cam, E., Oberto, J., Forterre, P., 2014. Extracellular membrane vesicles harbouring viral genomes. *Environ. Microbiol.* 16, 1167–1175.
- Gieg, L.M., Jack, T.R., Foght, J.M., 2011. Biological souring and mitigation in oil reservoirs. *Appl. Microbiol. Biotechnol.* 92, 263–282.
- Guo, M., Feng, C., Ren, J., Zhuang, X., Zhang, Y., Zhu, Y., Dong, K., He, P., Guo, X., Qin, J., 2017. A novel antimicrobial endolysin, LysPA26, against *Pseudomonas aeruginosa*. *Front. Microbiol.* 8, 293.
- Hammerl, J.A., Göllner, C., Al Dahouk, S., Nöckler, K., Reetz, J., Hertwig, S., 2016. Analysis of the first temperate broad host range Brucellaphage (BiPBO1) isolated from *B. inopinata*. *Front. Microbiol.* 7, 24.
- Hauser, L.J., Land, M.L., Brown, S.D., Larimer, F., Keller, K.L., Rapp-Giles, B.J., Price, M.N., Lin, M., Bruce, D.C., Detter, J.C., Tapia, R., Han, C.S., Goodwin, L.A., Cheng, J.-F., Pitluck, S., Copeland, A., Lucas, S., Nolan, M., Lapidus, A.L., Palumbo, A. V, Wall, J.D., 2011. Complete genome sequence and updated annotation of *Desulfovibrio alaskensis* G20. *J. Bacteriol.* 193, 4268–4269.
- Haynes, T.S., Klemm, D.J., Ruocco, J.J., Barton, L.L., 1995. Formate dehydrogenase activity in cells and outer membrane blebs of *Desulfovibrio gigas*. *Anaerobe* 1, 175–182.

- Ho, M.-H., Chen, C.-H., Goodwin, J.S., Wang, B.-Y., Xie, H., 2015. Functional advantages of *Porphyromonas gingivalis* vesicles. *PLoS One* 10, e0123448.
- Jin, T., Zhang, X., Zhang, Y., Hu, Z., Fu, Z., Fan, J., Wu, M., Wang, Y., Shen, P., Chen, X., 2014. Biological and genomic analysis of a PBSX-like defective phage induced from *Bacillus pumilus* AB94180. *Arch. Virol.* 159, 739–752.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modelling, prediction and analysis. *Nat. Protoc.* 10, 845–858.
- Klieve, A. V, Yokoyama, M.T., Forster, R.J., Ouwerkerk, D., Bain, P.A., Mawhinney, E.L., 2005. Naturally occurring DNA transfer system associated with membrane vesicles in cellulolytic *Ruminococcus* spp. of ruminal origin. *Appl. Environ. Microbiol.* 71, 4248–4253. h
- Kulp, A., Kuehn, M.J., 2010. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annu. Rev. Microbiol.* 64, 163–184.
- Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y., Drulis-Kawa, Z., 2017. Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Appl. Microbiol. Biotechnol.* 101, 3103–3119.
- Lovley, D.R., Ueki, T., Zhang, T., Malvankar, N.S., Shrestha, P.M., Flanagan, K.A., Aklujkar, M., Butler, J.E., Giloteaux, L., Rotaru, A.-E., Holmes, D.E., Franks, A.E., Orellana, R., Risso, C., Nevin, K.P., 2011. *Geobacter*: the microbe electric's physiology, ecology, and practical applications. *Adv. Microb. Physiol.* 59, 1–100.
- Manning, A.J., Kuehn, M.J., 2011. Contribution of bacterial outer membrane vesicles to innate bacterial defense. *BMC Microbiol.* 11, 258.
- McAuliffe, O., Ross, R.P., Hill, C., 2001. Lantibiotics: structure, biosynthesis and mode of action. *FEMS Microbiol. Rev.* 25, 285–308.
- Mobberley, J., Authement, R.N., Segall, A.M., Edwards, R.A., Slepecky, R.A., Paul, J.H., 2010. Lysogeny and sporulation in *Bacillus* isolates from the Gulf of Mexico. *Appl. Environ. Microbiol.* 76, 829–842. 9
- Nanda, A.M., Heyer, A., Kramer, C., Grunberger, A., Kohlheyer, D., Frunzke, J., 2014. Analysis of SOS-induced spontaneous prophage induction in *Corynebacterium glutamicum* at the single-cell level. *J. Bacteriol.* 196, 180–188.
- Nanda, A.M., Thormann, K., Frunzke, J., 2015. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions.

- J. Bacteriol. 197, 410–419.
- Nazina, T.N., Shestakova, N.M., Semenova, E.M., Korshunova, A. V, Kostrukova, N.K., Tourova, T.P., Min, L., Feng, Q., Poltarau, A.B., 2017. Diversity of metabolically active bacteria in water-flooded high-temperature heavy oil reservoir. *Front. Microbiol.* 8, 707.
- Persson, G., Pors, S.E., Thøfner, I.C.N., Bojesen, A.M., 2018. Vaccination with outer membrane vesicles and the fimbrial protein FliA offers improved protection against lesions following challenge with *Gallibacterium anatis*. *Vet. Microbiol.* 217, 104–111.
- Plugge, C.M., Henstra, A.M., Worm, P., Swarts, D.C., Paulitsch-Fuchs, A.H., Scholten, J.C.M., Lykidis, A., Lapidus, A.L., Goltsman, E., Kim, E., McDonald, E., Rohlin, L., Crable, B.R., Gunsalus, R.P., Stams, A.J.M., McInerney, M.J., 2012. Complete genome sequence of *Syntrophobacter fumaroxidans* strain (MPOB(T)). *Stand. Genomic Sci.* 7, 91–106.
- Postgate, J.R., 1965. Recent advances in the study of the sulfate-reducing bacteria. *Bacteriol. Rev.* 29, 425–441.
- Pratama, A.A., van Elsas, J.D., 2017. A novel inducible prophage from the mycosphere inhabitant *Paraburkholderia terrae* BS437. *Sci. Rep.* 7, 9156.
- Rajasekar, A., Anandkumar, B., Maruthamuthu, S., Ting, Y.-P., Rahman, P.K.S.M., 2010. Characterization of corrosive bacterial consortia isolated from petroleum-product-transporting pipelines. *Appl. Microbiol. Biotechnol.* 85, 1175–1188.
- Reyes-Robles, T., Dillard, R.S., Cairns, L.S., Silva-Valenzuela, C.A., Housman, M., Ali, A., Wright, E.R., Camilli, A., 2018. *Vibrio cholerae* outer membrane vesicles inhibit bacteriophage infection. *J. Bacteriol.*
- Rodriguez-Rubio, L., Martinez, B., Donovan, D.M., Rodriguez, A., Garcia, P., 2013. Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzymatics. *Crit. Rev. Microbiol.* 39, 427–434.
- Shi, Y., Yan, Y., Ji, W., Du, B., Meng, X., Wang, H., Sun, J., 2012. Characterization and determination of holin protein of *Streptococcus suis* bacteriophage SMP in heterologous host. *Viol. J.* 9, 70.
- Smith, M.J., Jeddloh, J.A., 2005. DNA methylation in lysogens of pathogenic *Burkholderia* spp. requires prophage induction and is restricted to excised phage DNA. *J. Bacteriol.* 187, 1196–1200.
- Stano, M., Beke, G., Klucar, L., 2016. viruSITE—integrated database for viral

- genomics. Database J. Biol. Databases Curation 2016, baw162.
- Tang, Y., Nielsen, L.N., Hvitved, A., Haaber, J.K., Wirtz, C., Andersen, P.S., Larsen, J., Wolz, C., Ingmer, H., 2017. Commercial biocides induce transfer of prophage  $\Phi$ 13 from human strains of *Staphylococcus aureus* to livestock CC398. *Front. Microbiol.* 8, 2418.
- Tavares, P., 2018. The bacteriophage head-to-tail interface. *Subcell. Biochem.* 88, 305–328.
- Wang, X., Kim, Y., Ma, Q., Hong, S.H., Pokusaeva, K., Sturino, J.M., Wood, T.K., 2010. Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* 1, 147.
- Yuan, Y., Gao, M., Peng, Q., Wu, D., Liu, P., Wu, Y., 2014. Genomic analysis of a phage and prophage from a *Bacillus thuringiensis* strain. *J. Gen. Virol.* 95, 751–761.

**Tabela S1: Regiões codificadoras do profago 1 (profago 2 de *D. alaskensis* DSM16109 descrito por Crispim e colaboradores (Crispim et al., 2018)).**

Gene	Localização	pb	Fita	Resíduos de aa	Homologia	Proteína correspondente (código NCBI)	E-value
1	317764-318528	764	-	254	Proteína de regulação transcricional hipotética	WP_051363901.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam07022 regulador transcricional	0.0
2	318793-319242	449	+	149	Proteína de regulação transcricional (CI) hipotética	WP_027181635.1 proteína hipotética ( <i>D. alaskensis</i> ); phyre2: regulador transcricional CI	4e-101
3	319239-319493	254	+	84	Regulador transcricional	WP_027181636.1 regulador transcricional ( <i>D. alaskensis</i> ); motivo: smart00530 regulador transcricional	7e-54
4	319436-319789	353	+	117	Proteína hipotética	WP_027181637.1 proteína hipotética ( <i>D. alaskensis</i> )	9e-68
5	319824-321470	1646	+	548	Proteína integrase hipotética	WP_051363902.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam00665 integrase	0.0
6	321502-321672	170	+	56	Proteína hipotética	WP_027181638.1 proteína hipotética ( <i>D. alaskensis</i> )	2e-29
7	321676-323496	1820	+	606	Proteína terminase hipotética	ACL07685.1 proteína não caracterizada de fago ( <i>D. vulgaris</i> Miyazaki F); motivo: TIGR01630 terminase; VirFam: ACL07685.1 proteína terminase (TermL)	0.0
8	323504-325084	1580	+	526	Proteína portal hipotética	WP_027181639.1 proteína hipotética ( <i>D. alaskensis</i> ); Motivo: COG4383 proteína gp29-like do profago Mu; virusite: YP_009013330 proteína portal; VirFam: WP_027181639.1 proteína portal	0.0
9	325086-326360	1274	+	424	Morfogênese	WP_027181640.1 proteína de morfogênese viral ( <i>D. alaskensis</i> ); motivo: TIGR01641 proteína de morfogênese; virusite: YP_009226538 morfogênese	0.0
10	326497-326988	491	+	163	Morfogênese (Ne1)	WP_027181641.1 proteína de morfogênese viral ( <i>D. alaskensis</i> ); motivo: pfam05069 proteína de morfogênese; virusite: NP_050635 morfogênese; VirFam: WP_027181641.1 proteína de conexão Ne1	1e-113
11	327160-328392	1232	+	410	Protease	WP_027181642.1 peptidase ( <i>D. alaskensis</i> ); motivo: pfam10123 proteína-like com profago Mu; virusite: YP_009168761 protease	0.0
12	328395-329324	929	+	309	Proteína do capsídeo	WP_027181643.1 capsídeo ( <i>D. alaskensis</i> ); motivo: pfam10124 capsídeo; virusite: YP_009125680 capsídeo; VirFam: WP_027181643.1 proteína do capsídeo (major)	0.0
13	329333-329626	293	+	97	Proteína viral hipotética	WP_027181644.1 proteína hipotética ( <i>D. alaskensis</i> ); phyre2: proteína-like gp35 do profago flumu	2e-39

14	329677-330030	353	+	117	Proteína Ad1 hipotética	WP_027181645.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: COG4387 proteína-like gp36 do profago Mu; VirFam: proteína de conexão Ad1	3e-76
15	330047-330655	608	+	202	Proteína Tc1 hipotética	WP_027181646.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam08873 pertence à família de proteínas de fagos (DUF1834) com função desconhecida; VirFam: WP_027181646.1 proteína de conexão Tc1	8e-144
16	330659-330868	209	+	69	Proteína viral hipotética	WP_027181647.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam10948 pertence à família de proteínas de fagos (DUF2635) com função desconhecida	3e-40
17	330885-332276	1391	+	463	Proteína contrátil da cauda	WP_027181648.1 Proteína viral contrátil ( <i>D. alaskensis</i> ); motivo: pfam04984 proteína viral contrátil; VirFam: WP_027181648.1 proteína contrátil da cauda	0.0
18	332416-333276	860	+	286	Proteína hipotética	WP_027181649.1 proteína hipotética ( <i>D. alaskensis</i> )	0.0
19	333260-333700	440	+	146	Proteína hipotética	WP_034618686.1 proteína hipotética ( <i>D. alaskensis</i> )	3e-35
20	333792-334166	374	+	124	Proteína da cauda hipotética	WP_027181651.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_009196087 proteína da cauda	5e-85
21	334184-334609	425	+	140	Proteína hipotética	WP_034618773.1 proteína hipotética ( <i>D. alaskensis</i> )	2e-64
22	334630-334791	161	+	53	Proteína hipotética	EGY27009.1 proteína hipotética ( <i>Desulfovibrio</i> sp. A2)	1e-04
23	334813-336780	1967	+	655	Controle do tamanho da cauda	WP_027181653.1 proteína de controle do tamanho cauda ( <i>D. alaskensis</i> ); motivo: pfam10145 proteína de controle do tamanho cauda; virusite: YP_009196090 proteína de controle do tamanho cauda	0.0
24	336790-338181	1391	+	463	Proteína de circularização do DNA hipotética	WP_027181654.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam07157 proteína de circularização do DNA; virusite: YP_009196091 proteína de circularização do DNA	0.0
25	338255-339265	1010	+	336	Protéina da cauda	WP_027181655.1 proteína da cauda ( <i>D. alaskensis</i> ); motivo: COG4379 proteína da cauda; virusite: YP_007002952 proteína da cauda	0.0
26	339286-339885	599	+	199	Proteína da placa basal	WP_034618687.1 proteína da placa basal ( <i>D. alaskensis</i> ); motivo: TIGR01644 proteína da placa basal; virusite: YP_007002953 proteína da placa basal	2e-144
27	339888-340145	257	-	85	Proteína hipotética	WP_027181657.1 proteína hipotética ( <i>D. alaskensis</i> )	7e-56

28	340197-340553	356	+	118	Proteína viral hipotética	WP_027181658.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: COG4381 proteína-like gp46 do profago Mu; phyre2: proteína-like gp46 do profago Mu; virusite: YP_009196094 proteína hipotética	1e-82
29	340573-341634	1061	+	353	Proteína da placa basal	WP_027181659.1 proteína da cauda ( <i>D. alaskensis</i> ); motivo: pfam04865 proteína da placa basal; virusite: YP_009196095 proteína da placa basal	0.0
30	341635-342216	581	+	193	Proteína da cauda	WP_027181660.1 proteína da cauda ( <i>D. alaskensis</i> ); motivo: pfam10076 pertence à família de proteínas de fagos (DUF2313) com função desconhecida; virusite: YP_009196096 proteína da cauda	3e-139

**Tabela S2: Regiões codificadoras do profago 2 (profago 3 de *D. alaskensis* G20 descrito por Crispim e colaboradores (Crispim et al., 2018)).**

Gene	Localização	pb	Fita	Resíduos de aa	Homologia	Proteína correspondente (código NCBI)	E-value
1	3332216-3332674	458	+	152	Proteína viral hipotética	WP_011369078.1 proteína GemA ( <i>D. alaskensis</i> ); motivo: pfam06252 pertence à família de proteínas de fagos (DUF1018) com função desconhecida;	6e-109
2	3333815-3334210	395	-	131	Proteína integrase	ABB40155.1 proteína integrase ( <i>D. alaskensis</i> ); motivo: pfam13683 proteína integrase; virusite: YP_001111115 transposase	1e-93
3	3334690-3334977	287	-	95	Proteína transposase	WP_011369082.1 proteína transposase ( <i>D. alaskensis</i> ); motivo: pfam01527 proteína transposase, virusite: YP_009289240 proteína transposase	1e-61
4	3335977-3336519	542	-	180	Proteína integrase	ABB40158.2 proteína integrase ( <i>D. alaskensis</i> ); motivo: pfam00665 proteína integrase; virusite: NP_958229 proteína transposase	4e-131
5	3336798-3336923	125	-	41	Proteína transposase	WP_041276950.1 proteína transposase ( <i>D. alaskensis</i> )	3e-18
6	3337535-3338032	497	+	165	Peptidase	WP_011369084.1 proteína peptidase M15 ( <i>D. alaskensis</i> ); motivo: pfam08291 peptidase M15; virusite: YP_008060500 peptidase	4e-118
7	3338036-3338644	608	+	202	Proteína hipotética	WP_011369085.1 proteína hipotética ( <i>D. alaskensis</i> )	2e-143
8	3338713-3338970	257	+	85	Proteína hipotética	WP_011369086.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam10805 pertence à família de proteínas (DUF2730) de função desconhecida	1e-52
9	3338967-3339263	296	+	98	Regulador transcricional hipotético	WP_011369087.1 proteína hipotética ( <i>D. alaskensis</i> ); phyre2: regulador transcricional	3e-65
10	3339268-3339864	596	+	198	Regulador transcricional hipotético	WP_011369088.1 proteína hipotética ( <i>D. alaskensis</i> ); phyre2: regulador transcricional	4e-137
11	3339877-3341349	1472	+	490	Proteína terminase hipotética	WP_011369089.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam03237 terminase; VirFam: WP_011369089.1 proteína terminase (TemrL)	0.0
12	3341396-3342931	1535	+	511	Proteína portal hipotética	WP_011369090.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: COG4383 proteína-like gp29 do profago Mu; virusite: YP_004934679 proteína portal	0.0
13	3342921-3344201	1280	+	426	Morfogênese do capsídeo (proteína portal)	WP_049753899.1 proteína de morfogênese do capsídeo ( <i>D. alaskensis</i> ); motivo: TIGR01641 proteína de morfogênese do capsídeo; virusite: YP_008129840 proteína de morfogênese do capsídeo; VirFam: WP_049753899.1 proteína portal	0.0

14	3344357-3345250	893	+	297	Proteína protease hipotética	WP_011369092.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam10123 proteína I-like com profago Um; virusite: YP_164072 protease	0.0
15	3345262-3345618	356	+	118	Proteína hipotética	WP_011369093.1 proteína hipotética ( <i>D. alaskensis</i> )	4e-78
16	3345622-3346611	989	+	329	Proteína do capsídeo	WP_011369094.1 proteína do capsídeo ( <i>D. alaskensis</i> ); motivo: pfam03864 proteína do capsídeo; phyre2: proteína do capsídeo; VirFam: WP_011369094.1 proteína do capsídeo (major)	0.0
17	3346725-3347156	431	+	143	Proteína Ad1 hipotética	WP_011369095.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam07030 proteína da família de proteína virais e bacterianas (DUF1320) de função desconhecida; VirFam: WP_011369095.1 proteína de conexão Ad1	6e-98
18	3347177-3347647	470	+	156	Morfogênese (Ne1)	WP_011369096.1 proteína de morfogênese viral ( <i>D. alaskensis</i> ); motivo: pfam05069 proteína de morfogênese viral; virusite: YP_009285927 proteína de morfogênese viral; VirFam: WP_011369096.1 proteína de conexão Ne1	3e-110
19	3347656-3348186	530	+	176	Proteína hipotética	WP_011369097.1 proteína hipotética ( <i>D. alaskensis</i> )	1e-126
20	3348179-3348394	215	+	70	Proteína contrátil da cauda hipotética	WP_011369098.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_001285652 proteína contrátil da cauda	1e-41
21	3348397-3350088	1691	+	563	Proteína contrátil da cauda hipotética	WP_011369099.1 proteína hipotética ( <i>D. alaskensis</i> ), phyre2: proteína da cauda; VirFam: WP_011369099.1 proteína contrátil da cauda (Sheath)	0.0
22	3350099-3350539	440	+	146	Proteína da cauda hipotética	WP_011369100.1 proteína hipotética ( <i>D. alaskensis</i> ); phyre2: proteína da cauda	2e-101
23	3350552-3350881	329	+	109	Proteína hipotética	WP_011369101.1 proteína hipotética ( <i>D. alaskensis</i> )	9e-76
24	3350899-3351084	185	+	61	Proteína hipotética	WP_011369102.1 proteína hipotética ( <i>D. alaskensis</i> )	4e-34
25	3351086-3353287	2201	+	733	Proteína de controle do tamanho da cauda	WP_011369103.1 proteína de controle do tamanho da cauda ( <i>D. alaskensis</i> ); motivo: TIGR01760 proteína de controle do tamanho da cauda; virusite: YP_007112502 proteína de controle do tamanho da cauda	0.0
26	3353289-3353804	515	+	171	Proteína hipotética	WP_011369104.1 proteína hipotética ( <i>D. alaskensis</i> )	4e-121
27	3353804-3354637	833	+	277	Proteína hipotética	WP_011369105.1 proteína hipotética ( <i>D. alaskensis</i> )	0.0
28	3354624-3355346	722	+	240	Proteína hipotética	WP_011369106.1 proteína hipotética ( <i>D. alaskensis</i> )	2e-176

29	3355768-3356229	461	+	153	Proteína hipotética da placa basal	WP_041276952.1 proteína hipotética ( <i>D. alaskensis</i> ), virusite: YP_009215885	1e-108
						proteína da placa basal	
30	3356201-3357364	1163	+	387	Proteína da placa basal	WP_011369108.1 proteína da placa basal ( <i>D. alaskensis</i> ); motivo: pfam04865	0.0
						proteína da placa basal, virusite: YP_007003523 proteína de montagem da placa basal	
31	3357364-3358011	647	+	215	Proteína da cauda hipotética	WP_011369109.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam09684	2e-156
						proteína da cauda, virusite: YP_001165271 proteína da cauda	
32	3358022-3359581	1559	+	519	Proteína da fimbria da cauda	WP_011369110.1 proteína da fimbria da cauda ( <i>D. alaskensis</i> ); motivo: pfam12571	0.0
						proteína da fimbria da cauda; virusite: YP_009284280 proteína da fimbria da cauda	
33	3359594-3360178	584	+	194	Proteína da fimbria da cauda hipotética	WP_011369111.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_008318502	8e-136
						fimbria da cauda	
34	3360417-3360602	185	+	61	Regulador transcricional (Com)	SHI61475.1 Proteína Com ( <i>D. alaskensis</i> ); motivo: pfam10122	7e-17
						proteína Com; virusite: YP_164084 regulador transcricional	
35	3360780-3361904	1124	+	374	Proteína envolvida no transporte do tipo ABC	WP_011369112.1 Proteína de ligação ao ATP no transporte ABC ( <i>D. alaskensis</i> ); motivo: pfam00005	0.0
						transportador ABC	
36	3362945-3364108	1163	+	387	Integrase	WP_011369114.1 integrase ( <i>D. alaskensis</i> ); motivo: pfam00589	0.0
						integrase; virusite: YP_009043902 integrase	
37	3364112-3365350	1238	+	412	Proteína hipotética	WP_011369115.1 proteína hipotética ( <i>D. alaskensis</i> )	0.0
38	3365917-3366162	245	+	81	Proteína hipotética	WP_011369116.1 proteína hipotética ( <i>D. alaskensis</i> )	1e-54
39	3366244-3366408	164	-	54	Proteína de transposição hipotética	WP_049753921.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: PHA02517	4e-29
						transposase; virusite: YP_008766887 transposase	
40	3366552-3366809	257	+	85	Proteína de regulação transcricional hipotética	WP_011369117.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: PLN02573	5e-57
						descarboxilase; phyre2: regulador transcricional	
41	3366806-3368296	1490	+	496	Proteína de replicação hipotética	WP_011369118.1 proteína hipotética ( <i>D. alaskensis</i> ); phyre2: replicação do DNA	0.0
42	3368632-3369381	749	-	249	Proteína de replicação hipotética	ABB39702.1 proteína de ligação ao ATP domínio IstB; motivo: COG1484	0.0
						replicação do DNA	

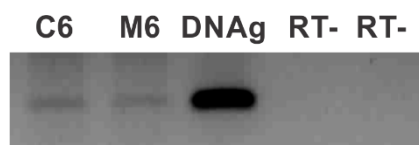
43	3369405-3370433	1028	-	342	Transposase	WP_011368690.1 transposase ( <i>D. alaskensis</i> ); motivo: pfam00665 integrase; virusite: YP_008766948 transposase	0.0
44	3370634-3370912	278	-	92	Proteína hipotética	WP_011369119.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_009303417 proteína hipotética	3e-49
45	3371229-3371360	131	+	43	Proteína hipotética	-	-
46	3371394-3371600	206	-	68	Proteína hipotética	WP_041276954.1 proteína hipotética ( <i>D. alaskensis</i> )	5e-43

**Tabela S3: Regiões codificadoras do profago 3 (profago 1 de *D. alaskensis* G20 descrito por Crispim e colaboradores (Crispim et al., 2018)).**

Gene	Localização	pb	Fita	Resíduos de aa	Homologia	Proteína correspondente (código NCBI)	E-value
1	939853-940767	914	-	300	Integrase	WP_011366957.1 integrase ( <i>D. alaskensis</i> ); motivo: pfam00589 integrase; virusite: YP_009276110 integrase	0.0
2	941490-941762	272	+	90	Proteína hipotética	WP_011366960.1 proteína hipotética ( <i>D. alaskensis</i> )	6e-60
3	941759-942037	278	+	92	Proteína hipotética	WP_011366961.1 proteína hipotética ( <i>D. alaskensis</i> )	0.0
4	942037-943560	1523	+	507	Proteína terminase hipotética	WP_011366962.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam03237 terminase; virusite: YP_008059213 terminase; VirFam: WP_011366962.1 proteína terminase (TermL)	0.0
5	943636-945183	1547	+	515	Proteína portal hipotética	WP_011366963.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam04860 proteína portal; virusite: YP_001429612 proteína portal	0.0
6	945183-949772	4598	+	1529	Morfogênese (portal)	WP_011366964.1 proteína de morfogênese ( <i>D. alaskensis</i> ); motivo: TIGR01641 proteína de morfogênese; virusite: YP_009285818 proteína de morfogênese; VirFam: WP_011366964.1 proteína portal	0.0
7	949769-950281	512	+	170	Proteína hipotética	WP_011366965.1 proteína hipotética ( <i>D. alaskensis</i> )	3e-121
8	950278-950631	353	+	117	Proteína hipotética	WP_011366966.1 proteína hipotética ( <i>D. alaskensis</i> )	6e-81
9	950735-953131	2396	+	798	DNA Metiltransferase	WP_011366967.1 metiltransferase ( <i>D. alaskensis</i> ); motivo: pfam02086 metiltransferase; virusite: YP_008061118 protease	0.0
10	953144-953506	362	+	120	Proteína hipotética	WP_011366968.1 proteína hipotética ( <i>D. alaskensis</i> )	2e-81
11	953525-954388	863	+	287	Proteína hipotética do capsídeo	WP_011366969.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_009204653 proteína do capsídeo; VirFam: WP_011366969.1 proteína do capsídeo (major)	0.0
12	954407-954766	359	+	119	Proteína hipotética	WP_011366970.1 proteína hipotética ( <i>D. alaskensis</i> )	3e-78
13	954753-955217	464	+	154	Proteína hipotética	WP_011366971.1 proteína hipotética ( <i>D. alaskensis</i> )	5e-107
14	955204-955563	359	+	119	Proteína hipotética	WP_011366972.1 proteína hipotética ( <i>D. alaskensis</i> )	3e-80
15	955553-956038	485	+	161	Proteína de morfogênese hipotética	WP_011366973.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_009168596 proteína de morfogênese; VirFam: WP_011366973.1 proteína de conexão Ne1	8e-114
16	956098-956721	623	+	207	Proteína hipotética	WP_011366974.1 proteína hipotética ( <i>D. alaskensis</i> )	4e-147

17	956734-956991	257	+	85	Proteína hipotética	WP_011366975.1 proteína hipotética ( <i>D. alaskensis</i> )	5e-53
18	956994-958523	1529	+	509	Proteína contrátil da cauda	WP_011366976.1 proteína da cauda ( <i>D. alaskensis</i> ); motivo: pfam04984 proteína contrátil da cauda; virusite: YP_009302431 proteína contrátil da cauda; VirFam: WP_011366976.1 proteína contrátil da cauda (sheath)	0.0
19	958539-959000	461	+	153	Proteína da cauda	WP_011366977.1 proteína da cauda ( <i>D. alaskensis</i> ); motivo: pfam06841 proteína da cauda; VirFam: WP_011366977.1 proteína da cauda (major)	5e-111
20	959016-959693	677	+	225	Proteína da placa basal hipotética	WP_011366978.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_009302433 proteína da placa basal	1e-161
21	959837-963472	3635	+	1211	Proteína de controle do tamanho da cauda	WP_011366979.1 proteína de controle do tamanho da cauda; motivo: TIGR01760 proteína de controle do tamanho da cauda; virusite: NP_536663 proteína de controle do tamanho da cauda	0.0
22	963457-963597	140	-	46	Proteína hipotética	-	-
23	963673-963792	119	+	39	Proteína hipotética	WP_011366980.1 proteína hipotética ( <i>D. alaskensis</i> )	7e-20
24	963796-964266	470	+	156	Proteína hipotética	WP_011366981.1 proteína hipotética ( <i>D. alaskensis</i> )	4e-113
25	964263-964553	290	+	96	Proteína hipotética	WP_011366982.1 proteína hipotética ( <i>D. alaskensis</i> )	1e-62
26	964620-965903	1283	+	427	Proteína da cauda hipotética	WP_011366983.1 proteína hipotética ( <i>D. alaskensis</i> )35; motivo: PHA02561 proteína da cauda; virusite: YP_001718718 proteína da cauda	0.0
27	965896-966270	374	+	124	Peptidase M15	WP_011366984.1 peptidase ( <i>D. alaskensis</i> ); motivo: pfam08291 peptidase; virusite: YP_009102792 peptidase	2e-87
28	966267-966686	419	+	139	Proteína de montagem da cauda hipotética	WP_011366985.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_009010153 proteína de montagem da cauda	8e-96
29	966679-967728	1049	+	349	Proteína da placa basal hipotética	WP_011366986.1 proteína hipotética ( <i>D. alaskensis</i> ); virusite: YP_009302439 proteína da placa basal	0.0
30	967741-968115	374	+	124	Proteína hipotética	WP_011366987.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: COG4653 proteína-like gp36 do fago phi-C31	8e-83
31	968112-968375	263	+	87	Proteína de membrana hipotética	WP_011366988.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam05488 proteína de membrana bacteriana	6e-54
32	968652-968936	284	-	94	Proteína hipotética	-	-

33	968898-969701	803	+	267	Proteína hipotética	WP_011366990.1 proteína hipotética ( <i>D. alaskensis</i> )	0.0
34	969714-970604	890	+	296	Proteína hipotética	WP_011366991.1 proteína hipotética ( <i>D. alaskensis</i> )	0.0
35	970604-971521	917	+	305	Proteína lisozima hipotética	WP_011366992.1 proteína hipotética ( <i>D. alaskensis</i> ); motivo: pfam04965 proteína lisozima-like; virusite: YP_717766 proteína da placa basal	0.0
36	971621-973031	1409	+	469	Proteína da placa basal	WP_011366993.1 proteína da placa basal ( <i>D. alaskensis</i> ); motivo: TIGR02243 proteína da placa basal; virusite: YP_007005933 proteína da placa basal	0.0
37	973030-974604	1574	+	524	Proteína da cauda	WP_011366994.1 proteína da cauda ( <i>D. alaskensis</i> ); motivo: TIGR02242 proteína da cauda	0.0
38	974615-975208	593	+	197	Proteína de conexão Tc2 hipotética	WP_011366995.1 proteína hipotética ( <i>D. alaskensis</i> ); VirFam: WP_011366995.1 proteína de conexão Tc2	2e-141
39	975219-976208	989	+	329	Proteína hipotética	WP_011366996.1 proteína hipotética ( <i>D. alaskensis</i> )	0.0
40	976469-976792	323	+	107	Proteína hipotética	WP_011366998.1 proteína hipotética ( <i>D. alaskensis</i> )	1e-67
41	976805-977339	533	+	177	Proteína hipotética	WP_011366999.1 proteína hipotética ( <i>D. alaskensis</i> )	4e-126
42	977354-980902	3548	+	1182	Proteína da cauda hipotética	WP_011367000.1 proteína com domínio DUF4815 ( <i>D. alaskensis</i> ); virusite: YP_009099081 proteína da cauda	0.0



Gene	Amostra	
	Controle	Mitomicina C
<i>cap1</i>	+	+
<i>cap2</i>	+	+
<i>cap3</i>	+	+

**Fig. S1.** Transcrição reversa dos mRNAs que codificam as proteínas do capsídeos dos profagos de *Desulfovibrio alaskensis*. A figura representa o resultado para o gene *cap2*, sendo o mesmo resultado encontrado para os genes *cap1* e *cap3* no tempo de 6 horas de crescimento da cultura, sem (controle) e com mitomicina C. DNA genômico (DNAg) foi utilizado como controle positivo da PCR e amostras C6 e M6 sem transcriptase reserva foram utilizadas como controle negativo (RT-).

## CAPÍTULO 3

### **Sequência genômica de *Desulfovibrio indonesiensis* P37SLT**

*Artigo de acordo com as normas da Microbiology Resource Announcements*

Josicelli Souza Crispim<sup>a</sup>, Pedro Marcus Pereira Vidigal<sup>b</sup>, Livia Carneiro Fidélis Silva<sup>a</sup>, Clara Nogueira Laguardia<sup>c</sup>, Larissa Cristina Araújo<sup>c</sup>, Roberto Sousa Dias<sup>c</sup>, Máira Paula de Sousa<sup>d</sup>, Sérgio Oliveira de Paula<sup>c</sup>, Cynthia Canêdo da Silva<sup>a</sup>

Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil<sup>a</sup>; Núcleo de Análise de Biomoléculas, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil<sup>b</sup>; Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil<sup>c</sup>; Centro de pesquisa e desenvolvimento Leopoldo Américo Miguez de Mello, CENPES, Rio de Janeiro, Brasil<sup>d</sup>.

**Palavras-chaves:** Bactéria redutora de sulfato, plataforma de petróleo, profagos

**Autor correspondente:** Cynthia Canêdo da Silva ([ccanedo@ufv.br](mailto:ccanedo@ufv.br)); telefone: +55 (31) 3899-2973

## Resumo

Este trabalho descreve o primeiro genoma da espécie *Desulfovibrio indonesiensis*. A linhagem *Desulfovibrio indonesiensis* P37SLT foi isolada de um tanque slop proveniente de uma plataforma de óleo e apresenta 4.2 Mb.

## Manuscrito

A espécie *Desulfovibrio indonesiensis* teve a primeira linhagem (Ind 1) isolada de um biofilme formado em estrutura de aço de origem marinha (1). A caracterização da linhagem Ind 1 como uma nova espécie de *Desulfovibrio* foi realizada por meio do gene rRNA 16S (NR\_044916.1) e a caracterização morfológica e fenotípica, em 1998 (1). A espécie possui forma vibrio, motilidade por flagelo unipolar e tamanho de 3-5  $\mu\text{m}$  de comprimento por 0,5-1,2  $\mu\text{m}$  de largura. Porém, nenhum genoma desta espécie de bactéria redutora de sulfato (BRS) tem sido publicado no banco de dados. As BRS tem a capacidade de reduzir sulfato e produzir  $\text{H}_2\text{S}$  como produto do metabolismo. O  $\text{H}_2\text{S}$  é extremamente tóxico para os trabalhadores das plataformas de petróleo e reduz a qualidade do petróleo devido ao processo de *souring* (2). Além disso, as BRS tem a capacidade de formação de biofilme, que leva a corrosão de tubulações de petróleo (3).

Neste trabalho, nós apresentamos o primeiro genoma de *Desulfovibrio indonesiensis*. O gene rRNA 16S da linhagem deste estudo (P37SLT) apresentou 99,79% de identidade com o gene rRNA 16S da linhagem *D. indonesiensis* OP12 (KP682307.1), o que nos levou a classificar nossa linhagem como pertencente a esta espécie, de acordo com Chun et al. (4). *D. indonesiensis* P37SLT foi isolada de um tanque *slop* proveniente de uma plataforma de extração de petróleo da Bacia de Campos, Rio de Janeiro, Brasil, que apresenta muitos problemas pela produção de  $\text{H}_2\text{S}$ .

A extração de DNA foi realizada pelo método de proteinase K (5) e enviada para o sequenciamento. A biblioteca gênômica foi preparada por meio do Kit Nextera DNA Sample (Illumina) após limpeza de pequenos fragmentos contaminantes (50-100pb), realizada com o Kit DNEasy PowerClean Pro Cleanup (Qiagen). A biblioteca foi sequenciada utilizando leituras *paired-end* (PE) de 2x300 bp pelo sistema HiSeq (Illumina). Inicialmente, as *reads* obtidas foram trimadas pelo programa Trimmomatic versão 0.36. Posteriormente, as *reads* foram montadas pelo método *de novo* utilizando o programa CLC Genomics Workbench (version 9.5), com variação da

bolha (b1000-5000) e o tamanho dos fragmentos (k23-63) para a construção dos grafos De Bruijn. A estatística descritiva das montagens foi realizada por meio do programa Assembly-Stats (<https://github.com/sanger-pathogens/assembly-stats>), com exclusão de *contigs* menores que 500 pb. O programa RAST (<http://rast.nmpdr.org/rast.cgi>) (6) foi usado para a anotação inicial e posteriormente a sequência foi depositada com a anotação do NCBI (*Prokaryotic Genome Annotation Pipeline* - PGAP). A presença de regiões relacionadas à profagos e à CRISPR-*cas* foram confirmadas pelos programas PHASTER (6) e CRISPRone (<http://omics.informatics.indiana.edu/CRISPRone/>), respectivamente.

A melhor montagem escolhida foi aquela que apresentou o menor número de *contigs* contendo também o maior *contig* em pb. A verificação de possíveis contaminações foi realizada por meio da análise do conteúdo G+C. Depois da verificação da qualidade, as sequências apresentaram 362,41 vezes de cobertura média do genoma de 4.293.765 pb, composto por 334 *contigs*. O tamanho médio em pb dos *contigs* foi de 12,86 Mb, sendo o maior deles com 458,85 kb. Apenas 7 *contigs* cobriram 50% do genoma (N50) e o conteúdo G+C encontrado foi de 60,5%. O PGAP permitiu a identificação de 4.026 genes classificados em 3.958 sequências codificadoras (CDSs), 4 regiões de arranjos do sistema CRISPR-Cas (*Clustered Regularly Interspaced Short Palindromic Repeat*), 10 rRNAs (4 genes 5S, 2 genes 16S, 4 genes 23S), 51 tRNAs, 7 RNAs não codificadores (ncRNAs) e 76 pseudogenes. O genoma apresenta regiões relacionadas à profagos, uma característica comum do gênero *Desulfovibrio*, como indentificado por Crispim e colaboradores (7). Além disso, possui dois *locus cas* completos dos tipos I-C e I-E, ambos também já indentificados em genomas de *Desulfovibrio* (7).

Neste trabalho, nós apresentamos o primeiro genoma completo de *Desulfovibrio indonensis*. A linhagem *D. indonensis* P37SLT está sendo utilizada como hospedeira para isolamento de bacteriófagos líticos, que serão usados no controle de BRS em ambientes relacionados á indústria de petróleo.

**Número de acesso.** O genoma foi depositado no GenBank (DDBJ/ENA/GenBank) com número de acesso QMIE00000000.

## AGRADECIMENTOS

Nós agradecemos à CNPq, FAPEMIG, Capes e Petrobrás pelo suporte financeiro. Ao Núcleo de Análise de Biomoléculas (NuBioMol) da Universidade Federal de Viçosa pelo suporte na análise dos dados. Aos pesquisadores Bruna Almeida Leão Ayupe, Jéssica Duarte da Silva, Vinícius da Silva Duarte, Michelle Dias de Oliveira, Kassiana Mafra Bicalho, Marcos Rogério Tótola e Hilário Cuquetto Mantovani pelo apoio no desenvolvimento do trabalho.

## Referências

1. Feio MJ, Beech IB, Carepo M, Lopes JM, Cheung CW, Franco R, Guezennec J, Smith JR, Mitchell JI, Moura JJ, Lino AR. 1998. Isolation and characterisation of a novel sulphate-reducing bacterium of the *Desulfovibrio* genus. *Anaerobe* 4:117–130.
2. Gieg LM, Jack TR, Foght JM. 2011. Biological souring and mitigation in oil reservoirs. *Appl Microbiol Biotechnol* 92:263–282.
3. Vigneron A, Alsop EB, Chambers B, Lomans BP, Head IM, Tsesmetzis N. 2016. Complementary microorganisms in highly corrosive biofilms from an offshore oil production facility. *Appl Environ Microbiol* 82:2545–2554.
4. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu X-W, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68:461–466.
5. Sambrook J, Russell DW. 2006. *The condensed protocols from molecular cloning: A laboratory manual* Cold Sprin. New York.
6. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21.
7. Crispim JS, Dias RS, Vidigal PMP, de Sousa MP, da Silva CC, Santana MF, de Paula SO. 2018. Screening and characterization of prophages in *Desulfovibrio* genomes. *Sci Rep* 8:9273.

## CAPÍTULO 4

### **Sequência genômica de *Desulfovibrio marinus* P48SEP, o primeiro genoma da espécie**

*Artigo de acordo com as normas da Microbiology Resource Announcements*

Josicelli Souza Crispim<sup>a</sup>, Pedro Marcus Pereira Vidigal<sup>b</sup>, Livia Carneiro Fidélis Silva<sup>a</sup>, Larissa Cristina Araújo<sup>c</sup>, Clara Nogueira Laguardia<sup>c</sup>, Roberto Sousa Dias<sup>c</sup>, Máira Paula de Sousa<sup>d</sup>, Sérgio Oliveira de Paula<sup>c</sup>, Cynthia Canêdo da Silva<sup>a</sup>

Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil<sup>a</sup>; Núcleo de Análise de Biomoléculas, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil<sup>b</sup>; Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil<sup>c</sup>; Centro de pesquisa e desenvolvimento Leopoldo Américo Miguez de Mello, CENPES, Rio de Janeiro, Brasil<sup>d</sup>.

**Palavras-chaves:** BRS, CRISPR-Cas, profagos

**Autor correspondente:** Cynthia Canêdo da Silva (ccanedo@ufv.br); telefone: +55 (31) 3899-2973

## Resumo

A linhagem *Desulfovibrio marinus* P48SEP foi isolada de uma amostra de separador de produção de uma plataforma de extração de petróleo e apresenta 5,1 Mb. Este trabalho descreve a sequência do primeiro genoma da espécie *Desulfovibrio marinus*.

## Manuscrito

A espécie *Desulfovibrio marinus* foi inicialmente caracterizada a partir de um isolado (DSM 18311) proveniente de sedimentos marinhos da Tunísia, em 2007 (1). A linhagem DSM18311 possui forma vibrio, motilidade, tamanho de 0,5 x 1,5-2,5 µm, é mesofílica e moderadamente halofílica. Estas características juntamente com a análise do gene rRNA 16S (NR\_043757.1) levaram a caracterização desta linhagem como sendo uma nova espécie do gênero *Desulfovibrio* (1). Esta espécie de bactéria redutora de sulfato (BRS) não possui nenhum genoma depositado no banco de dados, o que mostra a importância do seu estudo. As BRS produzem H<sub>2</sub>S como produto do metabolismo anaeróbico, que causa sérios danos na indústria de petróleo (2, 3).

Neste trabalho, nós apresentamos o primeiro genoma de *Desulfovibrio marinus*. Inicialmente foi encontrado o valor de 99,66 % de identidade entre o gene rRNA 16S da linhagem deste estudo com o gene 16S da linhagem *D. marinus* IMP-2 (DQ365925.1), o que nos levou a classificar a nossa linhagem (P48SEP) nesta espécie, de acordo com as normas de taxonomia (4). *D. marinus* P48SEP foi isolada de uma amostra de separador de produção de uma plataforma de extração de petróleo da Bacia de Campos, Rio de Janeiro, Brasil.

O método de proteinase K (5) foi utilizado para a extração de DNA, que foi enviado para o sequenciamento. Inicialmente, foi utilizado o Kit DNEasy PowerClean Pro Cleanup (Qiagen) para a limpeza de contaminantes com tamanho entre 50-100 pb. Em seguida, o Kit Nextera DNA Sample (Illumina) foi utilizado para a construção da biblioteca gênica. O sequenciamento foi realizado por meio do sistema HiSeq (Illumina) com leituras *paired-end* (PE) de 2x300 bp. As *reads* obtidas foram trimadas pelo programa Trimmomatic versão 0.36 e montadas pelo método *de novo* utilizando o programa CLC Genomics Workbench (version 9.5). Para a construção dos grafos De Bruijn, foi realizada a variação do tamanho da bolha (b1000-5000) e do tamanho dos fragmentos (k23-63) durante a montagem do genoma. O programa Assembly-Stats (<https://github.com/sanger-pathogens/assembly-stats>) foi usado para a análise

da estatística descritiva das montagens, tendo a exclusão de *contigs* menores que 500 pb. O programa RAST (<http://rast.nmpdr.org/rast.cgi>) (6) foi utilizado para a anotação genômica inicial. Posteriormente, a anotação final foi realizada por meio da anotação do NCBI (*Prokaryotic Genome Annotation Pipeline* - PGAP). Os programas PHASTER (7) e CRISPRone (<http://omics.informatics.indiana.edu/CRISPRone/>) foram utilizados para a confirmação de sequências relacionadas à profagos e à CRISPR-cas, respectivamente.

A partir dos dados da análise estatística das montagens, foi escolhida a montagem que apresentou o menor número de *contigs* e o maior *contig* em pb. A análise de conteúdo G+C das sequências foi utilizada como parâmetro de verificação de contaminação. Após a filtragem, as sequências levaram à 47,90 vezes à cobertura de um genoma de 5,14 Mb constituído por 392 *contigs*. O valor médio em pb dos *contigs* foi de 13,11 kb, sendo o maior deles com 634,84 kb. Apenas 7 *contigs* cobrem 50 % do genoma (N50). O conteúdo G+C encontrado foi de 62,1 %. O PGAP identificou 4.748 genes classificados em 4.680 sequências codificadoras (CDSs), 1 região de arranjo do sistema CRISPR-Cas (*Clustered Regularly Interspaced Short Palindromic Repeat*), 12 rRNAs (2 genes 5S, 7 genes 16S e 3 genes 23S), 52 tRNAs, 4 RNAs não codificadores (ncRNAs) e 79 pseudogenes. Como encontrado em outros genomas do gênero de *Desulfovibrio* (8), a linhagem deste trabalho apresenta regiões relacionadas à profagos e um locus *cas* completo do tipo I-C, o tipo mais abundante entre genomas completos de *Desulfovibrio* (8).

Neste trabalho, nós apresentamos o primeiro genoma completo de *Desulfovibrio marinus*. A linhagem *D. marinus* P48SEP está sendo usada em análises de formação de biofilme em ambientes associados à indústria de petróleo.

**Números de acesso.** O genoma foi depositado no GenBank (DDBJ/ENA/GenBank) com número de acesso QMIF00000000.

## **AGRADECIMENTOS**

Nós agradecemos ao Núcleo de Análises de Biomoléculas da Universidade Federal de Viçosa pelo suporte na análise de dados. À FAPEMIG, Capes, CNPq e Petrobrás pelo suporte financeiro. Aos pesquisadores Jéssica Duarte da Silva, Bruna Almeida Leão Ayupe, Michelle Dias de Oliveira, Kassiana Mafra Bicalho, Vinícius da Silva Duarte, Marcos Rogério Tótola e Hilário Cuquetto Mantovani pelo apoio no desenvolvimento do trabalho.

## Referências

1. Ben Dhia Thabet O, Fardeau M-L, Suarez-Nuñez C, Hamdi M, Thomas P, Ollivier B, Alazard D. 2007. *Desulfovibrio marinus* sp. nov., a moderately halophilic sulfate-reducing bacterium isolated from marine sediments in Tunisia. *Int J Syst Evol Microbiol* 57:2167–2170.
2. Gieg LM, Jack TR, Foght JM. 2011. Biological souring and mitigation in oil reservoirs. *Appl Microbiol Biotechnol* 92:263–282.
3. Vigneron A, Alsop EB, Chambers B, Lomans BP, Head IM, Tsesmetzis N. 2016. Complementary microorganisms in highly corrosive biofilms from an offshore oil production facility. *Appl Environ Microbiol* 82:2545–2554.
4. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu X-W, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68:461–466.
5. Sambrook J, Russell DW. 2006. *The condensed protocols from molecular cloning: A laboratory manual* Cold Spring. New York.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75.
7. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21.
8. Crispim JS, Dias RS, Vidigal PMP, de Sousa MP, da Silva CC, Santana MF, de Paula SO. 2018. Screening and characterization of prophages in *Desulfovibrio* genomes. *Sci Rep* 8:9273.

## CONCLUSÕES GERAIS E CONSIDERAÇÕES FINAIS

- ✓ Um total 128 elementos profago-*like* foram identificados em 46 dos 47 genomas de bactérias do gênero *Desulfovibrio* analisados, com linhagens apresentando mais de um elemento, o que caracteriza poli-lisogenia;
- ✓ A *workflow* desenvolvida permitiu a identificação de 53 profagos completos, sendo 69,82% pertencentes à família *Myoviridae*;
- ✓ Linhagens poli-lisogênicas e com profagos completos pertencentes à família *Myoviridae*, como *Desulfovibrio alaskensis*, representam alto potencial para estudos de indução de profagos que podem ser utilizados no controle de BRS;
- ✓ *Desulfovibrio alaskensis* DSM16109 apresentou profagos defectivos com a indução de mitomicina C, o que descarta seu uso no controle de BRS;
- ✓ Material genético relacionado à profagos e à bactéria hospedeira foram identificados no interior de vesículas de membrana externa, o que indica um possível envolvimento dessas sequências na transferência horizontal de genes;
- ✓ Os genomas das linhagens P37SLT e P48SEP são os primeiros genomas das espécies *Desulfovibrio indonesiensis* e *Desulfovibrio marinus* depositados no banco de dados;
- ✓ Os genomas de *D. indonesiensis* P37SLT e *D. marinus* P48SEP apresentam sequências relacionadas à profagos, uma característica comum do gênero identificado no trabalho anterior.

Os resultados do presente trabalho indicam que linhagens de *Desulfovibrio* apresentam grande potencial para estudo do controle de BRS por meio de profagos. Apesar dessa característica não ter sido satisfatória no estudo da expressão de profagos em *D. alaskensis*, os resultados desse trabalho demonstraram que a linhagem DSM16109 é produtora de vesículas de membrana externa capazes de transportar material genético bacteriano e relacionado à profago, estando assim, possivelmente relacionada com a transferência horizontal de genes. Os genomas caracterizados são os primeiros da espécie depositados no banco de dados, representando grande potencial para seu estudo por meio da comparação com genomas de outras espécies de BRS.