

UNIVERSIDADE FEDERAL DO PARANÁ

WILLIAN KLASSEN DE OLIVEIRA

Análise comparativa de genomas de isolados clínicos e ambientais do gênero *Herbaspirillum*

CURITIBA

2018

WILLIAN KLASSEN DE OLIVEIRA

ANÁLISE COMPARATIVA DE GENOMAS DE ISOLADOS CLÍNICOS E
AMBIENTAIS DO GÊNERO *HERBASPIRILLUM*

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Bioinformática, no Curso de Pós-Graduação em Bioinformática, Setor de Educação Profissional e Tecnológica, da Universidade Federal do Paraná.

Orientador: Prof. Dr. Helisson Faoro

CURITIBA

2018

Catálogo na publicação
Sistema de Bibliotecas UFPR
Biblioteca de Educação Profissional e Tecnológica

O48 Oliveira, Willian Klassen de
Análise comparativa de genomas de isolados clínicos e ambientais do gênero *Herbaspirillum* / Willian Klassen de Oliveira. - Curitiba, 2018.
141 p.: il., tabs, graf.

Orientador: Helisson Faoro
Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Educação Profissional e Tecnológica, Curso de Pós-Graduação em Bioinformática.
Inclui Bibliografia.

1. Herbaspirillum. 2. Genomas. 3. Bactérias. 4. Bioinformática.
I. Faoro, Helisson. II. Título. III. Universidade Federal do Paraná.

CDD 589.9

Elaboração: Angela Pereira de Farias Mengatto - CRB 9/1002



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA

Pós-Graduação em Bioinformática WWW.BIOINFO.UFPR.BR
E-mail: bioinfo@ufpr.br Tel: 41 33614906

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em BIOINFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **WILLIAN KLASSEN DE OLIVEIRA** intitulada: "**Análise comparativa de genomas de isolados clínicos e ambientais do gênero *Herbaspirillum***", após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua **APROVAÇÃO** no rito de defesa. A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 23 de fevereiro de 2018.

Dr. Helisson Fabro
Presidente

Programa de Pós-graduação em Bioinformática – UFPR
Instituto Carlos Chagas-FIOCRUZ/PR

Dr^a. Cyntia Maria Telles Fadel Picheth
Avaliadora Externa

Programa de Pós-graduação em Ciências Farmacêuticas - UFPR

Dr. Fabio de Oliveira Pedrosa
Avaliador Interno

Programa de Pós-graduação em Bioinformática
Departamento de Bioquímica - UFPR

Dedico este trabalho ao meu avô e
herói, Peter Klassen.

AGRADECIMENTOS

Agradeço primeiramente a Deus por me guiar até este momento. Sem Ele nada disto seria possível.

Ao meu orientador Prof. Dr. Helisson Faoro, pela confiança depositada, por todos os ensinamentos e paciência durante todo este tempo.

Ao Prof. Dr. Humberto Maciel da França Madeira, por me apresentar a bioinformática, me ensinar o que é a ciência e fazer com que eu me apaixonasse por ela.

Aos professores da Pós Graduação em Bioinformática da UFPR, pelo conhecimento adquirido e conversas de corredor.

Aos meus colegas do laboratório de Bioinformática - UFPR e do laboratório da Regulação da Expressão gênica – ICC/Fiocruz por tornarem o ambiente de trabalho sempre tão agradável e pelas discussões, por vezes científicas, e outras nem tanto.

Aos professores Dr. Fábio de Oliveira Pedrosa e Dra. Cyntia Maria Telles Fadel-Pichet por aceitarem participar da banca do presente do trabalho.

Ao Prof Dr. Vasco Ariston de Carvalho Azevedo e os colegas do laboratório de Bioinformática da UFMG por me receberem tão bem no laboratório e pelo conhecimento e experiênciãs vividas neste tempo.

À Universidade Federal do Paraná e à CAPES pela oportunidade e pela bolsa.

À secretária Suzana de A. Gobetti por estar sempre pronta a ajudar.

À minha família e a Elisa, minha namorada, pela dedicação, amor, grande apoio e incentivo.

“Ele fortalece ao cansado e dá grande vigor ao que está sem forças. Até os jovens se cansam e ficam exaustos, e os moços tropeçam e caem; mas aqueles que esperam no Senhor renovam as suas forças. Voam bem alto como águias; correm e não ficam exaustos, andam e não se cansam.”

(Isaías 40:29-31)

RESUMO

Bactérias do gênero *Herbaspirillum*, de origem ambiental, têm, recentemente, sido isoladas de amostras clínicas, em especial de pacientes imunocomprometidos. Para investigar como bactérias ambientais sobrevivem e causam doenças no organismo humano, foram sequenciados e anotados dois genomas de bactérias deste gênero, isolados de escarro de pacientes com fibrose cística, sendo um *Herbaspirillum frisingense* AU14559 e um inicialmente classificado como *Herbaspirillum* linhagem 2 AU13965. O genoma do organismo *H. frisingense* AU14559 foi finalizado em apenas um *contig* contendo 5.401.392 pb, conteúdo GC% de 63,1%, 3 operons de rRNAs, 59 tRNAs, 23 ncRNAs e 4.913 genes. Já o genoma do organismo *Herbaspirillum* linhagem 2 AU13965 foi fechado em um cromossomo de 5.350.014pb, com conteúdo GC% de 64,1%, 3 operons de rRNAs, 24 ncRNAs, 62 tRNAs e 4.720 CDSs, e um plasmídeo circular com 42.977pb, contendo 54 CDS e conteúdo GC% de 62,6%. Através de análises de identidade nucleotídica média realizadas, o genoma de *Herbaspirillum* lineage2 AU13965 foi classificado como pertencente a espécie *H. seropedicae*. Para identificação de genes possivelmente envolvidos na migração do solo para o organismo humano, foram realizadas comparações entre os genomas clínicos montados neste trabalho e isolados ambientais de suas respectivas espécies. Nestas análises foram identificadas perdas de pacotes gênicos como o operon *nif*, responsável pela fixação biológica de nitrogênio nas estirpes ambientais além da aquisição de genes potencialmente envolvidos na infecção de células humanas como o ciclic beta-1,2- glucan synthase, responsável pela criação de macromoléculas de monômeros de glucose, as quais estão, em outras bactérias patogênicas, envolvidas na evasão de lisossomos, sendo essencial para a sua sobrevivência intracelular. No entanto, apesar de vários genes serem encontrados exclusivamente nos isolados clínicos, a grande maioria dos fatores de virulência, interação patógeno-hospedeiro e genes de resistência, encontrados através de alinhamentos com bancos de dados são comuns entre os isolados clínicos e ambientais, levantando a hipótese de que estes organismos possuem o potencial para causar infecções no organismo humano.

Palavras-chave: Gênero *Herbaspirillum*, *H. frisingense*, *Herbaspirillum* lineage2, comparação de genomas, Bactérias oportunistas.

ABSTRACT

Bacteria from *Herbaspirillum* genus, from environmental origin, has, recently being related to clinical isolates, in especial from immunocompromised patients. To investigate how environmental bacteria survive and cause diseases to human organism, it was sequenced and annotated two genomes from this genus, recovered from sputum of patients with cystic fibrosis, the *Herbaspirillum frisingense* AU14559 and the strain initially classified as *Herbaspirillum lineage 2* AU13965. The genome of *Herbaspirillum frisingense* AU14559 was closed in only one contig of 5,401,393 bp, GC% content of 63.1%, 3 rRNA operons, 59 tRNAs, 23 ncRNAs and 4,913 genes. The genome of *Herbaspirillum lineage 2* AU13965 was closed in one chromosome of 5,350,014 bp, with a GC% content of 64.1%, 3 rRNAs operons, 24 ncRNAs, 62 tRNAs and 4.720 CDSs, and a circular plasmid with 42,977 bp, containing 54 CDSs and GC% content of 62.6%. Through Average Nucleotide Identity (ANI) analysis, the *H. lineage 2* AU13965 genome was classified as from the *Herbaspirillum seropedicae* specie. To identification of possible genes involved in this migration from soil to human organism, was performed comparisons between the assembled genomes from the clinical isolates and environmental genomes of their respective species. In these analyses was identified losses of genes, like *nif* operon, responsible for the biological nitrogen fixation in the environmental strains, besides the acquisition of genes potentially involved in the infection in human cells like beta-1,2-glucan synthase, a protein responsible for creation of macromolecules of glucose monomers, that are involved in the avoidance of lysosomes, wich are, in other bacteria, essential for its intracellular survival. However, despite being found several genes exclusively in the clinical isolates, the most of virulence, pathogen-host interaction factors and antibiotic resistance genes found through alignments performed against data bases are common to clinical and environmental isolates, raising the hypothesis that these organisms has a level of pre-adaptation to migrate from soil to the human organism.

Key-words: *Herbaspirillum* genus, *H. frisingense*, *Herbaspirillum lineage2*, genome comparison, opportunistic bacteria.

LISTA DE FIGURAS

FIGURA 1: ÁRVORE FILOGENÉTICA DE ISOLADOS CLÍNICOS DO GÊNERO HERBASPIRILLUM	24
---	----

ARTIGO A

FIGURA A 1: HEATMAP DE IDENTIDADE NUCLEOTÍDICA MÉDIA DOS ORGANISMOS DO GÊNERO <i>Herbaspirillum</i>	39
---	----

FIGURA A 2: REPRESENTAÇÃO CIRCULAR DO GENOMA DE <i>H. linhagem 2</i>	40
---	----

FIGURA A 3: REPRESENTAÇÃO GRÁFICA DA PERDA DO OPERON <i>nif</i> EM <i>H. linhagem 2</i>	42
---	----

ARTIGO B

FIGURA B 1: REPRESENTAÇÃO CIRCULAR DO GENOMA DE <i>H. frisingense</i> AU14559	91
---	----

FIGURA B 2: REPRESENTAÇÃO GRÁFICA DA PERDA DO OPERON <i>nif</i> EM <i>H. frisingense</i> AU14559	94
--	----

LISTA DE GRÁFICOS

ARTIGO A

GRÁFICO A 1: NÚMERO DE <i>HITS</i> NOS ALINHAMENTOS DOS GENES DO PLASMÍDEO DE AU13965.....	38
GRÁFICO A 2: GRÁFICO DE FUNÇÕES DE GENES PRESENTES NAS ILHAS DO GENOMA DE <i>Herbaspirillum seropedicae</i> AU13965	41

ARTIGO B

GRÁFICO B 1: GRÁFICO DE FUNÇÕES DE GENES PRESENTES NAS ILHAS DO GENOMA DE <i>H. frisingense</i> AU14559.....	92
---	----

LISTA DE TABELAS

TABELA 1: ISOLADOS CLÍNICOS DO GÊNERO HERBASPIRILLUM	22	
TABELA 2: COMPARAÇÃO ENTRE PLATAFORMAS DE SEQUENCIAMENTO DE DNA.....	28	
ARTIGO A		
TABELA A 1: GENES DE VIRULÊNCIA EXCLUSIVOS DA ESTIRPE CLÍNICA AU13965	43	
TABELA A 2: GENES DE INTERAÇÃO DA ESTIRPE CLÍNICA AU13965.....	45	
TABELA A 3: GENES DE RESISTÊNCIA A ANTIBIÓTICOS DA ESTIRPE AU13965	47	
TABELA AS 1: NÚMERO DE <i>HITS</i> POR GÊNERO DOS ALINHAMENTOS DOS GENES DO PLASMÍDEO.....		53
TABELA AS 2: GENES PRESENTES NA ILHA GENÔMICA 1 DE H. LINEAGE2 AU13965	54	
TABELA AS 3: GENES PRESENTES NA ILHA GENÔMICA 2 DE H. LINEAGE2 AU13965	54	
TABELA AS 4: GENES PRESENTES NA ILHA GENÔMICA 3 DE H. LINEAGE2 AU13965	55	
TABELA AS 5: GENES PRESENTES NA ILHA GENÔMICA 4 DE H. LINEAGE2 AU13965	55	
TABELA AS 6: GENES PRESENTES NA ILHA GENÔMICA 5 DE H. LINEAGE2 AU13965	55	
TABELA AS 7: GENES PRESENTES NA ILHA GENÔMICA 6 DE H. LINEAGE2 AU13965	56	
TABELA AS 8: GENES PRESENTES NA ILHA GENÔMICA 7 DE H. LINEAGE2 AU13965	56	
TABELA AS 9: GENES PRESENTES NA ILHA GENÔMICA 8 DE H. LINEAGE2 AU13965	57	
TABELA AS 10: GENES PRESENTES NA ILHA GENÔMICA 9 DE H. LINEAGE2 AU13965	57	

TABELA AS 11: GENES PRESENTES NA ILHA GENÔMICA 10 DE H. LINEAGE2 AU13965	58
TABELA AS 12: GENES PRESENTES NA ILHA GENÔMICA 11 DE H. LINEAGE2 AU13965	58
TABELA AS 13: GENES PRESENTES NA ILHA GENÔMICA 12 DE H. LINEAGE2 AU13965	59
TABELA AS 14: GENES PRESENTES NA ILHA GENÔMICA 13 DE H. LINEAGE2 AU13965	61
TABELA AS 15: GENES PRESENTES NA ILHA GENÔMICA 14 DE H. LINEAGE2 AU13965	61
TABELA AS 16: GENES PRESENTES NA ILHA GENÔMICA 15 DE H. LINEAGE2 AU13965	62
TABELA AS 17: GENES PRESENTES NA ILHA GENÔMICA 16 DE H. LINEAGE2 AU13965	62
TABELA AS 18: GENES PRESENTES NA ILHA GENÔMICA 17 DE H. LINEAGE2 AU13965	62
TABELA AS 19: GENES PRESENTES NA ILHA GENÔMICA 18 DE H. LINEAGE2 AU13965	63
TABELA AS 20: GENES PRESENTES NA ILHA GENÔMICA 19 DE H. LINEAGE2 AU13965	64
TABELA AS 21: GENES DE VIRULÊNCIA DA ESTIRPE H. LINHAGEM 2 AU13965	64
TABELA AS 22: GENES DE VIRULÊNCIA DA ESTIRPE <i>H. seropedicae</i> SmR1	70
TABELA AS 23: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO <i>H. linhagem 2</i> AU13965	75
TABELA AS 24: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO <i>H. seropedicae</i> SmR1	78
TABELA AS 25: GENES DE RESISTÊNCIA DA ESTIRPE <i>H. linhagem 2</i> AU13965	81
TABELA AS 26: GENES DE RESISTÊNCIA DA ESTIRPE <i>H. seropedicae</i> AU14040	82
TABELA AS 27: GENES DE RESISTÊNCIA DA ESTIRPE <i>H. seropedicae</i> SMR1 ..	83

TABELA B 1: IDENTIDADES NUCLEOTÍDICAS MÉDIAS ENTRE A ESTIRPE AU14559 E DEMAIS GENOMAS <i>H. FRISINGENSE</i>	90
TABELA B 2: GENES DE VIRULÊNCIA EXCLUSIVOS DA ESTIRPE CLÍNICA AU14559	95
TABELA B 3: GENES DE INTERAÇÃO PATÓGENO HOSPEDEIRO DA ESTIRPE <i>H. frisingense</i> AU14559	96
TABELA B 4: GENES DE RESISTÊNCIA A ANTIBIÓTICOS DA ESTIRPE AU14559	98
TABELA BS 1: GENES PRESENTES NA ILHA GENÔMICA 1 DE <i>H. FRISINGENSE</i> AU14559	105
TABELA BS 2: GENES PRESENTES NA ILHA GENÔMICA 2 DE <i>H. FRISINGENSE</i> AU14559	105
TABELA BS 3: GENES PRESENTES NA ILHA GENÔMICA 3 DE <i>H. FRISINGENSE</i> AU14559	106
TABELA BS 4: GENES PRESENTES NA ILHA GENÔMICA 4 DE <i>H. FRISINGENSE</i> AU14559	106
TABELA BS 5: GENES PRESENTES NA ILHA GENÔMICA 5 DE <i>H. FRISINGENSE</i> AU14559	107
TABELA BS 6: GENES PRESENTES NA ILHA GENÔMICA 6 DE <i>H. FRISINGENSE</i> AU14559	108
TABELA BS 7: GENES PRESENTES NA ILHA GENÔMICA 7 DE <i>H. FRISINGENSE</i> AU14559	108
TABELA BS 8: GENES PRESENTES NA ILHA GENÔMICA 8 DE <i>H. FRISINGENSE</i> AU14559	109
TABELA BS 9: GENES PRESENTES NA ILHA GENÔMICA 9 DE <i>H. FRISINGENSE</i> AU14559	110
TABELA BS 10: GENES PRESENTES NA ILHA GENÔMICA 10 DE <i>H. FRISINGENSE</i> AU14559	112
TABELA BS 11: GENES PRESENTES NA ILHA GENÔMICA 11 DE <i>H. FRISINGENSE</i> AU14559	113
TABELA BS 12: GENES PRESENTES NA ILHA GENÔMICA 12 DE <i>H. FRISINGENSE</i> AU14559	113

TABELA BS 13: GENES PRESENTES NA ILHA GENÔMICA 13 DE H. FRISINGENSE AU14559	113
TABELA BS 14: GENES PRESENTES NA ILHA GENÔMICA 14 DE H. FRISINGENSE AU14559	114
TABELA BS 15: GENES DE VIRULÊNCIA DA ESTIRPE H. FRISINGENSE AU14559	114
TABELA BS 16: GENES DE VIRULÊNCIA DA ESTIRPE H. FRISINGENSE GSF30	119
TABELA BS 17: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO H. FRISINGENSE AU14559	125
TABELA BS 18: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO H. FRISINGENSE GSF30.....	128
TABELA BS 19: GENES DE RESISTÊNCIA DA ESTIRPE H. FRISINGENSE AU14559	130
TABELA BS 20: GENES DE RESISTÊNCIA DA ESTIRPE H. FRISINGENSE.....	131
TABELA BS 21: GENES DE RESISTÊNCIA DA ESTIRPE H. FRISINGENSE GSF30	132

SUMÁRIO

1	INTRODUÇÃO	18
2	JUSTIFICATIVA	18
3	OBJETIVOS	19
4	REVISÃO DE LITERATURA	19
4.1	BACTÉRIAS OPORTUNISTAS	19
4.2	GÊNERO <i>HERBASPIRILLUM</i>	21
4.2.1	<i>Herbaspirillum seropedicae</i>	24
4.2.2	<i>Herbaspirillum frisingense</i>	25
4.3	TECNOLOGIAS DE SEQUENCIAMENTO	26
4.4	MONTAGEM E ANOTAÇÃO DE GENOMAS	29
4.4.1	Montagem de genomas	29
4.4.2	Anotação de genomas	30
4.5	GENÔMICA COMPARATIVA	30
5	COMPARAÇÃO GENÔMICA ENTRE ESTIRPES CLÍNICAS E AMBIENTAIS DE <i>HERBASPIRILLUM SEROPEDICAE</i>	32
5.1	RESUMO	32
5.2	INTRODUÇÃO	33
5.3	MATERIAIS E MÉTODOS	34
5.3.1	Purificação de DNA e sequenciamento	34
5.3.2	Montagem e anotação do genoma	35
5.3.3	Comparações genômicas	35
5.4	RESULTADOS	37
5.4.1	Genoma de <i>Herbaspirillum</i> linhagem 2 AU13965	37
5.4.2	Ilhas genômicas	39
5.4.3	Comparações entre os genomas SmR1 e AU13965	42
5.4.3.1	Fatores de virulência	43
5.4.3.2	Fatores de interação patógeno-hospedeiro	45
5.4.3.3	Características específicas do genoma de <i>H. seropedicae</i> AU13965	46
5.4.3.4	Resistência a antibióticos	47
5.5	DISCUSSÃO	48
5.6	REFERÊNCIAS	50
5.7	MATERIAL SUPLEMENTAR	53

6	ALTA SIMILARIDADE ENTRE OS GENOMAS DE ESTIRPES CLÍNICAS E AMBIENTAIS DE HERBASPIRILLUM FRISINGENSE SUGERE PRÉ-ADAPTAÇÃO A DIFERENTES HOSPEDEIROS	85
6.1	RESUMO.....	85
6.2	INTRODUÇÃO.....	86
6.3	MATERIAIS E MÉTODOS.....	87
6.3.1	Purificação de DNA e sequenciamento.....	87
6.3.2	Montagem do genoma.....	87
6.3.3	Comparações genômicas.....	88
6.4	RESULTADOS.....	89
6.4.1	Genoma <i>H. frisingense</i> AU14559.....	89
6.4.2	Ilhas genômicas.....	90
6.4.3	Comparações entre os genomas GSF30 e AU14559.....	92
6.4.3.1	Fatores de virulência.....	95
6.4.3.2	Fatores de interação patógeno-hospedeiro.....	96
6.4.3.3	Resistência a antibióticos.....	97
6.5	DISCUSSÃO.....	99
6.6	REFERÊNCIAS.....	101
6.7	MATERIAL SUPLEMENTAR.....	105
7	CONCLUSÕES	135
8	REFERÊNCIAS	136

1 INTRODUÇÃO

As bactérias do gênero *Herbaspirillum*, são bactérias gram negativas, vibrioides, com motilidade através de movimento flagelar e não fermentadoras. Elas representam grande importância para a economia, pois são capazes de fixar nitrogênio atmosférico na forma de íons amônio, além de se associarem a várias plantas de interesse econômico, em especial gramíneas, atuando como promotores de crescimento vegetal (BALDANI et al., 1986). Estas características tornaram estes organismos potenciais biofertilizantes para serem utilizados na agricultura. Atualmente o gênero *Herbaspirillum* compreende 11 espécies isoladas de diferentes ambientes como rizosfera (ROTHBALLER et al., 2006), solo (CARRO et al., 2012) e água (DE SOUZA et al., 2013), e apenas 3 são descritas como bactérias fixadoras de nitrogênio: *Herbaspirillum seropedicae*, *Herbaspirillum rubrisubalbicans* e *Herbaspirillum frisingense*.

No entanto, recentemente, bactérias do gênero *Herbaspirillum* vem sendo isoladas em amostras clínicas, principalmente de pacientes imunocomprometidos, acometidos por doenças como leucemia (ZIGA; DRULEY; BURNHAM, 2010) e fibrose cística (SPILKER et al., 2008), sugerindo que esse gênero possa apresentar um grupo de bactérias oportunistas capazes de colonizar hospedeiros humanos. Spilker e colaboradores (2008) reportaram a identificação de 28 isolados do gênero *Herbaspirillum* a partir de amostras de secreção pulmonar de pacientes com fibrose cística.

Estudos de comparação dos genomas de isolados clínicos e ambientais deste gênero podem auxiliar a compreender como bactérias de origem ambiental são capazes de colonizar e causar doenças ao organismo humano. Para isto, neste trabalho foram sequenciados, anotados e comprados os genomas dos isolados clínicos *Herbaspirillum frisingense* AU14559 e daquele classificado como *Herbaspirillum* linhagem 2 (AU13965).

2 JUSTIFICATIVA

A necessidade de uma melhor compreensão sobre os mecanismos utilizados por microorganismos da rizosfera para colonizar e causar doenças em humanos e

os eventos moleculares envolvidos neste processo são importantes para a prevenção destas doenças e dos fatores associados com a patogenicidade. Sendo a montagem e comparação de genomas ferramentas poderosas para um melhor esclarecimento sobre o tema.

3 OBJETIVOS

- Montar e anotar o genoma das bactérias *H. frisingense* AU14559 e *Herbaspirillum* linhagem 2 AU13964;
- Comparar a estrutura do cromossomo e a sequência nucleotídica dos genomas com outros genomas de espécies do gênero *Herbaspirillum* disponíveis em banco de dados públicos;
- Comparar as proteínas identificadas nos genomas montados com as proteínas codificados nos genomas disponíveis em bancos de dados públicos;
- Identificar os genes potencialmente envolvidos na interação, colonização e virulência em humanos;
- Caracterizar *in silico* as proteínas associadas a vias metabólicas exclusivas de cada organismo.

4 REVISÃO DE LITERATURA

4.1 BACTÉRIAS OPORTUNISTAS

Bactérias oportunistas são organismos normalmente encontrados no ambiente, mas que podem causar doença ao ser humano, principalmente quando este se encontra imunocomprometido (BERG et al., 2005), ou seja, quando o indivíduo possui seu sistema imune deprimido por imunossupressores ou doenças como a Síndrome da imunodeficiência adquirida (AIDS) e fibrose cística, aumentando assim a chance de ser acometido por uma infecção bacteriana, viral ou fúngica.

A fibrose cística é uma doença genética autossômica recessiva causada por uma mutação no gene que codifica a proteína reguladora da condutância transmembranar de fibrose cística (CFTR - Cystic Fibrosis Transmembrane

conductance Regulator) responsável, dentre outras funções, pelo transporte de íons cloreto (CUTTING, 2015). Esta doença foi identificada na década de 1930 quando, devido a complicações como obstrução da porção final do íleo por mecônio endurecido e desnutrição devido à insuficiência pancreática, os pacientes morriam prematuramente, possuindo uma expectativa de vida de apenas alguns meses. Hoje, com um diagnóstico precoce e o avanço da medicina a expectativa de vida aumentou para mais de 40 anos (ELBORN, 2016).

Dentre todas as implicações da doença, as que causam maior número de mortes são as pulmonares principalmente devido colonização e infecção das vias respiratórias levando a um dano tissular irreversível (CUTTING, 2015). Um dos fatores que favorece a infecção pulmonar em pacientes com fibrose cística é a maior viscosidade do muco pulmonar prejudicando o movimento mucociliar e levando a uma dilatação brônquica irreversível e função pulmonar diminuída (RIBAS, 2014). A diminuição da função pulmonar está relacionada com uma redução dos níveis séricos de MBL, uma proteína do sistema imune inato humano produzida no fígado e secretada no sangue durante fases agudas de infecções, atuando como proteína chave para a ativação do sistema complemento pela via das lectinas, realizando a ligação a microrganismos, ativando assim a cascata do complemento e destruição do patógeno através da formação do complexo lítico de membrana, além atuar na opsonização (RIBAS, 2014), sendo de grande importância no combate de infecções de bactérias oportunistas, que normalmente são encontradas no ambiente.

O sucesso de bactérias ambientais na colonização e infecção de seres humanos ainda não é totalmente esclarecido, no entanto, estudos apontam que a alta competição de bactérias em determinados ambientes, como a rizosfera, pode atuar como um fator de seleção para a aquisição de algumas características que podem beneficiar estas bactérias na competição pela colonização de outros nichos como o organismo humano (BERG et al., 2005; PARKE; GURIAN-SHERMAN, 2001).

A rizosfera é o ambiente adjacente às raízes das plantas e influenciado pelo seu metabolismo. O solo dessa região pode ser rico e diverso devido a exsudação, através da qual a planta pode liberar aminoácidos, ácidos orgânicos, polissacarídeos e proteínas, e esses compostos são potenciais fontes de energia para bactérias presentes no solo (BERG et al., 2005). A microbiota desta região é de

grande importância para a planta, pois pode atuar em diversos processos como ciclo de nutrientes, promoção do crescimento vegetal e supressão de patógenos vegetais. A competição por nutrientes neste ambiente é grande, o que beneficia bactérias com características como resistência a antibióticos, sistemas eficientes de aquisição de minerais como ferro e até mesmo produção de antibióticos (BERG et al., 2005; PARKE; GURIAN-SHERMAN, 2001). Este antagonismo à outras espécies são bastante comum na microbiota da rizosfera, e esses microorganismos podem ser benéficos às plantas inibindo, por exemplo, o crescimento de patógenos vegetais. No entanto, estes mesmos mecanismos, aliados a mecanismos de interação bactéria-planta, como a estrutura do antígeno-O de lipopolissacarídeos (LPS), parecem estar associados à capacidade de colonização do organismo humano. Outro fator que favorece o aparecimento de bactérias oportunistas é a densidade desta microbiota, o que favorece as transferências horizontais de genes (BERG, et al., 2005; PARQUE; GURIAN-SHERMAN, 2001).

Vários gêneros de bactérias possuem espécies associadas tanto às raízes de plantas, quanto em casos clínicos, como é o caso dos gêneros *Burkholderia*, *Enterobacter*, *Staphylococcus*, *Stenotrophomonas*, *Ochrobactrum*, *Pseudomonas*, *Ralstonia* e, mais recentemente, o gênero *Herbaspirillum* (BERG, et al., 2005).

4.2 GÊNERO *Herbaspirillum*

O gênero *Herbaspirillum* é um gênero de bactérias gram negativas, vibrióides, com motilidade através de movimento flagelar e não fermentadoras. Estas bactérias pertencem à família Oxalobacteraceae, ordem Burkholderiales da classe beta do filo Proteobacterias. Foram inicialmente isoladas de rizosfera de gramíneas como milho, arroz e sorgo (BALDANI et al., 1986). Algumas dessas bactérias se associam de forma endofítica às plantas e são capazes de secretar alguns fitohormônios como o ácido indolacético além de realizar a fixação biológica de nitrogênio (PEDROSA et al., 2011), desempenhando um importante papel biológico e econômico neste ambiente. Atualmente o gênero *Herbaspirillum* é composto por 10 espécies: *Herbaspirillum seropedicae*, *Herbaspirillum rubrisubalbicans*, *Herbaspirillum frisingense*, *Herbaspirillum hiltneri*, *Herbaspirillum lusitanum*, *Herbaspirillum huttiense*, *Herbaspirillum rhizosphaerae*, *Herbaspirillum chlorophenolicum*, *Herbaspirillum autotrophicum* e *Herbaspirillum aquaticum*.

Desses, somente *H. seropedicae*, *H. ribrisubalbicans* e *H. frisingense* são capazes de realizar a fixação de nitrogênio.

A fixação biológica de nitrogênio é um processo realizado por bactérias diazotróficas no qual utilizam o nitrogênio atmosférico para formar amônia, uma molécula nitrogenada assimilável pelas plantas (JAMES, 2000). Em *Herbaspirillum seropedicae* esse processo ocorre e é regulado, majoritariamente, por um cluster de genes nomeado *nif* composto por 46 genes (PEDROSA et al., 2011). Esta capacidade confere grande potencial biotecnológico a estes organismos, motivo pelo qual têm sido alvo de várias pesquisas visando um melhor entendimento e utilização deste processo na promoção do crescimento vegetal.

Apesar de mais comumente isolados de amostras ambientais como em rizosfera (ROTHBALLER et al., 2006), solo (CARRO et al., 2012) e água (DE SOUZA et al., 2013), essas bactérias vêm sendo isoladas de amostras clínicas (Tabela 1), principalmente de pacientes imunocomprometidos, acometidos por doenças como leucemia (CHEN et al., 2011; ZIGA; DRULEY; BURNHAM, 2010) e fibrose cística (COENYE et al., 2002; SPILKER et al., 2008) indicando que esse gênero possa apresentar um grupo de bactérias oportunistas capazes de colonizar e causar infecção em hospedeiros humanos suscetíveis.

TABELA 1: ISOLADOS CLÍNICOS DO GÊNERO HERBASPIRILLUM

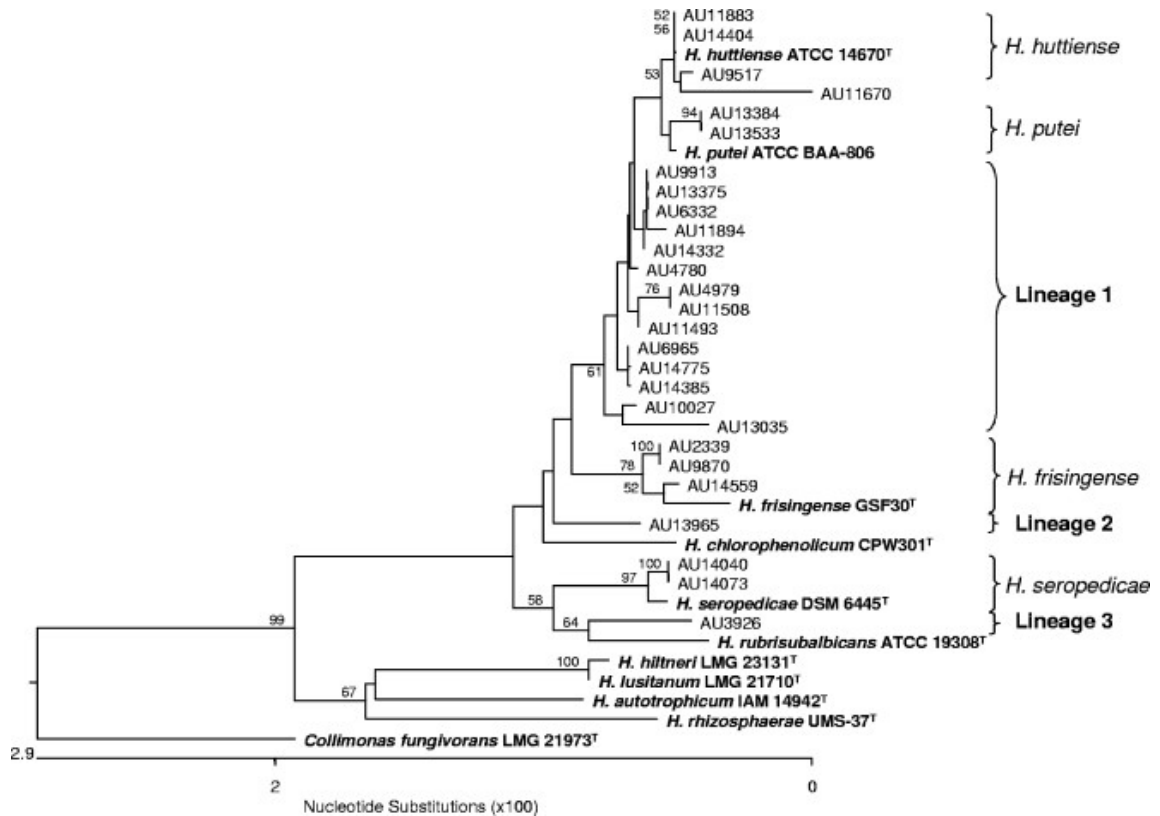
Espécie	Isolamento	Doença associada	Autor
<i>Herbaspirillum</i> sp. (2 isolados)	Secreção respiratória	Fibrose cística	Coenye, et al., 2002
<i>Herbaspirillum</i> sp.	Parede da arteria aorta	Aneurisma da aorta	Marques da Silva, et al., 2006
<i>Herbaspirillum</i> sp. (28 isolados)	Escarro	Fibrose cística	Spilker, et al., 2008
<i>Herbaspirillum</i> sp.	Sangue	Leucemia linfóide aguda	Ziga, et al., 2010
<i>Herbaspirillum</i> sp.	Sangue	Leucemia linfóide aguda	Chen et al. 2011
<i>H. aquaticum</i> ou <i>H. huttiense</i>	Sangue	Imunocompetente	Regunath, et al., 2015
<i>Herbaspirillum</i> sp.	Sangue, escarro, cateter venoso central e cateter implantado	Câncer	Chemaly, et al., 2015
<i>H. seropedicae</i>	Sangue	Mieloma multiplo	Suwuntarat, et al., 2015
<i>H. frisingense</i> VT-	Bexiga	Infecção do trato	Tetz; Tetz, 2017

Espécie	Isolamento	Doença associada	Autor
16-41		urinário	

Fonte: O Autor (2018)

Spilker e colaboradores (2008) reportaram a identificação de 28 isolados do gênero *Herbaspirillum* a partir de amostras de secreção pulmonar de pacientes com fibrose cística. Tendo como base o sequenciamento e comparação do gene 16S rRNA dos isolados e estirpes tipo para cada espécie, os autores classificaram esses isolados em 3 espécies e 3 linhagens diferentes: *H. seropedicae*, *H. frisingense*, *Herbaspirillum huttiense*, além de estirpes sem espécie definida, classificadas inicialmente como linhagens 1, 2 e 3 (Figura 1). Testes bioquímicos e fisiológicos aplicados aos isolados clínicos mostrou que elas pouco diferem bioquimicamente das isoladas do ambiente, porém, quando em contato com células HeLa geram respostas citopatogênicas como a condensação da cromatina, retração celular e a formação de vacúolos no citoplasma, no entanto em níveis muito inferiores às respostas geradas por outras bactérias patogênicas como as do gênero *Aeromonas*, reforçando a hipótese de que *Herbaspirillum* possa ser uma bactéria oportunista (MARQUES et al., 2015).

FIGURA 1: ÁRVORE FILOGENÉTICA DE ISOLADOS CLÍNICOS DO GÊNERO HERBASPIRILLUM



FONTE: Reproduzido a partir de Spilker et al., 2008.

Nota: Árvore filogenética construída a partir de sequências de genes 16S de bactérias do gênero *Herbaspirillum*, isoladas de escarro de pacientes com fibrose cística.

Atualmente o gênero *Herbaspirillum* possui genomas sequenciados de 2 isolados clínicos e 31 isolados ambientais, totalizando 33 genomas no gênero, dos quais 23 se apresentam depositados em forma de contigs, 3 em scaffolds e apenas 7 genomas fechados, ou seja, em um único contig.

4.2.1 *Herbaspirillum seropedicae*

Herbaspirillum seropedicae foi a primeira espécie desse gênero a ser descrita. Foi isolada pela primeira vez em 1986, por Baldani e colaboradores, de rizosfera de gramíneas durante contagens de *Azospirillum* sp. em Seropédica, Rio de Janeiro, Brasil. Desde de seu isolamento demonstrou capacidade de fixar nitrogênio e, por este motivo, despertou grande interesse biotecnológico, sendo alvo de diversas pesquisas (BALDANI et al., 1986).

A espécie *H. seropedicae* estirpe SmR1 teve seu genoma sequenciado e montado em 2011, sendo também a primeira espécie deste gênero a ter seu

genoma sequenciado e fechado, revelando genes envolvidos na interação planta-bactéria, fixação de nitrogênio, produção de fitohormônios, hemaglutininas, hemolisinas, entre outros (PEDROSA et al., 2011).

No entanto, em 2008, esta espécie foi associada com isolados clínicos, entre eles a estirpe AU14040 (SPILKER et al., 2008). Esta estirpe teve seu genoma sequenciado, montado, anotado e comparado com estirpes ambientais por Faoro e colaboradores (in prep.), demonstrando uma série de adaptações destas bactérias ao ambiente do organismo humano. Dentre as principais diferenças podemos citar a perda do cluster de genes *nif*, responsável pela fixação biológica de nitrogênio, a troca dos genes responsáveis pela biossíntese de LPS, principalmente os envolvidos na biossíntese de ácido acetil neuramínico (Neu5Ac), um monossacarídeo presente na superfície da células mamíferas e de algumas bactérias patogênicas, usado para evadir o sistema imune hospedeiro através da mimetização da sua superfície, além de outros genes relacionados principalmente à evasão do sistema imune (FAORO, et al., in prep).

4.2.2 *Herbaspirillum frisingense*

Bactérias da espécie *H. frisingense* foram inicialmente isoladas de plantas do gênero *Miscanthus* no sul da Alemanha (KIRCHHOF et al., 2001). O primeiro representante da espécie a ter seu genoma sequenciado e montado foi a estirpe GSF30, sendo também a estirpe referência desta espécie (STRAUB et al., 2013). A bactéria *H. frisingense* possui os sistemas de secreção Tipo I, Sec-SRP, Tat e pili tipo IV, envolvido na interação bactéria-hospedeiro em um relacionamento tanto simbiótico, quanto patogênico, porém não possui o sistema tipo III, frequentemente presente em bactérias associativas. Assim como *H. seropedicae*, também é capaz de fixar nitrogênio atmosférico (STRAUB et al., 2013).

No trabalho de Spilker e colaboradores (2008), os autores identificaram 3 isolados de *H. frisingense* através de sequenciamento do gene 16 S rRNA. Já em 2017, foi publicado o primeiro genoma de uma estirpe clínica desta espécie. Neste trabalho foi montado o genoma da bactéria *H. frisingense* ureolyticus VT-16-41 isolada de uma infecção de trato urinário (TETZ; TETZ, 2017).

4.3 TECNOLOGIAS DE SEQUENCIAMENTO

Em 1977, Sanger e colaboradores publicaram um método de sequenciamento de DNA baseado na utilização de dideoxynucleotídeos terminadores de cadeia (SANGER; NICKLEN; COULSON, 1977) que possibilitou, ainda em 1977, o sequenciamento do primeiro genoma viral (SANGER et al., 1977). Nas próximas décadas houveram grandes avanços na automatização do sequenciamento o que possibilitou a criação do projeto genoma humano, onde vários centros de pesquisa de todo o mundo realizavam sequenciamentos em larga escala com o objetivo de identificar toda a sequência de nucleotídeos do genoma humano. Este projeto terminou em 2003, levando 13 anos para ser completado e custou aproximadamente U\$3 bilhões (COLLINS; MORGAN; PATRINOS, 2003).

Apesar dos grandes avanços na tecnologia desenvolvida por Sanger, o custo para o sequenciamento em larga escala ainda era bastante elevado, o que restringia o seu uso apenas a grandes institutos. No entanto, em 2005, chegaram ao mercado os sequenciadores de segunda geração ou *next generation sequencing* (NGS) que simplificaram o processo de sequenciamento, além de reduzirem o custo e tempo necessários (LIU et al., 2012).

O primeiro método de sequenciamento de segunda geração foi o 454 Genome Sequencer, lançado em 2005 e desenvolvido pela *454 Life Sciences*, posteriormente adquirida pela Roche. Este sequenciador utiliza o método de pirosequenciamento baseado no uso do pirofosfato liberado durante a incorporação de um nucleotídeo e na atividade da enzima luciferase. De modo geral, o genoma é fragmentado em segmentos menores aos quais são ligados adaptadores em suas extremidades (GOODWIN; MCPHERSON; MCCOMBIE, 2016). Por complementaridade ao adaptador esses fragmentos são amplificados e imobilizados em nanoesferas (*beads*) em uma reação de PCR em emulsão originando a biblioteca de sequenciamento. A biblioteca pronta é depositada na célula de sequenciamento junto com a enzima DNA polimerase, sulfúrilase e luciferase, responsáveis pelas reações que emitirão o sinal de luz para o sequenciador. O processo ocorre em ciclos onde são adicionados cada um dos nucleotídeos em um ciclo independente (GOODWIN; MCPHERSON; MCCOMBIE, 2016; METZKER, 2010). Quando este é incorporado em uma das moléculas crescentes, libera um pirofosfato que, pela ação da enzima sulfúrilase, é convertido junto com adenosina

fosfosulfato (APS) em adenosina trifosfato (ATP) e usado pela enzima luciferase para oxidar a luciferina produzindo um sinal luminoso que é captado pelo sequenciador (LIU et al., 2012). O pirosequenciador da Roche, 454 pirosequencer, foi descontinuado em 2014, mas outra versão dessa tecnologia foi desenvolvida pela empresa Qiagen e incorporada no sequenciador *Gene Reader*. Outra tecnologia de sequenciamento NGS muito utilizada é a criada pela empresa Solexa, posteriormente adquirida pela Illumina. Por esse método o genoma também é inicialmente segmentado e ligado a adaptadores. Os fragmentos de DNA são imobilizados diretamente na célula de sequenciamento e amplificados através de PCR em ponte gerando *clusters* de moléculas de mesma sequência, dispensando o uso de nanoesferas (GOODWIN; MCPHERSON; MCCOMBIE, 2016). Assim como na tecnologia de pirosequenciamento, a tecnologia de síntese, criada pela Solexa, atua em ciclos, no entanto, todos os nucleotídeos são adicionados juntos à placa. Estes nucleotídeos são conjugados com uma molécula que atua como bloqueador, impedindo a continuação da síntese da cadeia de nucleotídeos, e um marcador fluorescente (GOODWIN; MCPHERSON; MCCOMBIE, 2016; METZKER, 2010). Após a fase de síntese, os nucleotídeos são lavados para eliminação dos não ligados às cadeias e são adicionadas enzimas que retiram as moléculas de bloqueio e o marcador fluorescente, gerando um sinal colorimétrico, que será detectado pelo sequenciador para cada um dos *clusters* (LIU et al., 2012; METZKER, 2010).

Em 2012, os mesmos criadores do pirosequenciador desenvolveram uma nova tecnologia de sequenciamento baseada na liberação de prótons durante a incorporação dos nucleotídeos na cadeia crescente de DNA. Este novo sequenciador foi denominado de Ion Torrent e posteriormente adquirido pela Life Technologies originando os sequenciadores PGM e Ion Proton. A construção das bibliotecas é semelhante aos métodos anteriores com fragmentação do genoma e ligação de adaptadores. O processo de sequenciamento é bastante semelhante ao utilizado pelo pirosequenciador, no entanto, ao invés de utilizar o pirofostato para detectar a incorporação de um nucleotídeo, é medida a variação de pH pela liberação de prótons (H⁺) (GOODWIN; MCPHERSON; MCCOMBIE, 2016; METZKER, 2010; VAN DIJK et al., 2014). Isso só é possível pela utilização de um chip semicondutor que atua como célula de sequenciamento e pHmetro.

Atualmente já existem no mercado também os sequenciadores de terceira geração, dentre os quais se destacam os sequenciadores PacBio e Nanopore. Estes não necessitam de etapa de PCR anterior ao sequenciamento, diminuindo consideravelmente o tempo de preparo das amostras. Possuem um sistema que faz captura e monitoração do sinal do sequenciamento em tempo real e apresentam leituras maiores que as tecnologias de segunda geração, facilitando assim o processo de montagem dos genomas (LIU et al., 2012).

Uma breve comparação entre diferentes tecnologias de sequenciamento está resumida na tabela 2.

TABELA 2: COMPARAÇÃO ENTRE PLATAFORMAS DE SEQUENCIAMENTO DE DNA

Sequenciador	Tamanho das leituras	Saída de dados	Leituras	Tempo por corrida	Taxa de erro	Preço por Gb (US\$)
Solid 5500 xl	75(SE)	240Gb	1,4 Bi	10 dias	≤0,1%	70
Illumina MiSeq v2	250(PE)	8,5Gb	30 Mi	39 horas	0,1%	142
Illumina MiSeq v3	300(PE)	15Gb	50 Mi	56 horas	0,1%	110
Illumina HiSeq2500 v4	125(PE)	500Gb	4 Bi	6 dias	0,1%	30
Illumina HiSeq 3000/4000	150(PE)	500Gb	2,5 Bi	3,5 dias	0,1%	22
Illumina HiSeqX	150(PE)	900 Gb p/ célula	3 Bi	3 dias	0,1%	7
Qiagen GeneReader	NA	12 genes; 1.250 mutações	NA	Alguns dias	NA	400-600 por painel
454 GS FLX Titanium XL+	1000(SE) 700(PE)	700Mb	1 Mi	23 horas	1%	9.500
Ion Proton	200 (SE)	10Gb	80 Mi	4 horas	1%	80
Ion S5 540	200(SE)	15Gb	80 Mi	2,5 horas	1%	300
Pacific Biosciences RSII	20kb	1Gb	55.000	4 horas	13%	1.000
Oxford Nanopore MinION	200kb	1,5Gb	>100.000	48 horas	12%	750

Fonte: Adaptado de GOODWIN; MCPHERSON; MCCOMBIE (2016)

4.4 MONTAGEM E ANOTAÇÃO DE GENOMAS

4.4.1 Montagem de genomas

Com o advento dos sequenciadores de DNA tornou-se possível identificar toda a sequência de nucleotídeos de um genoma, porém, como os dados de saída dos sequenciadores são pequenos fragmentos comparados ao tamanho total do genoma, são necessários *softwares* para unir e ordenar novamente estes fragmentos e gerar toda a sequência do DNA encontrado na célula. Estes *softwares* são chamados montadores de genomas (DOHM et al., 2007).

Estes programas geralmente buscam leituras com sequências altamente semelhantes em suas extremidades e as juntam, assumindo que estas são contíguas na molécula original. Este novo fragmento gerado através da união de leituras contíguas é chamado de contig (EKBLUM; WOLF, 2014). No entanto, nem sempre este fragmento gerado é correto pois em alguns casos como em regiões onde há uma sequência altamente repetitiva ou regiões que se repetem em diferentes porções do genoma, esta simples estratégia de unir leituras com extremidades iguais pode gerar erros na montagem (NAGARAJAN; POP, 2013). Outro fator que influencia nos erros de montagem é a taxa de erros da tecnologia de sequenciamento utilizada, por exemplo, se a taxa de erros da tecnologia for superior a taxa de divergência entre diferentes sequências, o montador terá dificuldade para diferenciar as duas, o que pode gerar erros. Assim, é desejável que os montadores consigam manipular e trabalhar com grandes quantidades de dados, além de saber lidar com os erros de sequenciamento da plataforma utilizada e a estrutura repetitiva do genoma trabalhado (NAGARAJAN; POP, 2013).

A utilização de leituras pareadas, ou seja, leituras *mate-pair* ou *paired-end*, facilita o processo de montagem pois estas possuem uma ligação de sequência desconhecida entre elas, auxiliando em regiões de repetição genômica, detectando erros estruturais e ajudando no posicionamento das leituras, podendo até mesmo criar ligações entre *contigs* com uma sequência desconhecida, formando os chamados “*scaffolds*” (NAGARAJAN; POP, 2013).

O processo de montagem pode ser realizado de duas formas, por referência ou *de novo*. O processo de montagem por referência utiliza o genoma de um organismo próximo, o qual serve como base para o alinhamento das leituras. Esta

estratégia é a mais simples, no entanto, pode gerar erros estruturais na sequência do genoma, pois as leituras são “forçadas” a se alinharem na mesma posição em que se encontram no genoma de referência, podendo haver perdas em regiões de translocação ou diferentes repetições em relação à referência. Já a estratégia *de novo* não utiliza de qualquer referência, utilizando unicamente dos alinhamentos entre as leituras para gerar a sequência do genoma. Montagens realizadas através da estratégia *de novo* costumam conter menos erros do que as realizadas por referência, no entanto demandam um custo computacional maior (EKBLUM; WOLF, 2014).

4.4.2 Anotação de genomas

Para enriquecer a informação gerada pela montagem do genoma e possibilitar uma melhor interpretação dos dados brutos, é necessário identificar regiões de interesse na sequência de nucleotídeos como regiões codificadoras de proteínas, rRNAs, tRNAs e ncRNAs.

A maioria dos algoritmos de anotação buscam sequências em fase aberta de leitura, as chamadas *open reading frames* ou ORFs, possuindo códon de início e final, além de regiões regulatórias à montante da ORF. Após a identificação destas regiões é necessária uma busca em banco de dados, identificando sequências ortólogas em outros organismos, para conferir uma informação de maior valor biológico como função, nome e processos biológicos nos quais este gene está envolvido. Este processo é chamado anotação do genoma (MÉDIGUE; MOSZER, 2007; STEIN, 2001).

4.5 GENÔMICA COMPARATIVA

Com o grande avanço em tecnologias de sequenciamento de DNA e a sua crescente acessibilidade devido às reduções no seu custo, houve um grande aumento nos depósitos de genomas em bancos de dados (NCBI, 2018). No entanto, ainda que a sequência anotada de um genoma traga consigo uma grande quantidade de informação, a sua comparação com outros genomas pode gerar informações que não seriam possíveis de outra maneira, dando origem à genômica comparativa. A genômica comparativa visa, através da similaridade e divergência,

determinados através de, principalmente, alinhamentos de genes e proteínas, inferir sequências homólogas, e assim, melhor entender a dinâmica evolutiva dos organismos (HARDISON, 2003).

Genomas procariotos, ainda que de organismos bastante próximos, como de estirpes diferentes de uma mesma espécie, podem apresentar regiões de grande divergência, oriundas principalmente de transferências horizontais de genes (HGT - Horizontal Gene Transfer), como transposons e plasmídeos provenientes de outros organismos (BINNEWIES et al., 2006). Estes fragmentos gênicos transferidos de um genoma a outro podem levar consigo características de um organismo A a um organismo B, acrescentando a ele funções antes inexistentes (BINNEWIES et al., 2006). Através de comparações genômicas é possível identificar regiões de um genoma oriundas de HGT e, através de homologia das proteínas dessas regiões com proteínas de outros organismos, inferir sua função, predizendo assim o seu possível impacto à bioquímica do organismo receptor.

Outro grande avanço proporcionado pela genômica comparativa é uma melhor classificação dos seres vivos através de seu material genético. Inicialmente a classificação dos organismos era realizada através de diferenças morfológicas e fisiológicas. No entanto com o advento da biologia molecular, novas formas de classificação começaram a surgir como a hibridização DNA-DNA, um experimento prático muito utilizado a partir dos anos de 1960, para determinar a relação entre bactérias (WAYNE et al., 1987) passando por comparações de genes codificantes o RNA ribossomal 16S de bactérias (MIZRAHI-MAN; DAVENPORT; GILAD, 2013) até, mais recentemente, métodos que se baseiam em toda a informação genética contida no genoma para a determinação da espécie de um organismo, como por exemplo a identidade nucleotídica média ou ANI (Average Nucleotide Identity), que analisa a taxa de mutação e similaridade entre genes em comum de dois organismos para a determinação se pertencem ou não a uma mesma espécie (GORIS et al., 2007).

5 COMPARAÇÃO GENÔMICA ENTRE ESTIRPES CLÍNICAS E AMBIENTAIS DE HERBASPIRILLUM SEROPEDICAE

Willian Klassen de Oliveira¹, Michelle Zibeti Tadra-Sfeir², Rodrigo Cardoso², Emanuel Maltempi de Souza², Fábio de Oliveira Pedrosa², Helisson Faoro^{1,3}

¹*Laboratório de Bioinformática, Setor de Educação Profissional e Tecnológica, Universidade Federal do Paraná, Curitiba, Brasil*

²*Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná, Curitiba, Brasil*

³*Laboratório da Regulação da Expressão Gênica, Instituto Carlos Chagas, Fiocruz-PR, Curitiba, Brasil*

Palavras Chave: *Herbaspirillum seropedicae*, isolados clínicos, bactérias oportunistas

5.1 RESUMO

Bactérias do gênero *Herbaspirillum* são normalmente encontradas em amostras ambientais, no entanto, recentemente, vêm sendo encontradas em isolados de amostras clínicas, em especial de pacientes imunocomprometidos. Spilker e colaboradores, em 2008, obtiveram 28 diferentes isolados de bactérias do gênero *Herbaspirillum*, de amostras de escarro de pacientes com fibrose cística. Dentre os isolados encontra-se uma estirpe inicialmente classificada como *Herbaspirillum* sp. linhagem 2 AU13965. Para melhor compreender como bactérias de origem ambiental passam a colonizar e infectar humanos, realizamos a montagem e comparação do genoma da bactéria *Herbaspirillum* lineage 2 AU13965 com outros genomas de representantes do gênero *Herbaspirillum*. O genoma foi sequenciado nas plataformas Illumina MiSeq e Ion Proton, montado usando o programa SPAdes e anotado usando o programa PROKKA. O genoma foi fechado, e contém um cromossomo circular de 5.350.014 pb e um plasmídeo de 42.977 pb. A comparação de genomas usando a identidade nucleotídica média revelou que o *Herbaspirillum* linhagem 2 pertencente a espécie *H. seropedicae*. Para investigar quais genes poderiam estar envolvidos na mudança de ambiente as proteínas de AU13965 foram comparadas com as de *H. seropedicae* SmR1, um isolado endofítico

desta espécie, além de bancos de dados de interação, virulência e resistência a antibióticos. Dentre as diferenças encontradas nestas análises estão a perda do *cluster* de genes *nif* e do sistema de secreção do tipo 3, que codificam respectivamente proteínas responsáveis pela fixação biológica de nitrogênio e interação com plantas, e a aquisição de genes que codificam proteínas potencialmente envolvidas na evasão do sistema imune como a responsável pela síntese de beta-1,2-glucana cíclica e a *ubiquitinin carboxyl-terminal hydrolase*. No entanto, apesar das diferenças encontradas entre os isolados clínico e ambiental, a grande maioria dos fatores de virulência e interação patógeno-hospedeiro encontrados são comuns entre os dois genomas, o que torna possível que esta espécie possua um potencial para colonizar e infectar o organismo humano podendo atuar como um patógeno oportunista.

5.2 INTRODUÇÃO

Herbaspirillum seropedicae é uma espécie da classe beta do filo Proteobacteria, isolada pela primeira vez no município de Seropédica no estado do Rio de Janeiro de rizosfera e raízes de gramíneas como milho, arroz e sorgo (BALDANI et al., 1986). Desde seu isolamento verificou-se que esta espécie era capaz de realizar a fixação biológica de nitrogênio (BALDANI et al., 1986), processo pelo qual bactérias diazotróficas convertem nitrogênio atmosférico (N₂) em íons amônio (NH₄⁺), uma molécula assimilável pelas plantas (JAMES, 2000). Somando-se a isto, estas bactérias são capazes de secretar fitohormônios (PEDROSA et al., 2011), promovendo assim, o crescimento vegetal, o que despertou grande interesse biotecnológico nesta espécie.

No entanto, recentemente, bactérias desta espécie começaram a ser isoladas em amostras clínicas, principalmente de pacientes imunocomprometidos acometidos por doenças como leucemia e fibrose cística. Spilker e colaboradores (2008) isolaram 28 estirpes diferentes, classificados como pertencentes ao gênero *Herbaspirillum* por comparação do gene 16S rRNA, de escarro de pacientes com fibrose cística. Dentre as estirpes isoladas por Spilker, se encontram 2 isolados da espécie *Herbaspirillum seropedicae* e um isolado inicialmente classificado como *H. lineage 2 AU13965*.

Um dos isolados de Spilker, a estirpe *H. seropedicae* AU14040, teve seu genoma montado e comparado com um isolado ambiental da mesma espécie mostrando que houveram aquisições e perdas de pacotes gênicos pelo isolado clínico. Entre as diferenças podemos citar a perda dos genes responsáveis pela fixação biológica de nitrogênio, a troca da região gênica de LPS e várias aquisições de genes possivelmente envolvidos na interação destas bactérias com células do organismo humano (FAORO et al., in prep). Marques e colaboradores (2015) realizaram testes bioquímicos com estas bactérias e comprovaram a perda de sua capacidade de fixar nitrogênio, no entanto não identificaram outras mudanças fisiológicas entre a estirpe clínica e ambiental.

Neste trabalho, realizamos o sequenciamento, a montagem e a anotação da estirpe clínica, inicialmente classificada como *H. lineage 2* AU13965, identificamos sua espécie e comparamos seu genoma com os demais genomas desta espécie.

5.3 MATERIAIS E MÉTODOS

5.3.1 Purificação de DNA e sequenciamento

As células de *Herbaspirillum* linhagem 2 AU13965 foram crescidas durante a noite em meio NFbHP (KLASSEN et al., 1997) a 37°C até atingir 1 unidade de absorvância em 600 nanômetros. O DNA foi então purificado através de extração por fenol clorofórmio (PEDROSA et al., 2011). O sequenciamento foi realizado em uma estratégia híbrida, utilizando duas tecnologias diferentes de sequenciamento, Ion Proton e Illumina MiSeq. A biblioteca construída para a plataforma de sequenciamento Ion Proton foi feita utilizando o kit *Ion Xpress Plus Fragment Library* kit (*ThermoFisher*) e sequenciado em leituras *single-end* de tamanho 200 pb. Já a biblioteca da plataforma Illumina MiSeq foi construída usando o kit *Nextera DNA sample preparation* (Illumina) e sequenciado com leituras pareadas (*paired-end*) na configuração 2x300 pb.

5.3.2 Montagem e anotação do genoma

A montagem do genoma foi realizada utilizando as leituras provenientes das diferentes plataformas de sequenciamento de forma individual (apenas leituras provenientes de uma tecnologia de sequenciamento) e de forma combinada (utilizando as leituras das duas plataformas de sequenciamento na mesma montagem). Foram utilizados os montadores Newbler v2.9 (454 Life Sciences, Brandford, CT), SPAdes v3.10.0 (BANKEVICH et al., 2012) e o montador CLC *de novo assembly* v10 (CLC bio, Qiagen). A avaliação das montagens foi realizada pelo programa QUAST (*Quality Assessment Tool for Genome Assemblies*) (GUREVICH et al., 2013) para a escolha da melhor montagem, a qual foi utilizada como montagem principal para a realização das próximas etapas. O fechamento dos *gaps* desta montagem foi realizada pelo programa GFinisher (GUIZELINI et al., 2016) e FGAP (PIRO et al., 2014) utilizando as demais montagens como conjunto de dados a ser utilizado pelos programas para efetuar o fechamento dos *gaps*. Os *gaps* restantes após este passo foram fechados manualmente buscando regiões de similaridade nas extremidades dos *contigs* e nas montagens realizadas pelos outros montadores utilizando o algoritmo Blastn (ALTSCHUL et al., 1990). A validação da montagem foi realizada através do mapeamento das leituras na montagem final a fim de identificar regiões de inconsistências na estrutura do genoma.

O genoma completo foi anotado através do servidor SILA, disponível em <http://200.236.3.34/SILA/login.jsp>, e do algoritmo Prokka v1.12 com a opção --rfam para realizar a busca de RNAs não codificadores (ncRNAs) (SEEMANN, 2014). As duas anotações foram mescladas e curadas manualmente.

5.3.3 Comparações genômicas

A comparação das sequências nucleotídicas dos genomas foi realizada através do algoritmo Pyani (PRITCHARD et al., 2016) que calcula a identidade nucleotídica média entre o genoma da estirpe sequenciada nesse trabalho e todos os genomas do gênero *Herbaspirillum* depositados no banco de dados NCBI. As similaridades e divergências entre os genomas do isolado clínico e o genoma da estirpe ambiental *H. seropedicae* SmR1 foram investigadas através de um alinhamento genoma-genoma realizado através do algoritmo *blastn* (ALTSCHUL et

al., 1990) e visualizadas através dos programas *Artemis Comparison Tool* - ACT (CARVER et al., 2005) e BRIG (ALIKHAN et al., 2011). Os genes presentes na linhagem clínica e ausentes na ambiental foram investigados de forma mais profunda através de buscas, realizadas pelo programa blastp (ALTSCHUL et al., 1997), das proteínas correspondentes a estes genes em bancos de dados para identificação de sua função, domínios e possível relação na interação e colonização no hospedeiro humano. As vias metabólicas da estirpe clínica *AU13965* e ambiental *H. seropedicae SmR1* foram anotadas através do algoritmo KAAS (MORIYA et al., 2007) e comparadas para identificar a presença de vias metabólicas exclusivas para cada um dos organismos. As possíveis regiões de transferência horizontal de genes foram identificadas através dos algoritmos *Alien Hunter* (VERNIKOS; PARKHILL, 2006) e GIPSy (SOARES et al., 2016). Para classificações de genes, quanto as suas funções, foi utilizado o algoritmo Blast KOALA (KANEHISA; SATO; MORISHIMA, 2016). Para a identificação de fatores de virulência e de interação patógeno-hospedeiro nos genomas analisados, foi utilizado o algoritmo Blastp (ALTSCHUL et al., 1997) para efetuar alinhamento de buscas contra o *Core dataset* (Apenas proteínas verificadas experimentalmente) do banco de dados VFDB (*Virulence Factors DataBase*) (CHEN et al., 2005) e PHI - *base* (*Pathogen Host Interactions base*) (URBAN et al., 2015) respectivamente. Para ambas as buscas nos bancos de dados, foi utilizado um *cut-off* de 50% de identidade, *e-value* de 10^{-10} e cobertura mínima de 80%. As proteínas encontradas nos organismos *AU13965* e *SmR1* foram copiadas para um novo arquivo multifasta e um novo alinhamento foi realizado entre elas para a identificação de fatores de interação e virulência exclusivas de cada um dos organismos. Já para realizar busca de genes que conferem resistência a antibióticos foi utilizado o algoritmo HMMer v3.1b2 para realizar buscas das proteínas dos organismos em um perfil HMM do banco curado de genes de resistência a antibióticos, ResFams (Core v1.2) (GIBSON; FORSBERG; DANTAS, 2015).

As proteínas codificadas pelos genes presentes no plasmídeo foram alinhadas contra o banco de dados do NCBI e foram retirados os 10 alinhamentos com maior *score*, para cada proteína, desde que estes apresentassem identidade maior a 50%, *e-value* inferior a 10^{-5} e cobertura superior a 80%, para contabilizar quais organismos compartilham ortólogos para estes genes.

5.4 RESULTADOS

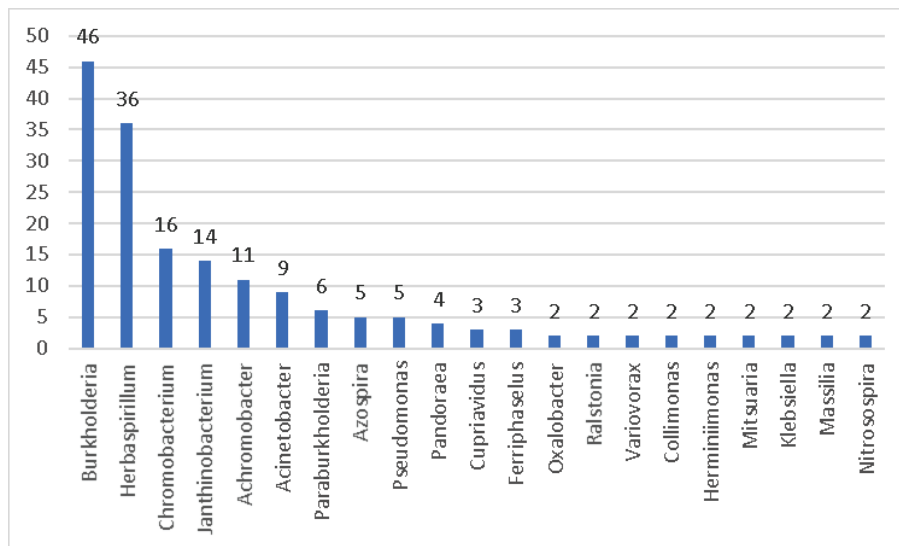
5.4.1 Genoma de *Herbaspirillum* linhagem 2 AU13965

Dentre as montagens realizadas e avaliadas pelo algoritmo QUAST, a que apresentou melhores parâmetros e, portanto, foi escolhida para se prosseguir com as próximas etapas foi a realizada pelo montador SPAdes. Esta montagem apresentou 38 *contigs* acima de 500 pb e um N50 de 274 kb. Após o fechamento automático dos *gaps*, restaram 12 contigs, os quais foram manualmente fechados com base em comparações entre as extremidades dos contigs e as demais montagens realizadas, até que restassem dois contigs, um representando toda a sequência do cromossomo e um plasmídeo circular. Após o mapeamento das leituras nas sequências foi verificado que estas não apresentam regiões de cobertura zero ou inconsistências estruturais.

O genoma fechado de *H. lineage 2 AU13965* possui um cromossomo de 5.350.014pb e com conteúdo GC% de 64,1%, 3 operons de rRNAs, 24 ncRNAs, 62 tRNAs e 4.720 CDSs. Já seu plasmídeo possui 42.977pb, contendo 54 CDSs e conteúdo GC% de 62,6%.

Os genes encontrados no plasmídeo apresentam ortólogos em diversos organismos, entre eles alguns gêneros que possuem bactérias oportunistas como *Burkholderia*, *Chromobacterium*, *Achromobacter*, *Acinetobacter*, *Pseudomonas*, *Pandoraea* e *Klebsiella*. O gráfico A1 representa os gêneros que apresentaram mais de 1 *hit* nos alinhamentos realizados contra o banco de dados NCBI. Os dados completos dos alinhamentos contra o banco de dados se encontram na tabela AS 1 do material suplementar.

GRÁFICO A 1: NÚMERO DE *HITS* NOS ALINHAMENTOS DOS GENES DO PLASMÍDEO DE AU13965

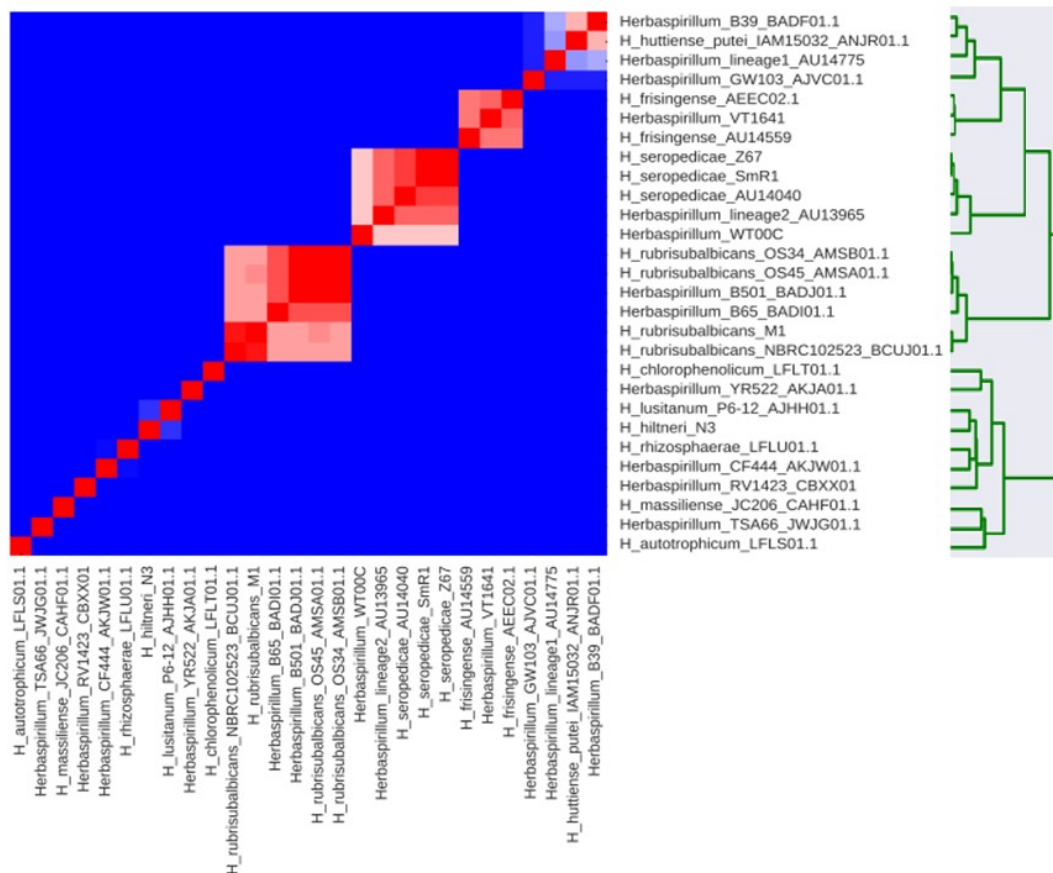


Fonte: O Autor, 2018

Nota: Gráfico de número de *hits* encontrado para cada gênero entre os alinhamentos realizados dos genes presentes no plasmídeo contra o banco de dados NCBI.

Através de cálculos de identidade nucleotídica média realizados pelo algoritmo Pyani, foi observado que este organismo se encaixa na espécie *Herbaspirillum seropedicae* (Figura A1), pois possui um ANI de 97,99% contra o genoma da estirpe ambiental *H. seropedicae SmR1* e 97,45% com o isolado clínico AU14040 desta mesma espécie.

FIGURA A 1: *HEATMAP* DE IDENTIDADE NUCLEOTÍDICA MÉDIA DOS ORGANISMOS DO GÊNERO *Herbaspirillum*



Fonte: O Autor, 2018.

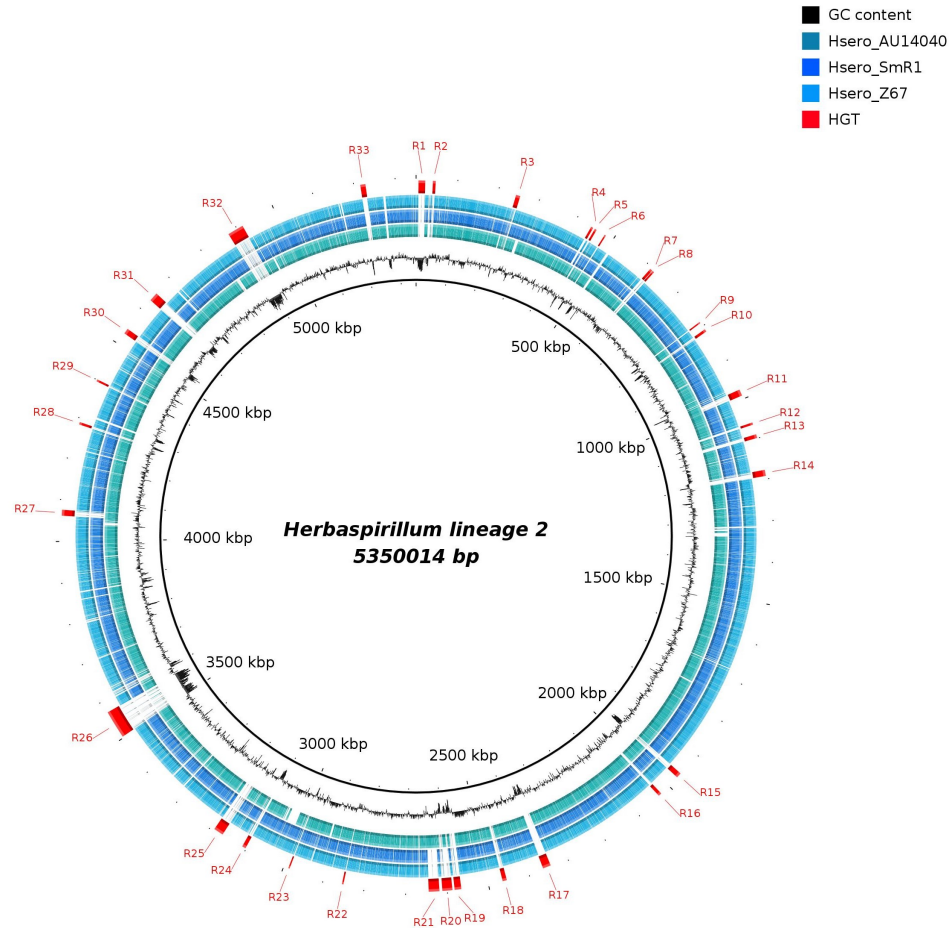
Nota: *Heatmap* e árvore filogenética construídos a partir de valores de identidade nucleotídica média entre genomas do gênero *Herbaspirillum*. Quanto mais vermelha a tonalidade dos pontos no *heatmap* maior a identidade entre os genomas.

5.4.2 Ilhas genômicas

Através do algoritmo *Alien Hunter* foram identificadas 45 ilhas genômicas na sequência do genoma de *AU13965*, sendo que, através de visualização no *Artemis Comparison Tool* (ACT) foi identificado que, 33 destas ilhas, não estavam presentes no genoma de *H. seropedicae* SmR1 (Figura A2). Já o algoritmo GIPSy encontrou 19 ilhas, dentre as quais classificou 5 como sendo de resistência (Ilhas genômicas 2, 3, 8, 15 e 19), 2 metabólicas (Ilhas genômicas 2 e 5), 2 de simbiose (Ilhas genômicas 18 e 19) e 7 de virulência (Ilhas genômicas 2, 3, 5,

8, 10, 11 e 18). Os 201 genes presentes nas ilhas encontradas pelo GIPSy se encontram nas tabelas AS 2 – AS 20.

FIGURA A 2: REPRESENTAÇÃO CIRCULAR DO GENOMA DE *H. linhagem 2*



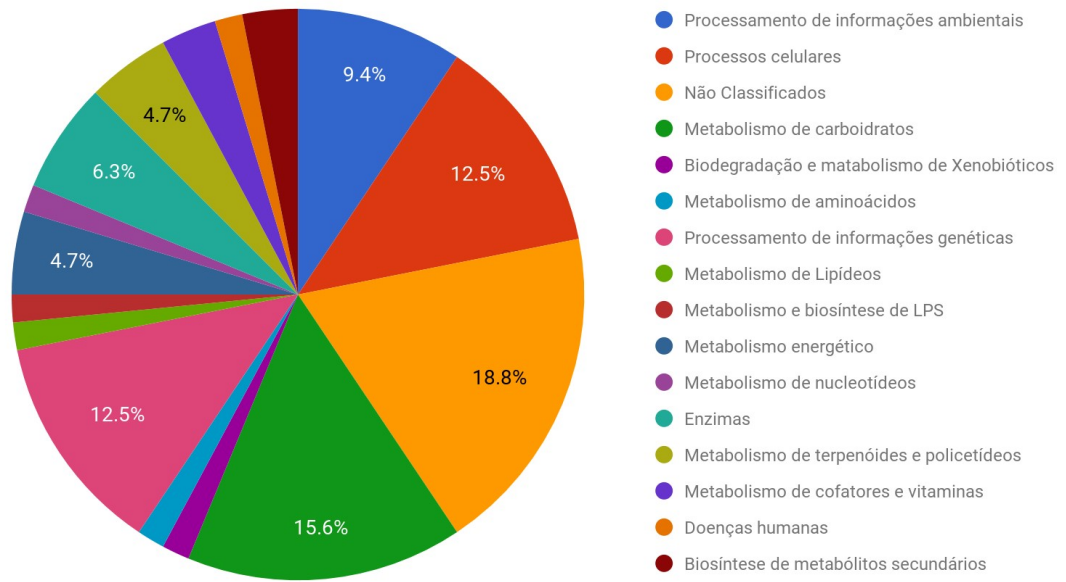
Fonte: O Autor, 2018.

Nota: Representação circular do genoma de *H. linhagem 2* AU13965. Os anéis da figura representam, do mais interno para o mais externo, o genoma da estirpe *H. linhagem 2* AU13965, conteúdo GC%, presença de respectivos genes ortólogos em *H. seropedicae* AU14040, presença de respectivos genes ortólogos em *H. seropedicae* SmR1, presença de respectivos genes ortólogos em *H. seropedicae* Z67, e as marcações R1-R33 representam as ilhas encontradas pelo algoritmo *AlienHunter* confirmadas através de visualização no ACT.

Os genes presentes nas ilhas genômicas exclusivas de AU13965 foram anotados pelo algoritmo Blast KOALA e classificados de acordo com suas funções.

Do total dos 201 genes presentes nas ilhas foram classificados 42. Os dados encontrados nesta classificação estão apresentados no gráfico A 2.

GRÁFICO A 2: GRÁFICO DE FUNÇÕES DE GENES PRESENTES NAS ILHAS DO GENOMA DE *Herbaspirillum seropedicae* AU13965



Fonte: O Autor, 2018

Nota: Gráfico construído a partir da predição da função dos genes presentes nas ilhas genômicas de *H. seropedicae* AU13965 a partir do algoritmo blast KOALA.

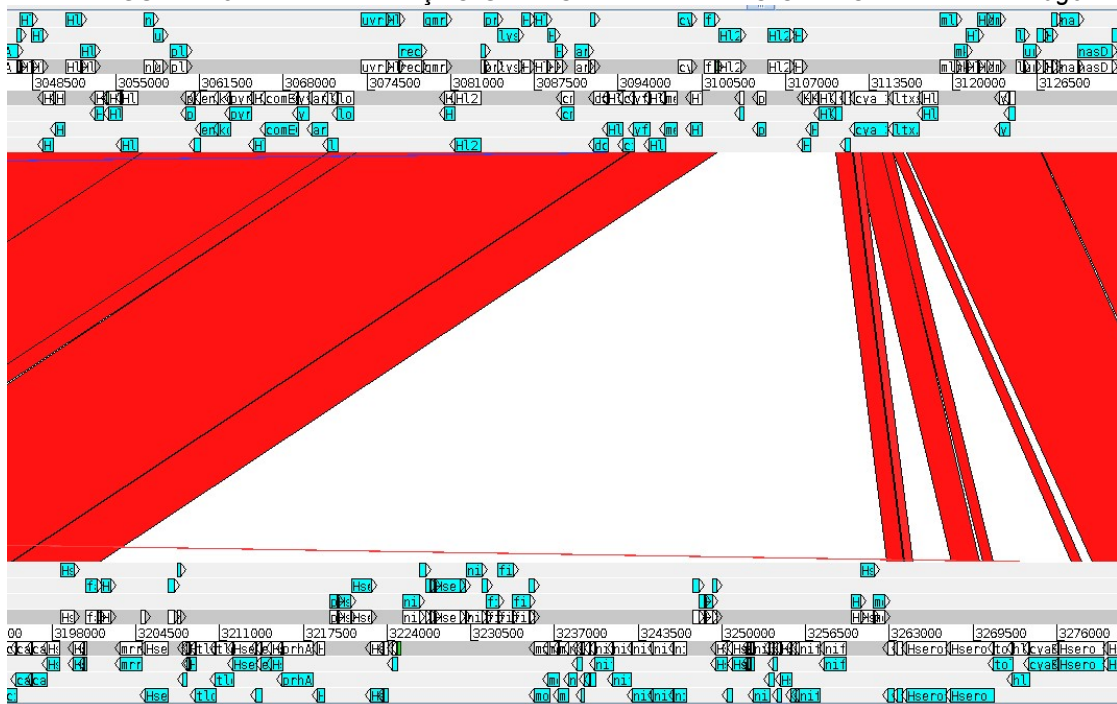
Dentre as seções que mais se destacam numericamente na classificação do Blast KOALA se encontram os genes envolvidos em processos celulares, metabolismo de carboidratos e processamento de informações genéticas. Houve também um gene (*HI2_08800*) classificado na categoria doenças humanas. Este gene foi anotado como uma chaperonina GroEL e podem desenvolver diversas funções em bactérias, atuando como toxina e protease (GARDUÑO et al., 2011), estando marcado no programa Blast KOALA nas vias de Legionellose e Tuberculose.

5.4.3 Comparações entre os genomas SmR1 e AU13965

90,95% das proteínas de *AU13965* possuem ortólogos em *SmR1* com um *cut-off* de 50% de identidade em análises de blastp recíproco. Alguns destes genes são elementos chave para a forma de vida destes organismos.

O operon *nif* codifica as proteínas estruturais da nitrogenase, que, por sua vez, realiza a conversão do nitrogênio atmosférico (N₂) em íons amônio (NH₄⁺), uma forma assimilável pelas plantas, através do processo chamado fixação biológica de nitrogênio realizado por bactérias diazotróficas (POZA-CARRIÓN et al., 2014). Entretanto, apenas as estirpes ambientais desta espécie possuem esses genes, sendo ausentes nas estirpes clínicas *AU13965* e *AU14040*, corroborando com o estilo de vida destas bactérias (figura A 3).

FIGURA A 3: REPRESENTAÇÃO GRÁFICA DA PERDA DO OPERON *nif* EM *H. linhagem 2*



Fonte: O Autor, 2018.

Nota: Representação gráfica do programa ACT, revelando a perda do operon de genes *nif* no organismo *H. seropedicae* AU13965 (genoma superior) em relação a estirpe ambiental SmR1 (genoma inferior). Os blocos vermelhos indicam regiões de identidade superior a 50% entre os genomas.

Além dos genes *nif* a estirpe AU13965 também perdeu os genes que codificam o sistema de secreção tipo 3 em relação ao isolado ambiental desta espécie. Esse sistema de secreção está normalmente associado a fatores de virulência em bactérias, atuando, principalmente, na injeção de proteínas bacterianas efetoras diretamente no citoplasma da célula hospedeira, efetuando mudanças nestas células e permitindo a entrada e colonização das mesmas (COBURN; SEKIROV; FINLAY, 2007). No entanto, apesar da tradicional função deste conjunto de genes, especula-se que a sua função em *SmR1* seja de secretar proteínas efetoras benéficas ao relacionamento simbiote bactéria-planta realizada por este organismo, função esta que já foi descrita em *Rhizobium* NGR234 (VIPREY et al., 1998), razão esta, possivelmente, de este conjunto de genes não ser mais necessário em uma interação com o organismo humano.

5.4.3.1 Fatores de virulência

Através de alinhamentos realizados pelo algoritmo blastp contra o banco de dados VFDB foi possível identificar em AU13965, 75 proteínas possivelmente envolvidas em alguma forma de virulência (Tabela AS 21), enquanto na estirpe ambiental *SmR1* foram identificados 83 possíveis fatores de virulência (Tabela AS 22). Dentre os fatores de virulência encontrados para AU13965, 12 não são encontrados nos resultados do alinhamento estirpe ambiental *SmR1* com o mesmo banco de dados (Tabela A1).

TABELA A 1: GENES DE VIRULÊNCIA EXCLUSIVOS DA ESTIRPE CLÍNICA AU13965

Gene em AU13965	Gene Banco de dados	Produto	Organismo	Identidade (%)
HI2_22670	<i>pvdL</i>	<i>peptide synthase PvdL [pyoverdine (IA001)]</i>	<i>Pseudomonas aeruginosa PAO1</i>	51,04
HI2_29250	<i>tsr</i>	<i>methyl-accepting chemotaxis protein I</i>	<i>Burkholderia pseudomallei K96243</i>	66,09
HI2_41260	<i>gmd</i>	<i>GDP-mannose 4,6-dehydratase [LPS (VF0367)]</i>	<i>Brucella melitensis bv. 1 str. 16M</i>	79,15
HI2_41290	<i>ddhA</i>	<i>glucose-1-phosphate cytidyltransferase [O-antigen (VF0392)]</i>	<i>Yersinia enterocolitica subsp. enterocolitica 8081</i>	70,70
HI2_41300	<i>ddhB</i>	<i>CDP-glucose 4,6-dehydratase</i>	<i>Yersinia enterocolitica</i>	59,73

Gene em AU13965	Gene Banco de dados	Produto	Organismo	Identidade (%)
		[O-antigen (VF0392)]	<i>subsp. enterocolitica</i> 8081	
HI2_41430	<i>cap8E</i>	<i>capsular polysaccharide synthesis enzyme Cap8E</i>	<i>Staphylococcus aureus subsp. aureus</i> MW2	66,57
HI2_41440	<i>cap8G</i>	<i>capsular polysaccharide synthesis enzyme Cap8G</i>	<i>Staphylococcus aureus subsp. aureus</i> MW2	52,01
HI2_43270	<i>rffG</i>	<i>dTDP-glucose 46-dehydratase</i> [LOS (CVF494)]	<i>Haemophilus influenzae</i> Rd KW20	59,09
HI2_46720	<i>ureA</i>	<i>urease alpha subunit ureA</i>	<i>Helicobacter pylori</i> 26695	63,00
HI2_46730	<i>ureA</i>	<i>urease alpha subunit UreA</i>	<i>Helicobacter pylori</i> 26695	55,67
HI2_46740	<i>ureB</i>	<i>urease beta subunit UreB - urea amidohydrolase</i>	<i>Helicobacter pylori</i> 26695	64,50
HI2_46780	<i>ureG</i>	<i>urease accessory protein</i>	<i>Helicobacter pylori</i> 26695	64,65

Fonte: O Autor (2018).

Dentre os fatores de virulência encontrados, através dos alinhamentos contra o banco de dados VFDB, apenas em AU13965 está um conjunto de genes ureAB, o qual apresenta em *Helicobacter pylori* a função de hidrolisar ureia em amônia e CO₂, aumentando o pH do meio estomacal e permitindo a colonização destas bactérias neste epitélio (EATON et al., 1991). Outra função destes genes é a utilização de ureia como fonte de nitrogênio quando em condições limitantes deste elemento (FONG et al., 2013). No entanto, estes genes codificam para metaloproteínas, as quais necessitam de dois íons níquel para seu correto funcionamento na célula, sendo a proteína codificada pelo gene *ureG*, essencial para a entrega destes metais no sítio ativo da enzima tornando-a funcional (FONG et al., 2013).

Os genes encontrados entre os *locus* HI2_41260 e HI2_41440 se encontram em um operon responsável pela síntese de lipopolissacarídeo (LPS), o qual foi modificado em relação ao presente na estirpe ambiental. Os lipopolissacarídeos são moléculas presentes na membrana externa de bactérias

gram negativas e estão muitas vezes associados com a interação com o hospedeiro e com a proteção das células bacterianas ao sistema imune durante uma infecção. Estas moléculas podem agir na prevenção da ligação do sistema complemento ou mimetizando a superfície das células do hospedeiro através da ligação de carboidratos normalmente presente nestas células (MATSUURA, 2013). Portanto, a troca das proteínas envolvidas na síntese de LPS por proteínas encontradas em patógenos humanos pode beneficiar a colonização da bactéria no organismo humano.

Um grande número de proteínas encontradas no alinhamento contra o banco de dados de virulência esteve presente em ambos os genomas, sendo em grande parte, proteínas relacionadas a quimiotaxia e formação do flagelo, as quais também podem atuar na adesão à superfície das células epiteliais e em muitas bactérias patogênicas, estas proteínas têm demonstrado função de adesinas, exportação de proteínas e formação de biofilme (HAIKO; WESTERLUND-WIKSTRÖM, 2013).

5.4.3.2 Fatores de interação patógeno-hospedeiro

Através de alinhamentos com o banco de dados de interação patógeno-hospedeiro PHI-base foi possível identificar 98 proteínas possivelmente envolvidas na interação com o hospedeiro em *AU13965* (Tabela AS 23) e 101 em *SmR1* (Tabela AS24), das quais 2 são exclusivas da estirpe clínica (Tabela A2).

TABELA A 2: GENES DE INTERAÇÃO DA ESTIRPE CLÍNICA AU13965

Gene em <i>AU13965</i>	Gene Banco de dados	Produto	Organismo	Identidade (%)
<i>HI2_03110</i>	<i>hupA</i>	Nucleoid - associated protein (NAP) HU	<i>Edwardsiella tarda</i>	53,93
<i>HI2_25120</i>	<i>hupA</i>	Nucleoid - associated protein (NAP) HU	<i>Edwardsiella tarda</i>	55,56

Fonte: O Autor, 2018

As proteínas codificadas pelos genes *HI2_03110* e *HI2_25120* apresentam alta similaridade com a proteína codificada pelo gene *hupA* de *Edwardsiella tarda*, um patógeno de vários animais, em especial de peixes, nos quais pode causar septicemia hemorrágica sistêmica (WANG et al., 2014). Esta

proteína é um regulador transcricional que quando inativada em patógenos, como *Salmonella enterica* serovar typhimurium, reduzem a expressão de genes de virulência, reduzindo significativamente a capacidade do patógeno invadir e se multiplicar intracelularmente no hospedeiro (MANGAN et al., 2011; SCHECHTER et al., 2003). Além disto a expressão de genes relacionados a síntese de LPS também são modulados por esta proteína (ALBERTI-SEGUI et al., 2010).

5.4.3.3 Características específicas do genoma de *H. seropedicae* AU13965

Foram também encontrados alguns genes com possível envolvimento na forma de vida dessas bactérias que, porém, não foram identificadas nos alinhamentos contra os bancos de dados de interação e virulência, sendo identificadas através de alinhamentos dos genes presentes nas ilhas contra o banco de dados NCBI.

O gene *HI2_23070* codifica para a proteína beta-1,2-glucana cíclica sintase, a qual é responsável pela síntese de uma estrutura macromolecular construída através de monômeros de D-glucose, a glucana cíclica (BOHIN, 2000). Estas proteínas podem estar envolvidas tanto na interação bactéria-planta como bactéria-humano (RIGANO et al., 2007a). Em alguns patógenos como *Brucella abortus*, este gene é necessário para sobrevivência intracelular deste patógeno, permitindo a replicação intracelular e evasão dos lisossomos da célula hospedeira (ARELLANO-REYNOSO et al., 2005). Este gene também é encontrado nos patógenos *Salmonella typhi*, *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Yersinia pestis*, *Neisseria gonorrhoeae* (BOHIN, 2000), além de *H. seropedicae* AU14040 e *H. frisingense* VT-16-41, isolados clínicos dos gênero *Herbaspirillum*.

A proteína *HI2_08870*, identificada como proteína hipotética através do alinhamento gerado pelo algoritmo Blastp, possui ortólogo em *SmR1* e vários domínios *Se1*. Proteínas com multidomínios *Se1* podem estar associadas com a montagem de complexos macromoleculares por vezes envolvidos na interação com células eucarióticas (MITTL; SCHNEIDER-BRACHERT, 2007). Entre estas proteínas podemos citar as codificadas pelos genes *EnhC* e *LpnE* do patógeno humano *Legionella pneumophila*, os quais estão associados com a entrada destas bactérias em monócitos e células epiteliais humanas (MITTL; SCHNEIDER-BRACHERT, 2007).

O gene *HI2_41750* codifica a enzima *ubiquitinin carboxyl-terminal hydrolase*, uma proteína contendo o domínio Peptidase C19B, normalmente encontrado apenas em proteínas eucarióticas e que remove marcações de ubiquitina de polipeptídeos marcados para degradação. Embora a ubiquitina seja mais comumente utilizada na marcação de polipeptídeos, células eucarióticas podem utilizar estas proteínas para realizar a marcação de patógenos, os encaminhando para lisossomos para posterior degradação (WATSON; MANZANILLO; COX, 2012). No entanto, algumas bactérias adquiriram genes que realizam a desmarcação de ubiquitina, as *ubiquitinin carboxyl-terminal hydrolases*, as quais se tornam importantes adaptações na evasão do sistema imune humano (CATIC et al., 2007).

5.4.3.4 Resistência a antibióticos

O genoma da estirpe *AU13965* apresentou, através das buscas realizadas através do perfil HMM no banco de dados de genes de resistência a antibióticos ResFams, um total de 26 genes de resistência, sendo a maioria destes de resistência a beta-lactâmicos. Já os genomas das estirpes *AU14040* e *SmR1* apresentaram um total de 27 e 25 genes de resistência respectivamente (Tabela A3). De modo geral, no que diz respeito a genes de resistência a antibióticos, as estirpes clínicas *AU13965* e *AU14040* pouco diferem da *SmR1*, alterando apenas o número de cópias dos genes de determinadas famílias de resistência, porém sem possuir nenhum gene exclusivo.

Os *locus* dos genes encontrados nesta análise encontram-se nas tabelas AS 25 – AS 27 presentes no material suplementar.

TABELA A 3: GENES DE RESISTÊNCIA A ANTIBIÓTICOS DA ESTIRPE AU13965

Descrição	ResFam Family Name	ResFam ID	Número de genes encontrados nos organismos		
			<i>AU13965</i>	<i>AU14040</i>	<i>SmR1</i>
<i>ATP binding cassette</i> (ABC) - bomba de efluxo de antibióticos	ABC <i>antibiotic efflux pump</i>	RF0007	6	8	6
Fator de membrana externa de complexo de efluxo de drogas	<i>AdeC-AdeK- OprM</i>	RF0021	2	1	1

Descrição	ResFam Family Name	ResFam ID	Número de genes encontrados nos organismos		
			AU13965	AU14040	SmR1
Beta lactamases de Classe A	<i>Class A</i>	RF0053	1	1	1
Beta lactamases de Classe B	<i>Class B</i>	RF0054	1	1	1
Beta Lactamase <i>Exo</i> (Classe A)	<i>Exo</i>	RF0072	1	1	1
Subunidade de bomba de efluxo que confere resistência a antibióticos	<i>Mac B</i>	RF0089	1	1	1
Proteína de fusão de membrana do complexo de efluxo multidrogas MexEF-OprN	<i>Mex E</i>	RF0098	2	2	2
<i>ATP binding cassette</i> (ABC) - bomba de efluxo de antibióticos	<i>MsbA</i>	RF0107	1	1	1
resistance-nodulation-cell division (RND) - Bomba de efluxo	<i>RND efflux</i>	RF0115	8	8	8
(Metallo-) Beta lactamases de subclasse B1 - Hidrolizam penicilinas, cefalosporinas e carbapenêmicos	<i>Subclass B1</i>	RF0123	1	1	1
(Metallo-) Beta lactamases de subclasse B2 - seletivas a carbapenêmicos	<i>Subclass B2</i>	RF0124	1	1	1
	<i>MexX</i>	RF0101	1	1	1
Total			26	27	25

Fonte: O Autor, 2018

5.5 DISCUSSÃO

A bactéria inicialmente identificada como *Herbaspirillum seropedicae* AU13965, isolada de escarro de pacientes com fibrose cística, teve seu genoma sequenciado e montado, sendo classificada, através de sua sequência de nucleotídeos, como pertencente a espécie *H. seropedicae*, através da análise de identidade nucleotídica média. Este organismo sofreu perdas de regiões gênicas como o operon *nif*, responsável pela fixação biológica de nitrogênio e o sistema de secreção tipo 3, ambos importantes na interação bactéria-planta. Este organismo também recebeu diversas transferências horizontais de genes, quando comparando seu genoma ao genoma do isolado ambiental *SmR1*. No entanto, apesar desta

estirpe ser um isolado clínico, a maior parte dos genes encontrados nestas regiões de ilhas genômicas estão relacionados a diversos processos metabólicos e celulares. Entre os genes presentes nas ilhas também há alguns com uma possível ligação à nova forma de vida desta bactéria como o responsável pela síntese da beta-1,2-glucana cíclica (*HI2_23070*) que, aliada ao gene *ubiquitin carboxyl-terminal hydrolase* (*HI2_41750*) podem auxiliar esta estirpe na defesa contra o sistema imune humano, evadindo a fusão de lisossomos em uma possível colonização intracelular. No entanto também há vários genes em comum entre os dois isolados como o grande número de fatores de virulência, interação patógeno-hospedeiro e de resistência a antibióticos compartilhados pelos dois organismos.

O grande número de regiões provenientes de transferências horizontais de genes destaca a grande plasticidade destes genomas, o que também beneficia a adaptação destes organismos a diferentes nichos como o solo e o organismo humano. Outro fator que beneficia esta migração é o fato de um de seus habitats ser a rizosfera. Este ambiente propicia uma grande competição entre microrganismos pelos nutrientes presentes nos exsudatos liberados pelas plantas, atuando assim como um fator de seleção para microrganismos com mecanismos de captação de nutrientes mais eficientes, produção de defesas, como resistência a antibióticos, além de mecanismos de infecção de células eucarióticas presentes no meio e secreção de moléculas de defesa. Todas estas características citadas, além dos genes de interação bactéria-planta, presentes nos genomas da espécie *H. seropedicae* podem beneficiar estes organismos a sobreviverem no complexo meio que é o organismo humano, evadindo os mecanismos de defesa do hospedeiro, auxiliando na infecção de células humanas, além de auxiliar na defesa contra fármacos que por muitas vezes são isolados de microrganismos presentes no solo.

5.6 REFERÊNCIAS

ALBERTI-SEGUI, C. *et al.* HU protein affects transcription of surface polysaccharide synthesis genes in *Porphyromonas gingivalis*. **Journal of Bacteriology**, v. 192, n. 23, p. 6217–6229, 2010.

ALIKHAN, N. F. *et al.* BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. **BMC Genomics**, v. 12, n. 1, p. 402, 2011.

ALTSCHUL, S. F. *et al.* Basic local alignment search tool. **Journal of molecular biology**, v. 215, n. 3, p. 403–10, 1990.

ALTSCHUL, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997.

ARELLANO-REYNOSO, B. *et al.* Cyclic beta-1,2-glucan is a *Brucella* virulence factor required for intracellular survival. **Nature immunology**, v. 6, n. 6, p. 618–625, 2005.

BALDANI, J. I. *et al.* Characterization of *Herbaspirillum seropedicae* gen. nov. sp. nov. a Root-Associated Nitrogen-Fixing Bacterium. **International journal of systematic bacteriology**, p. 86–93, 1986.

BANKEVICH, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455–477, 2012.

BOHIN, J. P. Osmoregulated periplasmic glucans in Proteobacteria. **FEMS Microbiology Letters**, v. 186, n. 1, p. 11–19, 2000.

CARVER, T. J. *et al.* ACT: The Artemis comparison tool. **Bioinformatics**, v. 21, n. 16, p. 3422–3423, 2005.

CATIC, A. *et al.* ElaD, a deubiquitinating protease expressed by *E. coli*. **PLoS ONE**, v. 2, n. 4, p. 1–8, 2007.

CHEN, L. *et al.* VFDB: A reference database for bacterial virulence factors. **Nucleic Acids Research**, v. 33, n. DATABASE ISS., p. 325–328, 2005.

COBURN, B.; SEKIROV, I.; FINLAY, B. B. Type III secretion systems and disease. **Clinical Microbiology Reviews**, v. 20, n. 4, p. 535–549, 2007.

EATON, K. A. *et al.* Essential Role of Urease in Pathogenesis of Gastritis Induced by *Helicobacter pylori* in Gnotobiotic Piglets. **Infection and Immunity**, v. 59, n. 7, p. S681–S685, 1991.

FONG, Y. H. *et al.* Structure of UreG/UreF/UreH Complex Reveals How Urease Accessory Proteins Facilitate Maturation of *Helicobacter pylori* Urease. **PLoS Biology**, v. 11, n. 10, 2013.

GARDUÑO, R. A. *et al.* The Legionella pneumophila chaperonin - An unusual multifunctional protein in unusual locations. **Frontiers in Microbiology**, v. 2, n. JUNE, p. 1–12, 2011.

GIBSON, M. K.; FORSBERG, K. J.; DANTAS, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. **ISME Journal**, v. 9, n. 1, p. 207–216, 2015.

GUIZELINI, D. *et al.* GFinisher: A new strategy to refine and finish bacterial genome assemblies. **Scientific Reports**, v. 6, n. October, p. 1–8, 2016.

GUREVICH, A. *et al.* QUAST: Quality assessment tool for genome assemblies. **Bioinformatics**, v. 29, n. 8, p. 1072–1075, 2013.

HAIKO, J.; WESTERLUND-WIKSTRÖM, B. The Role of the Bacterial Flagellum in Adhesion and Virulence. **Biology**, v. 2, n. 4, p. 1242–1267, 2013.

JAMES, E. K. Nitrogen fixation in endophytic and associative symbiosis. **Field Crops Research**, v. 65, n. 2–3, p. 197–209, 2000.

KANEHISA, M.; SATO, Y.; MORISHIMA, K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. **Journal of Molecular Biology**, v. 428, n. 4, p. 726–731, 2016.

KLASSEN, G. *et al.* Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae* SMRI. v. 891, p. 887–891, 1997.

MANGAN, M. W. *et al.* Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in *Salmonella enterica* serovar Typhimurium. **Microbiology**, v. 157, n. 4, p. 1075–1087, 2011.

MARQUES, A. C. Q. *et al.* Biochemical characteristics, adhesion, and cytotoxicity of environmental and clinical isolates of *Herbaspirillum* spp. **Journal of Clinical Microbiology**, v. 53, n. 1, p. 302–308, 2015.

MATSUURA, M. Structural modifications of bacterial lipopolysaccharide that facilitate gram-negative bacteria evasion of host innate immunity. **Frontiers in Immunology**, v. 4, n. MAY, p. 1–9, 2013.

MITTL, P. R. E.; SCHNEIDER-BRACHERT, W. Sel1-like repeat proteins in signal transduction. **Cellular Signalling**, v. 19, n. 1, p. 20–31, 2007.

MORIYA, Y. *et al.* KAAS: An automatic genome annotation and pathway reconstruction server. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. 182–185, 2007.

PEDROSA, F. O. *et al.* Genome of *herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. **PLoS Genetics**, v. 7, n. 5, 2011.

PIRO, V. C. *et al.* FGAP: an automated gap closing tool. **BMC research notes**, v. 7, n. 1, p. 371, 2014.

- POZA-CARRIÓN, C. *et al.* Kinetics of *nif* gene expression in a nitrogen-fixing bacterium. **Journal of Bacteriology**, v. 196, n. 3, p. 595–603, 2014.
- PRITCHARD, L. *et al.* Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. **Anal. Methods**, v. 8, n. 1, p. 12–24, 2016.
- RIGANO, L. A. *et al.* Bacterial cyclic beta-(1,2)-glucan acts in systemic suppression of plant immune responses. **The Plant cell**, v. 19, n. 6, p. 2077–2089, 2007.
- SCHECHTER, L. M. *et al.* The Small Nucleoid-Binding Proteins H-NS , HU , and Fis Affect *hilA* Expression in *Salmonella enterica* Serovar Typhimurium. v. 71, n. 9, p. 5432–5435, 2003.
- SEEMANN, T. Prokka: Rapid prokaryotic genome annotation. **Bioinformatics**, v. 30, n. 14, p. 2068–2069, 2014.
- SOARES, S. C. *et al.* GIPSy: Genomic island prediction software. **Journal of Biotechnology**, v. 232, p. 2–11, 2016.
- SPIPKER, T. *et al.* Recovery of *Herbaspirillum* species from persons with cystic fibrosis. **Journal of Clinical Microbiology**, v. 46, n. 8, p. 2774–2777, 2008.
- URBAN, M. *et al.* The Pathogen-Host Interactions database (PHI-base): Additions and future developments. **Nucleic Acids Research**, v. 43, n. D1, p. D645–D655, 2015.
- VERNIKOS, G. S.; PARKHILL, J. Interpolated variable order motifs for identification of horizontally acquired DNA: Revisiting the *Salmonella* pathogenicity islands. **Bioinformatics**, v. 22, n. 18, p. 2196–2203, 2006.
- WANG, L. *et al.* **HU-induced polymorphous filamentation in fish pathogen *Edwardsiella tarda* leading to reduced invasion and virulence in zebrafish.** [s.l.] Elsevier B.V., 2014. v. 171
- WATSON, R. O.; MANZANILLO, P. S.; COX, J. S. Extracellular *M. tuberculosis* DNA targets bacteria for autophagy by activating the host DNA-sensing pathway. **Cell**, v. 150, n. 4, p. 803–815, 2012.

5.7 MATERIAL SUPLEMENTAR

TABELA AS 1: NÚMERO DE *HITS* POR GÊNERO DOS ALINHAMENTOS DOS GENES DO PLASMÍDEO

Organismo	<i>hits</i>
Burkholderia	46
Herbaspirillum	36
Chromobacterium	16
Janthinobacterium	14
Achromobacter	11
Acinetobacter	9
Paraburkholderia	6
Azospira	5
Pseudomonas	5
Pandoraea	4
Cupriavidus	3
Ferriphaselus	3
Oxalobacter	2
Ralstonia	2
Variovorax	2
Collimonas	2
Herminiimonas	2
Mitsuaria	2
Klebsiella	2
Massilia	2
Nitrospira	2
Hyphomicrobium	1
Marinobacterium	1
Methylobacillus	1
Chloroflexi	1
Nitrosomonas	1
Selenomonas	1
Nitrospira	1
Bordetella	1
Giesbergeria	1
Sphingomonas	1
Ideonella	1
Bacillus	1
Rhizobacter	1

Organismo	hits
Bradyrhizobium	1
Enterobacter	1
Anaeromusa	1
Duganella	1
Alkanindiges	1
Raoultella	1
Vogesella	1

Fonte: O Autor, 2018

TABELA AS 2: GENES PRESENTES NA ILHA GENÔMICA 1 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_00040	hypothetical protein 5867:6541 reverse MW:26044
HI2_00041	Hypothetical Protein 6810:7220 forward MW:15198
HI2_00050	hypothetical protein 7416:8168 forward MW:28310
HI2_00060	tnsB Transposon Tn7 transposition protein TnsB 8161:10134 forward MW:75086
HI2_00070	hypothetical protein 10134:11102 forward MW:35909
HI2_00080	hypothetical protein 11156:12877 forward MW:66085
HI2_00090	hypothetical protein 13006:14370 forward MW:51761
HI2_00100	hypothetical protein 14794:15342 forward MW:20323
HI2_00110	hypothetical protein 15528:17270 forward MW:65084
HI2_00120	hypothetical protein 17478:18227 forward MW:28094
HI2_00130	hypothetical protein 19359:19652 forward MW:10914
HI2_00140	hypothetical protein 19619:20539 forward MW:34370

Fonte: O Autor, 2018

TABELA AS 3: GENES PRESENTES NA ILHA GENÔMICA 2 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_02350	cynR_2 HTH-type transcriptional regulator CynR 243468:244349 reverse MW:33430
HI2_02360	4-sulfomuconolactone hydrolase 244651:245424 forward MW:28265
HI2_02370	ttuB_1 Putative tartrate transporter 245435:246760 forward MW:47393
HI2_02380	Lactonase drp35 246783:247706 forward MW:33326

Locus	Produto
HI2_02390	4-hydroxy-4-methyl-2-oxoglutarate aldolase/4-carboxy-4-hydroxy-2-oxoadipate aldolase 247758:248450 forward MW:23863
HI2_02400	Hydroxypyruvate reductase 248447:249391 forward MW:33156
HI2_02410	ttuB_2 Putative tartrate transporter 249455:250777 forward MW:48134
HI2_02420	ligI 2-pyrone-4,6-dicarbaxylate hydrolase 251059:251958 forward MW:32913

Fonte: O Autor, 2018

TABELA AS 4: GENES PRESENTES NA ILHA GENÔMICA 3 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_05720	hypothetical protein 613724:614386 reverse MW:24311
HI2_05730	Thermostable monoacylglycerol lipase 614528:615391 reverse MW:32261
HI2_05731	site-specific recombinase XerD 615634:615933 forward MW:11445
HI2_05740	hpxO_1 FAD-dependent urate hydroxylase 616014:617243 reverse MW:44024
HI2_05750	dmlR_2 HTH-type transcriptional regulator DmlR 617384:618196 forward MW:30192
HI2_05760	hypothetical protein 618504:619496 forward MW:37048
HI2_05761	transposase, partial 619562:619753 forward MW:7184
HI2_05770	amtB_2 Ammonia channel 620067:621269 reverse MW:42269
HI2_05780	2-(acetamidomethylene)succinate hydrolase 621503:622297 reverse MW:28329
HI2_05790	dmlR_3 HTH-type transcriptional regulator DmlR 622411:623316 forward MW:32833
HI2_05800	btr_1 HTH-type transcriptional activator Btr 623373:624275 reverse MW:33309

Fonte: O Autor, 2018

TABELA AS 5: GENES PRESENTES NA ILHA GENÔMICA 4 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_07490	vgrG1_4 Actin cross-linking toxin VgrG1 806238:809102 forward MW:100716
HI2_07500	hypothetical protein 809121:810098 forward MW:36860
HI2_07510	hypothetical protein 810103:812922 forward MW:103637
HI2_07520	hypothetical protein 812923:813654 forward MW:27461
HI2_07530	hypothetical protein 813668:814411 forward MW:27712
HI2_07540	hypothetical protein 814477:815052 forward MW:21175

Fonte: O Autor, 2018

TABELA AS 6: GENES PRESENTES NA ILHA GENÔMICA 5 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_11760	ndx1 Diadenosine hexaphosphate hydrolase 1312479:1312919 forward MW:15904

Locus	Produto
HI2_11770	hypothetical protein 1312916:1313434 forward MW:18733
HI2_11780	gabR_2 HTH-type transcriptional regulatory protein GabR 1313480:1314907 reverse MW:52193
HI2_11790	hypothetical protein 1315087:1315824 forward MW:25406
HI2_11800	hypothetical protein 1315881:1316426 forward MW:18937

Fonte: O Autor, 2018

TABELA AS 7: GENES PRESENTES NA ILHA GENÔMICA 6 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_17490	hypothetical protein 1959750:1960871 reverse MW:43049
HI2_17500	3'-5' exoribonuclease 1961971:1962453 reverse MW:18712
HI2_17510	hypothetical protein 1962466:1962918 reverse MW:17119
HI2_17520	hypothetical protein 1963517:1964371 forward MW:31769
HI2_17530	lexA_1 LexA repressor 1964471:1965094 forward MW:22301
HI2_17540	hypothetical protein 1965105:1965884 forward MW:29441
HI2_17550	hypothetical protein 1965951:1967213 forward MW:48142
HI2_17560	hypothetical protein 1968282:1968530 reverse MW:9009
HI2_17570	hypothetical protein 1968703:1969503 forward MW:28816

Fonte: O Autor, 2018

TABELA AS 8: GENES PRESENTES NA ILHA GENÔMICA 7 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_21090	intA Prophage integrase IntA 2343818:2345014 forward MW:45677
HI2_21100	hypothetical protein 2345438:2346418 reverse MW:37309
HI2_21110	hypothetical protein 2346725:2347609 forward MW:32502
HI2_21111	hypothetical protein 2347652:2347993 forward MW:12559
HI2_21121	conjugal transfer protein TraJ 2348279:2348665 forward MW:13982
HI2_21121	conjugal transfer protein TraJ 2348279:2348665 forward MW:13982
HI2_21130	hypothetical protein 2348718:2349026 reverse MW:12119
HI2_21140	hin DNA-invertase hin 2349374:2349997 forward MW:22795
HI2_21150	hypothetical protein 2349994:2352780 forward MW:103247
HI2_21160	hypothetical protein 2352777:2354831 forward MW:75867
HI2_21170	hypothetical protein 2354824:2356008 forward MW:43613
HI2_21180	hypothetical protein 2356480:2357676 forward MW:44173
HI2_21190	hypothetical protein 2358066:2358914 reverse MW:32341

Locus	Produto
HI2_21200	rnIA mRNA endoribonuclease LS 2360132:2361205 forward MW:40105
HI2_21210	rnIB Antitoxin RnIB 2361206:2361544 forward MW:12287

Fonte: O Autor, 2018

TABELA AS 9: GENES PRESENTES NA ILHA GENÔMICA 8 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_22800	tsr_16 Methyl-accepting chemotaxis protein I 2566033:2567601 reverse MW:54909
HI2_22810	citA Putative citrate synthase 2 2567742:2568512 reverse MW:27319
HI2_22820	Long-chain-fatty-acid--CoA ligase FadD13 2568523:2570094 reverse MW:56917
HI2_22830	hdhA 7-alpha-hydroxysteroid dehydrogenase 2570171:2571019 reverse MW:29137
HI2_22840	ttuB_7 Putative tartrate transporter 2571019:2572353 reverse MW:48944
HI2_22850	kipR_1 HTH-type transcriptional regulator KipR 2572707:2573525 forward MW:29757
HI2_22860	hypothetical protein 2573803:2573940 reverse MW:5022
HI2_22861	transposase 2574207:2574458 forward MW:9562
HI2_22870	hypothetical protein 2574550:2575641 forward MW:41608
HI2_22880	hypothetical protein 2575811:2575993 reverse MW:6295
HI2_22890	ybjG Putative undecaprenyl-diphosphatase YbjG 2576276:2576878 forward MW:22189
HI2_22900	hypothetical protein 2576939:2577709 forward MW:28048
HI2_22910	hypothetical protein 2577779:2578537 forward MW:26703
HI2_22920	qseB_1 Transcriptional regulatory protein QseB 2578559:2579224 forward MW:24159
HI2_22930	qseC_2 Sensor protein QseC 2579200:2580585 forward MW:50467
HI2_22940	yceJ_2 Cytochrome b561 2580595:2581143 forward MW:20116
HI2_22950	3'-5' exoribonuclease 2581439:2581915 reverse MW:18586

Fonte: O Autor, 2018

TABELA AS 10: GENES PRESENTES NA ILHA GENÔMICA 9 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_23020	fabI_1 Enoyl-[acyl-carrier-protein] reductase [NADH] FabI 2586968:2587339 forward MW:13189
HI2_23030	hypothetical protein 2587361:2589661 reverse MW:87163
HI2_23040	hypothetical protein 2589943:2590107 reverse MW:6059
HI2_23050	hypothetical protein 2591095:2591679 forward MW:21341
HI2_23060	3'-5' exoribonuclease 2591676:2592182 forward MW:19654
HI2_23070	cbpA Cellobiose phosphorylase 2592497:2601058 reverse MW:312516
HI2_23071	hypothetical protein 2601414:2601719 reverse MW:11499

Locus	Produto
HI2_23072	Hypothetical Protein 2602221:2602418 reverse MW:7059
HI2_23073	Hypothetical Protein 2602909:2603229 reverse MW:11778
HI2_23074	hypotetical protein 2604445:2604684 reverse MW:8708
HI2_23080	hypothetical protein 2604808:2605071 forward MW:9784
HI2_23090	hypothetical protein 2605275:2605913 forward MW:24428
HI2_23100	xerC_1 Tyrosine recombinase XerC 2606175:2607179 reverse MW:36499
HI2_23110	smc_3 Chromosome partition protein Smc 2607350:2608411 forward MW:38778
HI2_23120	arsH_1 NADPH-dependent FMN reductase ArsH 2608760:2609497 reverse MW:27634
HI2_23130	arsC_2 Arsenate reductase 2609463:2609891 reverse MW:15550
HI2_23140	arsB_1 Arsenical pump membrane protein 2609915:2611195 reverse MW:45107

Fonte: O Autor, 2018

TABELA AS 11: GENES PRESENTES NA ILHA GENÔMICA 10 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_27520	hypothetical protein 3101804:3103447 forward MW:62036
HI2_27521	transcriptional regulator 3103454:3103699 reverse MW:8802
HI2_27530	pinR Serine recombinase PinR 3104850:3105434 reverse MW:21003
HI2_27540	hypothetical protein 3105608:3107278 forward MW:62418
HI2_27550	hypothetical protein 3107678:3108229 forward MW:20441
HI2_27560	hypothetical protein 3108304:3108921 reverse MW:23050
HI2_27561	hypothetical protein 3108924:3109436 reverse MW:19156
HI2_27570	hypothetical protein 3109643:3110866 reverse MW:45885
HI2_27580	Virulence protein 3110900:3111286 reverse MW:14429
HI2_27590	DNA-binding protein Bv3F 3111649:3111993 reverse MW:12755
HI2_27600	cya_3 Bifunctional hemolysin/adenylate cyclase 3112392:3114902 reverse MW:87982

Fonte: O Autor, 2018

TABELA AS 12: GENES PRESENTES NA ILHA GENÔMICA 11 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_28050	hypothetical protein 3166790:3168610 forward MW:67415
HI2_28060	hypothetical protein 3168817:3169536 forward MW:24007
HI2_28070	hypothetical protein 3169606:3170247 forward MW:22091
HI2_28080	hypothetical protein 3170244:3170966 forward MW:25364
HI2_28090	hypothetical protein 3170982:3171461 forward MW:17590

Locus	Produto
HI2_28100	bfpB Outer membrane lipoprotein BfpB 3171480:3173195 forward MW:59622
HI2_28110	hypothetical protein 3173188:3174567 forward MW:49864
HI2_28120	hypothetical protein 3174545:3175273 forward MW:25532
HI2_28130	tcpT Toxin coregulated pilus biosynthesis protein T 3175270:3176970 forward MW:61770
HI2_28140	tcpE Toxin coregulated pilus biosynthesis protein E 3176967:3178055 forward MW:40567
HI2_28150	hypothetical protein 3178055:3180355 forward MW:81653
HI2_28160	hypothetical protein 3180359:3180682 forward MW:11530
HI2_28170	hypothetical protein 3180683:3181135 forward MW:16367
HI2_28180	outO_2 Type 4 prepilin-like proteins leader peptide-processing enzyme 3181189:3182067 forward MW:31337
HI2_28190	hypothetical protein 3181976:3182275 reverse MW:11593
HI2_28210	cph2_7 Phytochrome-like protein cph2 3182628:3183782 forward MW:42427
HI2_28220	bvgS Virulence sensor protein BvgS 3183662:3186625 reverse MW:105584
HI2_28230	hypothetical protein 3187305:3188348 forward MW:38468

Fonte: O Autor, 2018

TABELA AS 13: GENES PRESENTES NA ILHA GENÔMICA 12 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_31140	hypothetical protein 3503200:3503904 reverse MW:26714
HI2_31150	soj_1 Sporulation initiation inhibitor protein Soj 3504383:3505276 reverse MW:32876
HI2_31160	hypothetical protein 3505912:3506316 forward MW:13558
HI2_31170	hypothetical protein 3506728:3507168 forward MW:16590
HI2_31180	hypothetical protein 3507168:3507818 forward MW:24484
HI2_31190	hypothetical protein 3507889:3508554 forward MW:24812
HI2_31200	hypothetical protein 3508791:3509510 forward MW:26360
HI2_31210	hypothetical protein 3509606:3510331 forward MW:26646
HI2_31220	hypothetical protein 3510984:3511193 forward MW:8006
HI2_31221	hypothetical protein 3511197:3511409 forward MW:8013
HI2_31222	AlpA family phage regulatory protein 3511513:3511707 forward MW:7181
HI2_31230	hypothetical protein 3511704:3512501 reverse MW:30380
HI2_31240	hypothetical protein 3512714:3513634 reverse MW:34370
HI2_31250	hypothetical protein 3513601:3513894 reverse MW:10914
HI2_31251	hypothetical protein 3514086:3514373 reverse MW:9626
HI2_31260	hypothetical protein 3514590:3514922 forward MW:12165
HI2_31261	hypothetical protein 3515204:3515656 reverse MW:16450
HI2_31262	hypothetical protein 3515757:3516152 reverse MW:14223

Locus	Produto
HI2_31270	hypothetical protein 3517297:3518139 forward MW:30492
HI2_31280	hypothetical protein 3519344:3520138 forward MW:28789
HI2_31290	hypothetical protein 3520548:3521702 forward MW:43257
HI2_31300	hypothetical protein 3521728:3522081 reverse MW:12865
HI2_31310	oatA_5 O-acetyltransferase OatA 3522467:3523582 forward MW:42251
HI2_31320	hypothetical protein 3523642:3524286 reverse MW:23643
HI2_31330	hypothetical protein 3524354:3525007 reverse MW:21653
HI2_31340	hypothetical protein 3525038:3525652 reverse MW:22214
HI2_31350	sppA_2 Putative signal peptide peptidase SppA 3525673:3526569 reverse MW:31294
HI2_31360	hypothetical protein 3526566:3528212 reverse MW:61143
HI2_31370	hypothetical protein 3528212:3528475 reverse MW:9795
HI2_31380	hypothetical protein 3528478:3528702 reverse MW:8016
HI2_31390	hypothetical protein 3528711:3530828 reverse MW:78825
HI2_31400	hypothetical protein 3530767:3531381 reverse MW:22400
HI2_31410	hypothetical protein 3531423:3532037 reverse MW:23043
HI2_31420	hypothetical protein 3532045:3532707 reverse MW:23827
HI2_31421	hypothetical protein 3532883:3533182 reverse MW:11071
HI2_31440	hypothetical protein 3534461:3535015 reverse MW:21141
HI2_31450	hypothetical protein 3535012:3535308 reverse MW:10932
HI2_31460	hypothetical protein 3535311:3536249 reverse MW:35484
HI2_31470	hypothetical protein 3536240:3536713 reverse MW:17505
HI2_31480	hypothetical protein 3536710:3537600 reverse MW:31861
HI2_31490	hypothetical protein 3538538:3538759 forward MW:7977
HI2_31500	hypothetical protein 3538883:3539089 forward MW:7799
HI2_31510	hypothetical protein 3539074:3539343 reverse MW:10116
HI2_31511	hypothetical protein 3539337:3539582 reverse MW:9287
HI2_31520	hypothetical protein 3539678:3540406 forward MW:27304
HI2_31530	hypothetical protein 3541160:3541639 reverse MW:18314
HI2_31540	hypothetical protein 3541760:3541948 forward MW:7186
HI2_31550	hypothetical protein 3542022:3542267 forward MW:8631
HI2_31560	hypothetical protein 3542264:3542959 forward MW:25504
HI2_31561	hypothetical protein 3542956:3543252 forward MW:11352
HI2_31570	hypothetical protein 3543249:3543458 forward MW:7719
HI2_31580	hypothetical protein 3543455:3543703 forward MW:9295
HI2_31590	hypothetical protein 3543700:3544377 forward MW:24755
HI2_31600	hypothetical protein 3544374:3544568 forward MW:7234
HI2_31610	hypothetical protein 3544720:3546228 forward MW:55656

Locus	Produto
HI2_31620	hypothetical protein 3547522:3547716 reverse MW:7443
HI2_31630	hypothetical protein 3547729:3548502 reverse MW:28915
HI2_31640	pcrA ATP-dependent DNA helicase PcrA 3548615:3550447 reverse MW:68940
HI2_31650	recF_1 DNA replication and repair protein RecF 3550447:3552300 reverse MW:68505
HI2_31660	hypothetical protein 3553007:3553438 forward MW:16844
HI2_31670	hypothetical protein 3553533:3554957 forward MW:54707
HI2_31680	traD Coupling protein TraD 3555037:3556920 reverse MW:71073
HI2_31690	hypothetical protein 3557114:3557707 forward MW:23261
HI2_31691	hypothetical protein 3557742:3558173 reverse MW:17174
HI2_31700	xerC_2 Tyrosine recombinase XerC 3558613:3559731 forward MW:42446
HI2_31710	tral Multifunctional conjugation protein Tral 3559834:3562758 forward MW:109165
HI2_31720	hypothetical protein 3562770:3563111 forward MW:12656
HI2_31730	hypothetical protein 3563200:3564756 forward MW:59308
HI2_31740	hypothetical protein 3564846:3566057 reverse MW:45470
HI2_31750	hypothetical protein 3566449:3567066 forward MW:23045
HI2_31760	bvgA Virulence factors putative positive transcription regulator BvgA 3567081:3567740 forward MW:24532
HI2_31770	hypothetical protein 3567820:3568782 forward MW:35608
HI2_31780	hypothetical protein 3568816:3569511 reverse MW:26438

Fonte: O Autor, 2018

TABELA AS 14: GENES PRESENTES NA ILHA GENÔMICA 13 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_34150	hypothetical protein 3852106:3852450 reverse MW:13451
HI2_34160	hypothetical protein 3852986:3853825 forward MW:30667
HI2_34170	hypothetical protein 3853874:3854317 forward MW:15670
HI2_34180	hypothetical protein 3854372:3854626 forward MW:9095

Fonte: O Autor, 2018

TABELA AS 15: GENES PRESENTES NA ILHA GENÔMICA 14 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_38080	hypothetical protein 4284902:4285267 reverse MW:13769
HI2_38081	hypothetical protein 4285501:4285773 reverse MW:10487
HI2_38090	hypothetical protein 4285927:4286592 reverse MW:25131
HI2_38091	hypothetical protein 4286623:4287072 reverse MW:17037

Locus	Produto
HI2_38100	hypothetical protein 4287083:4287211 reverse MW:5058
HI2_38110	hypothetical protein 4287408:4287614 reverse MW:7676
HI2_38111	hypothetical protein 4287625:4288161 reverse MW:19262
HI2_38120	hypothetical protein 4288602:4289042 reverse MW:16171
HI2_38130	hypothetical protein 4289055:4289954 reverse MW:31504

Fonte: O Autor, 2018

TABELA AS 16: GENES PRESENTES NA ILHA GENÔMICA 15 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_39060	hypothetical protein 4396095:4396244 reverse MW:5103
HI2_39070	shlB_4 Hemolysin transporter protein ShlB 4396453:4397934 forward MW:54984
HI2_39080	hypothetical protein 4397960:4398070 forward MW:4006
HI2_39090	cdiA2_3 tRNA nuclease CdiA-2 4398375:4399655 forward MW:46530
HI2_39100	hypothetical protein 4399652:4400290 forward MW:23987
HI2_39110	hypothetical protein 4400300:4400641 forward MW:12481

Fonte: O Autor, 2018

TABELA AS 17: GENES PRESENTES NA ILHA GENÔMICA 16 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_40320	rapA_1 RNA polymerase-associated protein RapA 4529169:4532399 forward MW:121437
HI2_40330	hypothetical protein 4532396:4533181 forward MW:29353
HI2_40340	hypothetical protein 4533191:4535314 forward MW:81674
HI2_40350	hypothetical protein 4535354:4537462 forward MW:79657
HI2_40360	hypothetical protein 4537473:4538498 forward MW:39168
HI2_40370	hypothetical protein 4538498:4538938 forward MW:16554

Fonte: O Autor, 2018

TABELA AS 18: GENES PRESENTES NA ILHA GENÔMICA 17 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_43450	hypothetical protein 4875924:4876271 reverse MW:12378
HI2_43460	hypothetical protein 4876497:4877045 reverse MW:20046
HI2_43470	hypothetical protein 4878087:4879436 forward MW:49629
HI2_43480	hypothetical protein 4879456:4880991 forward MW:59109
HI2_43490	hypothetical protein 4881232:4882866 forward MW:63076

Locus	Produto
HI2_43500	hypothetical protein 4882921:4883943 forward MW:38747
HI2_43510	hypothetical protein 4884055:4886316 forward MW:84874
HI2_43520	cynR_7 HTH-type transcriptional regulator CynR 4886789:4887130 reverse MW:12741
HI2_43530	phnO Aminoalkylphosphonate N-acetyltransferase 4887200:4887670 reverse MW:16775
HI2_43540	lexA_3 LexA repressor 4887758:4888231 forward MW:17532
HI2_43550	dinB1 DNA polymerase IV 1 4888222:4889520 forward MW:48322
HI2_43560	hypothetical protein 4889523:4889882 forward MW:13336
HI2_43570	hypothetical protein 4889983:4890858 reverse MW:33081
HI2_43580	hypothetical protein 4891036:4891680 reverse MW:24266
HI2_43590	hypothetical protein 4891796:4892518 forward MW:27309
HI2_43600	hypothetical protein 4892598:4893059 forward MW:17701
HI2_43610	hypothetical protein 4893128:4894369 reverse MW:45830
HI2_43620	hypothetical protein 4894516:4899789 reverse MW:196665
HI2_43630	hypothetical protein 4899934:4901118 reverse MW:45227
HI2_43640	hypothetical protein 4901422:4902384 reverse MW:34992
HI2_43650	hypothetical protein 4902890:4903249 forward MW:13783
HI2_43651	Hypothetical Protein 4903588:4904121 forward MW:19810
HI2_43652	Hypothetical Protein 4904167:4904658 reverse MW:17255
HI2_43653	hypothetical protein 4904691:4905302 reverse MW:22132
HI2_43660	hypothetical protein 4905308:4905631 reverse MW:11941
HI2_43670	hypothetical protein 4906077:4907174 forward MW:42555
HI2_43680	dacA_2 D-alanyl-D-alanine carboxypeptidase DacA 4907998:4908879 reverse MW:31929
HI2_43690	Aspartate racemase 4908884:4910350 reverse MW:52690
HI2_43700	gltP Proton/glutamate-aspartate symporter 4910613:4911749 reverse MW:40329
HI2_43710	aspA_2 Aspartate ammonia-lyase 4911790:4913217 reverse MW:50802

Fonte: O Autor, 2018

TABELA AS 19: GENES PRESENTES NA ILHA GENÔMICA 18 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_41250	algA Alginate biosynthesis protein AlgA 4631138:4632238 forward MW:40490
HI2_41260	gmd GDP-mannose 4,6-dehydratase 4632245:4633351 forward MW:41742
HI2_41270	colC GDP-L-colitose synthase 4633357:4634295 forward MW:34925
HI2_41280	colD_1 GDP-4-keto-6-deoxy-D-mannose-3-dehydratase / pyridoxamine-phosphate transaminase 4634285:4635451 forward MW:43160
HI2_41290	rfbF Glucose-1-phosphate cytidyltransferase 4635455:4636228 forward MW:28994
HI2_41300	rfbG_1 CDP-glucose 4,6-dehydratase 4636396:4637316 forward MW:34084
HI2_41310	colD_2 GDP-4-keto-6-deoxy-D-mannose-3-dehydratase / pyridoxamine-phosphate transaminase 4637332:4638654 forward MW:48506

Locus	Produto
HI2_41320	rmd GDP-6-deoxy-D-mannose reductase 4638695:4639744 forward MW:37777
HI2_41330	fas2_2 Ferredoxin fas2 4639741:4640547 forward MW:28992
HI2_41340	dxs_4 1-deoxy-D-xylulose-5-phosphate synthase 4640544:4641488 forward MW:33849
HI2_41350	hypothetical protein 4641581:4642957 forward MW:50165
HI2_41360	ykvP Spore protein YkvP 4642954:4643952 forward MW:38735
HI2_41370	epsE_2 Putative glycosyltransferase EpsE 4643945:4644976 forward MW:38342
HI2_41380	fucT Alpha-(1,3)-fucosyltransferase FucT 4644994:4646046 forward MW:40295
HI2_41390	hypothetical protein 4646043:4647461 forward MW:52163

Fonte: O Autor, 2018

TABELA AS 20: GENES PRESENTES NA ILHA GENÔMICA 19 DE H. LINEAGE2 AU13965

Locus	Produto
HI2_08800	groL 60 kDa chaperonin 973596:975245 reverse MW:57557
HI2_08810	groS 10 kDa chaperonin 975369:975659 reverse MW:10539
HI2_08820	hypothetical protein 976052:976798 forward MW:27282
HI2_08830	hypothetical protein 976815:977501 forward MW:25089
HI2_08840	hypothetical protein 977516:979636 reverse MW:79739
HI2_08850	hypothetical protein 979646:981046 reverse MW:51802
HI2_08860	chiA1 Chitinase A1 981043:986262 reverse MW:175247
HI2_08870	esiB_1 Secretory immunoglobulin A-binding protein EsiB 986335:987276 reverse MW:33258
HI2_08880	hypothetical protein 987371:988294 reverse MW:30492
HI2_08890	mepA Multidrug/solvent efflux pump periplasmic linker protein MepA 988336:989166 reverse MW:29636
HI2_08900	hypothetical protein 989177:990718 reverse MW:56850

Fonte: O Autor, 2018

TABELA AS 21: GENES DE VIRULÊNCIA DA ESTIRPE H. LINHAGEM 2 AU13965

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
HI2_02870	Cj1427c	sugar-nucleotide epimerase/dehydratease [Capsule (VF0323)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.70
HI2_03290	lpxC	UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	54.30
HI2_03410	pilD	type IV prepilin leader peptide type M1 [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	51.89
HI2_03410	xcpA/pilD	type 4 prepilin peptidase PilD [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	50.71

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
HI2_03430	pilF	type IV pilus assembly ATPase protein [Type IV pili (VF0075)]	Neisseria meningitidis MC58	57.69
HI2_03430	pilB	type 4 fimbrial biogenesis protein PilB [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	56.15
HI2_03430	pilB	type IV pilus assembly protein [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	57.18
HI2_03430	pilB	type IV pilus assembly protein PilB [Type IV pili (VF0155)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	50.90
HI2_04710	algC	phosphomannomutase AlgC [Alginate biosynthesis (CVF522)]	Pseudomonas aeruginosa PAO1	54.70
HI2_04790	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Neisseria meningitidis MC58	55.01
HI2_07480	clpV1	type VI secretion system AAA+ family ATPase [HSI-I (VF0334)]	Pseudomonas aeruginosa PAO1	50.44
HI2_08060	lpg2359	Dot/Icm type IV secretion system effector [Dot/Icm (VF0156)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	54.42
HI2_08800	htpB	Hsp60, 60K heat shock protein HtpB [Hsp60 (VF0159)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	72.41
HI2_10520	cyaB	cyclolysin secretion ATP-binding protein [Cya (VF0028)]	Bordetella pertussis Tohama I	66.00
HI2_10520	hlyB	Hemolysin B [Hemolysin (VF0225)]	Escherichia coli CFT073	55.59
HI2_10520	hlyB	hemolysin transport protein [Hemolysin (VF0207)]	Escherichia coli O157:H7 str. EDL933	55.71
HI2_10600	sodB	superoxide dismutase [SodB (VF0169)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	65.45
HI2_11430	katA	catalase/(hydro)peroxidase [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	60.38
HI2_11430	katB	catalase-peroxidase KatB [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	56.12
HI2_14500	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	61.72
HI2_14500	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.16
HI2_15800	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.84
HI2_15800	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	57.14
HI2_18290	acpXL	acyl carrier protein [LPS (CVF383)]	Brucella melitensis bv. 1 str. 16M	61.11
HI2_18320	algU	alginate biosynthesis protein AlgZ/FimS [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	64.36
HI2_18740	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.65
HI2_18740	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.82
HI2_19000	manC	GDP-mannose pyrophosphorylase [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	58.14
HI2_19000	algA	phosphomannose isomerase / guanosine 5'-diphospho-D-mannose pyrophosphorylase [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	55.65

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
HI2_19000	manC	mannose-1-phosphate guanylyltransferase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	55.72
HI2_19090	galE	UDP-glucose 4-epimerase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	63.00
HI2_19090	galE	UDP-glucose 4-epimerase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.79
HI2_19230	cheR	chemotaxis protein methyltransferase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.86
HI2_19230	cheR	chemotaxis methyltransferase CheR [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.14
HI2_19240	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.76
HI2_19250	cheB	chemotaxis-specific methylesterase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.75
HI2_19250	cheB	chemotaxis-specific methylesterase CheB [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	65.55
HI2_19260	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	78.91
HI2_19260	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	77.60
HI2_19280	flhB	flagellar biosynthesis protein FlhB [Flagella (VF0430)]	Burkholderia pseudomallei K96243	52.14
HI2_19290	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	69.42
HI2_19290	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.83
HI2_19290	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	55.98
HI2_19290	flhA	flagellar biosynthetic protein FlhA [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	54.82
HI2_19320	fliA	flagellar biosynthesis sigma factor [Flagella (VF0430)]	Burkholderia pseudomallei K96243	58.85
HI2_19320	fliA	flagellar biosynthesis sigma factor [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	50.42
HI2_19370	flgC	flagellar basal body rod protein FlgC [Flagella (VF0430)]	Burkholderia pseudomallei K96243	65.87
HI2_19370	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	61.16
HI2_19370	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	51.13
HI2_19400	flgF	flagellar basal body rod protein FlgF [Flagella (VF0430)]	Burkholderia pseudomallei K96243	52.17
HI2_19410	flgG	flagellar basal body rod protein FlgG [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.92
HI2_19410	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	71.43
HI2_19410	flgG	flagellar basal body rod protein FlgG [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	59.46
HI2_19410	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	59.85

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
HI2_19410	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.26
HI2_19410	flgG	flagellar basal-body rod protein (flgG) [Pse5Ac7Ac (AI150)]	Helicobacter pylori 26695	51.54
HI2_19430	flgl	flagellar P-ring protein precursor Flgl [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.11
HI2_19430	flgl	flagellar P-ring protein precursor Flgl [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.81
HI2_19430	flgl	flagellar P-ring protein precursor Flgl [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	56.03
HI2_19510	fliQ	flagellar biosynthetic protein FliQ [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.95
HI2_19510	fliQ	flagellar biosynthesis protein FliQ [Flagella (VF0430)]	Burkholderia pseudomallei K96243	64.10
HI2_19520	fliP	flagellar biosynthesis protein FliP [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.89
HI2_19520	fliP	flagellar biosynthetic protein FliP [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	69.91
HI2_19520	fliP	flagellar biosynthesis protein FliP [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	51.54
HI2_19540	fliN	flagellar motor switch protein FliN [Flagella (VF0430)]	Burkholderia pseudomallei K96243	69.34
HI2_19540	fliN	flagellar motor switch protein FliN [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.14
HI2_19550	fliM	flagellar motor switch protein FliM [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.55
HI2_19590	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0430)]	Burkholderia pseudomallei K96243	67.82
HI2_19590	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	67.81
HI2_19590	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	62.92
HI2_19590	fliI	flagellum-specific ATP synthase FliI [polar flagella (AI149)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	57.39
HI2_19610	fliG	flagellar motor switch protein G [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.14
HI2_19610	fliG	flagellar motor switch protein G [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.22
HI2_19680	fliS	flagellar protein FliS [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.69
HI2_20140	icl	Isocitrate lyase Icl (isocitrase) (isocitratase) [Isocitrate lyase (VF0253)]	Mycobacterium tuberculosis H37Rv	65.79
HI2_20330	orfM	deoxyribonucleotide triphosphate pyrophosphatase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	50.52
HI2_20370	sodCI	Gifsy-2 prophage: superoxide dismutase precursor (Cu-Zn) [SodCI (VF0109)]	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	55.11

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
HI2_20890	lpxA	UDP-N-acetylglucosamine acyltransferase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	52.87
HI2_21610	msrA/BpilB	trifunctional thioredoxin/methionine sulfoxide reductase A/B protein [MsrAB (VF0456)]	Neisseria meningitidis MC58	50.81
HI2_22630	mbtH-like	MbtH-like protein from the pyoverdine cluster [pyoverdine (IA001)]	Pseudomonas aeruginosa PAO1	63.38
HI2_22670	pvdL	peptide synthase PvdL [pyoverdine (IA001)]	Pseudomonas aeruginosa PAO1	51.74
HI2_22740	pvdA	L-ornithine N5-oxygenase PvdA [Pyoverdine (VF0094)]	Pseudomonas aeruginosa PAO1	50.12
HI2_25160	clpP	ATP-dependent Clp protease proteolytic subunit [ClpP (VF0074)]	Listeria monocytogenes EGD-e	65.28
HI2_26480	farA	fatty acid efflux system protein FarA [FarAB (VF0450)]	Neisseria meningitidis MC58	50.43
HI2_26490	farB	fatty acid efflux system protein FarB [FarAB (VF0450)]	Neisseria meningitidis MC58	51.29
HI2_26810	tviB	Vi polysaccharide biosynthesis protein, UDP-glucose/GDP-mannose dehydrogenase TviB [Vi antigen (VF0101)]	Salmonella enterica subsp. enterica serovar Typhi str. CT18	52.59
HI2_26820	tviC	Vi polysaccharide biosynthesis protein, epimerase TviC [Vi antigen (VF0101)]	Salmonella enterica subsp. enterica serovar Typhi str. CT18	61.06
HI2_27180	kdsA	2-dehydro-3-deoxyphosphooctonate aldolase [LPS (CVF383)]	Brucella melitensis bv. 1 str. 16M	59.92
HI2_28720	cheW	chemotaxis protein CheW [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.79
HI2_28720	cheW	purine-binding chemotaxis protein CheW [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	64.67
HI2_28730	cheA	chemotaxis protein CheA [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	50.41
HI2_28740	cheY1	chemotaxis two-component response regulator CheY1 [Flagella (VF0430)]	Burkholderia pseudomallei K96243	50.88
HI2_28760	motA	flagellar motor protein MotA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	67.13
HI2_28760	motA	flagellar motor protein MotA [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	64.56
HI2_28760	motA	flagellar motor protein [Deoxyhexose linking sugar, 209 Da capping structure (A1138)]	Pseudomonas aeruginosa PAO1	51.25
HI2_28780	flhC	flagellar biosynthesis transcription activator FlhC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.56
HI2_29250	tsr	methyl-accepting chemotaxis protein I [Flagella (VF0430)]	Burkholderia pseudomallei K96243	66.09
HI2_29250	cheD	methyl-accepting chemotaxis protein CheD [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	62.61
HI2_30630	mntB	Manganese transport system membrane protein MntB [MntABC (VF0455)]	Neisseria meningitidis MC58	52.23
HI2_33400	mgtB	Mg ²⁺ transport protein [MgtBC (VF0106)]	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	55.64
HI2_36290	rfaE	ADP-heptose synthase [LOS (VF0044)]	Haemophilus influenzae Rd KW20	51.16

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
HI2_37530	pilT	twitching motility protein PilT [Type IV pili (VF0075)]	Neisseria meningitidis MC58	54.30
HI2_37530	pilT	type IV pilus twitching motility protein PilT [type IV pili (A1111)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	52.40
HI2_37530	pilT	twitching motility protein PilT [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	50.45
HI2_39170	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Neisseria meningitidis MC58	51.54
HI2_41250	algA	phosphomannose isomerase / guanosine 5'-diphospho-D-mannose pyrophosphorylase [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	52.57
HI2_41260	gmd	GDP-mannose 4,6-dehydratase [LPS (VF0367)]	Brucella melitensis bv. 1 str. 16M	79.15
HI2_41260	gmd	GDP-mannose 4,6-dehydratase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	70.44
HI2_41260	rfbD	GDP-D-mannose dehydratase [LPS (VF0056)]	Helicobacter pylori 26695	58.60
HI2_41290	ddhA	glucose-1-phosphate cytidyltransferase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	70.70
HI2_41300	ddhB	CDP-glucose 4,6-dehydratase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.73
HI2_41430	cap8E	capsular polysaccharide synthesis enzyme Cap8E [Capsule (VF0003)]	Staphylococcus aureus subsp. aureus MW2	66.57
HI2_41430	cps4J	capsular polysaccharide biosynthesis protein Cps4J [Capsule (VF0144)]	Streptococcus pneumoniae TIGR4	58.04
HI2_41440	cap8G	capsular polysaccharide synthesis enzyme Cap8G [Capsule (VF0003)]	Staphylococcus aureus subsp. aureus MW2	52.01
HI2_41500	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Burkholderia pseudomallei K96243	55.29
HI2_43260	galE	UDP-glucose 4-epimerase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	57.86
HI2_43260	galE	UDP-glucose 4-epimerase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	55.49
HI2_43270	rffG	dTDP-glucose 4-epimerase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	59.09
HI2_46720	ureA	urease alpha subunit UreA [Urease (VF0050)]	Helicobacter pylori 26695	63.00
HI2_46730	ureA	urease alpha subunit UreA [Urease (VF0050)]	Helicobacter pylori 26695	55.67
HI2_46740	ureB	urease beta subunit UreB, urea amidohydrolase [Urease (VF0050)]	Helicobacter pylori 26695	64.50
HI2_46780	ureG	urease accessory protein (ureG) [Urease (VF0050)]	Helicobacter pylori 26695	64.65

Fonte: O Autor, 2018

TABELA AS 22: GENES DE VIRULÊNCIA DA ESTIRPE *H. seropedicae* SmR1

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hsero_0305	Cj1427c	sugar-nucleotide epimerase/dehydratase [Capsule (VF0323)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.70
Hsero_0343	lpxC	UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	54.64
Hsero_0355	xcpA/pilD	type 4 prepilin peptidase PilD [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	51.77
Hsero_0355	pilD	type IV prepilin leader peptide type M1 [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	51.52
Hsero_0357	pilF	type IV pilus assembly ATPase protein [Type IV pili (VF0075)]	Neisseria meningitidis MC58	57.69
Hsero_0357	pilB	type 4 fimbrial biogenesis protein PilB [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	56.15
Hsero_0357	pilB	type IV pilus assembly protein [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	57.18
Hsero_0357	pilB	type IV pilus assembly protein PilB [Type IV pili (VF0155)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	50.90
Hsero_0472	algC	phosphomannomutase AlgC [Alginate biosynthesis (CVF522)]	Pseudomonas aeruginosa PAO1	54.70
Hsero_0483	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Neisseria meningitidis MC58	55.10
Hsero_0747	clpV1	type VI secretion system AAA+ family ATPase [HSI-I (VF0334)]	Pseudomonas aeruginosa PAO1	50.22
Hsero_0790	vscN	type III secretion system ATPase VscN [T3SS1 (VF0408)]	Vibrio parahaemolyticus RIMD 2210633	52.26
Hsero_0790	pscN	type III secretion system ATPase PscN [TTSS (VF0083)]	Pseudomonas aeruginosa PAO1	52.97
Hsero_0790	yscN	type III secretion system ATPase YscN [TTSS (VF0140)]	Yersinia pestis CO92	51.90
Hsero_0790	bscN	ATP synthase in type III secretion system [TTSS (VF0035)]	Bordetella pertussis Tohama I	52.13
Hsero_0810	pilQ	type IV pilus biosynthesis protein [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	60.56
Hsero_0811	pilR	type IV pilus biosynthesis protein PilR [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	61.76
Hsero_0816	pilT	twitching motility protein [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	62.15
Hsero_0867	lpg2359	Dot/Icm type IV secretion system effector [Dot/Icm (VF0156)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	54.42
Hsero_0996	htpB	Hsp60, 60K heat shock protein HtpB [Hsp60 (VF0159)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	72.41
Hsero_1165	sodB	superoxide dismutase [SodB (VF0169)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	64.92
Hsero_1242	katA	catalase/(hydro)peroxidase [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	60.92
Hsero_1242	katB	catalase-peroxidase KatB [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	56.25
Hsero_1555	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	61.72
Hsero_1555	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.16

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hsero_1697	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.84
Hsero_1697	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	57.14
Hsero_1926	acpXL	acyl carrier protein [LPS (CVF383)]	Brucella melitensis bv. 1 str. 16M	61.11
Hsero_1929	algU	alginate biosynthesis protein AlgZ/FimS [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	64.36
Hsero_1976	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	60.94
Hsero_1976	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.98
Hsero_2001	manC	GDP-mannose pyrophosphorylase [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	58.35
Hsero_2001	algA	phosphomannose isomerase / guanosine 5'-diphospho-D-mannose pyrophosphorylase [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	55.44
Hsero_2001	manC	mannose-1-phosphate guanylyltransferase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	55.51
Hsero_2010	galE	UDP-glucose 4-epimerase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	63.00
Hsero_2010	galE	UDP-glucose 4-epimerase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.79
Hsero_2020	cheR	chemotaxis protein methyltransferase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.86
Hsero_2020	cheR	chemotaxis methyltransferase CheR [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.53
Hsero_2021	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.76
Hsero_2022	cheB	chemotaxis-specific methylesterase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	72.02
Hsero_2022	cheB	chemotaxis-specific methylesterase CheB [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	65.36
Hsero_2023	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	78.91
Hsero_2023	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	77.60
Hsero_2025	flhB	flagellar biosynthesis protein FlhB [Flagella (VF0430)]	Burkholderia pseudomallei K96243	52.14
Hsero_2026	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	69.57
Hsero_2026	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.97
Hsero_2026	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	56.12
Hsero_2026	flhA	flagellar biosynthetic protein FlhA [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	54.68
Hsero_2029	fliA	flagellar biosynthesis sigma factor [Flagella (VF0430)]	Burkholderia pseudomallei K96243	58.85
Hsero_2029	fliA	flagellar biosynthesis sigma factor [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	50.42

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hsero_2034	flgC	flagellar basal body rod protein FlgC [Flagella (VF0430)]	Burkholderia pseudomallei K96243	66.67
Hsero_2034	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	61.94
Hsero_2034	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	52.74
Hsero_2037	flgF	flagellar basal body rod protein FlgF [Flagella (VF0430)]	Burkholderia pseudomallei K96243	52.17
Hsero_2038	flgG	flagellar basal body rod protein FlgG [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.92
Hsero_2038	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	71.43
Hsero_2038	flgG	flagellar basal body rod protein FlgG [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	59.46
Hsero_2038	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	59.85
Hsero_2038	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.26
Hsero_2038	flgG	flagellar basal-body rod protein (flgG) [Pse5Ac7Ac (A1150)]	Helicobacter pylori 26695	51.54
Hsero_2040	flgI	flagellar P-ring protein precursor FlgI [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.11
Hsero_2040	flgI	flagellar P-ring protein precursor FlgI [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.81
Hsero_2040	flgI	flagellar P-ring protein precursor FlgI [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	56.03
Hsero_2048	fliQ	flagellar biosynthetic protein FliQ [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.95
Hsero_2048	fliQ	flagellar biosynthesis protein FliQ [Flagella (VF0430)]	Burkholderia pseudomallei K96243	64.10
Hsero_2049	fliP	flagellar biosynthesis protein FliP [Flagella (VF0430)]	Burkholderia pseudomallei K96243	72.06
Hsero_2049	fliP	flagellar biosynthetic protein FliP [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	69.91
Hsero_2049	fliP	flagellar biosynthetic protein FliP [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	64.29
Hsero_2049	fliP	flagellar biosynthetic protein FliP [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	53.18
Hsero_2049	fliP	flagellar biosynthesis protein FliP [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	51.54
Hsero_2051	fliN	flagellar motor switch protein FliN [Flagella (VF0430)]	Burkholderia pseudomallei K96243	69.34
Hsero_2051	fliN	flagellar motor switch protein FliN [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.14
Hsero_2052	fliM	flagellar motor switch protein FliM [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.55
Hsero_2056	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0430)]	Burkholderia pseudomallei K96243	68.03
Hsero_2056	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	68.04

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hsero_2056	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	61.54
Hsero_2056	fliI	flagellum-specific ATP synthase FliI [polar flagella (AI149)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	57.39
Hsero_2058	fliG	flagellar motor switch protein G [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.14
Hsero_2058	fliG	flagellar motor switch protein G [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.22
Hsero_2066	fliS	flagellar protein FliS [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.69
Hsero_2114	icl	Isocitrate lyase Icl (isocitrase) (isocitratase) [Isocitrate lyase (VF0253)]	Mycobacterium tuberculosis H37Rv	65.79
Hsero_2134	orfM	deoxyribonucleotide triphosphate pyrophosphatase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	50.52
Hsero_2138	sodCI	Gifsy-2 prophage: superoxide dismutase precursor (Cu-Zn) [SodCI (VF0109)]	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	55.11
Hsero_2190	lpxA	UDP-N-acetylglucosamine acyltransferase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	52.87
Hsero_2246	msrA/BpilB	trifunctional thioredoxin/methionine sulfoxide reductase A/B protein [MsrAB (VF0456)]	Neisseria meningitidis MC58	50.81
Hsero_2339	mbtH-like	MbtH-like protein from the pyoverdine cluster [pyoverdine (IA001)]	Pseudomonas aeruginosa PAO1	64.79
Hsero_2346	pvdA	L-ornithine N5-oxygenase PvdA [Pyoverdine (VF0094)]	Pseudomonas aeruginosa PAO1	50.12
Hsero_2582	clpP	ATP-dependent Clp protease proteolytic subunit [ClpP (VF0074)]	Listeria monocytogenes EGD-e	65.28
Hsero_2704	farA	fatty acid efflux system protein FarA [FarAB (VF0450)]	Neisseria meningitidis MC58	50.43
Hsero_2705	farB	fatty acid efflux system protein FarB [FarAB (VF0450)]	Neisseria meningitidis MC58	51.49
Hsero_2736	tviB	Vi polysaccharide biosynthesis protein, UDP-glucose/GDP-mannose dehydrogenase TviB [Vi antigen (VF0101)]	Salmonella enterica subsp. enterica serovar Typhi str. CT18	52.84
Hsero_2737	tviC	Vi polysaccharide biosynthesis protein, epimerase TviC [Vi antigen (VF0101)]	Salmonella enterica subsp. enterica serovar Typhi str. CT18	60.77
Hsero_2771	kdsA	2-dehydro-3-deoxyphosphooctonate aldolase [LPS (CVF383)]	Brucella melitensis bv. 1 str. 16M	59.92
Hsero_2881	cyaB	cyclolysin secretion ATP-binding protein [Cya (VF0028)]	Bordetella pertussis Tohama I	63.17
Hsero_2881	hlyB	Hemolysin B [Hemolysin (VF0225)]	Escherichia coli CFT073	53.82
Hsero_2881	hlyB	hemolysin transport protein [Hemolysin (VF0207)]	Escherichia coli O157:H7 str. EDL933	53.55
Hsero_2981	cheW	chemotaxis protein CheW [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.79
Hsero_2981	cheW	purine-binding chemotaxis protein CheW [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	64.67
Hsero_2982	cheA	chemotaxis protein CheA [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	50.87
Hsero_2983	cheY1	chemotaxis two-component response regulator CheY1 [Flagella	Burkholderia pseudomallei K96243	50.88

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hsero_2985	motB	(VF0430) flagellar motor protein MotB [Flagella (VF0430)]	Burkholderia pseudomallei K96243	56.81
Hsero_2985	motB	flagellar motor protein MotB [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	53.71
Hsero_2986	motA	flagellar motor protein MotA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	67.13
Hsero_2986	motA	flagellar motor protein MotA [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	64.56
Hsero_2986	motA	flagellar motor protein [Deoxyhexose linking sugar, 209 Da capping structure (A1138)]	Pseudomonas aeruginosa PAO1	51.25
Hsero_2988	flhC	flagellar biosynthesis transcription activator FlhC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.56
Hsero_3165	mntB	Manganese transport system membrane protein MntB [MntABC (VF0455)]	Neisseria meningitidis MC58	52.58
Hsero_3166	mntA	ABC transporter ATP-binding protein MntA [MntABC (VF0455)]	Neisseria meningitidis MC58	52.42
Hsero_3372	mgtB	Mg ²⁺ transport protein [MgtBC (VF0106)]	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	55.87
Hsero_3684	rfaE	ADP-heptose synthase [LOS (VF0044)]	Haemophilus influenzae Rd KW20	51.16
Hsero_3804	pilT	twitching motility protein PilT [Type IV pili (VF0075)]	Neisseria meningitidis MC58	54.01
Hsero_3804	pilT	type IV pilus twitching motility protein PilT [type IV pili (A1111)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	52.10
Hsero_3804	pilT	twitching motility protein PilT [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	50.15
Hsero_3957	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Neisseria meningitidis MC58	51.64
Hsero_4036	hsiC1/vipB	type VI secretion system tubule-forming protein VipB [HSI-I (VF0334)]	Pseudomonas aeruginosa PAO1	51.33
Hsero_4052	algW	AlgW protein [Alginate regulation (CVF523)]	Pseudomonas aeruginosa PAO1	58.20
Hsero_4193	tviB	Vi polysaccharide biosynthesis protein, UDP-glucose/GDP-mannose dehydrogenase TviB [Vi antigen (VF0101)]	Salmonella enterica subsp. enterica serovar Typhi str. CT18	69.45
Hsero_4194	tviC	Vi polysaccharide biosynthesis protein, epimerase TviC [Vi antigen (VF0101)]	Salmonella enterica subsp. enterica serovar Typhi str. CT18	63.58
Hsero_4195	bplA	probable oxidoreductase [LPS (VF0033)]	Bordetella pertussis Tohama I	64.77
Hsero_4196	bplB	probable acetyltransferase [LPS (VF0033)]	Bordetella pertussis Tohama I	62.84
Hsero_4197	bplC	lipopolysaccharide biosynthesis protein [LPS (VF0033)]	Bordetella pertussis Tohama I	74.38
Hsero_4219	cap8E	capsular polysaccharide synthesis enzyme Cap8E [Capsule (VF0003)]	Staphylococcus aureus subsp. aureus MW2	64.18
Hsero_4219	cps4J	capsular polysaccharide biosynthesis protein Cps4J [Capsule	Streptococcus pneumoniae TIGR4	58.93

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hsero_4220	cap8G	(VF0144)] capsular polysaccharide synthesis enzyme Cap8G [Capsule (VF0003)]	Staphylococcus aureus subsp. aureus MW2	51.47
Hsero_4227	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Burkholderia pseudomallei K96243	55.29
Hsero_4409	galE	UDP-glucose 4-epimerase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	57.86
Hsero_4409	galE	UDP-glucose 4-epimerase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	55.49
Hsero_4410	rffG	dTDP-glucose 46-dehydratase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	59.09
Hsero_4441	htpB	Hsp60, 60K heat shock protein HtpB [Hsp60 (VF0159)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	70.40
Hsero_4719	ureA	urease alpha subunit UreA [Urease (VF0050)]	Helicobacter pylori 26695	63.00
Hsero_4720	ureA	urease alpha subunit UreA [Urease (VF0050)]	Helicobacter pylori 26695	56.70
Hsero_4721	ureB	urease beta subunit UreB, urea amidohydrolase [Urease (VF0050)]	Helicobacter pylori 26695	64.15
Hsero_4724	ureG	urease accessory protein (ureG) [Urease (VF0050)]	Helicobacter pylori 26695	65.15

Fonte: O Autor, 2018.

TABELA AS 23: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO *H. linhagem 2 AU13965*

Locus	Gene do banco de dados	Organismo	Identidade
HI2_00860	PA2537	Pseudomonas_aeruginosa	59.62
HI2_01190	GzOB039	Fusarium_graminearum	54.96
HI2_01540	MoPRX1	Magnaporthe_oryzae	52.53
HI2_01830	lrp	Xenorhabdus_nematophila	59.60
HI2_02710	PCC21__023220	Pectobacterium_carotovorum	57.21
HI2_02970	oqxB	Klebsiella_pneumoniae	61.44
HI2_03110	hupA	Edwardsiella_tarda	53.93
HI2_03320	secA	Acinetobacter_baumannii	54.88
HI2_03410	VcpD	Vibrio_cholerae	53.31
HI2_04790	acrB	Klebsiella_pneumoniae	58.15
HI2_04790	AcrB	Erwinia_amylovora	57.66
HI2_04790	AcrD	Erwinia_amylovora	54.84
HI2_04790	adeB	Acinetobacter_baumannii	52.66
HI2_04790	adeB	Acinetobacter_baumannii	52.66
HI2_04790	adeJ	Acinetobacter_baumannii	51.67
HI2_04970	CspV	Vibrio_cholerae	58.46

Locus	Gene do banco de dados	Organismo	Identidade
HI2_04970	cspA	Brucella_melitensis	50.70
HI2_05450	xrp8	Xanthomonas_oryzae	62.58
HI2_06090	Fur	Escherichia_coli	57.89
HI2_06090	Fur	Pseudomonas_fluorescens	61.34
HI2_06090	fur	Stenotrophomonas_maltophilia	52.80
HI2_06550	recA	Streptococcus_pneumoniae	63.25
HI2_06961	XC_2466	Xanthomonas_campestris	55.65
HI2_07990	thioredoxin_1	Salmonella_enterica	65.74
HI2_08470	glpD	Yersinia_pestis	50.73
HI2_08800	GroEL	Porphyromonas_gingivalis	63.12
HI2_09600	Sdh1	Parastagonospora_nodorum	59.18
HI2_09600	MoSSADH	Magnaporthe_oryzae	58.93
HI2_09710	GGT	Helicobacter_pylori	54.32
HI2_10190	Sdh1	Parastagonospora_nodorum	52.36
HI2_10190	MoSSADH	Magnaporthe_oryzae	53.62
HI2_11210	glyA	Edwardsiella_ictaluri	61.78
HI2_11430	katG	Acinetobacter_nosocomialis	63.18
HI2_11430	CPXB	Magnaporthe_oryzae	60.00
HI2_11590	pstB	Xanthomonas_citri	61.28
HI2_12560	acrB	Klebsiella_pneumoniae	65.26
HI2_12560	AcrB	Erwinia_amylovora	64.18
HI2_12560	adeJ	Acinetobacter_baumannii	57.52
HI2_12560	AcrD	Erwinia_amylovora	57.79
HI2_12990	CspV	Vibrio_cholerae	66.13
HI2_12990	CspR	Enterococcus_faecalis	61.54
HI2_12990	cspA	Brucella_melitensis	51.56
HI2_13000	clpS/clpA	Pseudomonas_aeruginosa	53.47
HI2_13370	Orn	Pseudomonas_aeruginosa	63.84
HI2_15660	Irp	Xenorhabdus_nematophila	59.86
HI2_16420	pnp	Salmonella_enterica	65.57
HI2_16420	pnp	Pseudomonas_aeruginosa	61.82
HI2_16580	NOS1	Fusarium_graminearum	57.35
HI2_16890	htpG	Escherichia_coli	59.52
HI2_17090	HspL	Agrobacterium_tumefaciens	54.61
HI2_17090	hspAT1	Agrobacterium_tumefaciens	52.59
HI2_17090	hspAT2	Agrobacterium_tumefaciens	50.37
HI2_18040	rarA	Klebsiella_pneumoniae	55.30
HI2_18280	fabG1	Ralstonia_solanacearum	72.98
HI2_18320	rpoE	Salmonella_enterica	61.58
HI2_18320	RpoE	Proteus_mirabilis	58.42
HI2_19250	cheB	Salmonella_enterica	64.71
HI2_19480	MacB	Salmonella_enterica	56.57
HI2_20530	ompR	Salmonella_enterica	64.38
HI2_20670	galU	Xanthomonas_campestris	56.06
HI2_20670	galU	Listeria_monocytogenes	50.90
HI2_20890	PA3644	Pseudomonas_aeruginosa	50.76
HI2_21040	smpB	Escherichia_coli	52.03
HI2_21040	smpB	Salmonella_enterica	50.68
HI2_21070	IMPDH	Streptococcus_suis	55.12
HI2_22210	nhaA	Escherichia_coli	53.14
HI2_22920	PmrA	Coxiella_burnetii	50.23
HI2_23800	Pst-Lov	Pseudomonas_syringae	57.66
HI2_23910	XAC1258	Xanthomonas_citri	51.93
HI2_24430	MqsRA	Xylella_fastidiosa	65.00
HI2_25120	hupA	Edwardsiella_tarda	55.56
HI2_25140	LonA	Actinobacillus_pleuropneumoniae	64.63
HI2_25150	ClpX	Staphylococcus_aureus	59.56
HI2_25160	ClpP	Salmonella_enterica	70.98
HI2_25160	clpP	Pseudomonas_aeruginosa	75.00

Locus	Gene do banco de dados	Organismo	Identidade
HI2_25160	ClpP	Staphylococcus_aureus	66.32
HI2_26130	galU	Xanthomonas_campestris	59.29
HI2_26580	MoSFA1	Magnaporthe_oryzae	60.87
HI2_26580	gnoA	Aspergillus_fumigatus	62.40
HI2_26580	Fdh3	Candida_albicans	58.97
HI2_26580	GNO1	Cryptococcus_neoformans	56.79
HI2_26820	A0A0971B07	Riemerella_anatipestifer	57.36
HI2_28301	purA	Salmonella_enterica	60.23
HI2_28360	Hfq	Bordetella_pertussis	85.90
HI2_28360	Hfq	Vibrio_cholerae	82.09
HI2_28360	hfq	Yersinia_enterocolitica	84.85
HI2_28360	Hfq	Yersinia_pestis	84.85
HI2_28360	hfq	Salmonella_enterica	83.33
HI2_28360	hfq	Haemophilus_ducreyi	72.97
HI2_28440	XC_2203	Xanthomonas_campestris	74.47
HI2_28440	Ndk	Pseudomonas_aeruginosa	73.05
HI2_28470	rpoS	Salmonella_enterica	53.82
HI2_28470	rpoS	Erwinia_amylovora	53.45
HI2_28720	cheW	Escherichia_coli	64.71
HI2_28760	motA	Pantoea_ananatis	66.32
HI2_28760	motA	Escherichia_coli	60.00
HI2_28790	flhD	Burkholderia_glumae	52.94
HI2_29070	dgcB	Burkholderia_glumae	55.38
HI2_29210	CspV	Vibrio_cholerae	66.67
HI2_29210	CspR	Enterococcus_faecalis	57.81
HI2_29210	cspA	Brucella_melitensis	53.12
HI2_29220	AoMIs	Arthrotrichy_oligospora	52.96
HI2_29220	Q5J4D6	Parastagonospora_nodorum	53.26
HI2_29980	raxP_(cysD)	Xanthomonas_oryzae	50.68
HI2_30430	lscR	Yersinia_pseudotuberculosis	50.93
HI2_32530	TRPS	Aspergillus_fumigatus	61.05
HI2_34060	opgG	Dickeya_dadantii	50.58
HI2_36030	FVEG_12529	Fusarium_verticillioides	54.61
HI2_36160	Sdh1	Parastagonospora_nodorum	52.37
HI2_36160	MoSSADH	Magnaporthe_oryzae	53.29
HI2_36290	rfaE	Haemophilus_parasuis	50.17
HI2_36330	himD	Salmonella_enterica	61.54
HI2_36360	aroA	Burkholderia_glumae	66.06
HI2_36360	P07637	Salmonella_enterica	56.74
HI2_36400	GyrA	Burkholderia_glumae	79.70
HI2_36400	GyrA	Burkholderia_glumae	79.70
HI2_36400	GyrA	Burkholderia_glumae	79.58
HI2_36400	GyrA	Burkholderia_glumae	79.46
HI2_36400	GyrA	Burkholderia_glumae	79.46
HI2_36630	fabG2	Ralstonia_solanacearum	62.45
HI2_37250	kdpB	Pseudomonas_aeruginosa	61.24
HI2_37390	Rv3232c	Mycobacterium_tuberculosis	54.42
HI2_37540	G4N178	Magnaporthe_oryzae	57.96
HI2_37540	gpx31	Candida_albicans	54.04
HI2_37540	AaGPx3	Alternaria_alternata	54.94
HI2_38590	Mollv1	Magnaporthe_oryzae	54.60
HI2_38710	FTT0673p/prsAp	Francisella_tularensis	57.78
HI2_39170	adeB	Acinetobacter_baumannii	62.52
HI2_39170	adeB	Acinetobacter_baumannii	62.52
HI2_39170	acrB	Klebsiella_pneumoniae	51.08
HI2_39170	AcrB	Erwinia_amylovora	51.76
HI2_39710	fcs	Ralstonia_solanacearum	56.30
HI2_40070	tatC	Vibrio_cholerae	50.42
HI2_40560	paaA	Burkholderia_cenocepacia	69.93

Locus	Gene do banco de dados	Organismo	Identidade
HI2_40570	paaB	Burkholderia_cenocepacia	60.64
HI2_40590	paaD	Burkholderia_cenocepacia	53.18
HI2_41030	rfaE	Haemophilus_parasuis	55.22
HI2_41490	Mqo	Staphylococcus_aureus	50.82
HI2_41970	MntE	Streptococcus_pyogenes	50.42
HI2_42110	MGG_00383	Magnaporthe_oryzae	54.67
HI2_43140	PmrA	Coxiella_burnetii	53.21
HI2_45460	MqsRA	Xylella_fastidiosa	62.63
HI2_45570	AcrB	Erwinia_amylovora	61.08
HI2_45570	acrB	Klebsiella_pneumoniae	59.51
HI2_45570	AcrD	Erwinia_amylovora	58.59
HI2_45570	adeJ	Acinetobacter_baumannii	54.45
HI2_46720	URE1	Cryptococcus_gattii	56.00
HI2_46720	URE1	Cryptococcus_neoformans	56.00
HI2_46740	urease	Helicobacter_pylori	64.50
HI2_46740	URE1	Cryptococcus_gattii	60.88
HI2_46740	URE1	Cryptococcus_neoformans	59.47
HI2_46740	GzZC144	Fusarium_graminearum	55.68
HI2_47200	lipA	Acinetobacter_baumannii	62.50
HI2_47310	opgG	Dickeya_dadantii	57.79

Fonte: O Autor, 2018.

TABELA AS 24: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO *H. seropedicae* SmR1

Locus	Gene do banco de dados	Organismo	Identidade
Hsero_0077	PA2537	Pseudomonas_aeruginosa	60.10
Hsero_0109	GzOB039	Fusarium_graminearum	54.96
Hsero_0144	MoPRX1	Magnaporthe_oryzae	52.53
Hsero_0173	Irp	Xenorhabdus_nematophila	59.60
Hsero_0288	PCC21_023220	Pectobacterium_carotovorum	57.71
Hsero_0315	oqxB	Klebsiella_pneumoniae	61.44
Hsero_0346	secA	Acinetobacter_baumannii	54.99
Hsero_0355	VcpD	Vibrio_cholerae	53.31
Hsero_0483	acrB	Klebsiella_pneumoniae	58.35
Hsero_0483	AcrB	Erwinia_amylovora	57.76
Hsero_0483	AcrD	Erwinia_amylovora	54.94
Hsero_0483	adeB	Acinetobacter_baumannii	52.76
Hsero_0483	adeB	Acinetobacter_baumannii	52.76
Hsero_0483	adeJ	Acinetobacter_baumannii	51.67
Hsero_0505	CspV	Vibrio_cholerae	58.46
Hsero_0505	cspA	Brucella_melitensis	50.70
Hsero_0549	xrp8	Xanthomonas_oryzae	62.58
Hsero_0608	Fur	Escherichia_coli	57.89
Hsero_0608	Fur	Pseudomonas_fluorescens	61.34
Hsero_0608	fur	Stenotrophomonas_maltophilia	52.80
Hsero_0655	recA	Streptococcus_pneumoniae	63.25
Hsero_0674	ADE2	Saccharomyces_cerevisiae	57.42
Hsero_0674	ADE2	Candida_albicans	55.70
Hsero_0674	ADE2	Cryptococcus_neoformans	55.84
Hsero_0697	XC_2466	Xanthomonas_campestris	55.65
Hsero_0790	bscN	Bordetella_bronchiseptica	51.42
Hsero_0860	thioredoxin_1	Salmonella_enterica	65.74
Hsero_0966	glpD	Yersinia_pestis	50.94
Hsero_0996	GroEL	Porphyromonas_gingivalis	63.12
Hsero_1077	Sdh1	Parastagonospora_nodorum	59.59
Hsero_1077	MoSSADH	Magnaporthe_oryzae	59.14
Hsero_1089	GGT	Helicobacter_pylori	54.51
Hsero_1125	Sdh1	Parastagonospora_nodorum	52.36

Locus	Gene do banco de dados	Organismo	Identidade
Hsero_1125	MoSSADH	Magnaporthe_oryzae	53.40
Hsero_1223	glyA	Edwardsiella_ictaluri	61.54
Hsero_1242	katG	Acinetobacter_nosocomialis	63.31
Hsero_1242	CPXB	Magnaporthe_oryzae	60.27
Hsero_1258	pstB	Xanthomonas_citri	61.19
Hsero_1358	acrB	Klebsiella_pneumoniae	65.55
Hsero_1358	AcrB	Erwinia_amylovora	64.38
Hsero_1358	adeJ	Acinetobacter_baumannii	57.62
Hsero_1358	AcrD	Erwinia_amylovora	57.88
Hsero_1397	CspV	Vibrio_cholerae	66.13
Hsero_1397	CspR	Enterococcus_faecalis	61.54
Hsero_1397	cspA	Brucella_melitensis	51.56
Hsero_1398	clpS/clpA	Pseudomonas_aeruginosa	53.47
Hsero_1434	Orn	Pseudomonas_aeruginosa	63.84
Hsero_1685	lrp	Xenorhabdus_nematophila	59.86
Hsero_1755	pnp	Salmonella_enterica	65.57
Hsero_1755	pnp	Pseudomonas_aeruginosa	62.11
Hsero_1770	NOS1	Fusarium_graminearum	57.11
Hsero_1801	htpG	Escherichia_coli	59.52
Hsero_1824	HspL	Agrobacterium_tumefaciens	54.61
Hsero_1824	hspAT1	Agrobacterium_tumefaciens	52.59
Hsero_1824	hspAT2	Agrobacterium_tumefaciens	50.37
Hsero_1906	rarA	Klebsiella_pneumoniae	55.30
Hsero_1925	fabG1	Ralstonia_solanacearum	73.39
Hsero_1929	rpoE	Salmonella_enterica	61.58
Hsero_1929	RpoE	Proteus_mirabilis	58.42
Hsero_2022	cheB	Salmonella_enterica	64.53
Hsero_2045	MacB	Salmonella_enterica	56.64
Hsero_2154	ompR	Salmonella_enterica	64.38
Hsero_2168	galU	Xanthomonas_campestris	56.06
Hsero_2168	galU	Listeria_monocytogenes	50.90
Hsero_2190	PA3644	Pseudomonas_aeruginosa	50.76
Hsero_2202	smxB	Escherichia_coli	52.03
Hsero_2202	smxB	Salmonella_enterica	50.68
Hsero_2205	IMPDH	Streptococcus_suis	54.92
Hsero_2291	nhaA	Escherichia_coli	53.14
Hsero_2453	Pst-Lov	Pseudomonas_syringae	57.31
Hsero_2464	XAC1258	Xanthomonas_citri	52.28
Hsero_2493	stoS	Xanthomonas_oryzae	50.60
Hsero_2514	MqsRA	Xylella_fastidiosa	66.00
Hsero_2580	LonA	Actinobacillus_pleuropneumoniae	64.50
Hsero_2581	ClpX	Staphylococcus_aureus	59.56
Hsero_2582	ClpP	Salmonella_enterica	70.98
Hsero_2582	clpP	Pseudomonas_aeruginosa	75.00
Hsero_2582	ClpP	Staphylococcus_aureus	66.32
Hsero_2666	galU	Xanthomonas_campestris	54.68
Hsero_2713	MoSFA1	Magnaporthe_oryzae	60.87
Hsero_2713	gnoA	Aspergillus_fumigatus	62.40
Hsero_2713	Fdh3	Candida_albicans	58.97
Hsero_2713	GNO1	Cryptococcus_neoformans	56.79
Hsero_2737	A0A0971B07	Riemerella_anatipestifer	58.18
Hsero_2843	modA	Pseudomonas_aeruginosa	61.40
Hsero_2943	purA	Salmonella_enterica	60.23
Hsero_2948	Hfq	Bordetella_pertussis	85.90
Hsero_2948	Hfq	Vibrio_cholerae	82.09
Hsero_2948	hfq	Yersinia_enterocolitica	84.85
Hsero_2948	Hfq	Yersinia_pestis	84.85
Hsero_2948	hfq	Salmonella_enterica	83.33
Hsero_2948	hfq	Haemophilus_ducreyi	72.97

Locus	Gene do banco de dados	Organismo	Identidade
Hsero_2956	XC_2203	Xanthomonas_campestris	75.18
Hsero_2956	Ndk	Pseudomonas_aeruginosa	73.76
Hsero_2959	rpoS	Salmonella_enterica	53.82
Hsero_2959	rpoS	Erwinia_amylovora	53.45
Hsero_2981	cheW	Escherichia_coli	64.71
Hsero_2986	motA	Pantoea_ananatis	66.32
Hsero_2986	motA	Escherichia_coli	60.00
Hsero_2989	flhD	Burkholderia_glumae	52.94
Hsero_3016	dgcB	Burkholderia_glumae	55.69
Hsero_3028	CspV	Vibrio_cholerae	66.67
Hsero_3028	CspR	Enterococcus_faecalis	57.81
Hsero_3028	cspA	Brucella_melitensis	53.12
Hsero_3029	AoMIs	Arthrobotrys_oligospora	52.96
Hsero_3029	Q5J4D6	Parastagonospora_nodorum	53.26
Hsero_3102	raxP_(cysD)	Xanthomonas_oryzae	50.68
Hsero_3145	lscR	Yersinia_pseudotuberculosis	50.93
Hsero_3286	ADE4	Saccharomyces_cerevisiae	50.10
Hsero_3292	TRPS	Aspergillus_fumigatus	61.05
Hsero_3449	opgG	Dickeya_dadantii	50.39
Hsero_3669	FVEG_12529	Fusarium_verticillioides	54.37
Hsero_3684	rfaE	Haemophilus_parasuis	50.17
Hsero_3688	himD	Salmonella_enterica	61.54
Hsero_3691	aroA	Burkholderia_glumae	66.06
Hsero_3691	P07637	Salmonella_enterica	56.51
Hsero_3695	GyrA	Burkholderia_glumae	79.70
Hsero_3695	GyrA	Burkholderia_glumae	79.70
Hsero_3695	GyrA	Burkholderia_glumae	79.58
Hsero_3695	GyrA	Burkholderia_glumae	79.46
Hsero_3695	GyrA	Burkholderia_glumae	79.46
Hsero_3715	fabG2	Ralstonia_solanacearum	62.45
Hsero_3778	kdpB	Pseudomonas_aeruginosa	61.24
Hsero_3790	Rv3232c	Mycobacterium_tuberculosis	57.09
Hsero_3805	G4N178	Magnaporthe_oryzae	57.96
Hsero_3805	gpx31	Candida_albicans	54.72
Hsero_3805	AaGPx3	Alternaria_alternata	54.60
Hsero_3908	Mollv1	Magnaporthe_oryzae	54.79
Hsero_3920	FTT0673p/prsAp	Francisella_tularensis	57.46
Hsero_3957	adeB	Acinetobacter_baumannii	62.33
Hsero_3957	adeB	Acinetobacter_baumannii	62.33
Hsero_3957	acrB	Klebsiella_pneumoniae	51.08
Hsero_3957	AcrB	Erwinia_amylovora	51.76
Hsero_4007	fcs	Ralstonia_solanacearum	56.63
Hsero_4057	tatC	Vibrio_cholerae	50.42
Hsero_4125	paaA	Burkholderia_cenocepacia	70.26
Hsero_4126	paaB	Burkholderia_cenocepacia	60.64
Hsero_4128	paaD	Burkholderia_cenocepacia	53.76
Hsero_4171	rfaE	Haemophilus_parasuis	55.22
Hsero_4194	A0A097IB07	Riemerella_anatipestifer	55.95
Hsero_4225	Mqo	Staphylococcus_aureus	50.82
Hsero_4275	MntE	Streptococcus_pyogenes	50.42
Hsero_4290	MGG_00383	Magnaporthe_oryzae	55.20
Hsero_4397	PmrA	Coxiella_burnetii	53.21
Hsero_4441	GroEL	Porphyromonas_gingivalis	62.64
Hsero_4600	MqsRA	Xylella_fastidiosa	62.63
Hsero_4611	AcrB	Erwinia_amylovora	60.88
Hsero_4611	acrB	Klebsiella_pneumoniae	59.90
Hsero_4611	AcrD	Erwinia_amylovora	59.45
Hsero_4611	adeJ	Acinetobacter_baumannii	54.31
Hsero_4719	URE1	Cryptococcus_gattii	56.00

Locus	Gene do banco de dados	Organismo	Identidade
Hsero_4719	URE1	Cryptococcus_neoformans	56.00
Hsero_4721	urease	Helicobacter_pylori	64.15
Hsero_4721	URE1	Cryptococcus_gattii	60.70
Hsero_4721	URE1	Cryptococcus_neoformans	59.30
Hsero_4721	GzZC144	Fusarium_graminearum	55.50
Hsero_4771	lipA	Acinetobacter_baumannii	62.17
Hsero_4784	opgG	Dickeya_dadantii	57.99

Fonte: O Autor, 2018

TABELA AS 25: GENES DE RESISTÊNCIA DA ESTIRPE *H. linhagem 2 AU13965*

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
HI2_33930	ABC_efflux	RF0007	1.4e-94	HI2_33930 Putative multidrug export ATP-binding/permease protein 3827056:3828900 forward MW:65380
HI2_20720	ABC_efflux	RF0007	6.1e-86	HI2_20720 putative ABC transporter ATP-binding protein 2304870:2306210 forward MW:49866
HI2_10520	ABC_efflux	RF0007	3.9e-84	apxIB_1 Toxin RTX-I translocation ATP-binding protein 1180446:1182644 reverse MW:80130
HI2_38050	ABC_efflux	RF0007	8.7e-83	atm1 ATM1-type heavy metal exporter 4281198:4283102 reverse MW:70416
HI2_04800	ABC_efflux	RF0007	1.7e-79	mexA_1 Multidrug resistance protein MexA 516086:517294 reverse MW:42208
HI2_12570	adeC-adeK-oprM	RF0021	3.7e-182	oprM_4 Outer membrane protein OprM 1412402:1413898 forward MW:53151
HI2_45560	adeC-adeK-oprM	RF0021	1.2e-176	oprM_12 Outer membrane protein OprM 5116915:5118393 reverse MW:52726
HI2_33170	ClassA	RF0053	2.1e-104	bla Beta-lactamase CTX-M-97 3739087:3739986 forward MW:31837
HI2_05390	ClassB	RF0054	2.8e-34	ccrA Metallo-beta-lactamase type 2 581626:582585 reverse MW:34954
HI2_33170	Exo	RF0072	4.4e-77	bla Beta-lactamase CTX-M-97 3739087:3739986 forward MW:31837
HI2_19480	macB	RF0089	5.9e-229	macB Macrolide export ATP-binding/permease protein MacB 2165435:2167405 reverse MW:70554
HI2_02980	MexE	RF0098	2.6e-105	bepF Efflux pump periplasmic linker BepF 326486:327700 reverse MW:42107
HI2_04800	MexE	RF0098	2.3e-58	mexA_1 Multidrug resistance protein MexA 516086:517294 reverse MW:42208
HI2_04800	MexX	RF0101	1.8e-98	mexA_1 Multidrug resistance protein MexA 516086:517294 reverse MW:42208
HI2_35060	msbA	RF0107	3.8e-174	msbA Lipid A export ATP-binding/permease protein MsbA 3952416:3954152 forward MW:63564
HI2_02970	RND_efflux	RF0115	0	bepE Efflux pump membrane transporter BepE 323171:326386 reverse MW:115220
HI2_04790	RND_efflux	RF0115	0	acrB_1 Multidrug efflux pump subunit AcrB 512928:516074 reverse MW:111883
HI2_45570	RND_efflux	RF0115	0	mexB Multidrug resistance protein MexB 5118393:5121548 reverse MW:112690
HI2_12560	RND_efflux	RF0115	0	acrB_2 Multidrug efflux pump subunit AcrB 1409251:1412400 forward MW:112111

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
HI2_14180	RND_efflux	RF0115	0	mdtB_1 Multidrug resistance protein MdtB 1604417:1607563 forward MW:112283
HI2_14190	RND_efflux	RF0115	0	mdtC_1 Multidrug resistance protein MdtC 1607576:1610812 forward MW:115738
HI2_39170	RND_efflux	RF0115	0	acrF Multidrug export protein AcrF 4407093:4410194 reverse MW:110020
HI2_24110	RND_efflux	RF0115	1.2e-287	mdtB_2 Multidrug resistance protein MdtB 2721341:2724496 forward MW:113480
HI2_05390	SubclassB1	RF0123	8.3e-29	ccrA Metallo-beta-lactamase type 2 581626:582585 reverse MW:34954
HI2_05390	SubclassB2	RF0124	1.8e-24	ccrA Metallo-beta-lactamase type 2 581626:582585 reverse MW:34954

Fonte: O Autor, 2018

TABELA AS 26: GENES DE RESISTÊNCIA DA ESTIRPE *H. seropedicae* AU14040

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
WP_069374526.1	ABC_efflux	RF0007	2.3e-94	ABC transporter [Herbaspirillum seropedicae]
WP_013234152.1	ABC_efflux	RF0007	5.1e-92	ABC transporter [Herbaspirillum seropedicae]
WP_069374232.1	ABC_efflux	RF0007	6.4e-88	peptidase C39 [Herbaspirillum seropedicae]
WP_083247235.1	ABC_efflux	RF0007	2.9e-84	type I secretion system permease/ATPase [Herbaspirillum seropedicae]
WP_069374824.1	ABC_efflux	RF0007	3e-82	metal ABC transporter permease [Herbaspirillum seropedicae]
WP_083247135.1	ABC_efflux	RF0007	1.5e-81	hypothetical protein [Herbaspirillum seropedicae]
WP_069372937.1	ABC_efflux	RF0007	1.2e-79	efflux transporter periplasmic adaptor subunit [Herbaspirillum seropedicae]
WP_069373384.1	adeC-adeK-oprM	RF0021	1.6e-181	multidrug transporter [Herbaspirillum seropedicae]
WP_069374479.1	ClassA	RF0053	1.6e-104	class A beta-lactamase [Herbaspirillum seropedicae]
WP_069372971.1	ClassB	RF0054	2.7e-34	MBL fold metallo-hydrolase [Herbaspirillum seropedicae]
WP_069374479.1	Exo	RF0072	1.3e-77	class A beta-lactamase [Herbaspirillum seropedicae]
WP_069373820.1	macB	RF0089	3e-227	macrolide ABC transporter permease/ATP-binding protein MacB [Herbaspirillum seropedicae]
WP_069372841.1	MexE	RF0098	5.8e-105	efflux transporter periplasmic adaptor subunit [Herbaspirillum seropedicae]
WP_069372937.1	MexE	RF0098	4.3e-59	efflux transporter periplasmic adaptor subunit [Herbaspirillum seropedicae]
WP_069372937.1	MexX	RF0101	4.6e-99	efflux transporter periplasmic adaptor subunit [Herbaspirillum seropedicae]
WP_069374637.1	msbA	RF0107	3.8e-174	lipid A export permease/ATP-binding protein MsbA [Herbaspirillum seropedicae]
WP_069372840.1	RND_efflux	RF0115	0	hydrophobe/amphiphile efflux-1 family RND transporter [Herbaspirillum seropedicae]
WP_013232521.1	RND_efflux	RF0115	0	multidrug efflux RND transporter permease subunit [Herbaspirillum seropedicae]
WP_069375283.1	RND_efflux	RF0115	0	hydrophobe/amphiphile efflux-1 family RND transporter [Herbaspirillum seropedicae]
WP_013233379.1	RND_efflux	RF0115	0	multidrug efflux RND transporter permease subunit [Herbaspirillum seropedicae]
WP_069373564.1	RND_efflux	RF0115	0	multidrug transporter subunit MdtB [Herbaspirillum seropedicae]
WP_069373565.1	RND_efflux	RF0115	0	multidrug transporter subunit MdtC [Herbaspirillum seropedicae]
WP_069374879.1	RND_efflux	RF0115	0	hydrophobe/amphiphile efflux-1 family RND transporter [Herbaspirillum seropedicae]

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
WP_069374055.1	RND_efflux	RF0115	2.9e-288	acriflavine resistance protein B [Herbaspirillum seropedicae]
WP_069372971.1	SubclassB1	RF0123	8.1e-29	MBL fold metallo-hydrolase [Herbaspirillum seropedicae]
WP_069372971.1	SubclassB2	RF0124	1.8e-24	MBL fold metallo-hydrolase [Herbaspirillum seropedicae]

Fonte: O Autor, 2018.

TABELA AS 27: GENES DE RESISTÊNCIA DA ESTIRPE *H. seropedicae* SMR1

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
Hsero_3437	ABC_efflux	RF0007	4.6e-94	Hsero_3437 ABC-type multidrug transport system, ATPase and permease components protein 3941740:3943584 forward MW:65334
Hsero_2173	ABC_efflux	RF0007	5.2e-92	mdlB ABC-type multidrug transport system, ATPase and permease components protein 2446051:2448327 forward MW:84176
Hsero_2881	ABC_efflux	RF0007	3.3e-87	cyaB ABC-type bacteriocin/lantibiotic (cyclolysin-type) secretion transporter protein 3273870:3276050 reverse MW:79849
Hsero_3860	ABC_efflux	RF0007	3.9e-83	Hsero_3860 ABC-type transport system involved in Fe-S cluster assembly, permease and ATPase components, protein 4418799:4420703 reverse MW:70447
Hsero_0484	ABC_efflux	RF0007	1.7e-79	acrA acriflavin resistance lipoprotein A protein 530092:531300 reverse MW:42222
Hsero_1359	adeC-adeK-oprM	RF0021	4.7e-181	tolC outer membrane drug efflux lipoprotein 1549806:1551302 forward MW:53195
Hsero_3349	ClassA	RF0053	7.8e-104	penP beta-lactamase class A protein 3840276:3841175 forward MW:31836
Hsero_0544	ClassB	RF0054	1.9e-34	Hsero_0544 Zn-dependent hydrolase (glyoxylase) protein 599528:600487 reverse MW:34985
Hsero_3349	Exo	RF0072	4.9e-77	penP beta-lactamase class A protein 3840276:3841175 forward MW:31836
Hsero_2045	macB	RF0089	3.1e-228	Hsero_2045 ABC-type antimicrobial peptide transport system, ATPase component protein 2309877:2311850 reverse MW:70754
Hsero_0316	MexE	RF0098	6.6e-106	Hsero_0316 HlyD-family secretion protein 347390:348589 reverse MW:41665
Hsero_0484	MexE	RF0098	1.9e-58	acrA acriflavin resistance lipoprotein A protein 530092:531300 reverse MW:42222
Hsero_0484	MexX	RF0101	2e-98	acrA acriflavin resistance lipoprotein A protein 530092:531300 reverse MW:42222
Hsero_3566	msbA	RF0107	2.8e-174	Hsero_3566 ABC-type multidrug transport system, ATPase and permease components protein 4088785:4090521 forward MW:63497
Hsero_0315	RND_efflux	RF0115	0	Hsero_0315 cation/multidrug efflux pump protein 344075:347290 reverse MW:115278
Hsero_0483	RND_efflux	RF0115	0	acrB acriflavin resistance transmembrane protein 526934:530080 reverse MW:111898
Hsero_4611	RND_efflux	RF0115	0	Hsero_4611 cation/multidrug efflux pump protein 5288823:5291978 reverse MW:112831
Hsero_1358	RND_efflux	RF0115	0	Hsero_1358 cation/multidrug efflux pump protein 1546655:1549804 forward MW:112130
Hsero_1521	RND_efflux	RF0115	0	Hsero_1521 cation/multidrug efflux pump transmembrane protein 1743666:1746812 forward MW:112194
Hsero_1522	RND_efflux	RF0115	0	Hsero_1522 cation/multidrug efflux pump protein 1746825:1750061 forward MW:115881

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
Hsero_3957	RND_efflux	RF0115	0	acrF multidrug efflux system transmembrane protein 4535647:4538748 reverse MW:110018
Hsero_2483	RND_efflux	RF0115	2e-288	Hsero_2483 cation/multidrug efflux pump protein 2819695:2822850 forward MW:113492
Hsero_0544	SubclassB1	RF0123	2e-28	Hsero_0544 Zn-dependent hydrolase (glyoxylase) protein 599528:600487 reverse MW:34985
Hsero_0544	SubclassB2	RF0124	2.7e-24	Hsero_0544 Zn-dependent hydrolase (glyoxylase) protein 599528:600487 reverse MW:34985

Fonte: O Autor, 2018.

6 ALTA SIMILARIDADE ENTRE OS GENOMAS DE ESTIRPES CLÍNICAS E AMBIENTAIS DE *HERBASPIRILLUM FRISINGENSE* SUGERE PRÉ-ADAPTAÇÃO A DIFERENTES HOSPEDEIROS

Willian Klassen de Oliveira¹, Michelle Zibeti Tadra-Sfeir², Rodrigo Cardoso², Emanuel Maltempi de Souza², Fábio de Oliveira Pedrosa², Helisson Faoro^{1,3}

¹*Laboratório de Bioinformática, Setor de Educação Profissional e Tecnológica, Universidade Federal do Paraná, Curitiba, Brazil*

²*Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná, Curitiba, Brazil*

³*Laboratório da Regulação da Expressão Gênica, Instituto Carlos Chagas, Fiocruz-PR, Curitiba, Brazil*

Palavras Chave: *Herbaspirillum frisingense*, isolados clínicos, bactérias oportunistas

6.1 RESUMO

O gênero *Herbaspirillum*, amplamente estudado por conter espécies que realizam fixação biológica de nitrogênio, recentemente começou a ter representantes isolados de amostras clínicas, principalmente recuperadas de pacientes imunocomprometidos. Dentre esses casos está o *Herbaspirillum frisingense*. Inicialmente isolado de gramíneas, essa espécie teve um representante isolado de infecção urinária, nomeado como estirpe VT-16-41, e outro isolado de escarro de pacientes com fibrose cística, nomeado como estirpe AU14559. Na busca para esclarecer como essas bactérias ambientais trocam de ambiente, se tornando capazes de colonizar um meio tão complexo como o organismo humano, fizemos o sequenciamento, montagem e anotação do genoma da bactéria *H. frisingense* AU14559, isolada de escarro de pacientes com fibrose cística, e sua comparação com os genomas da estirpe clínica VT-16-41 e a ambiental *H. frisingense* GSF30 com o objetivo de identificar genes adquiridos e perdidos por esta bactéria que possibilitem sua interação e resiliência em humanos. Os genes adquiridos por meio de transferências horizontais por estes isolados clínicos são, em sua maioria

relacionados com metabolismo e processos celulares. Os genes responsáveis pela fixação biológica de nitrogênio foram perdidos por esta estirpe, corroborando com o seu novo estilo vida. Já os fatores de virulência e interação com hospedeiro são, em sua maioria, compartilhados com GSF30. Esses dados nos levam à hipótese de que esta espécie adquiriu características que lhe proporcionaram diversidade metabólica para que se tornasse capaz de ampliar seus habitats e colonizar e infectar hospedeiros humanos suscetíveis.

6.2 INTRODUÇÃO

O gênero *Herbaspirillum* é um gênero da classe beta do filo Proteobacterias. Atualmente, há 11 espécies neste gênero, isoladas de diferentes ambientes como água (DE SOUZA et al., 2013), rizosfera (BALDANI et al., 1986), tecidos vegetais (GUIZELINI et al., 2015), entre outros. Três espécies deste gênero, *H. frisingense*, *H. seropedicae* e *H. rubrisubalbicans*, são descritas como bactérias fixadoras de nitrogênio, ou seja, bactérias capazes de converter nitrogênio atmosférico (N_2) em íons amônio (NH_3), uma forma assimilável pelas plantas, através do complexo enzimático da nitrogenase, codificado pelos genes *nifHDK* (STRAUB et al., 2013). Essas bactérias estabelecem associação de forma endofítica com plantas de interesse econômico como trigo, milho, cana de açúcar e sorgo, promovendo o crescimento destes vegetais através da fixação de nitrogênio, além da secreção de fitohormônios como o ácido indol acético (STRAUB et al., 2013).

A primeira identificação de uma estirpe clínica pertencente ao gênero *Herbaspirillum* ocorreu em 1996 quando bactérias, previamente descrita como *Pseudomonas*, foram reclassificada como *Herbaspirillum* (BALDANI et al., 1986). Recentemente, isolados deste gênero passaram a ser encontrados mais frequentemente em humanos, em especial pacientes imunocomprometidos (REGUNATH et al., 2015; SUWANTARAT et al., 2015). O primeiro isolado clínico de *H. frisingense*, cujo genoma foi sequenciado, foi a estirpe VT-16-41, isolada de um caso de infecção do trato urinário (TETZ; TETZ, 2017). Em outro estudo, Spilker e colaboradores (2008) isolaram 28 estirpes, classificadas no gênero *Herbaspirillum* através de análises do gene 16S, de escarro de pacientes com fibrose cística. Estas estirpes foram distribuídas em diferentes espécies deste gênero, incluindo *H. frisingense*, identificada como estirpe AU14559. Entretanto, ainda não é totalmente

claro se estas bactérias são apenas oportunistas, encontradas ocasionalmente em amostras clínicas ou se são patógenos podendo causar infecções secundárias e piorar o estado dos pacientes.

Atualmente há 3 genomas de *Herbaspirillum frisingense* depositados, sendo eles pertencentes às estirpes ambientais GSF30 isolado de plantas do gênero *Miscanthus* no sul da Alemanha (KIRCHHOF et al., 2001) e AA6 isolado de raízes (NIU; KOLTER, 2017), além de uma estirpe clínica VT-16-41 isolado de um caso de infecção de trato urinário (TETZ; TETZ, 2017). Neste trabalho descrevemos o genoma de *H. frisingense* AU14559, isolado de escarro de pacientes com fibrose cística e sua comparação com *H. frisingense* GSF30, *H. frisingense* AA6 e *H. frisingense* ureolyticus VT-16-41, buscando elucidar como uma bactéria inicialmente ambiental começa a infectar humanos.

6.3 MATERIAIS E MÉTODOS

6.3.1 Purificação de DNA e sequenciamento

As células de *H. frisingense* AU14559 foram crescidas *overnight* em meio NFbHP (KLASSEN et al., 1997) a 37°C até densidade óptica em 600 nanômetros de 1,0 unidade de absorvância. O DNA cromossomal foi então purificado através do método de fenol clorofórmio (PEDROSA et al., 2011) e o genoma foi sequenciado em uma estratégia híbrida através das tecnologias de sequenciamento Miseq (Illumina) e Ion Proton (ThermoFisher). Para a plataforma Miseq, a biblioteca genômica foi construída usando “Nextera DNA sample preparation” (Illumina) e sequenciado em 2x300 pb *paired-end*. Para a plataforma Ion Proton, a biblioteca genômica foi construída usando usando o kit “Ion Xpress Plus Fragment” (ThermoFisher) e sequenciado em leituras *single-end* de 200 pb.

6.3.2 Montagem do genoma

O genoma foi montado individualmente e de forma combinada para as diferentes plataformas de sequenciamento através dos montadores Newbler v2.9 (454 Life Sciences, Branford, CT), SPAdes v3.10.0 (BANKEVICH et al., 2012) e o montador CLC de novo assembly v10 (CLC bio, Dinamarca). As montagens foram avaliadas através do algoritmo QUAST (*Quality Assessment Tool for Genome*

Assemblies) (GUREVICH et al., 2013) para escolher a melhor montagem, com a qual foi procedidos os próximos passos. Os *contigs* foram ordenados primeiramente usando o genoma de de *H. frisingense* GSF30 como referência, genoma mais próximo da estirpe AU14559. Como o genoma da estirpe GSF30 não estava fechado, os seus contigs foram ordenados usando como referência o genoma de *Herbaspirillum seropedicae* SmR1, espécie mais próxima com o genoma fechado, através da visualização dos contigs no Artemis Comparison Tool - ACT (CARVER et al., 2005). O fechamento dos gaps foi realizado através dos algoritmos FGAP (PIRO et al., 2014), GFinisher (GUIZELINI et al., 2016) e por inspeção manual, buscando por regiões de similaridade nas extremidades dos contigs e nas montagens realizadas por outros montadores utilizando o algoritmo Blastn (ALTSCHUL et al., 1990). Para validar a montagem, todas as *leituras* foram mapeadas no genoma montado com o propósito de identificar regiões sem cobertura e inconsistências no genoma. O genoma completo foi anotado através do servidor SILVA, disponível em <http://200.236.3.34/SILA/login.jsp>, e do algoritmo Prokka (SEEMANN, 2014). As duas anotações foram mescladas e curadas manualmente.

6.3.3 Comparações genômicas

A identidade nucleotídica média entre os genomas clínicos *H. frisingense* AU14550, *H. frisingense ureolyticus* VT-16-41 (GCA_001994935.1) e os genomas das estirpes ambientais *H. frisingense* GSF30 (GCA_000300975.2) e *H. frisingense* AA6 (GCA_002025725.1) foi calculado através do ANI *calculator*, disponível em <http://enve-omics.ce.gatech.edu/ani/>, baseado no trabalho de Goris (2007). As similaridades e divergências foram investigadas usando um alinhamento genoma-genoma realizado através do blastn (ALTSCHUL et al., 1990) e visualizado através dos *softwares* ACT (CARVER et al., 2005) e BRIG (ALIKHAN et al., 2011). Os genes presentes na linhagem clínica e ausentes na linhagem ambiental foram investigados de forma mais profunda, através de sua função, domínios de sua proteína codificada e a relação destes com a interação e resiliência em humanos. Este processo foi realizado através do algoritmo blastp (ALTSCHUL et al., 1997), banco de dados Pfam (FINN et al., 2010) e buscas no PubMed. As vias metabólicas foram anotadas através do algoritmo KAAS (MORIYA et al., 2007) e comparadas para identificar vias exclusivas em cada um dos organismos. As possíveis regiões de transferência

horizontal de genes foram determinadas pelos *softwares* Alien Hunter (VERNIKOS; PARKHILL, 2006) e GIPSy (SOARES et al., 2016). Para uma visão geral do conteúdo destas ilhas, os genes presentes nestas foram enviados para o programa Blast KOALA para classificação dos genes adquiridos por HGT. Para a identificação de fatores de virulência e de interação patógeno-hospedeiro nos genomas analisados, foi utilizado o algoritmo blastp para efetuar alinhamento de buscas contra o *Core dataset* (Apenas proteínas verificadas experimentalmente) do banco de dados VFDB (*Virulence Factors DataBase*) (CHEN et al., 2005) e PHI - *base* (*Pathogen Host Interactions base*) (URBAN et al., 2015) respectivamente. Para ambas as buscas nos bancos de dados, foi utilizado um *cut-off* de 50% de identidade, *e-value* de 10^{-10} e cobertura mínima de 80%. As proteínas encontradas nos organismos AU14559 e GSF30 foram copiadas para um novo arquivo multifasta e um novo alinhamento foi realizado entre elas para a identificação de fatores de interação e virulência exclusivas de cada um dos organismos. Já para realizar busca de genes que conferem resistência a antibióticos foi utilizado o algoritmo HMMer v3.1b2 para realizar buscas das proteínas dos organismos em um perfil HMM do banco curado, de genes de resistência a antibióticos, ResFams (Core v1.2) (GIBSON; FORSBERG; DANTAS, 2015).

6.4 RESULTADOS

6.4.1 Genoma *H. frisingense* AU14559

As diferentes montagens de genoma de *H. frisingense* AU14559 foram avaliadas pelo programa QCAST e a montagem proveniente do SPAdes apresentou melhores resultados, sendo então escolhida a continuidade das análises. Esta montagem apresentou 42 contigs e um N50 de 492 kb. Após o fechamento automático dos gaps, usando a ferramenta FGAP, restaram 7 contigs, os quais foram manualmente fechados com base em comparações com o genoma de *H. frisingense* GSF30, ordenado através do genoma de *H. seropedicae* SmR1, até se obter um único contig representando todo o genoma. A estrutura e confiabilidade do genoma foi checada mapeando as leituras oriundas do sequenciamento na montagem. Nenhuma região sem cobertura ou inconsistências entre as *leituras paired-end* foram encontradas.

O genoma fechado de *H. frisingense* AU14559 tem uma única molécula de DNA com 5.401.392 pb, conteúdo GC% de 63,1%, 3 operons de rRNAs, 59 tRNAs, 23 ncRNAs e 4.913 genes codificadores de proteínas cobrindo 87,7% do tamanho total do genoma. O valor das identidades nucleotídicas médias calculadas entre o genoma *H. frisingense* AU14559 e os demais genomas classificados como pertencentes a esta espécie, depositados nos bancos de dados, são apresentados na tabela B1.

TABELA B 1: IDENTIDADES NUCLEOTÍDICAS MÉDIAS ENTRE A ESTIRPE AU14559 E DEMAIS GENOMAS *H. FRISINGENSE*

Estirpe	Identidade Nucleotídica Média
<i>H. frisingense</i> GSF30	97,49%
<i>H. frisingense</i> VT-16-41	97,45%
<i>H. frisingense</i> AA6	84,76%

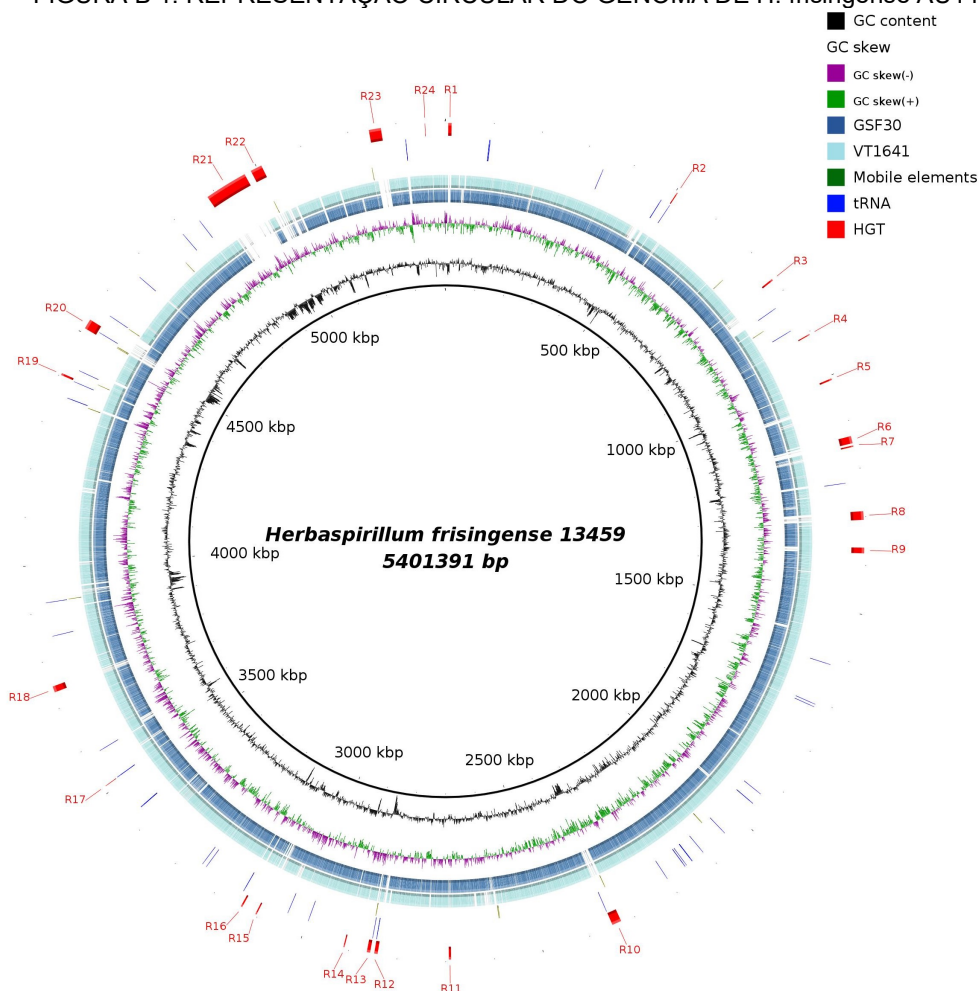
Fonte: O Autor, 2018

Como o valor de ANI para que dois genomas sejam considerados de mesma espécie é 95% podemos verificar que o organismo *H. frisingense* AU14559 pertencente a esta espécie, no entanto a estirpe *H. frisingense* AA6 apresentou um ANI abaixo deste limite, portanto optamos por não incluí-la nas análises.

6.4.2 Ilhas genômicas

O algoritmo *Alien Hunter* identificou 67 ilhas genômicas em AU14559, sendo que, através de visualização no *Artemis Comparison Tool* foi identificado que, 24 destas ilhas, não estavam presentes no genoma de *H. frisingense* GSF30 (Figura B1). Já o algoritmo GIPSY identificou 14 ilhas genômicas, das quais classificou 3 como possíveis ilhas de patogenicidade (Ilhas genômicas 2, 12 e 14), 1 de resistência (Ilha genômica 14) e 2 ilhas como metabólicas (Ilhas genômicas 2 e 9). Os 243 genes presentes nas ilhas genômicas encontradas pelo GIPSY se encontram nas tabelas BS1 a BS14 do material suplementar.

FIGURA B 1: REPRESENTAÇÃO CIRCULAR DO GENOMA DE *H. frisingense* AU14559

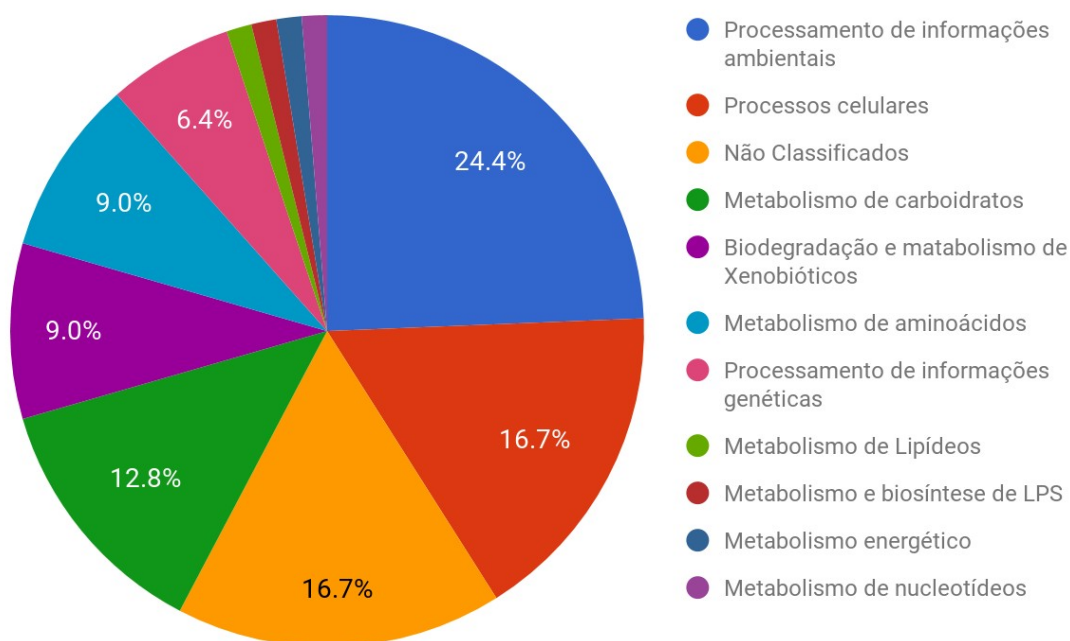


Fonte: O Autor, 2018.

Nota: Representação circular do genoma de *H. frisingense* AU14559. Os anéis da figura representam, do mais interno para o mais externo, o genoma da estirpe *H. frisingense* AU14559, conteúdo GC%, GC skew, presença de respectivos genes ortólogos em *H. frisingense* GSF30, presença de respectivos genes ortólogos em *H. frisingense* VT-16-41, e as marcações R1-R24 representam as ilhas encontradas pelo algoritmo *AlienHunter* e confirmadas através de visualização no ACT

Os genes presentes nestas ilhas foram então classificados pelo Blast KOALA, tendo 57 genes, dos 243 totais, anotados, sendo as porções mais significativas para os genes relacionados a processamento de informações ambientais, processos celulares e metabolismo de carboidratos.

GRÁFICO B 1: GRÁFICO DE FUNÇÕES DE GENES PRESENTES NAS ILHAS DO GENOMA DE *H. frisingense* AU14559



Fonte: O Autor, 2018

Nota: Gráfico construído a partir da predição da função dos genes presentes nas ilhas genômicas de *H. frisingense* AU14559 a partir do algoritmo blast KOALA.

6.4.3 Comparações entre os genomas GSF30 e AU14559

80,48% das proteínas de *AU14559* possuem ortólogos em *GSF30* com um *cut-off* de 50% de identidade em análises de blastp recíproco. Alguns destes genes são elementos chave para a forma de vida destes organismos.

O operon *wss* (*Hfc_11920-Hfc_12000*), composto por 9 genes, é encontrados nos dois organismos e codificam proteínas relacionadas à síntese e degradação de celulose. Estes genes são importantes para a colonização de plantas por patógenos vegetais, no entanto também estão envolvidos na formação de biofilme e criação de agregado de bactérias (MONTEIRO et al., 2012; SPIERS et al., 2003). A formação de biofilme é um ponto importante na sobrevivência de bactérias patogênicas no organismo humano, já que este atua como um escudo contra as defesas imunológicas e os antibióticos durante a infecção no hospedeiro (COSTERTON; STEWART; GREENBERG, 1999). Os dados de infecções humanas

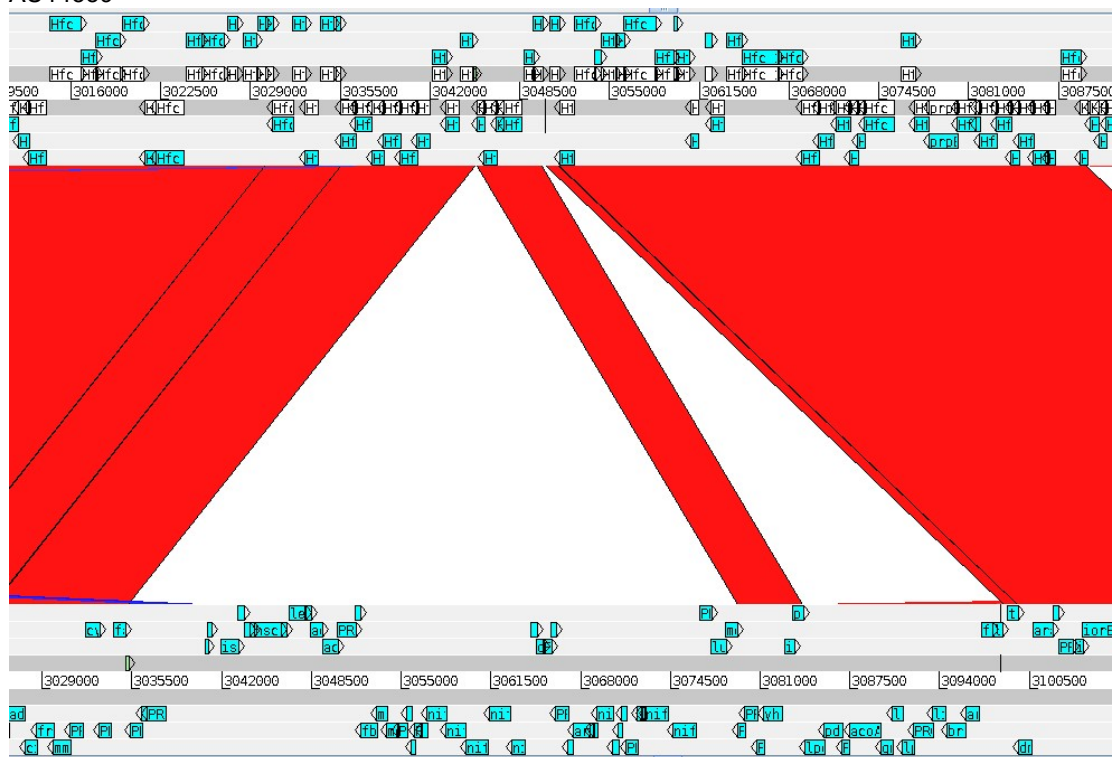
relacionadas a produção de biofilme variam de entre as fontes, sendo relatado uma taxa de aproximadamente 65% de acordo com o *Center for Disease Control* (CDC) e 80% de acordo com o *National Institutes of Health* (NIH) (JOO; OTTO, 2013).

O gene *Hfc_35550* codifica para a proteína beta-1,2-glucana cíclica sintase. Esta proteína é responsável por criar glucanas cíclicas, uma estrutura de homopolissacarídeo cíclico formada por moléculas de D-glucose (BOHIN, 2000). Estas proteínas estão relacionadas com a interação bactéria-plantas e fatores de virulência de patógenos humanos (RIGANO et al., 2007b). Em *Brucella abortus* este gene é necessário para a sobrevivência intracelular do patógeno, permitindo replicação intracelular e evasão da fusão de lisossomos (ARELLANO-REYNOSO et al., 2005). Este gene também é encontrado em *Salmonella typhi*, *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Yersinia pestis*, *Neisseria gonorrhoeae* (BOHIN, 2000) e em *Herbaspirillum seropedicae* AU14040, *H. frisingense* VT16-41, ambos isolados de humanos, como também em GSF30, isolado ambiental da espécie.

Os sistemas de secreção das estirpes GSF30 e AU14559 são idênticos, possuindo os sistemas tipo I, VI, Sec-SRP e Tat (*Twin-arginine translocation*) e a ausência do sistema de secreção tipo IV, relacionado à patogenicidade, atuando principalmente na exportação de elementos genéticos e moléculas efetoras para células eucarióticas (CASCALES; CHRISTIE, 2003).

O operon *nif* codifica várias proteínas que atuam juntas na montagem dos componentes da nitrogenase, responsável por converter nitrogênio atmosférico (N₂) em amônia (NH₃), através do processo de fixação biológica de nitrogênio (POZA-CARRIÓN et al., 2014). Entretanto, o operon *nif* é encontrado apenas na estirpe ambiental GSF30 estando ausente nas estirpes clínicas AU14559 e VT-16-41, o que corrobora com o estilo de vida destas bactérias (figura B1).

FIGURA B 2: REPRESENTAÇÃO GRÁFICA DA PERDA DO OPERON *nif* EM *H. frisingense* AU14559



Fonte: O Autor, 2018.

Nota: Representação gráfica do programa ACT, revelando a perda do operon de genes *nif* no organismo *H. frisingense* AU14559 (genoma superior) em relação a estirpe ambiental GSF30 (genoma inferior). Os blocos vermelhos indicam regiões de identidade superior a 50% entre os genomas.

O genoma de *AU14559* apresenta cópias extras de genes relacionados com a aquisição de nutrientes, como do complexo *TonB* (*Hfc_47470*, *Hfc_47480* e *Hfc_47550*), relacionado com a aquisição de ferro e essencial para a sobrevivência de patógenos como *Neisseria meningitidis* (HAGEN; CORNELISSEN, 2006) e *Neisseria gonorrhoeae* (LARSON et al., 2002) no organismo humano, contribuindo para a persistência da infecção (HAGEN; CORNELISSEN, 2006). Este complexo também foi encontrada em *VT-16-41*, sendo citada pelo autor do trabalho referente a esta estirpe como de importante característica deste organismo (TETZ; TETZ, 2017).

A proteína hipotética codificada pelo gene *Hfc_35570* possui várias repetições do domínio *Se11*. Proteínas com repetições deste domínio são solenóides e com conformação α -hélice em bactérias. Estas proteínas parecem estar associadas com a montagem de complexos macromoleculares e por vezes envolvidos na interação

com células eucarióticas (MITTL; SCHNEIDER-BRACHERT, 2007). Exemplos de proteínas com repetições *Sel1* atuando na interação com células eucarióticas são encontrados nos produtos dos genes *EnhC* e *LpnE* de *Legionella pneumophila* e estão associados com a entrada destas bactérias em monócitos humanos e células de linhagens epiteliais (MITTL; SCHNEIDER-BRACHERT, 2007).

6.4.3.1 Fatores de virulência

Através de alinhamentos com o banco de dados VFDB foi possível identificar 71 proteínas que, possivelmente, são envolvidas de alguma forma na virulência de *GSF30*, e 73 proteínas em *AU14559*, dentre as quais 4 são exclusivas da estirpe clínica (Tabela B 2). Os dados completos de fatores de virulência encontrados nestes organismos se encontram nas tabelas BS 15 e BS 16 do material suplementar.

TABELA B 2: GENES DE VIRULÊNCIA EXCLUSIVOS DA ESTIRPE CLÍNICA AU14559

Gene <i>AU14559</i>	em Gene Banco de dados	Produto	Organismo	Identidade (%)
Hfc_07430	WcbC	capsular polysaccharide biosynthesis/export protein	Burkholderia pseudomallei K96243	50,41
Hfc_10100	tsr	methyl-accepting chemotaxis protein I	Burkholderia pseudomallei K96243	57,14
Hfc_10480	tsr	methyl-accepting chemotaxis protein I	Burkholderia pseudomallei K96243	65,81
Hfc_45880	cheD	methyl-accepting chemotaxis protein CheD	Yersinia enterocolitica subsp. enterocolitica 8081	58,08

Fonte: O Autor, 2018

Dentre os fatores de virulência encontrados nos dois organismos podemos citar hemolisinas (*Hfc_28310*), proteínas que normalmente atuam contra eritrócitos (RAMARAO; SANCHIS, 2013), além de outras proteínas que podem atuar na defesa das células bacterianas durante a infecção no hospedeiro como uma catalase peroxidase (*Hfc_11490*) que possui 60,52% de identidade com a enzima bifuncional catalase-peroxidase codificada pelo gene *katA* de *Legionella pneumophila*, atuando, principalmente, na fase estacionária desta bactéria, como

antioxidante, na resposta ao *stress* por falta de nutrientes, além de ser de grande importância durante a infecção destas bactérias em macrófagos (BANDYOPADHYAY; STEINMAN, 2000) e a superóxido dismutase (*Hfc_10630*), com 64,4% de identidade com o produto do gene *sodB*, também de *L. pneumophila*, normalmente produzida por bactérias para neutralizar espécies reativas de oxigênio, mas que também possuem importante papel na infecção de várias bactérias patogênicas, principalmente quanto infectando intracelularmente seu hospedeiro, normalmente em células fagocitárias, necessitando assim de proteção contra as substâncias altamente oxidativas produzidas por estas células (LYNCH; KURAMITSU, 2000). Entretanto a grande maioria das proteínas encontradas por esta análise foram relativas a quimiotaxia e flagelo, com um total de 31 proteínas. Estas proteínas são normalmente associadas à motilidade bacteriana, no entanto, quando a bactéria atinge a célula hospedeira, utiliza diretamente a flagelina para aderir e colonizar a superfície das células epiteliais e em muitas bactérias patogênicas, estas proteínas têm demonstrado função de adesinas (HAIKO; WESTERLUND-WIKSTRÖM, 2013).

6.4.3.2 Fatores de interação patógeno-hospedeiro

Através de alinhamentos com o banco de dados de interação patógeno-hospedeiro PHI-base foi possível identificar 97 proteínas possivelmente envolvidas na interação com o hospedeiro em *GSF30* (Tabela BS 18) e 98 em *AU14559* (Tabela BS17), das quais 11 são exclusivas desta estirpe (Tabela B 3).

TABELA B 3: GENES DE INTERAÇÃO PATÓGENO HOSPEDEIRO DA ESTIRPE *H. frisingense* AU14559

Gene <i>AU14559</i>	em dados	Gene Banco de dados	Produto	Organismo	Identidade (%)
Hfc_10780	<i>nucB</i>		DNase <i>nucB</i>	<i>Ralstonia solanacearum</i>	52,43
Hfc_18920	<i>Trr1</i>		Thioredoxin reductase	<i>Cryptococcus neoformans</i>	53,12
Hfc_29400	<i>argGH</i>		argininosuccinate	<i>Streptococcus</i>	50,14

Gene em <i>AU14559</i>	Gene Banco de dados	Produto	Organismo	Identidade (%)
		synthase	<i>pneumoniae</i>	
Hfc_43080	<i>MGG_00383</i>	S-adenosyl methionine synthase	<i>Magnaporthe oryzae</i>	55,73
Hfc_44040	<i>PmrA</i>	<i>two-component system</i>	<i>Coxiella burnetii</i>	54,32
Hfc_44920	<i>BaTdp</i>	Toll/IL-1 receptor	<i>Bacillus anthracis</i>	61,00
Hfc_46380	<i>MqsRA</i>	toxin-antitoxin system	<i>Xylella fastidiosa</i>	55,56
Hfc_47840	URE1	urease	<i>Cryptococcus neoformans</i>	56,00
Hfc_47860	URE1	urease	<i>Cryptococcus neoformans</i>	59,30
Hfc_48420	<i>lipA</i>	<i>Lipoyl synthase</i>	<i>Acinetobacter baumannii</i>	61,84
Hfc_48540	<i>opgG</i>	<i>periplasmic glucan biosynthesis protein</i>	<i>Dickeya dadantii</i>	57,70

Fonte: O Autor, 2018

Apesar de *AU14559* ser um isolado clínico, este também apresenta entre seus genes exclusivos, genes de interação provenientes de bactérias que realizam interação com plantas como os *nucB*, *MGG_00383*, *MqsRA* e *opgG*, o que pode ser resultado de uma recente migração desta bactéria para o organismo humano ou uma dualidade no seu modo de vida, podendo habitar tanto a rizosfera, quanto o corpo humano.

6.4.3.3 Resistência a antibióticos

O genoma da estirpe *AU14559* apresentou, através das buscas realizadas através do perfil HMM no banco de dados de genes de resistência a antibióticos ResFams, um total de 25 genes (2 vezes o HMM colocou o mesmo gene em duas classificações diferentes) de resistência, sendo a maioria destes de resistência a beta-lactâmicos. Já os genomas das estirpes *VT-16-41* e *GSF30* apresentaram um total de 31 e 29 genes de resistência respectivamente. As principais diferenças entre os genes de resistência encontradas em *AU14559* e nas outras duas depositadas no banco de dados foi a redução dos genes envolvidos na resistência a Nodulation Cell division (RND) e a ausência de genes envolvidos nas

vias de acetiltransferase de Cloranfenicol e MexW-MexI (Tabela B4). Os *locus* dos genes encontrados nesta análise encontram-se nas tabelas BS 19, BS20 e BS21 presentes no material suplementar.

TABELA B 4: GENES DE RESISTÊNCIA A ANTIBIÓTICOS DA ESTIRPE AU14559

Descrição	ResFam Name	Family	ResFam ID	Número de genes encontrados nos organismos		
				AU14559	VT-16-41	GSF30
ATP binding cassette (ABC) - bomba de efluxo de antibióticos	ABC	antibiotic efflux pump	RF0007	6	6	6
Fator de membrana externa de complexo de efluxo de drogas	AdeC-AdeK-OprM		RF0021	1	2	2
Beta lactamases de Classe A	Class A		RF0053	2	2	2
Beta lactamases de Classe B	Class B		RF0054	1	1	1
Beta Lactamase Exo (Classe A)	Exo		RF0072	1	1	1
Subunidade de bomba de efluxo que confere resistência a antibióticos	Mac B		RF0089	2	2	2
Proteína de fusão de membrana do complexo de efluxo multidrogas MexEF-OprN	Mex E		RF0098	2	3	2
ATP binding cassette (ABC) - bomba de efluxo de antibióticos	MsbA		RF0107	1	1	1
resistance-nodulation-cell division (RND) - Bomba de efluxo	RND efflux		RF0115	6	9	8
(Metallo-) Beta lactamases de subclasse B1 - Hidrolizam penicilinas, cefalosporinas e carbapenêmicos	Subclass B1		RF0123	1	1	1
(Metallo-) Beta lactamases de subclasse B2 - seletivas a carbapenêmicos	Subclass B2		RF0124	1	1	1
Cloranfenicol acetil-transferase (CAT)	Chlor_Acetyltrans CAT		RF0050	0	0	1
Agrupamento de subunidades mexal e mexi relacionadas de bombas de efluxo que confere resistência a antibióticos	MexW-MexI		RF0100	0	1	0
Total				25	31	29

6.5 DISCUSSÃO

A estirpe *AU14559* de *Herbaspirillum frisingense* apresenta diferenças em seu genoma, quando comparada ao isolado ambiental *H. frisingense GSF30*, como a perda dos genes *nif*, além da aquisição de outros genes, como a adenylate kinase e proteínas com repetições do domínio *Sel1*, que podem estar envolvidas na interação e evasão do sistema imune humano, proteínas de aquisição de minerais, fatores de virulência e interação com o hospedeiro. No entanto, também apresenta várias características em comum com a estirpe ambiental e que podem estar envolvidos em sua resiliência no hospedeiro, como o fato de os sistemas de secreção das duas estirpes serem idênticos, a maioria de seus fatores de virulência e interação, e a presença de genes como o que codifica para a proteína responsável pela síntese da beta-1,2-glucana cíclica, envolvida na evasão do sistema imune humano e requisito para sobrevivência intracelular de outros patógenos, e o operon *wss*, possivelmente envolvidos na produção de biofilme, uma importante estratégia de patógenos para a sobrevivência no organismo humano. A estirpe *AU14559* também apresenta menor número de genes de resistência a antibióticos (25) do que a *GSF30* (29).

Diante dos dados apresentados, devemos avaliar a possibilidade de a espécie *H. frisingense* possuir um potencial de colonizar e infectar diferentes hospedeiros. Isto pode ocorrer pelo fato de que o ambiente “nativo” destas bactérias, a rizosfera, facilitar o surgimento de bactérias oportunistas. Este ambiente é um ambiente rico em nutrientes, devido aos exsudatos liberados pelas plantas, o que gera uma grande competição entre as bactérias presentes neste meio, atuando assim como um fator de seleção para as células com mecanismos de captação de nutrientes e minerais mais eficientes, o que é uma característica comum entre patógenos humanos. Além disto, este ambiente de competição, muitas vezes também torna necessário o desenvolvimento de mecanismos de defesa contra outros microrganismos, tanto na produção de defesas com a aquisição de genes de resistência a antibióticos, como no ataque, com a produção de substâncias antibióticas e infecção de outros organismos eucarióticos presentes neste ambiente como fungos e protozoários.

Estas características, aliadas aos genes de interação bactéria-planta, que também podem estar associados na interação bactéria-humano, podem estar contribuindo para os recentes casos de identificação de *H. frisingense* em amostras clínicas.

6.6 REFERÊNCIAS

- ALIKHAN, N. F. *et al.* BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. **BMC Genomics**, v. 12, n. 1, p. 402, 2011.
- ALTSCHUL, S. F. *et al.* Basic local alignment search tool. **Journal of molecular biology**, v. 215, n. 3, p. 403–10, 1990.
- ALTSCHUL, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997.
- ARELLANO-REYNOSO, B. *et al.* Cyclic beta-1,2-glucan is a Brucella virulence factor required for intracellular survival. **Nature immunology**, v. 6, n. 6, p. 618–625, 2005.
- BALDANI, J. I. *et al.* Characterization of Herbaspirillum seropedicae gen. nov. sp. nov. a Root-Associated Nitrogen-Fixing Bacterium. **International journal of systematic bacteriology**, p. 86–93, 1986.
- BANDYOPADHYAY, P.; STEINMAN, H. M. Catalase-peroxidases of Legionella pneumophila: Cloning of the katA gene and studies of KatA function. **Journal of Bacteriology**, v. 182, n. 23, p. 6679–6686, 2000.
- BANKEVICH, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455–477, 2012.
- BOHIN, J. P. Osmoregulated periplasmic glucans in Proteobacteria. **FEMS Microbiology Letters**, v. 186, n. 1, p. 11–19, 2000.
- CARVER, T. J. *et al.* ACT: The Artemis comparison tool. **Bioinformatics**, v. 21, n. 16, p. 3422–3423, 2005.
- CASCALES, E.; CHRISTIE, P. J. The versatile bacterial type IV secretion systems. **Nature reviews. Microbiology**, v. 1, n. 2, p. 137–49, 2003.
- CHEN, L. *et al.* VFDB: A reference database for bacterial virulence factors. **Nucleic Acids Research**, v. 33, n. DATABASE ISS., p. 325–328, 2005.
- COSTERTON, J. W.; STEWART, P. S.; GREENBERG, E. P. Bacterial biofilms: A common cause of persistent infections. **Science**, v. 284, n. 5418, p. 1318–1322, 1999.
- FINN, R. D. *et al.* The Pfam protein families database. **Nucleic Acids Res 40: D290-D301.**, v. 38, n. 1, 2010.
- GIBSON, M. K.; FORSBERG, K. J.; DANTAS, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. **ISME**

Journal, v. 9, n. 1, p. 207–216, 2015.

GORIS, J. *et al.* DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. **International Journal of Systematic and Evolutionary Microbiology**, v. 57, n. 1, p. 81–91, 2007.

GUIZELINI, D. *et al.* Complete Genome Sequence of *Herbaspirillum hiltneri* N3 (DSM 17495), Isolated from Surface-Sterilized Wheat Roots. **Genome Announcements**, v. 3, n. 5, p. e01288-15, 2015.

_____. GFinisher: A new strategy to refine and finish bacterial genome assemblies. **Scientific Reports**, v. 6, n. October, p. 1–8, 2016.

GUREVICH, A. *et al.* QUAST: Quality assessment tool for genome assemblies. **Bioinformatics**, v. 29, n. 8, p. 1072–1075, 2013.

HAGEN, T. A.; CORNELISSEN, C. N. *Neisseria gonorrhoeae* requires expression of TonB and the putative transporter TdfF to replicate within cervical epithelial cells. **Molecular Microbiology**, v. 62, n. 4, p. 1144–1157, 2006.

HAIKO, J.; WESTERLUND-WIKSTRÖM, B. The Role of the Bacterial Flagellum in Adhesion and Virulence. **Biology**, v. 2, n. 4, p. 1242–1267, 2013.

JOO, H.-S.; OTTO, M. Molecular basis of in-vivo biofilm formation by bacterial pathogens. **Chem Biol.**, v. 19, n. 12, p. 1503–1513, 2013.

KIRCHHOF, G. *et al.* *Herbaspirillum frisingense* sp. nov., a new nitrogen-fixing bacterial species that occurs in C4-fibre plants. **International Journal of Systematic and Evolutionary Microbiology**, v. 51, n. 1, p. 157–168, 2001.

KLASSEN, G. *et al.* Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae* SMRI. v. 891, p. 887–891, 1997.

LARSON, J. A. *et al.* Replication of *Neisseria meningitidis* within Epithelial Cells Requires TonB-Dependent Acquisition of Host Cell Iron Replication of *Neisseria meningitidis* within Epithelial Cells Requires TonB-Dependent Acquisition of Host Cell Iron. v. 70, n. 3, p. 1461–1467, 2002.

LYNCH, M.; KURAMITSU, H. Expression and role of superoxide dismutases (SOD) in pathogenic bacteria. **Microbes and infection / Institut Pasteur**, v. 2, n. 10, p. 1245–1255, 2000.

MITTL, P. R. E.; SCHNEIDER-BRACHERT, W. Sel1-like repeat proteins in signal transduction. **Cellular Signalling**, v. 19, n. 1, p. 20–31, 2007.

MONTEIRO, R. A. *et al.* Genomic comparison of the endophyte *Herbaspirillum seropedicae* SmR1 and the phytopathogen *Herbaspirillum rubrisubalbicans* M1 by suppressive subtractive hybridization and partial genome sequencing. **FEMS Microbiology Ecology**, v. 80, n. 2, p. 441–451, 2012.

MORIYA, Y. *et al.* KAAS: An automatic genome annotation and pathway reconstruction server. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. 182–185,

2007.

NIU, B.; KOLTER, R. crossm Complete Genome Sequences of Seven Strains Composing a Model Bacterial Community of Maize Roots. **Genome Announcements**, v. 5, n. 36, p. 35–36, 2017.

PEDROSA, F. O. *et al.* Genome of herbaspirillum seropedicae strain SmR1, a specialized diazotrophic endophyte of tropical grasses. **PLoS Genetics**, v. 7, n. 5, 2011.

PIRO, V. C. *et al.* FGAP: an automated gap closing tool. **BMC research notes**, v. 7, n. 1, p. 371, 2014.

POZA-CARRIÓN, C. *et al.* Kinetics of nif gene expression in a nitrogen-fixing bacterium. **Journal of Bacteriology**, v. 196, n. 3, p. 595–603, 2014.

RAMARAO, N.; SANCHIS, V. The pore-forming haemolysins of bacillus cereus: a review. **Toxins**, v. 5, n. 6, p. 1119–1139, 2013.

REGUNATH, H. *et al.* Severe community-acquired pneumonia with bacteremia caused by Herbaspirillum aquaticum or Herbaspirillum huttiense in an immune-competent adult. **Journal of Clinical Microbiology**, v. 53, n. 9, p. 3086–3088, 2015.

RIGANO, L. A. *et al.* Bacterial Cyclic β -(1,2)-Glucan Acts in Systemic Suppression of Plant Immune Responses. **The Plant Cell**, v. 19, n. 6, p. 2077–2089, 2007.

SEEMANN, T. Prokka: Rapid prokaryotic genome annotation. **Bioinformatics**, v. 30, n. 14, p. 2068–2069, 2014.

SOARES, S. C. *et al.* GIPSy: Genomic island prediction software. **Journal of Biotechnology**, v. 232, p. 2–11, 2016.

SOUZA, V. DE *et al.* Draft Genome Sequence of Herbaspirillum huttiense subsp. putei IAM 15032, a Strain Isolated from Well Water. **Genome Announcements**, v. 1, n. 1, p. e00252-12-e00252-12, 2013.

SPIERS, A. J. *et al.* Biofilm formation at the air-liquid interface by the Pseudomonas fluorescens SBW25 wrinkly spreader requires an acetylated form of cellulose. **Molecular Microbiology**, v. 50, n. 1, p. 15–27, 2003.

SPIPKER, T. *et al.* Recovery of Herbaspirillum species from persons with cystic fibrosis. **Journal of Clinical Microbiology**, v. 46, n. 8, p. 2774–2777, 2008.

STRAUB, D. *et al.* The genome of the endophytic bacterium H. frisingense GSF30T identifies diverse strategies in the Herbaspirillum genus to interact with plants. **Frontiers in Microbiology**, v. 4, n. JUN, p. 1–10, 2013.

SUWANTARAT, N. *et al.* Fatal case of Herbaspirillum seropedicae bacteremia secondary to pneumonia in an end-stage renal disease patient with multiple myeloma. **Diagnostic Microbiology and Infectious Disease**, v. 82, n. 4, p. 331–333, 2015.

TETZ, V.; TETZ, G. Draft Genome Sequence of the. **Genome Announcements**, v. 1, n. 1, p. e00032-13, 2017.

URBAN, M. *et al.* The Pathogen-Host Interactions database (PHI-base): Additions and future developments. **Nucleic Acids Research**, v. 43, n. D1, p. D645–D655, 2015.

VERNIKOS, G. S.; PARKHILL, J. Interpolated variable order motifs for identification of horizontally acquired DNA: Revisiting the *Salmonella* pathogenicity islands. **Bioinformatics**, v. 22, n. 18, p. 2196–2203, 2006.

6.7 MATERIAL SUPLEMENTAR

TABELA BS 1: GENES PRESENTES NA ILHA GENÔMICA 1 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_10240	NmrA family protein 1131834:1132730 forward MW:33238
Hfc_10250	LysR family transcriptional regulator 1133032:1133937 reverse MW:33399
Hfc_10260	chemotaxis protein 1134250:1135884 forward MW:57436
Hfc_10270	spermidine/putrescine transport system ATP-binding protein 1136529:1137581 forward MW:38067
Hfc_10280	spermidine/putrescine transport protein, ABC transporter permease protein 1137578:1138438 forward MW:30962
Hfc_10290	ABC transporter membrane spanning protein (polyamine) 1138435:1139229 forward MW:28352
Hfc_10300	FAD dependent oxidoreductase 1139226:1140335 forward MW:39865
Hfc_10310	putative sarcosine oxidase subunit alpha 1140328:1140639 forward MW:11130
Hfc_10320	BFD-like protein 1140636:1142045 forward MW:49412
Hfc_10330	endoribonuclease L-PSP 1142035:1142400 forward MW:12777
Hfc_10340	2-methylcitrate dehydratase 1142530:1143879 forward MW:47100
Hfc_10350	extracellular solute-binding protein 1144181:1145173 forward MW:35000

Fonte: O Autor, 2018

TABELA BS 2: GENES PRESENTES NA ILHA GENÔMICA 2 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_10430	AraC family transcriptional regulator 1152263:1153105 forward MW:31436
Hfc_10440	low specificity L-threonine aldolase 1153228:1154235 forward MW:35540
Hfc_10450	2,5-didehydrogluconate reductase B 1154287:1155108 reverse MW:29713
Hfc_10460	amino acid ABC transporter periplasmic protein 1155493:1156290 forward MW:27912
Hfc_10470	hypothetical protein 1156789:1157181 reverse MW:13984
Hfc_10480	methyl-accepting chemotaxis protein 1157254:1158069 forward MW:27491
Hfc_10490	transcriptional regulator 1158066:1158995 reverse MW:34311
Hfc_10500	Major Facilitator Superfamily transporter 1159102:1160388 forward MW:45615
Hfc_10510	amidohydrolase 1160492:1161697 forward MW:43433

Fonte: O Autor, 2018

TABELA BS 3: GENES PRESENTES NA ILHA GENÔMICA 3 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_21240	integrase 2328809:2330026 forward MW:46010
Hfc_21241	anti-repressor protein 2330178:2330588 forward MW:14708
Hfc_21242	hypothetical protein 2331631:2332092 reverse MW:17212
Hfc_21250	hypothetical protein 2332105:2332857 reverse MW:26880
Hfc_21260	hypothetical protein 2332913:2333452 reverse MW:21012
Hfc_21270	hypothetical protein 2333445:2334275 reverse MW:31151
Hfc_21271	hypothetical protein 2334566:2334847 forward MW:10267
Hfc_21290	hypothetical protein 2335133:2335474 forward MW:12116
Hfc_21300	hypothetical protein 2335471:2335743 forward MW:9655
Hfc_21310	hypothetical protein 2335740:2336111 forward MW:13807
Hfc_21320	hypothetical protein 2336195:2336815 forward MW:22732
Hfc_21330	hypothetical protein 2337029:2337787 forward MW:27562
Hfc_21331	hypothetical protein 2337844:2338245 forward MW:14848
Hfc_21340	hypothetical protein 2338495:2338770 reverse MW:10359
Hfc_21350	hypothetical protein 2338899:2339369 reverse MW:17484
Hfc_21360	hypothetical protein 2339773:2340246 forward MW:17279
Hfc_21370	hypothetical protein 2340310:2340552 forward MW:9165
Hfc_21380	Phage DNA binding protein Roi 2340564:2341322 forward MW:27795
Hfc_21390	Hypothetical Protein 2341322:2342110 forward MW:29109
Hfc_21400	dnaB replicative DNA helicase protein 2342107:2343477 forward MW:49698
Hfc_21410	hypothetical protein 2343761:2344546 forward MW:28285
Hfc_21420	hypothetical protein 2344896:2345843 forward MW:32053
Hfc_21430	hypothetical protein 2345856:2346038 forward MW:6787
Hfc_21440	hypothetical protein 2346102:2346719 forward MW:22945
Hfc_21441	hypothetical protein 2347121:2347354 forward MW:8207

Fonte: O Autor, 2018

TABELA BS 4: GENES PRESENTES NA ILHA GENÔMICA 4 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_28080	putative flavoprotein 3089777:3090340 reverse MW:21157
Hfc_28090	arsenate reductase (glutaredoxin) 3090480:3090908 reverse MW:15638
Hfc_28100	arsenical efflux pump membrane protein ArsB 3090935:3092215 reverse

Locus	Produto
Hfc_28110	MW:44933 transcriptional regulator 3092228:3092575 reverse MW:12087
Hfc_28120	oxidoreductase domain-containing protein 3092982:3093869 forward MW:31039
Hfc_28130	arginine ABC transporter ATP-binding protein 3094095:3094829 reverse MW:26891
Hfc_28140	glutamate/aspartate transporter permease GltK 3094826:3095509 reverse MW:25154
Hfc_28150	glutamate/aspartate ABC transporter permease GltJ 3095521:3096246 reverse MW:26402
Hfc_28160	4-carboxymuconolactone decarboxylase 3096268:3096843 reverse MW:20916
Hfc_28170	ABC transporter 3096882:3097799 reverse MW:33714
Hfc_28180	aspartate aminotransferase family protein 3097839:3099161 reverse MW:47791
Hfc_28190	DNA-binding protein 3099476:3100375 reverse MW:32607
Hfc_28200	hypothetical protein 3100896:3101192 forward MW:10973
Hfc_28210	putative esterase of the alpha/beta hydrolase 3101244:3101804 reverse MW:20079
Hfc_28220	LysR family transcriptional regulator 3103111:3104055 reverse MW:33856
Hfc_28230	alpha/beta hydrolase 3104168:3104962 forward MW:28441
Hfc_28240	hypothetical protein 3105833:3106456 forward MW:20994
Hfc_28250	thioredoxin reductase 3106467:3107081 forward MW:22377
Hfc_28260	molybdopterin-binding protein 3107078:3107854 forward MW:29054
Hfc_28270	epimerase 3107868:3108917 reverse MW:37625

Fonte: O Autor, 2018

TABELA BS 5: GENES PRESENTES NA ILHA GENÔMICA 5 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_32610	hypothetical protein 3585974:3586117 reverse MW:5269
Hfc_32620	gamma-glutamylcyclotransferase 3586332:3586931 forward MW:22310
Hfc_32630	NAD-GH domain containing protein 3586971:3587390 forward MW:14863
Hfc_32640	lysophospholipase 3588311:3589159 reverse MW:30741
Hfc_32650	MFS transporter 3589156:3590577 reverse MW:52634
Hfc_32660	transmembrane protein 3590635:3591504 forward MW:30945
Hfc_32670	chemotaxis protein 3591731:3592303 forward MW:20735

Locus	Produto
Hfc_32680	chemotaxis protein 3592255:3593565 forward MW:44209
Hfc_32690	metal dependent phosphohydrolase 3593638:3594525 reverse MW:31682
Hfc_32700	hypothetical protein 3594588:3595229 reverse MW:23079
Hfc_32710	RNA polymerase sigma factor 3595226:3595855 reverse MW:23499
Hfc_32720	hypothetical protein 3595889:3596554 reverse MW:23263
Hfc_32730	integral membrane protein 3596798:3597076 forward MW:9425
Hfc_32740	hypothetical protein 3597088:3597939 forward MW:31050
Hfc_32750	hypothetical protein 3597936:3598706 forward MW:27601

Fonte: O Autor, 2018

TABELA BS 6: GENES PRESENTES NA ILHA GENÔMICA 6 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_33850	LysR family transcription regulator protein 3730489:3731388 reverse MW:33073
Hfc_33860	sugar ABC transporter 3731549:3732094 forward MW:20229
Hfc_33870	hypothetical protein 3732099:3732425 forward MW:12184
Hfc_33880	Zn-dependent beta-lactamase domain hydrolase 3732453:3733283 forward MW:29910
Hfc_33890	alpha-amylase 3734288:3737653 forward MW:125373
Hfc_33900	alpha-amylase 3737655:3741176 forward MW:132200
Hfc_33910	glycogen-branching enzyme 3741242:3743683 forward MW:90607
Hfc_33920	2-nitropropane dioxygenase 3743907:3744983 reverse MW:37394

Fonte: O Autor, 2018

TABELA BS 7: GENES PRESENTES NA ILHA GENÔMICA 7 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_35520	PHB depolymerase family esterase 3932019:3933050 forward MW:37146
Hfc_35530	tri-BON domain-containing protein 3933360:3934010 forward MW:23334
Hfc_35540	hypothetical protein 3934954:3935172 forward MW:8691
Hfc_35541	hypothetical protein 3935236:3935538 forward MW:11674
Hfc_35542	hypothetical protein 3936051:3936248 forward MW:7184
Hfc_35543	hypothetical protein 3936440:3936676 forward MW:9206
Hfc_35544	hypothetical protein 3936751:3937056 forward MW:11542
Hfc_35550	Cyclic beta-1,2-glucan synthase 3937410:3945968 forward MW:312869
Hfc_35560	hypothetical protein 3946656:3947258 forward MW:22467

Locus	Produto
Hfc_35570	SEL1 subfamily TPR repeat containing protein 3947261:3948496 forward MW:44344
Hfc_35580	esterase/lipase 3948885:3949811 reverse MW:33648
Hfc_35590	putative flavoprotein 3950154:3950717 reverse MW:21230
Hfc_35600	arsenate reductase (glutaredoxin) 3950857:3951285 reverse MW:15760
Hfc_35610	arsenical efflux pump membrane protein ArsB 3951313:3952593 reverse MW:45061
Hfc_35620	transcriptional regulator 3952607:3952942 reverse MW:11686

Fonte: O Autor, 2018

TABELA BS 8: GENES PRESENTES NA ILHA GENÔMICA 8 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_40650	phage-related integrase 4510357:4511715 forward MW:50196
Hfc_40660	hypothetical protein 4511800:4512018 forward MW:7832
Hfc_40670	hypothetical protein 4512247:4512972 forward MW:27071
Hfc_40680	hypothetical protein 4513087:4514292 forward MW:41336
Hfc_40681	hypothetical protein 4514411:4514602 forward MW:6771
Hfc_40682	hypothetical protein 4514599:4514973 forward MW:13820
Hfc_40683	hypothetical protein 4514966:4515655 forward MW:25768
Hfc_40684	hypothetical protein 4515645:4516109 forward MW:17404
Hfc_40690	hypothetical protein 4516106:4516531 forward MW:15050
Hfc_40700	hypothetical protein 4516528:4516785 forward MW:9440
Hfc_40710	Hypothetical Protein 4516936:4517616 reverse MW:24201
Hfc_40711	hypothetical protein 4517613:4517822 reverse MW:7919
Hfc_40730	hypothetical protein 4518150:4518842 reverse MW:25925
Hfc_40740	hypothetical protein 4518854:4519540 reverse MW:25733
Hfc_40750	hypothetical protein 4519527:4520216 reverse MW:25816
Hfc_40760	hypothetical protein 4520640:4522364 reverse MW:64595
Hfc_40770	chromosome partitioning protein ParA 4522364:4523824 reverse MW:53865
Hfc_40780	DNA (cytosine-5-)-methyltransferase 4523811:4525283 reverse MW:54237
Hfc_40790	hypothetical protein 4525294:4525395 reverse MW:3907
Hfc_40800	bacteriophage capsid portal protein 4526185:4526556 reverse MW:13694
Hfc_40801	hypothetical protein 4526879:4527190 forward MW:11788
Hfc_40810	hypothetical protein 4528329:4528574 reverse MW:9158
Hfc_40811	response regulator 4529618:4530025 reverse MW:14728

Fonte: O Autor, 2018

TABELA BS 9:GENES PRESENTES NA ILHA GENÔMICA 9 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_44350	GntR family transcriptional regulator 4882356:4882781 reverse MW:16586
Hfc_44360	aldo/keto reductase 4883124:4884164 reverse MW:37624
Hfc_44370	siderophore-interacting protein 4884174:4885412 reverse MW:46398
Hfc_44380	inosine-5'-monophosphate dehydrogenase 4885427:4886890 reverse MW:51438
Hfc_44390	methyl-accepting chemotaxis protein 4887514:4889490 reverse MW:69172
Hfc_44400	hypothetical protein 4889693:4890463 reverse MW:28848
Hfc_44410	aryldialkylphosphatase-related protein 4890643:4891986 reverse MW:48803
Hfc_44420	alpha/beta hydrolase 4892011:4892790 reverse MW:27105
Hfc_44430	5-carboxymethyl-2-hydroxyuconate delta-isomerase 4892790:4893668 reverse MW:31640
Hfc_44440	GntR family transcriptional regulator 4893829:4894521 forward MW:26051
Hfc_44450	putative exported protein 4894543:4895373 reverse MW:30171
Hfc_44460	4-carboxyuconolactone decarboxylase 4895693:4896094 forward MW:14104
Hfc_44470	major facilitator superfamily protein 4896129:4897328 forward MW:41606
Hfc_44480	transcription regulator protein 4897411:4898040 forward MW:23702
Hfc_44490	hypothetical protein, possibly involved in utilization of glycolate and propanediol 4898095:4898499 reverse MW:13738
Hfc_44500	FAD-binding protein 4898638:4899546 reverse MW:33301
Hfc_44510	AraC family transcriptional regulator 4899647:4900480 reverse MW:30765
Hfc_44520	transcriptional regulator LysR 4900577:4901470 reverse MW:32293
Hfc_44530	short-chain dehydrogenase 4901575:4902339 forward MW:26307
Hfc_44540	NADH:flavin oxidoreductase 4902655:4903761 reverse MW:41026
Hfc_44550	alpha/beta superfamily hydrolase 4904371:4905267 forward MW:32974
Hfc_44560	nanM N-acetylneuraminate epimerase 4905841:4906515 reverse MW:23486
Hfc_44570	transposase protein 4906555:4907271 reverse MW:26365
Hfc_44580	MFS transporter 4907894:4909093 reverse MW:41439
Hfc_44590	meta-pathway of phenol degradation protein 4909118:4910149 reverse MW:37309
Hfc_44600	AraC family transcription regulator protein 4910438:4911409 reverse MW:35679
Hfc_44610	transcription regulator protein 4911514:4912290 reverse MW:28126
Hfc_44620	glyoxalase 4912322:4913155 reverse MW:31068

Locus	Produto
Hfc_44630	3-(3-hydroxyphenyl)propionate hydroxylase 4913176:4914852 reverse MW:61289
Hfc_44640	5-carboxymethyl-2-hydroxymuconate isomerase 4914925:4915764 reverse MW:30238
Hfc_44650	LysR family transcriptional regulator 4915898:4916824 reverse MW:34996
Hfc_44660	arabinose efflux permease family protein 4917018:4918340 forward MW:46336
Hfc_44670	9-cis-epoxycarotenoid dioxygenase 4918382:4919827 reverse MW:53504
Hfc_44680	betaine-aldehyde dehydrogenase 4919860:4921338 reverse MW:52781
Hfc_44690	transcription regulator protein 4921551:4922150 forward MW:22002
Hfc_44700	hypothetical protein 4922156:4922980 forward MW:29773
Hfc_44710	benzaldehyde dehydrogenase (NAD) protein 4923890:4925389 reverse MW:53308
Hfc_44720	hypothetical protein 4926398:4926523 forward MW:4733
Hfc_44730	carbohydrate-selective porin OprB 4926625:4928166 forward MW:54528
Hfc_44740	major facilitator superfamily transporter 4928763:4930034 forward MW:43875
Hfc_44750	hypothetical protein 4930109:4931368 forward MW:45803
Hfc_44760	dehydrogenase of unknown specificity, short-chain alcohol dehydrogenase 4932180:4932953 reverse MW:26600
Hfc_44770	9-cis-epoxycarotenoid dioxygenase 4932955:4934424 reverse MW:55384
Hfc_44780	hypothetical protein 4934518:4935369 reverse MW:31480
Hfc_44790	putative Acyl-CoA dehydrogenase 4935366:4936601 reverse MW:44720
Hfc_44800	hypothetical protein 4938051:4938293 reverse MW:8925
Hfc_44810	hypothetical protein 4938698:4939525 reverse MW:31072
Hfc_44820	Hypothetical Protein 4939575:4941350 reverse MW:67506
Hfc_44821	hypothetical protein 4941555:4941833 forward MW:10037
Hfc_44830	hypothetical protein 4941901:4943001 forward MW:40950
Hfc_44840	tnsB Transposon Tn7 transposition protein TnsB 4942998:4945013 forward MW:76105
Hfc_44850	hypothetical protein 4945010:4946359 forward MW:50945
Hfc_44860	transcription regulator protein 4946778:4947149 reverse MW:14171
Hfc_44870	Zn-dependent hydrolase 4947290:4948273 forward MW:35044
Hfc_44880	transcription regulator protein 4948298:4949185 reverse MW:32267
Hfc_44890	MFS transporter 4949359:4950597 forward MW:44593
Hfc_44900	N-acyl-D-glutamate deacylase 4950630:4952096 forward MW:52821
Hfc_44910	aspartate tRNA ^{Asn} /Glu-tRNA ^{Gln} amidotransferase subunit A-like protein

Locus	Produto
Hfc_44920	4952147:4954102 forward MW:69073
Hfc_44930	hypothetical protein 4954750:4955493 reverse MW:26128
Hfc_44940	hypothetical protein 4956762:4957922 forward MW:44258
Hfc_44950	hypothetical protein 4958403:4959320 forward MW:34426
Hfc_44960	high-affinity gluconate transporter (GNT-III system) transmembrane protein 4961097:4962419 reverse MW:45412
Hfc_44970	gluconolactonase 4962783:4964060 reverse MW:46623
Hfc_44980	MFS transporter 4964222:4966264 forward MW:71801
Hfc_44990	dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex 4967382:4967939 forward MW:20257
Hfc_44990	dihydrolipoamide dehydrogenase 4968013:4970076 forward MW:72809

Fonte: O Autor, 2018

TABELA BS 10: GENES PRESENTES NA ILHA GENÔMICA 10 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_45140	enolase superfamily enzyme related to L-alanine-DL-glutamate epimerase 4986251:4987345 forward MW:40288
Hfc_45150	short-chain dehydrogenase 4987577:4988338 reverse MW:26584
Hfc_45160	hypothetical protein 4989038:4989706 reverse MW:25426
Hfc_45170	hypothetical protein 4990098:4991141 reverse MW:38126
Hfc_45180	dehydrogenase of unknown specificity, short-chain alcohol dehydrogenase 4991151:4991918 reverse MW:27220
Hfc_45190	oxidoreductase 4991954:4993138 reverse MW:42209
Hfc_45200	iolE Inosose dehydratase 4993151:4994206 reverse MW:39498
Hfc_45210	hypothetical protein 4994263:4994982 reverse MW:26287
Hfc_45220	3-oxoacyl-ACP synthase 4995063:4995842 reverse MW:27028
Hfc_45230	aspartate carbamoyltransferase 4996493:4997773 forward MW:47069
Hfc_45240	nucS Endonuclease NucS 4998767:4999507 reverse MW:27382
Hfc_45250	transposase 5000056:5000763 forward MW:26366
Hfc_45260	hypothetical protein 5001463:5003625 reverse MW:80305
Hfc_45270	hypothetical protein 5003641:5004093 reverse MW:16087
Hfc_45280	hypothetical protein 5004277:5004726 reverse MW:16562
Hfc_45290	hypothetical protein 5004716:5005645 reverse MW:33989
Hfc_45300	VGR-related protein 5005642:5008590 reverse MW:102808
Hfc_45310	hypothetical protein 5008665:5009537 reverse MW:29925
Hfc_45320	peptidase M15 5010144:5010983 forward MW:31273

Locus	Produto
Hfc_45330	hypothetical protein 5011070:5011327 forward MW:8309
Fonte: O Autor, 2018	

TABELA BS 11: GENES PRESENTES NA ILHA GENÔMICA 11 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_45540	Rhs element Vgr protein 5040711:5043557 forward MW:101425
Hfc_45550	hypothetical protein 5043590:5044591 forward MW:38565
Hfc_45560	hypothetical protein 5044609:5045478 forward MW:32831
Fonte: O Autor, 2018	

TABELA BS 12: GENES PRESENTES NA ILHA GENÔMICA 12 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_47420	hypothetical protein 5242999:5243457 reverse MW:16925
Hfc_47430	hypothetical protein 5243490:5244023 reverse MW:18857
Hfc_47440	parvulin-like peptidyl-prolyl isomerase 5244040:5244960 reverse MW:32470
Hfc_47450	hypothetical protein 5245045:5245710 reverse MW:24035
Hfc_47460	hypothetical protein 5245720:5247459 reverse MW:63235
Hfc_47470	hypothetical protein 5247533:5248237 reverse MW:25051
Hfc_47480	biopolymer transport ExbD/TolR protein 5248239:5248658 reverse MW:14921
Hfc_47490	protein tolQ 5248669:5250435 reverse MW:62415
Hfc_47500	hemolysin activation/secretion protein 5250485:5252161 reverse MW:60033
Hfc_47510	hypothetical protein 5252534:5252869 forward MW:11716
Hfc_47520	hypothetical protein 5252972:5265430 reverse MW:422324
Hfc_47530	transmembrane sensor protein 5265646:5266602 reverse MW:34001
Hfc_47540	RNA polymerase sigma factor, sigma-70 family 5266710:5267222 reverse MW:19279
Fonte: O Autor, 2018	

TABELA BS 13: GENES PRESENTES NA ILHA GENÔMICA 13 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_25710	Transcriptional regulator, GntR family / Aspartate aminotransferase 2840628:2842001 reverse MW:48694

Locus	Produto
Hfc_25720	pht5 4,5-dihydroxyphthalate decarboxylase 2842101:2842988 forward MW:33498
Hfc_25730	Putative amidase 2842988:2843809 forward MW:29691
Hfc_25740	Periplasmic binding protein 2843821:2844777 forward MW:33630
Hfc_25750	amine acid ABC transporter, permease protein, 3-TM region, His/Glu/Gln/Arg/opine family 2844774:2845697 forward MW:33397
Hfc_25760	ABC-type amino acid transport system, ATPase component protein 2845694:2846500 forward MW:29424

Fonte: O Autor, 2018

TABELA BS 14: GENES PRESENTES NA ILHA GENÔMICA 14 DE H. FRISINGENSE AU14559

Locus	Produto
Hfc_11730	outer membrane protein 1288796:1290472 forward MW:62275
Hfc_11740	RND family efflux transporter, MFP subunit 1290516:1291265 forward MW:26792
Hfc_11750	Hypothetical Protein 1291419:1297997 forward MW:217672
Hfc_11760	membrane-fusion protein 1297982:1299367 forward MW:51157
Hfc_11770	membrane-fusion protein 1299364:1301469 forward MW:78303
Hfc_11780	dehydrogenase 1301656:1302426 reverse MW:27817
Hfc_11790	arabinose efflux permease family protein 1302484:1303683 reverse MW:40830
Hfc_11800	hydroxy-atrazine ethyl amino hydrolase 1303774:1305102 reverse MW:47359

Fonte: O Autor, 2018

TABELA BS 15: GENES DE VIRULÊNCIA DA ESTIRPE H. FRISINGENSE AU14559

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hfc_02980	lpxC	UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	55.30
Hfc_03100	pilD	type IV pilus prepilin peptidase [Type IV pili (VF0075)]	Neisseria meningitidis MC58	50.38
Hfc_03100	xcpA/pilD	type 4 prepilin peptidase PilD [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	54.65
Hfc_03100	pilD	type IV prepilin leader peptide type M1 [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	50.76
Hfc_03120	pilF	type IV pilus assembly ATPase protein [Type IV pili (VF0075)]	Neisseria meningitidis MC58	56.30
Hfc_03120	pilB	type 4 fimbrial biogenesis protein PilB [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	56.02

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hfc_03120	pilB	type IV pilus assembly protein [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	57.11
Hfc_03120	pilB	type IV pilus assembly protein PilB [Type IV pili (VF0155)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	50.90
Hfc_03120	gspE	general secretion pathway protein E [T2SS (VF0333)]	Shigella dysenteriae Sd197	50.79
Hfc_04340	rffG	dTDP-glucose 46-dehydratase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	54.68
Hfc_04520	gmd	GDP-mannose 4,6-dehydratase [LPS (VF0367)]	Brucella melitensis bv. 1 str. 16M	75.14
Hfc_04520	gmd	GDP-mannose 4,6-dehydratase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	68.12
Hfc_04520	rfbD	GDP-D-mannose dehydratase [LPS (VF0056)]	Helicobacter pylori 26695	60.85
Hfc_04530	algC	phosphomannomutase AlgC [Alginate biosynthesis (CVF522)]	Pseudomonas aeruginosa PAO1	55.09
Hfc_04620	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Neisseria meningitidis MC58	54.55
Hfc_07420	wcbT	acyl-CoA transferase [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	60.58
Hfc_07430	wcbC	capsular polysaccharide biosynthesis/export protein [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	50.41
Hfc_07470	cps4I	UDP-N-acetylglucosamine-2-epimerase [Capsule (VF0144)]	Streptococcus pneumoniae TIGR4	53.80
Hfc_07470	cap8P	capsular polysaccharide synthesis enzyme Cap8P [Capsule (VF0003)]	Staphylococcus aureus subsp. aureus MW2	50.82
Hfc_08040	lpg2359	Dot/Icm type IV secretion system effector [Dot/Icm (VF0156)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	53.06
Hfc_08640	htpB	Hsp60, 60K heat shock protein HtpB [Hsp60 (VF0159)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	72.80
Hfc_10100	tsr	methyl-accepting chemotaxis protein I [Flagella (VF0430)]	Burkholderia pseudomallei K96243	57.14
Hfc_10480	tsr	methyl-accepting chemotaxis protein I [Flagella (VF0430)]	Burkholderia pseudomallei K96243	65.81
Hfc_10480	cheD	methyl-accepting chemotaxis protein CheD [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	62.82
Hfc_10630	sodB	superoxide dismutase [SodB (VF0169)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	64.40
Hfc_11490	katA	catalase/(hydro)peroxidase [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	60.52
Hfc_11490	katB	catalase-peroxidase KatB [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	57.05
Hfc_15470	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (A1145)]	Yersinia enterocolitica subsp. enterocolitica 8081	63.28
Hfc_15470	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	58.91
Hfc_19170	acpXL	acyl carrier protein [LPS (CVF383)]	Brucella melitensis bv. 1 str. 16M	61.11

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hfc_19200	algU	alginate biosynthesis protein AlgZ/FimS [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	64.36
Hfc_19630	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	60.94
Hfc_19630	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.98
Hfc_19920	manC	GDP-mannose pyrophosphorylase [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	57.32
Hfc_19920	algA	phosphomannose isomerase / guanosine 5'-diphospho-D-mannose pyrophosphorylase [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	54.04
Hfc_19920	manC	mannose-1-phosphate guanylyltransferase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.96
Hfc_19950	kdsA	2-dehydro-3-deoxyphosphooctonate aldolase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	67.63
Hfc_19960	kpsF	D-arabinose 5-phosphate isomerase [Capsule (VF0323)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.44
Hfc_20090	cheR	chemotaxis protein methyltransferase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.29
Hfc_20090	cheR	chemotaxis methyltransferase CheR [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.14
Hfc_20100	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Burkholderia pseudomallei K96243	61.59
Hfc_20110	cheB	chemotaxis-specific methylesterase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.75
Hfc_20110	cheB	chemotaxis-specific methylesterase CheB [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	65.55
Hfc_20120	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	78.91
Hfc_20120	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	77.60
Hfc_20140	flhB	flagellar biosynthesis protein FlhB [Flagella (VF0430)]	Burkholderia pseudomallei K96243	53.48
Hfc_20150	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	67.94
Hfc_20150	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	60.94
Hfc_20150	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	56.69
Hfc_20150	flhA	flagellar biosynthetic protein FlhA [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	54.48
Hfc_20180	fliA	flagellar biosynthesis sigma factor [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.84
Hfc_20180	fliA	flagellar biosynthesis sigma factor [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	51.45
Hfc_20230	flgC	flagellar basal body rod protein FlgC [Flagella (VF0430)]	Burkholderia pseudomallei K96243	67.39
Hfc_20230	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	61.19
Hfc_20230	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	52.05

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hfc_20260	flgF	flagellar basal body rod protein FlgF [Flagella (VF0430)]	Burkholderia pseudomallei K96243	51.38
Hfc_20270	flgG	flagellar basal body rod protein FlgG [Flagella (VF0430)]	Burkholderia pseudomallei K96243	70.77
Hfc_20270	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	71.89
Hfc_20270	flgG	flagellar basal body rod protein FlgG [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	60.23
Hfc_20270	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	59.07
Hfc_20270	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.26
Hfc_20270	flgG	flagellar basal-body rod protein (flgG) [Pse5Ac7Ac (AI150)]	Helicobacter pylori 26695	51.92
Hfc_20290	flgl	flagellar P-ring protein precursor Flgl [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.39
Hfc_20290	flgl	flagellar P-ring protein precursor Flgl [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.03
Hfc_20290	flgl	flagellar P-ring protein precursor Flgl [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	57.47
Hfc_20370	fliQ	flagellar biosynthetic protein FliQ [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.95
Hfc_20370	fliQ	flagellar biosynthesis protein FliQ [Flagella (VF0430)]	Burkholderia pseudomallei K96243	62.96
Hfc_20380	fliP	flagellar biosynthesis protein FliP [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.66
Hfc_20380	fliP	flagellar biosynthesis protein FliP [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	50.22
Hfc_20400	fliN	flagellar motor switch protein FliN [Flagella (VF0430)]	Burkholderia pseudomallei K96243	69.34
Hfc_20400	fliN	flagellar motor switch protein FliN [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.86
Hfc_20410	fliM	flagellar motor switch protein FliM [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.55
Hfc_20450	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0430)]	Burkholderia pseudomallei K96243	68.25
Hfc_20450	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	68.26
Hfc_20450	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	61.32
Hfc_20450	fliI	flagellum-specific ATP synthase FliI [polar flagella (AI149)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	57.39
Hfc_20470	fliG	flagellar motor switch protein G [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.14
Hfc_20470	fliG	flagellar motor switch protein G [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.22
Hfc_20540	fliS	flagellar protein FliS [Flagella (VF0430)]	Burkholderia pseudomallei K96243	51.13
Hfc_20540	fliS	flagellar protein FliS [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	53.60

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hfc_20990	icl	Isocitrate lyase Icl (isocitrase) (isocitratase) [Isocitrate lyase (VF0253)]	Mycobacterium tuberculosis H37Rv	66.27
Hfc_21180	orfM	deoxyribonucleotide triphosphate pyrophosphatase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	51.55
Hfc_21220	sodCI	Gifsy-2 prophage: superoxide dismutase precursor (Cu-Zn) [SodCI (VF0109)]	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	59.22
Hfc_21960	lpxA	UDP-N-acetylglucosamine acyltransferase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	51.34
Hfc_22540	msrA/BpilB	trifunctional thioredoxin/methionine sulfoxide reductase A/B protein [MsrAB (VF0456)]	Neisseria meningitidis MC58	50.81
Hfc_25840	clpP	ATP-dependent Clp protease proteolytic subunit [ClpP (VF0074)]	Listeria monocytogenes EGD-e	63.92
Hfc_27070	farA	fatty acid efflux system protein FarA [FarAB (VF0450)]	Neisseria meningitidis MC58	50.14
Hfc_27080	farB	fatty acid efflux system protein FarB [FarAB (VF0450)]	Neisseria meningitidis MC58	50.69
Hfc_27840	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	51.64
Hfc_28310	cyaB	cyclolysin secretion ATP-binding protein [Cya (VF0028)]	Bordetella pertussis Tohama I	62.70
Hfc_28310	hlyB	Hemolysin B [Hemolysin (VF0225)]	Escherichia coli CFT073	52.98
Hfc_28310	hlyB	hemolysin transport protein [Hemolysin (VF0207)]	Escherichia coli O157:H7 str. EDL933	52.47
Hfc_30040	cheW	chemotaxis protein CheW [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.79
Hfc_30040	cheW	purine-binding chemotaxis protein CheW [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	64.67
Hfc_30060	cheY1	chemotaxis two-component response regulator CheY1 [Flagella (VF0430)]	Burkholderia pseudomallei K96243	50.88
Hfc_30080	motB	flagellar motor protein MotB [Flagella (VF0430)]	Burkholderia pseudomallei K96243	56.77
Hfc_30080	motB	flagellar motor protein MotB [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	53.71
Hfc_30090	motA	flagellar motor protein MotA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	68.61
Hfc_30090	motA	flagellar motor protein MotA [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	65.09
Hfc_30090	motA	flagellar motor protein [Deoxyhexose linking sugar, 209 Da capping structure (AI138)]	Pseudomonas aeruginosa PAO1	50.37
Hfc_30110	flhC	flagellar biosynthesis transcription activator FlhC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.14
Hfc_30720	xcpA/pilD	type 4 prepilin peptidase PilD [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	51.94
Hfc_32050	mntB	Manganese transport system membrane protein MntB [MntABC (VF0455)]	Neisseria meningitidis MC58	53.42
Hfc_34550	mgtB	Mg ²⁺ transport protein [MgtBC (VF0106)]	Salmonella enterica subsp. enterica	56.11

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
Hfc_38510	pilT	twitching motility protein PilT [Type IV pili (VF0075)]	serovar Typhimurium str. LT2	
Hfc_38510	pilT	type IV pilus twitching motility protein PilT [type IV pili (AI111)]	Neisseria meningitidis MC58	55.36
Hfc_38510	pilT	twitching motility protein PilT [Type IV pili (VF0082)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	53.71
Hfc_39920	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Pseudomonas aeruginosa PAO1	50.72
Hfc_41260	algW	AlgW protein [Alginate regulation (CVF523)]	Neisseria meningitidis MC58	58.51
Hfc_42400	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Pseudomonas aeruginosa PAO1	54.71
Hfc_43880	rffG	dTDP-glucose 4,6-dehydratase [LOS (CVF494)]	Burkholderia pseudomallei K96243	57.58
Hfc_44130	ddhB	CDP-glucose 4,6-dehydratase [O-antigen (VF0392)]	Haemophilus influenzae Rd KW20	52.39
Hfc_44140	galE	UDP-glucose 4-epimerase [LOS (CVF494)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.28
Hfc_44140	galE	UDP-glucose 4-epimerase [O-antigen (VF0392)]	Haemophilus influenzae Rd KW20	56.85
Hfc_44210	gmhA2	phosphoheptose isomerase [Capsule (VF0323)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.38
Hfc_44210	gmhA	phosphoheptose isomerase [Capsule I (VF0436)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.85
Hfc_44300	ddhA	glucose-1-phosphate cytidyltransferase [O-antigen (VF0392)]	Burkholderia pseudomallei K96243	61.33
Hfc_45880	cheD	methyl-accepting chemotaxis protein CheD [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.08
Hfc_47840	ureA	urease alpha subunit UreA [Urease (VF0050)]	Yersinia enterocolitica subsp. enterocolitica 8081	63.00
Hfc_47850	ureA	urease alpha subunit UreA [Urease (VF0050)]	Helicobacter pylori 26695	54.64
Hfc_47860	ureB	urease beta subunit UreB, urea amidohydrolase [Urease (VF0050)]	Helicobacter pylori 26695	64.67
Hfc_47900	ureG	urease accessory protein (ureG) [Urease (VF0050)]	Helicobacter pylori 26695	64.97

Fonte: O Autor, 2018.

TABELA BS 16: GENES DE VIRULÊNCIA DA ESTIRPE H. FRISINGENSE GSF30

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
GSF30_00288	lpxC	UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	55.30

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
GSF30_00300	pilD	type IV pilus prepilin peptidase [Type IV pili (VF0075)]	Neisseria meningitidis MC58	50.77
GSF30_00300	xcpA/pilD	type 4 prepilin peptidase PilD [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	53.88
GSF30_00300	pilD	type IV prepilin leader peptide type M1 [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	50.76
GSF30_00302	pilF	type IV pilus assembly ATPase protein [Type IV pili (VF0075)]	Neisseria meningitidis MC58	56.30
GSF30_00302	pilB	type 4 fimbrial biogenesis protein PilB [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	56.02
GSF30_00302	pilB	type IV pilus assembly protein [Type IV pili (VF0431)]	Burkholderia pseudomallei K96243	57.36
GSF30_00302	pilB	type IV pilus assembly protein PilB [Type IV pili (VF0155)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	50.90
GSF30_00302	gspE	general secretion pathway protein E [T2SS (VF0333)]	Shigella dysenteriae Sd197	50.52
GSF30_00426	rffG	dTDP-glucose 46-dehydratase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	54.98
GSF30_00444	gmd	GDP-mannose 4,6-dehydratase [LPS (VF0367)]	Brucella melitensis bv. 1 str. 16M	75.21
GSF30_00444	gmd	GDP-mannose 4,6-dehydratase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	68.12
GSF30_00444	rfbD	GDP-D-mannose dehydratase [LPS (VF0056)]	Helicobacter pylori 26695	61.05
GSF30_00445	algC	phosphomannomutase AlgC [Alginate biosynthesis (CVF522)]	Pseudomonas aeruginosa PAO1	55.41
GSF30_00451	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Neisseria meningitidis MC58	54.46
GSF30_00723	wcbT	acyl-CoA transferase [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	60.58
GSF30_00728	cps4I	UDP-N-acetylglucosamine-2-epimerase [Capsule (VF0144)]	Streptococcus pneumoniae TIGR4	53.80
GSF30_00728	cap8P	capsular polysaccharide synthesis enzyme Cap8P [Capsule (VF0003)]	Staphylococcus aureus subsp. aureus MW2	50.82
GSF30_00785	lpg2359	Dot/lcm type IV secretion system effector [Dot/lcm (VF0156)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	53.06
GSF30_00864	htpB	Hsp60, 60K heat shock protein HtpB [Hsp60 (VF0159)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	72.80
GSF30_01040	sodB	superoxide dismutase [SodB (VF0169)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	64.92
GSF30_01123	katA	catalase/(hydro)peroxidase [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	60.11
GSF30_01123	katB	catalase-peroxidase KatB [KatAB (VF0168)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	56.78

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
GSF30_01508	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	pneumophila str. Philadelphia 1 Yersinia enterocolitica subsp. enterocolitica 8081	63.28
GSF30_01508	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	58.91
GSF30_01889	acpXL	acyl carrier protein [LPS (CVF383)]	Brucella melitensis bv. 1 str. 16M	61.11
GSF30_01892	algU	alginate biosynthesis protein AlgZ/FimS [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	64.36
GSF30_01930	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	60.94
GSF30_01930	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.98
GSF30_01958	manC	GDP-mannose pyrophosphorylase [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	57.32
GSF30_01958	algA	phosphomannose isomerase / guanosine 5'-diphospho-D-mannose pyrophosphorylase [Alginate (VF0091)]	Pseudomonas aeruginosa PAO1	54.26
GSF30_01958	manC	mannose-1-phosphate guanylyltransferase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.96
GSF30_01961	kdsA	2-dehydro-3-deoxyphosphooctonate aldolase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	67.63
GSF30_01962	kpsF	D-arabinose 5-phosphate isomerase [Capsule (VF0323)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.44
GSF30_01975	cheR	chemotaxis protein methyltransferase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.64
GSF30_01975	cheR	chemotaxis methyltransferase CheR [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	59.14
GSF30_01976	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.98
GSF30_01977	cheB	chemotaxis-specific methylesterase [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.75
GSF30_01977	cheB	chemotaxis-specific methylesterase CheB [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	65.55
GSF30_01978	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	78.91
GSF30_01978	cheY	chemotaxis protein CheY [Flagella (VF0430)]	Burkholderia pseudomallei K96243	77.60
GSF30_01980	flhB	flagellar biosynthesis protein FlhB [Flagella (VF0430)]	Burkholderia pseudomallei K96243	53.48
GSF30_01981	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	69.20
GSF30_01981	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	60.94

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
GSF30_01981	flhA	flagellar biosynthesis protein FlhA [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	56.64
GSF30_01981	flhA	flagellar biosynthetic protein FlhA [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	54.48
GSF30_01983	flhF	flagellar biosynthesis regulator FlhF [Flagella (VF0430)]	Burkholderia pseudomallei K96243	54.10
GSF30_01985	fliA	flagellar biosynthesis sigma factor [Flagella (VF0430)]	Burkholderia pseudomallei K96243	59.84
GSF30_01985	fliA	flagellar biosynthesis sigma factor [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	51.45
GSF30_01990	flgC	flagellar basal body rod protein FlgC [Flagella (VF0430)]	Burkholderia pseudomallei K96243	67.39
GSF30_01990	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	61.19
GSF30_01990	flgC	flagellar basal-body rod protein FlgC [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	52.05
GSF30_01993	flgF	flagellar basal body rod protein FlgF [Flagella (VF0430)]	Burkholderia pseudomallei K96243	51.38
GSF30_01994	flgG	flagellar basal body rod protein FlgG [Flagella (VF0430)]	Burkholderia pseudomallei K96243	70.77
GSF30_01994	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	71.89
GSF30_01994	flgG	flagellar basal body rod protein FlgG [Flagella (VF0157)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	60.23
GSF30_01994	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	59.07
GSF30_01994	flgG	flagellar basal-body rod protein FlgG [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	53.26
GSF30_01994	flgG	flagellar basal-body rod protein (flgG) [Pse5Ac7Ac (AI150)]	Helicobacter pylori 26695	51.92
GSF30_01996	flgI	flagellar P-ring protein precursor FlgI [Flagella (VF0430)]	Burkholderia pseudomallei K96243	60.39
GSF30_01996	flgI	flagellar P-ring protein precursor FlgI [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.03
GSF30_01996	flgI	flagellar P-ring protein precursor FlgI [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	57.47
GSF30_02004	fliQ	flagellar biosynthetic protein FliQ [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.95
GSF30_02004	fliQ	flagellar biosynthesis protein FliQ [Flagella (VF0430)]	Burkholderia pseudomallei K96243	64.20
GSF30_02005	fliP	flagellar biosynthesis protein FliP [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.66
GSF30_02005	fliP	flagellar biosynthesis protein FliP [Flagella (VF0114)]	Campylobacter jejuni subsp. jejuni NCTC 11168	50.22
GSF30_02007	fliN	flagellar motor switch protein FliN [Flagella (VF0430)]	Burkholderia pseudomallei K96243	69.34
GSF30_02007	fliN	flagellar motor switch protein FliN [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	57.86

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
GSF30_02008	fliM	flagellar motor switch protein FliM [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.55
GSF30_02013	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0430)]	Burkholderia pseudomallei K96243	68.03
GSF30_02013	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	66.96
GSF30_02013	fliI	flagellum-specific ATP synthase FliI [Flagella (VF0273)]	Pseudomonas aeruginosa PAO1	61.54
GSF30_02013	fliI	flagellum-specific ATP synthase FliI [polar flagella (AI149)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	57.39
GSF30_02015	fliG	flagellar motor switch protein G [Flagella (VF0430)]	Burkholderia pseudomallei K96243	63.14
GSF30_02015	fliG	flagellar motor switch protein G [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	54.22
GSF30_02022	fliS	flagellar protein FliS [Flagella (VF0430)]	Burkholderia pseudomallei K96243	50.38
GSF30_02022	fliS	flagellar protein FliS [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	52.80
GSF30_02068	icl	Isocitrate lyase Icl (isocitrase) (isocitratase) [Isocitrate lyase (VF0253)]	Mycobacterium tuberculosis H37Rv	66.27
GSF30_02087	orfM	deoxyribonucleotide triphosphate pyrophosphatase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	51.55
GSF30_02091	sodCI	Gifsy-2 prophage: superoxide dismutase precursor (Cu-Zn) [SodCI (VF0109)]	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	59.09
GSF30_02144	lpxA	UDP-N-acetylglucosamine acyltransferase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	51.72
GSF30_02214	msrA/BpilB	trifunctional thioredoxin/methionine sulfoxide reductase A/B protein [MsrAB (VF0456)]	Neisseria meningitidis MC58	51.61
GSF30_02586	clpP	ATP-dependent Clp protease proteolytic subunit [ClpP (VF0074)]	Listeria monocytogenes EGD-e	63.92
GSF30_02703	farA	fatty acid efflux system protein FarA [FarAB (VF0450)]	Neisseria meningitidis MC58	50.14
GSF30_02704	farB	fatty acid efflux system protein FarB [FarAB (VF0450)]	Neisseria meningitidis MC58	50.69
GSF30_02844	cheY	chemotaxis regulatory protein CheY [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	51.64
GSF30_02881	cyaB	cyclolysin secretion ATP-binding protein [Cya (VF0028)]	Bordetella pertussis Tohama I	62.41
GSF30_02881	hlyB	Hemolysin B [Hemolysin (VF0225)]	Escherichia coli CFT073	52.70
GSF30_02881	hlyB	hemolysin transport protein [Hemolysin (VF0207)]	Escherichia coli O157:H7 str. EDL933	53.31
GSF30_03009	cheW	chemotaxis protein CheW [Flagella (VF0430)]	Burkholderia pseudomallei K96243	71.79
GSF30_03009	cheW	purine-binding chemotaxis protein CheW [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	64.67

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
GSF30_03011	cheY1	chemotaxis two-component response regulator CheY1 [Flagella (VF0430)]	Burkholderia pseudomallei K96243	50.88
GSF30_03013	motB	flagellar motor protein MotB [Flagella (VF0430)]	Burkholderia pseudomallei K96243	56.77
GSF30_03013	motB	flagellar motor protein MotB [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	53.71
GSF30_03014	motA	flagellar motor protein MotA [Flagella (VF0430)]	Burkholderia pseudomallei K96243	66.43
GSF30_03014	motA	flagellar motor protein MotA [peritrichous flagella (AI145)]	Yersinia enterocolitica subsp. enterocolitica 8081	64.91
GSF30_03014	motA	flagellar motor protein [Deoxyhexose linking sugar, 209 Da capping structure (AI138)]	Pseudomonas aeruginosa PAO1	50.18
GSF30_03016	flhC	flagellar biosynthesis transcription activator FlhC [Flagella (VF0394)]	Yersinia enterocolitica subsp. enterocolitica 8081	58.14
GSF30_03070	xcpA/pilD	type 4 prepilin peptidase PilD [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	52.94
GSF30_03192	mntB	Manganese transport system membrane protein MntB [MntABC (VF0455)]	Neisseria meningitidis MC58	53.42
GSF30_03485	mgtB	Mg ²⁺ transport protein [MgtBC (VF0106)]	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	56.11
GSF30_03793	rfaE	ADP-heptose synthase [LOS (VF0044)]	Haemophilus influenzae Rd KW20	50.33
GSF30_03907	pilT	twitching motility protein PilT [Type IV pili (VF0075)]	Neisseria meningitidis MC58	55.36
GSF30_03907	pilT	type IV pilus twitching motility protein PilT [type IV pili (AI111)]	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	52.52
GSF30_03907	pilT	twitching motility protein PilT [Type IV pili (VF0082)]	Pseudomonas aeruginosa PAO1	53.12
GSF30_04036	mtrD	multiple transferable resistance system protein MtrD [MtrCDE (VF0451)]	Neisseria meningitidis MC58	50.38
GSF30_04153	algW	AlgW protein [Alginate regulation (CVF523)]	Pseudomonas aeruginosa PAO1	58.51
GSF30_04270	cheD	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)]	Burkholderia pseudomallei K96243	54.71
GSF30_04431	rffG	dTDP-glucose 46-dehydratase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	57.58
GSF30_04456	ddhB	CDP-glucose 4,6-dehydratase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	52.39
GSF30_04457	galE	UDP-glucose 4-epimerase [LOS (CVF494)]	Haemophilus influenzae Rd KW20	58.28
GSF30_04457	galE	UDP-glucose 4-epimerase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	56.85
GSF30_04463	gmhA2	phosphoheptose isomerase [Capsule (VF0323)]	Campylobacter jejuni subsp. jejuni NCTC 11168	57.38

Locus	Gene no banco de dados	Produto no banco de dados	Organismo	Identidade
GSF30_04463	gmhA	phosphoheptose isomerase [Capsule I (VF0436)]	Burkholderia pseudomallei K96243	53.85
GSF30_04472	ddhA	glucose-1-phosphate cytidylyltransferase [O-antigen (VF0392)]	Yersinia enterocolitica subsp. enterocolitica 8081	61.33
GSF30_04765	ureA	urease alpha subunit UreA [Urease (VF0050)]	Helicobacter pylori 26695	63.00
GSF30_04766	ureA	urease alpha subunit UreA [Urease (VF0050)]	Helicobacter pylori 26695	53.61
GSF30_04767	ureB	urease beta subunit UreB, urea amidohydrolase [Urease (VF0050)]	Helicobacter pylori 26695	64.67
GSF30_04770	ureG	urease accessory protein (ureG) [Urease (VF0050)]	Helicobacter pylori 26695	64.97

Fonte: O Autor, 2018.

TABELA BS 17: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO H. FRISINGENSE AU14559

Locus	Gene do banco de dados	Organismo	Identidade
Hfc_01020	GzOB039	Fusarium_graminearum	54.96
Hfc_01370	MoPRX1	Magnaporthe_oryzae	51.15
Hfc_01660	Irp	Xenorhabdus_nematophila	58.94
Hfc_02390	PCC21_023220	Pectobacterium_carotovorum	58.67
Hfc_02610	BP1026B_II1587	Burkholderia_pseudomallei	60.87
Hfc_02650	oqxB	Klebsiella_pneumoniae	61.59
Hfc_02660	oqxB	Klebsiella_pneumoniae	61.75
Hfc_02800	hupA	Edwardsiella_tarda	53.93
Hfc_03010	secA	Acinetobacter_baumannii	55.50
Hfc_03100	VcpD	Vibrio_cholerae	55.00
Hfc_04620	acrB	Klebsiella_pneumoniae	57.96
Hfc_04620	AcrB	Erwinia_amylovora	57.56
Hfc_04620	AcrD	Erwinia_amylovora	54.38
Hfc_04620	adeB	Acinetobacter_baumannii	52.56
Hfc_04620	adeB	Acinetobacter_baumannii	52.56
Hfc_04620	adeJ	Acinetobacter_baumannii	51.67
Hfc_04840	CspV	Vibrio_cholerae	61.19
Hfc_04840	cspA	Brucella_melitensis	51.43
Hfc_06080	Fur	Escherichia_coli	57.89
Hfc_06080	Fur	Pseudomonas_fluorescens	61.34
Hfc_06080	fur	Stenotrophomonas_maltophilia	52.80

Locus	Gene do banco de dados	Organismo	Identidade
Hfc_06540	recA	Streptococcus_pneumoniae	63.25
Hfc_06730	ADE2	Saccharomyces_cerevisiae	58.06
Hfc_06730	ADE2	Candida_albicans	55.41
Hfc_06730	ADE2	Cryptococcus_neoformans	57.82
Hfc_06960	XC_2466	Xanthomonas_campestris	55.65
Hfc_07420	wcbT	Burkholderia_pseudomallei	60.58
Hfc_07960	thioredoxin_1	Salmonella_enterica	65.74
Hfc_08320	glpD	Yersinia_pestis	51.15
Hfc_08640	GroEL	Porphyromonas_gingivalis	63.12
Hfc_08670	ecf6	Pseudomonas_syringae	51.23
Hfc_08830	ecf18	Pseudomonas_syringae	59.56
Hfc_09330	Sdh1	Parastagonospora_nodorum	59.63
Hfc_09330	MoSSADH	Magnaporthe_oryzae	59.34
Hfc_09450	GGT	Helicobacter_pylori	53.57
Hfc_09920	Sdh1	Parastagonospora_nodorum	52.26
Hfc_09920	MoSSADH	Magnaporthe_oryzae	52.77
Hfc_10780	Q8XWM2	Ralstonia_solanacearum	52.43
Hfc_11270	glyA	Edwardsiella_ictaluri	61.78
Hfc_11490	katG	Acinetobacter_nosocomialis	62.91
Hfc_11490	CPXB	Magnaporthe_oryzae	59.87
Hfc_11640	pstB	Xanthomonas_citri	60.90
Hfc_13270	CspV	Vibrio_cholerae	66.13
Hfc_13270	CspR	Enterococcus_faecalis	61.54
Hfc_13270	cspA	Brucella_melitensis	51.56
Hfc_13280	clpS/clpA	Pseudomonas_aeruginosa	52.58
Hfc_13730	Q9HWS6	Pseudomonas_aeruginosa	58.51
Hfc_13790	Orn	Pseudomonas_aeruginosa	63.84
Hfc_16650	Irp	Xenorhabdus_nematophila	59.86
Hfc_17300	pnp	Salmonella_enterica	65.42
Hfc_17300	pnp	Pseudomonas_aeruginosa	61.51
Hfc_17470	NOS1	Fusarium_graminearum	57.59
Hfc_17790	htpG	Escherichia_coli	59.05
Hfc_18010	HspL	Agrobacterium_tumefaciens	53.90
Hfc_18010	hspAT1	Agrobacterium_tumefaciens	51.85
Hfc_18920	Trr1	Beauveria_bassiana	55.21
Hfc_18920	TRR1	Cryptococcus_neoformans	53.12
Hfc_18960	rarA	Klebsiella_pneumoniae	54.61
Hfc_19160	fabG1	Ralstonia_solanacearum	75.40
Hfc_19200	rpoE	Salmonella_enterica	61.58
Hfc_19200	RpoE	Proteus_mirabilis	58.42
Hfc_20110	cheB	Salmonella_enterica	64.99
Hfc_20340	MacB	Salmonella_enterica	55.79
Hfc_21600	ompR	Salmonella_enterica	64.81
Hfc_21750	galU	Xanthomonas_campestris	55.71
Hfc_22070	smpB	Escherichia_coli	54.41
Hfc_22070	smpB	Salmonella_enterica	52.94
Hfc_22100	IMPDH	Streptococcus_suis	54.99
Hfc_23200	nhaA	Escherichia_coli	53.23
Hfc_24300	Pst-Lov	Pseudomonas_syringae	57.17
Hfc_24410	XAC1258	Xanthomonas_citri	51.93
Hfc_25800	hupA	Edwardsiella_tarda	53.33
Hfc_25820	LonA	Actinobacillus_pleuropneumoniae	64.76
Hfc_25830	ClpX	Staphylococcus_aureus	59.56
Hfc_25840	ClpP	Salmonella_enterica	72.02
Hfc_25840	clpP	Pseudomonas_aeruginosa	75.52
Hfc_25840	ClpP	Staphylococcus_aureus	66.32
Hfc_26710	galU	Xanthomonas_campestris	58.80
Hfc_27130	MoSFA1	Magnaporthe_oryzae	61.14
Hfc_27130	gnoA	Aspergillus_fumigatus	62.13

Hfc_27130	Fdh3	Candida_albicans	58.70
Hfc_27130	GNO1	Cryptococcus_neoformans	56.79
Hfc_29400	argGH	Streptococcus_pneumoniae	50.14
Hfc_29630	purA	Salmonella_enterica	60.00
Hfc_29680	Hfq	Bordetella_pertussis	85.90
Hfc_29680	Hfq	Vibrio_cholerae	82.09
Hfc_29680	hfq	Yersinia_enterocolitica	84.85
Hfc_29680	Hfq	Yersinia_pestis	84.85
Hfc_29680	hfq	Salmonella_enterica	83.33
Hfc_29680	hfq	Haemophilus_ducreyi	72.97
Hfc_29760	XC_2203	Xanthomonas_campestris	75.89
Hfc_29760	Ndk	Pseudomonas_aeruginosa	75.18
Hfc_29790	rpoS	Salmonella_enterica	50.16
Hfc_30040	cheW	Escherichia_coli	64.71
Hfc_30090	motA	Pantoea_ananatis	66.55
Hfc_30090	motA	Escherichia_coli	60.00
Hfc_30120	flhD	Burkholderia_glumae	53.92
Hfc_30340	dgcB	Burkholderia_glumae	55.87
Hfc_30470	CspV	Vibrio_cholerae	66.67
Hfc_30470	CspR	Enterococcus_faecalis	57.81
Hfc_30470	cspA	Brucella_melitensis	53.12
Hfc_30480	AoMIs	Arthrotrichia_oligospora	52.39
Hfc_30480	Q5J4D6	Parastagonospora_nodorum	53.07
Hfc_30720	VcpD	Vibrio_cholerae	50.76
Hfc_31470	raxP_(cysD)	Xanthomonas_oryzae	50.68
Hfc_31830	lscR	Yersinia_pseudotuberculosis	50.31
Hfc_33320	ADE4	Saccharomyces_cerevisiae	50.10
Hfc_37430	himD	Salmonella_enterica	61.54
Hfc_37460	aroA	Burkholderia_glumae	66.51
Hfc_37460	P07637	Salmonella_enterica	56.28
Hfc_37500	GyrA	Burkholderia_glumae	79.46
Hfc_37500	GyrA	Burkholderia_glumae	79.46
Hfc_3750	GyrA	Burkholderia_glumae	79.35
Hfc_37500	GyrA	Burkholderia_glumae	79.23
Hfc_37500	GyrA	Burkholderia_glumae	79.23
Hfc_38160	kdpB	Pseudomonas_aeruginosa	60.40
Hfc_38370	Rv3232c	Mycobacterium_tuberculosis	53.82
Hfc_38520	gpx31	Candida_albicans	52.17
Hfc_38520	G4N178	Magnaporthe_oryzae	54.14
Hfc_38520	AaGPx3	Alternaria_alternata	53.09
Hfc_39580	Mollv1	Magnaporthe_oryzae	54.99
Hfc_39700	FTT0673p/prsAp	Francisella_tularensis	57.78
Hfc_39920	adeB	Acinetobacter_baumannii	61.58
Hfc_39920	adeB	Acinetobacter_baumannii	61.58
Hfc_39920	acrB	Klebsiella_pneumoniae	50.88
Hfc_39920	AcrB	Erwinia_amylovora	50.88
Hfc_40500	fcs	Ralstonia_solanacearum	56.63
Hfc_41280	tatC	Vibrio_cholerae	51.26
Hfc_41620	paaA	Burkholderia_cenocepacia	66.56
Hfc_41630	paaB	Burkholderia_cenocepacia	60.64
Hfc_42080	rfaE	Haemophilus_parasuis	52.35
Hfc_42390	Mqo	Staphylococcus_aureus	50.31
Hfc_42930	MntE	Streptococcus_pyogenes	50.42
Hfc_43080	MGG_00383	Magnaporthe_oryzae	55.73
Hfc_44040	PmrA	Coxiella_burnetii	53.21
Hfc_44380	IMPDH	Streptococcus_suis	54.32
Hfc_44920	BaTdp	Bacillus_anthraxis	61.00
Hfc_46380	MqsRA	Xylella_fastidiosa	55.56
Hfc_46500	AcrB	Erwinia_amylovora	60.78
Hfc_46500	acrB	Klebsiella_pneumoniae	59.41

Hfc_46500	AcrD	Erwinia_amylovora	59.67
Hfc_46500	adeJ	Acinetobacter_baumannii	54.22
Hfc_47840	URE1	Cryptococcus_gattii	56.00
Hfc_47840	URE1	Cryptococcus_neoformans	56.00
Hfc_47860	urease	Helicobacter_pylori	64.67
Hfc_47860	URE1	Cryptococcus_gattii	60.70
Hfc_47860	URE1	Cryptococcus_neoformans	59.30
Hfc_47860	GzZC144	Fusarium_graminearum	55.68
Hfc_48420	lipA	Acinetobacter_baumannii	61.84
Hfc_48540	opgG	Dickeya_dadantii	57.70

Fonte: O Autor, 2018

TABELA BS 18: GENES DE INTERAÇÃO PATÓGENO-HOSPEDEIRO NO ORGANISMO H. FRISINGENSE GSF30

Locus	Gene do banco de dados	Organismo	Identidade
GSF30_00094	GzOB039	Fusarium_graminearum	54.96
GSF30_00130	MoPRX1	Magnaporthe_oryzae	51.15
GSF30_00159	lrp	Xenorhabdus_nematophila	58.94
GSF30_00232	PCC21_023220	Pectobacterium_carotovorum	58.67
GSF30_00254	BP1026B_II1587	Burkholderia_pseudomallei	60.87
GSF30_00258	oqxB	Klebsiella_pneumoniae	61.25
GSF30_00270	hupA	Edwardsiella_tarda	53.93
GSF30_00291	secA	Acinetobacter_baumannii	55.82
GSF30_00300	VcpD	Vibrio_cholerae	54.23
GSF30_00451	acrB	Klebsiella_pneumoniae	57.96
GSF30_00451	AcrB	Erwinia_amylovora	57.56
GSF30_00451	AcrD	Erwinia_amylovora	54.47
GSF30_00451	adeJ	Acinetobacter_baumannii	51.77
GSF30_00451	adeB	Acinetobacter_baumannii	52.47
GSF30_00451	adeB	Acinetobacter_baumannii	52.47
GSF30_00470	CspV	Vibrio_cholerae	61.19
GSF30_00470	cspA	Brucella_melitensis	51.43
GSF30_00591	Fur	Escherichia_coli	57.89
GSF30_00591	Fur	Pseudomonas_fluorescens	61.34
GSF30_00591	fur	Stenotrophomonas_maltophilia	52.80
GSF30_00637	recA	Streptococcus_pneumoniae	63.25
GSF30_00656	ADE2	Saccharomyces_cerevisiae	58.71
GSF30_00656	ADE2	Candida_albicans	54.78
GSF30_00656	ADE2	Cryptococcus_neoformans	57.14
GSF30_00679	XC_2466	Xanthomonas_campestris	55.65
GSF30_00723	wcbT	Burkholderia_pseudomallei	60.58
GSF30_00778	thioredoxin_1	Salmonella_enterica	65.74
GSF30_00831	glpD	Yersinia_pestis	50.94
GSF30_00864	GroEL	Porphyromonas_gingivalis	63.12
GSF30_00869	ecf6	Pseudomonas_syringae	51.23
GSF30_00885	ecf18	Pseudomonas_syringae	60.24
GSF30_00944	Sdh1	Parastagonospora_nodorum	60.04
GSF30_00944	MoSSADH	Magnaporthe_oryzae	59.75
GSF30_00956	GGT	Helicobacter_pylori	53.76
GSF30_00992	Sdh1	Parastagonospora_nodorum	52.06
GSF30_00992	MoSSADH	Magnaporthe_oryzae	52.98
GSF30_01103	glyA	Edwardsiella_ictaluri	61.78
GSF30_01123	katG	Acinetobacter_nosocomialis	63.18
GSF30_01123	CPXB	Magnaporthe_oryzae	59.60
GSF30_01137	pstB	Xanthomonas_citri	60.90
GSF30_01231	acrB	Klebsiella_pneumoniae	65.85
GSF30_01231	AcrB	Erwinia_amylovora	65.06
GSF30_01231	AcrD	Erwinia_amylovora	58.59

Locus	Gene do banco de dados	Organismo	Identidade
GSF30_01231	adeJ	Acinetobacter_baumannii	57.90
GSF30_01288	CspV	Vibrio_cholerae	66.13
GSF30_01288	CspR	Enterococcus_faecalis	61.54
GSF30_01288	cspA	Brucella_melitensis	51.56
GSF30_01289	clpS/clpA	Pseudomonas_aeruginosa	52.58
GSF30_01334	Q9HWS6	Pseudomonas_aeruginosa	57.88
GSF30_01340	Orn	Pseudomonas_aeruginosa	63.84
GSF30_01371	LpdA	Legionella_pneumophila	58.00
GSF30_01641	lrp	Xenorhabdus_nematophila	59.86
GSF30_01707	pnp	Salmonella_enterica	65.42
GSF30_01707	pnp	Pseudomonas_aeruginosa	61.54
GSF30_01723	NOS1	Fusarium_graminearum	57.59
GSF30_01757	htpG	Escherichia_coli	58.73
GSF30_01779	HspL	Agrobacterium_tumefaciens	53.90
GSF30_01779	hspAT1	Agrobacterium_tumefaciens	51.85
GSF30_01869	rarA	Klebsiella_pneumoniae	54.61
GSF30_01888	fabG1	Ralstonia_solanacearum	75.81
GSF30_01892	rpoE	Salmonella_enterica	61.58
GSF30_01892	RpoE	Proteus_mirabilis	58.42
GSF30_01977	cheB	Salmonella_enterica	64.99
GSF30_02001	MacB	Salmonella_enterica	55.79
GSF30_02108	ompR	Salmonella_enterica	64.81
GSF30_02123	galU	Xanthomonas_campestris	55.71
GSF30_02161	smxB	Escherichia_coli	52.03
GSF30_02161	smxB	Salmonella_enterica	50.68
GSF30_02164	IMPDH	Streptococcus_suis	54.99
GSF30_02284	nhaA	Escherichia_coli	53.49
GSF30_02434	Pst-Lov	Pseudomonas_syringae	56.98
GSF30_02445	XAC1258	Xanthomonas_citri	51.93
GSF30_02513	MqsRA	Xylella_fastidiosa	64.00
GSF30_02583	hupA	Edwardsiella_tarda	53.33
GSF30_02584	LonA	Actinobacillus_pleuropneumonia e	64.76
GSF30_02585	ClpX	Staphylococcus_aureus	59.56
GSF30_02586	ClpP	Salmonella_enterica	72.02
GSF30_02586	clpP	Pseudomonas_aeruginosa	75.52
GSF30_02586	ClpP	Staphylococcus_aureus	66.32
GSF30_02664	galU	Xanthomonas_campestris	58.80
GSF30_02709	MoSFA1	Magnaporthe_oryzae	61.14
GSF30_02709	gnoA	Aspergillus_fumigatus	62.13
GSF30_02709	Fdh3	Candida_albicans	58.70
GSF30_02709	GNO1	Cryptococcus_neoformans	56.79
GSF30_02782	modA	Pseudomonas_aeruginosa	62.72
GSF30_02969	purA	Salmonella_enterica	60.00
GSF30_02974	Hfq	Bordetella_pertussis	85.90
GSF30_02974	Hfq	Vibrio_cholerae	82.09
GSF30_02974	hfq	Yersinia_enterocolitica	84.85
GSF30_02974	Hfq	Yersinia_pestis	84.85
GSF30_02974	hfq	Salmonella_enterica	83.33
GSF30_02974	hfq	Haemophilus_ducreyi	72.97
GSF30_02982	XC_2203	Xanthomonas_campestris	75.89
GSF30_02982	Ndk	Pseudomonas_aeruginosa	75.18
GSF30_02985	rpoS	Salmonella_enterica	50.16
GSF30_03009	cheW	Escherichia_coli	64.71
GSF30_03014	motA	Pantoea_ananatis	65.96
GSF30_03014	motA	Escherichia_coli	59.30
GSF30_03017	flhD	Burkholderia_glumae	53.92
GSF30_03039_	dgcB	Burkholderia_glumae	55.56

cph2_16

Locus	Gene do banco de dados	Organismo	Identidade
GSF30_03051	CspV	Vibrio_cholerae	66.67
GSF30_03051	CspR	Enterococcus_faecalis	57.81
GSF30_03051	cspA	Brucella_melitensis	53.12

Fonte: O Autor, 2018

TABELA BS 19: GENES DE RESISTÊNCIA DA ESTIRPE H. FRISINGENSE AU14559

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
Hfc_35050	ABC_efflux	RF0007	8.2e-94	Hfc_35050 ABC transporter 3871869:3873752 forward MW:66807
Hfc_21790	ABC_efflux	RF0007	1.7e-92	Hfc_21790 ABC-type multidrug transport system, ATPase and permease component 2388483:2390759 forward MW:84000
Hfc_28310	ABC_efflux	RF0007	1e-86	Hfc_28310 type I secretion system permease/ATPase 3113134:3115335 reverse MW:79639
Hfc_39050	ABC_efflux	RF0007	3.7e-85	Hfc_39050 Fe-S cluster assembly ABC transporter permease/ATPase 4331678:4333582 reverse MW:70773
Hfc_04630	ABC_efflux	RF0007	3.9e-76	Hfc_04630 acriflavin resistance lipoprotein A 492927:494135 reverse MW:42296
Hfc_04610	adeC-adeK-oprM	RF0021	2.1e-175	Hfc_04610 drug efflux lipoprotein 488316:489767 reverse MW:52047
Hfc_34340	ClassA	RF0053	8.8e-104	Hfc_34340 class A beta-lactamase 3787159:3788058 forward MW:31725
Hfc_33810	ClassA	RF0053	1.8e-80	Hfc_33810 class A beta-lactamase 3725968:3726846 reverse MW:31982
Hfc_05190	ClassB	RF0054	5e-31	Hfc_05190 Zn-dependent hydrolase (glyoxylase) protein 552489:553448 reverse MW:35106
Hfc_34340	Exo	RF0072	5.9e-77	Hfc_34340 class A beta-lactamase 3787159:3788058 forward MW:31725
Hfc_20340	macB	RF0089	3.2e-226	Hfc_20340 antimicrobial peptide ABC transporter ATPase 2231961:2233940 reverse MW:70496
Hfc_22180	macB	RF0089	1.3e-188	Hfc_22180 macrolide ABC transporter permease/ATP-binding protein MacB 2434916:2436877 forward MW:69591
Hfc_02670	MexE	RF0098	2.5e-106	Hfc_02670 hemolysin secretion protein D 286300:287499 reverse MW:41707
Hfc_04630	MexE	RF0098	7.5e-58	Hfc_04630 acriflavin resistance lipoprotein A 492927:494135 reverse MW:42296
Hfc_04630	MexX	RF0101	1.1e-97	Hfc_04630 acriflavin resistance lipoprotein A 492927:494135 reverse MW:42296
Hfc_36040	msbA	RF0107	1.6e-175	Hfc_36040 multidrug ABC transporter ATPase/permease 3996966:3998708 forward MW:63716
Hfc_04620	RND_efflux	RF0115	0	Hfc_04620 multidrug efflux RND transporter permease subunit 489760:492915 reverse MW:112162
Hfc_46500	RND_efflux	RF0115	0	Hfc_46500 multidrug transporter 5141658:5144813 reverse MW:112693
Hfc_14970	RND_efflux	RF0115	0	Hfc_14970 multidrug transporter subunit MdtB 1670031:1673186 forward MW:112777
Hfc_14980	RND_efflux	RF0115	0	Hfc_14980 cation/multidrug efflux pump

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
Hfc_39920	RND_efflux	RF0115	0	protein 1673199:1676435 forward MW:115879 Hfc_39920 multidrug efflux system transmembrane protein 4426876:4429983 reverse MW:110470
Hfc_24710	RND_efflux	RF0115	1.6e-289	Hfc_24710 acriflavine resistance protein B 2711663:2714830 forward MW:114159
Hfc_05190	SubclassB1	RF0123	5.7e-28	Hfc_05190 Zn-dependent hydrolase (glyoxylase) protein 552489:553448 reverse MW:35106
Hfc_05190	SubclassB2	RF0124	1.9e-22	Hfc_05190 Zn-dependent hydrolase (glyoxylase) protein 552489:553448 reverse MW:35106

Fonte: O Autor, 2018.

TABELA BS 20: GENES DE RESISTÊNCIA DA ESTIRPE H. FRISINGENSE

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
WP_077238380.1	ABC_efflux	RF0007	8.1e-93	ABC transporter [Herbaspirillum sp. VT-16-41]
WP_077237117.1	ABC_efflux	RF0007	1.8e-92	ABC transporter [Herbaspirillum sp. VT-16-41]
WP_077235217.1	ABC_efflux	RF0007	3.9e-87	type I secretion system permease/ATPase [Herbaspirillum sp. VT-16-41]
WP_077234794.1	ABC_efflux	RF0007	4.2e-85	metal ABC transporter permease [Herbaspirillum sp. VT-16-41]
WP_077236881.1	ABC_efflux	RF0007	1.2e-76	efflux transporter periplasmic adaptor subunit [Herbaspirillum sp. VT-16-41]
WP_077237765.1	adeC-adeK-oprM	RF0021	5.6e-180	multidrug transporter [Herbaspirillum sp. VT-16-41]
WP_006465537.1	adeC-adeK-oprM	RF0021	3.2e-176	MULTISPECIES: adeC/adeK/oprM family multidrug efflux complex outer membrane factor [Herbaspirillum]
WP_077238329.1	ClassA	RF0053	8.5e-105	class A beta-lactamase [Herbaspirillum sp. VT-16-41]
WP_077236308.1	ClassA	RF0053	1.3e-79	class A beta-lactamase [Herbaspirillum sp. VT-16-41]
WP_077236915.1	ClassB	RF0054	1e-30	MBL fold metallo-hydrolase [Herbaspirillum sp. VT-16-41]
WP_077238329.1	Exo	RF0072	9.3e-78	class A beta-lactamase [Herbaspirillum sp. VT-16-41]
WP_077237044.1	macB	RF0089	3.4e-227	macrolide ABC transporter permease/ATP-binding protein MacB [Herbaspirillum sp. VT-16-41]
WP_077237143.1	macB	RF0089	3.5e-187	macrolide ABC transporter permease/ATP-binding protein MacB [Herbaspirillum sp. VT-16-41]
WP_077235818.1	MexE	RF0098	2.6e-106	efflux transporter periplasmic adaptor subunit [Herbaspirillum sp. VT-16-41]
WP_077237544.1	MexE	RF0098	3.6e-83	efflux transporter periplasmic adaptor subunit [Herbaspirillum sp. VT-16-41]
WP_077236881.1	MexE	RF0098	1.5e-58	efflux transporter periplasmic

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
WP_077237543.1	MexW-MexI	RF0100	6.9e-211	adaptor subunit [Herbaspirillum sp. VT-16-41] multidrug efflux RND transporter permease subunit [Herbaspirillum sp. VT-16-41]
WP_077236881.1	MexX	RF0101	8.1e-98	efflux transporter periplasmic adaptor subunit [Herbaspirillum sp. VT-16-41]
WP_077237594.1	msbA	RF0107	3.2e-177	lipid A export permease/ATP-binding protein MsbA [Herbaspirillum sp. VT-16-41]
WP_006464058.1	RND_efflux	RF0115	0	MULTISPECIES: multidrug efflux RND transporter permease subunit [Herbaspirillum]
WP_077237543.1	RND_efflux	RF0115	0	multidrug efflux RND transporter permease subunit [Herbaspirillum sp. VT-16-41]
WP_077236880.1	RND_efflux	RF0115	0	multidrug efflux RND transporter permease subunit [Herbaspirillum sp. VT-16-41]
WP_077237764.1	RND_efflux	RF0115	0	hydrophobe/amphiphile efflux-1 family RND transporter [Herbaspirillum sp. VT-16-41]
WP_077234866.1	RND_efflux	RF0115	0	hydrophobe/amphiphile efflux-1 family RND transporter [Herbaspirillum sp. VT-16-41]
WP_077237902.1	RND_efflux	RF0115	0	multidrug transporter subunit MdtB [Herbaspirillum sp. VT-16-41]
WP_006462423.1	RND_efflux	RF0115	0	MULTISPECIES: cation/multidrug efflux pump protein [Herbaspirillum]
WP_077235962.1	RND_efflux	RF0115	0	multidrug efflux RND transporter permease subunit [Herbaspirillum sp. VT-16-41]
WP_077238033.1	RND_efflux	RF0115	5e-290	acriflavine resistance protein B [Herbaspirillum sp. VT-16-41]
WP_077236915.1	SubclassB1	RF0123	1.8e-27	MBL fold metallo-hydrolase [Herbaspirillum sp. VT-16-41]
WP_077236915.1	SubclassB2	RF0124	3e-22	MBL fold metallo-hydrolase [Herbaspirillum sp. VT-16-41]

Fonte: O Autor, 2018.

TABELA BS 21: GENES DE RESISTÊNCIA DA ESTIRPE H. FRISINGENSE GSF30

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
GSF30_03539	ABC_efflux	RF0007	3.8e-93	Putative_multidrug_export_ATP-binding/permease_protein
GSF30_02127	ABC_efflux	RF0007	1.8e-92	putative_ABC_transporter_ATP-binding_protein
GSF30_02881	ABC_efflux	RF0007	6.7e-87	apxIB_2_Toxin_RTX-I_translocation_ATP-binding_protein
GSF30_03960	ABC_efflux	RF0007	3.7e-85	Putative_multidrug_export_ATP-binding/permease_protein
GSF30_00452	ABC_efflux	RF0007	4.1e-76	ttgA_putative_efflux_pump_periplasmic_linker_TtgA_precursor

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
GSF30_01232	adeC-adeK-oprM	RF0021	9e-179	oprM_5_Outer_membrane_protein_OprM_precursor
GSF30_00450	adeC-adeK-oprM	RF0021	3.1e-176	oprM_5_Outer_membrane_protein_OprM_precursor
GSF30_02154	Chlor_Acetyltrans_CAT	RF0050	8.8e-84	cat_Chlor_Acetyltrans_CAT
GSF30_03463	ClassA	RF0053	1.6e-103	bla_2_Beta-lactamase_Toho-1_precursor
GSF30_03415	ClassA	RF0053	9.6e-80	bla_1_Carbepenem-hydrolyzing_beta-lactamase_KPC_precursor
GSF30_00501	ClassB	RF0054	8.7e-31	ccrA_Beta-lactamase_type_II_precursor
GSF30_03463	Exo	RF0072	5.2e-76	-
GSF30_02001	macB	RF0089	4.2e-227	bla_2_Beta-lactamase_Toho-1_precursor
GSF30_02175	macB	RF0089	4.1e-188	macB_2_Macrolide_export_ATP-binding/permease_protein_MacB
GSF30_00259	MexE	RF0098	1.6e-106	bepF_Efflux_pump_periplasmic_linker_BepF
GSF30_00452	MexE	RF0098	8.1e-58	ttgA_putative_efflux_pump_periplasmic_linker_TtgA_precursor
GSF30_00452	MexX	RF0101	7.1e-98	ttgA_putative_efflux_pump_periplasmic_linker_TtgA_precursor
GSF30_03617	msbA	RF0107	6.2e-176	sbA_Lipid_A_export_ATP-binding/permease_protein_MsbA
GSF30_00258	RND_efflux	RF0115	0	sbA_Lipid_A_export_ATP-binding/permease_protein_MsbA
GSF30_00451	RND_efflux	RF0115	0	acrB_1_Multidrug_efflux_pump_subunit_AcrB
GSF30_01231	RND_efflux	RF0115	0	acrB_2_Multidrug_efflux_pump_subunit_AcrB
GSF30_04576	RND_efflux	RF0115	0	mexB_Multidrug_resistance_protein_MexB
GSF30_01459	RND_efflux	RF0115	0	mdtB_1_Multidrug_resistance_protein_MdtB
GSF30_01460	RND_efflux	RF0115	0	mdtC_Multidrug_resistance_protein_MdtC
GSF30_04036	RND_efflux	RF0115	0	acrB_4_Multidrug_efflux_pump_subunit_AcrB
GSF30_02471	RND_efflux	RF0115	8.9e-290	mdtB_2_Multidrug_resistance_protein_MdtB
GSF30_00501	SubclassB1	RF0123	9.4e-28	ccrA_Beta-lactamase_type_II_precursor
GSF30_00501	SubclassB2	RF0124	5e-23	ccrA_Beta-

Locus	Classe do gene de resistência	Código ResFam	e-value	Anotação do gene
				lactamase_type_II_precursor

Fonte: O Autor, 2018

7 CONCLUSÕES

- O isolado *Herbaspirillum* linhagem 2 AU13965 pertence à espécie *Herbaspirillum seropedicae* teve seu genoma sequenciado, montado, fechado e anotado;
- O organismo *H. frisingense* AU14559 teve seu genoma sequenciado, montado, fechado e anotado, sendo o primeiro genoma desta espécie a ter seu genoma representado em um único *contig*;
- A comparação entre os genomas de isolados clínicos e ambientais revelou alguns genes exclusivos que podem estar envolvidos na sua interação e sobrevivência no organismo humano, no entanto, a maior parte dos fatores de resistência, interação patógeno-hospedeiro e genes de resistência são comuns a ambas sugerindo que estes organismos apresentam um potencial de interação e colonização em diferentes hospedeiros.

8 REFERÊNCIAS

ALBERTI-SEGUI, C. et al. HU protein affects transcription of surface polysaccharide synthesis genes in *Porphyromonas gingivalis*. **Journal of Bacteriology**, v. 192, n. 23, p. 6217–6229, 2010.

ALIKHAN, N. F. et al. BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. **BMC Genomics**, v. 12, n. 1, p. 402, 2011.

ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of molecular biology**, v. 215, n. 3, p. 403–10, 1990.

ALTSCHUL, S. F. et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997.

ARELLANO-REYNOSO, B. et al. Cyclic beta-1,2-glucan is a *Brucella* virulence factor required for intracellular survival. **Nature immunology**, v. 6, n. 6, p. 618–625, 2005.

BALDANI, J. I. et al. Characterization of *Herbaspirillum seropedicae* gen. nov. sp. nov. a Root-Associated Nitrogen-Fixing Bacterium. **International journal of systematic bacteriology**, p. 86–93, 1986.

BANDYOPADHYAY, P.; STEINMAN, H. M. Catalase-peroxidases of *Legionella pneumophila*: Cloning of the katA gene and studies of KatA function. **Journal of Bacteriology**, v. 182, n. 23, p. 6679–6686, 2000.

BANKEVICH, A. et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455–477, 2012.

BERG, G. et al. The Rhizosphere as a Reservoir for Opportunistic Human Pathogenic Bacteria. **Molecular Microbial Ecology of the Rhizosphere**, v. 2, p. 1209–1216, 2005.

BINNEWIES, T. T. et al. Ten years of bacterial genome sequencing: Comparative-genomics-based discoveries. **Functional and Integrative Genomics**, v. 6, n. 3, p. 165–185, 2006.

BOHIN, J. P. Osmoregulated periplasmic glucans in Proteobacteria. **FEMS Microbiology Letters**, v. 186, n. 1, p. 11–19, 2000.

CARRO, L. et al. *Herbaspirillum canariense* sp. nov., *Herbaspirillum aurantiacum* sp. nov. and *Herbaspirillum soli* sp. nov., isolated from volcanic mountain soil, and emended description of the genus *Herbaspirillum*. **International Journal of Systematic and Evolutionary Microbiology**, v. 62, n. 6, p. 1300–1306, 2012.

CARVER, T. J. et al. ACT: The Artemis comparison tool. **Bioinformatics**, v. 21, n. 16, p. 3422–3423, 2005.

CASCALES, E.; CHRISTIE, P. J. The versatile bacterial type IV secretion systems.

Nature reviews. Microbiology, v. 1, n. 2, p. 137–49, 2003.

CATIC, A. et al. ElaD, a deubiquitinating protease expressed by *E. coli*. **PLoS ONE**, v. 2, n. 4, p. 1–8, 2007.

CHEN, J. et al. *Herbaspirillum* species: A potential pathogenic bacteria isolated from acute lymphoblastic leukemia patient. **Current Microbiology**, v. 62, n. 1, p. 331–333, 2011.

CHEN, L. et al. VFDB: A reference database for bacterial virulence factors. **Nucleic Acids Research**, v. 33, n. DATABASE ISS., p. 325–328, 2005.

COBURN, B.; SEKIROV, I.; FINLAY, B. B. Type III secretion systems and disease. **Clinical Microbiology Reviews**, v. 20, n. 4, p. 535–549, 2007.

COENYE, T. et al. Characterization of unusual bacteria isolated from respiratory secretions of cystic fibrosis patients and description of *Inquilinus limosus* gen. nov., sp. nov. **Journal of clinical microbiology**, v. 40, n. 6, p. 2062–2069, 2002.

COLLINS, F. S.; MORGAN, M.; PATRINOS, A. The Human Genome Project: Lessons from large-scale biology. **Science**, v. 300, n. 5617, p. 286–290, 2003.

COSTERTON, J. W.; STEWART, P. S.; GREENBERG, E. P. Bacterial biofilms: A common cause of persistent infections. **Science**, v. 284, n. 5418, p. 1318–1322, 1999.

CUTTING, G. R. Cystic fibrosis genetics: from molecular understanding to clinical application. **Nat Rev Genet**, v. 16, n. 1, p. 45–56, 2015.

DE SOUZA, V. et al. Draft Genome Sequence of *Herbaspirillum huttiense* subsp. putei IAM 15032, a Strain Isolated from Well Water. **Genome Announcements**, v. 1, n. 1, p. e00252-12-e00252-12, 2013.

DOHM, J. C. et al. SHARCGS, a fast and highly accurate short-read assembly algorithm for de novo genomic sequencing. **Genome Research**, v. 17, n. 11, p. 1697–1706, 2007.

EATON, K. A. et al. Essential Role of Urease in Pathogenesis of Gastritis Induced by *Helicobacter pylori* in Gnotobiotic Piglets. **Infection and Immunity**, v. 59, n. 7, p. S681–S685, 1991.

EKBLOM, R.; WOLF, J. B. W. A field guide to whole-genome sequencing, assembly and annotation. **Evolutionary Applications**, v. 7, n. 9, p. 1026–1042, 2014.

ELBORN, J. S. Cystic fibrosis. **The Lancet**, v. 388, n. 10059, p. 2519–2531, 2016.

FINN, R. D. et al. The Pfam protein families database. **Nucleic Acids Res** 40: D290–D301., v. 38, n. 1, 2010.

FONG, Y. H. et al. Structure of *UreG/UreF/UreH* Complex Reveals How Urease Accessory Proteins Facilitate Maturation of *Helicobacter pylori* Urease. **PLoS Biology**, v. 11, n. 10, 2013.

- GARDUÑO, R. A. et al. The *Legionella pneumophila* chaperonin - An unusual multifunctional protein in unusual locations. **Frontiers in Microbiology**, v. 2, n. JUNE, p. 1–12, 2011.
- GIBSON, M. K.; FORSBERG, K. J.; DANTAS, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. **ISME Journal**, v. 9, n. 1, p. 207–216, 2015.
- GOODWIN, S.; MCPHERSON, J. D.; MCCOMBIE, W. R. Coming of age: Ten years of next-generation sequencing technologies. **Nature Reviews Genetics**, v. 17, n. 6, p. 333–351, 2016.
- GORIS, J. et al. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. **International Journal of Systematic and Evolutionary Microbiology**, v. 57, n. 1, p. 81–91, 2007.
- GUIZELINI, D. et al. Complete Genome Sequence of *Herbaspirillum hiltneri* N3 (DSM 17495), Isolated from Surface-Sterilized Wheat Roots. **Genome Announcements**, v. 3, n. 5, p. e01288-15, 2015.
- GUIZELINI, D. et al. GFinisher: A new strategy to refine and finish bacterial genome assemblies. **Scientific Reports**, v. 6, n. October, p. 1–8, 2016.
- GUREVICH, A. et al. QUAST: Quality assessment tool for genome assemblies. **Bioinformatics**, v. 29, n. 8, p. 1072–1075, 2013.
- HAGEN, T. A.; CORNELISSEN, C. N. *Neisseria gonorrhoeae* requires expression of *TonB* and the putative transporter TdfF to replicate within cervical epithelial cells. **Molecular Microbiology**, v. 62, n. 4, p. 1144–1157, 2006.
- HAIKO, J.; WESTERLUND-WIKSTRÖM, B. The Role of the Bacterial Flagellum in Adhesion and Virulence. **Biology**, v. 2, n. 4, p. 1242–1267, 2013.
- HARDISON, R. C. Comparative genomics. **PLoS Biology**, v. 1, n. 2, p. 156–160, 2003.
- JAMES, E. K. Nitrogen fixation in endophytic and associative symbiosis. **Field Crops Research**, v. 65, n. 2–3, p. 197–209, 2000.
- JOO, H.-S.; OTTO, M. Molecular basis of in-vivo biofilm formation by bacterial pathogens. **Chem Biol.**, v. 19, n. 12, p. 1503–1513, 2013.
- KANEHISA, M.; SATO, Y.; MORISHIMA, K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. **Journal of Molecular Biology**, v. 428, n. 4, p. 726–731, 2016.
- KIRCHHOF, G. et al. *Herbaspirillum frisingense* sp. nov., a new nitrogen-fixing bacterial species that occurs in C4-fibre plants. **International Journal of Systematic and Evolutionary Microbiology**, v. 51, n. 1, p. 157–168, 2001.
- KLASSEN, G. et al. Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae* SMRI. v. 891, p. 887–891, 1997.

LARSON, J. A. et al. Replication of *Neisseria meningitidis* within Epithelial Cells Requires TonB-Dependent Acquisition of Host Cell Iron. *Journal of Bacteriology*, v. 174, n. 10, p. 3113–3121, 2002.

LIU, L. et al. Comparison of next-generation sequencing systems. **Journal of Biomedicine and Biotechnology**, v. 2012, 2012.

LYNCH, M.; KURAMITSU, H. Expression and role of superoxide dismutases (SOD) in pathogenic bacteria. **Microbes and infection / Institut Pasteur**, v. 2, n. 10, p. 1245–1255, 2000.

MANGAN, M. W. et al. Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in *Salmonella enterica* serovar *Typhimurium*. **Microbiology**, v. 157, n. 4, p. 1075–1087, 2011.

MARQUES, A. C. Q. et al. Biochemical characteristics, adhesion, and cytotoxicity of environmental and clinical isolates of *Herbaspirillum* spp. **Journal of Clinical Microbiology**, v. 53, n. 1, p. 302–308, 2015.

MATSUURA, M. Structural modifications of bacterial lipopolysaccharide that facilitate gram-negative bacteria evasion of host innate immunity. **Frontiers in Immunology**, v. 4, n. MAY, p. 1–9, 2013.

MÉDIGUE, C.; MOSZER, I. Annotation, comparison and databases for hundreds of bacterial genomes. **Research in Microbiology**, v. 158, n. 10, p. 724–736, 2007.

METZKER, M. L. Sequencing technologies the next generation. **Nature Reviews Genetics**, v. 11, n. 1, p. 31–46, 2010.

MITTL, P. R. E.; SCHNEIDER-BRACHER, W. Sel1-like repeat proteins in signal transduction. **Cellular Signalling**, v. 19, n. 1, p. 20–31, 2007.

MIZRAHI-MAN, O.; DAVENPORT, E. R.; GILAD, Y. Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. **PLoS ONE**, v. 8, n. 1, p. 18–23, 2013.

MONTEIRO, R. A. et al. Genomic comparison of the endophyte *Herbaspirillum seropedicae* SmR1 and the phytopathogen *Herbaspirillum rubrisubalbicans* M1 by suppressive subtractive hybridization and partial genome sequencing. **FEMS Microbiology Ecology**, v. 80, n. 2, p. 441–451, 2012.

MORIYA, Y. et al. KAAS: An automatic genome annotation and pathway reconstruction server. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. 182–185, 2007.

NAGARAJAN, N.; POP, M. Sequence assembly demystified. **Nature Reviews Genetics**, v. 14, n. 3, p. 157–167, 2013.

NCBI, Genbank and GWS Statistics. Disponível em:
<https://www.ncbi.nlm.nih.gov/genbank/statistics/> Acesso em: 16 jan 2018.

- NIU, B.; KOLTER, R. crossm Complete Genome Sequences of Seven Strains Composing a Model Bacterial Community of Maize Roots. **Genome Announcements**, v. 5, n. 36, p. 35–36, 2017.
- PARKE, J. L.; GURIAN-SHERMAN, D. Diversity of the *burkholderia cepacia* complex and implications for risk assessment of biological control strains. **Environmental Protection**, v. 84, n. 5, p. 1229–1236, 2001.
- PEDROSA, F. O. et al. Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. **PLoS Genetics**, v. 7, n. 5, 2011.
- PIRO, V. C. et al. FGAP: an automated gap closing tool. **BMC research notes**, v. 7, n. 1, p. 371, 2014.
- POZA-CARRIÓN, C. et al. Kinetics of nif gene expression in a nitrogen-fixing bacterium. **Journal of Bacteriology**, v. 196, n. 3, p. 595–603, 2014.
- PRITCHARD, L. et al. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. **Anal. Methods**, v. 8, n. 1, p. 12–24, 2016.
- RAMARAO, N.; SANCHIS, V. The pore-forming haemolysins of bacillus cereus: a review. **Toxins**, v. 5, n. 6, p. 1119–1139, 2013.
- REGUNATH, H. et al. Severe community-acquired pneumonia with bacteremia caused by *Herbaspirillum aquaticum* or *Herbaspirillum huttiense* in an immune-competent adult. **Journal of Clinical Microbiology**, v. 53, n. 9, p. 3086–3088, 2015.
- RIBAS, Danieli Isabel Romanovitch, Estudo de mutações do gene cftr e da concentração sérica da lectina ligante de manose (MBL) em crianças com fibrose cística identificadas pela triagem neonatal. 2014. 97. Tese. Universidade Federal do Paraná, Curitiba, 2014. Disponível em: <<http://acervodigital.ufpr.br/bitstream/handle/1884/37683/R%20-%20T%20%DANIELI%20ISABEL%20ROMANOVITCH%20RIBAS.pdf?sequence=3&isAllowed=y>> Acesso em 13 de janeiro de 2018.
- RIGANO, L. A. et al. Bacterial cyclic beta-(1,2)-glucan acts in systemic suppression of plant immune responses. **The Plant cell**, v. 19, n. 6, p. 2077–2089, 2007a.
- RIGANO, L. A. et al. Bacterial Cyclic β -(1,2)-Glucan Acts in Systemic Suppression of Plant Immune Responses. **The Plant Cell**, v. 19, n. 6, p. 2077–2089, 2007b.
- ROTHBALLER, M. et al. *Herbaspirillum hiltneri* sp. nov., isolated from surface-sterilized wheat roots. **International Journal of Systematic and Evolutionary Microbiology**, v. 56, n. 6, p. 1341–1348, 2006.
- SANGER, F. et al. Nucleotide sequence of bacteriophage phi X174 DNA. **Nature**, v. 265, n. 5596, p. 687–695, 1977.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences**, v. 74, n.

12, p. 5463–5467, 1977.

SCHECHTER, L. M. et al. The Small Nucleoid-Binding Proteins H-NS , HU , and Fis Affect hilA Expression in *Salmonella enterica* Serovar Typhimurium. v. 71, n. 9, p. 5432–5435, 2003.

SEEMANN, T. Prokka: Rapid prokaryotic genome annotation. **Bioinformatics**, v. 30, n. 14, p. 2068–2069, 2014.

SOARES, S. C. et al. GIPSY: Genomic island prediction software. **Journal of Biotechnology**, v. 232, p. 2–11, 2016.

SPIERS, A. J. et al. Biofilm formation at the air-liquid interface by the *Pseudomonas fluorescens* SBW25 wrinkly spreader requires an acetylated form of cellulose. **Molecular Microbiology**, v. 50, n. 1, p. 15–27, 2003.

SPIPKER, T. et al. Recovery of *Herbaspirillum* species from persons with cystic fibrosis. **Journal of Clinical Microbiology**, v. 46, n. 8, p. 2774–2777, 2008.

STEIN, L. Genome annotation: From sequence to biology. **Nature Reviews Genetics**, v. 2, n. 7, p. 493–503, 2001.

STRAUB, D. et al. The genome of the endophytic bacterium *H. frisingense* GSF30T identifies diverse strategies in the *Herbaspirillum* genus to interact with plants. **Frontiers in Microbiology**, v. 4, n. JUN, p. 1–10, 2013.

SUWANTARAT, N. et al. Fatal case of *Herbaspirillum seropedicae* bacteremia secondary to pneumonia in an end-stage renal disease patient with multiple myeloma. **Diagnostic Microbiology and Infectious Disease**, v. 82, n. 4, p. 331–333, 2015.

TETZ, V.; TETZ, G. Draft Genome Sequence of the. **Genome Announcements**, v. 1, n. 1, p. e00032-13, 2017.

URBAN, M. et al. The Pathogen-Host Interactions database (PHI-base): Additions and future developments. **Nucleic Acids Research**, v. 43, n. D1, p. D645–D655, 2015.

VAN DIJK, E. L. et al. Ten years of next-generation sequencing technology. **Trends in Genetics**, v. 30, n. 9, p. 418–426, 2014.

VERNIKOS, G. S.; PARKHILL, J. Interpolated variable order motifs for identification of horizontally acquired DNA: Revisiting the *Salmonella* pathogenicity islands. **Bioinformatics**, v. 22, n. 18, p. 2196–2203, 2006.

VIPREY, V. et al. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. **Molecular Microbiology**, v. 28, n. 6, p. 1381–1389, 1998.

WANG, L. et al. **HU-induced polymorphous filamentation in fish pathogen *Edwardsiella tarda* leading to reduced invasion and virulence in zebrafish.** [s.l.] Elsevier B.V., 2014. v. 171

WATSON, R. O.; MANZANILLO, P. S.; COX, J. S. Extracellular *M. tuberculosis* DNA targets bacteria for autophagy by activating the host DNA-sensing pathway. **Cell**, v. 150, n. 4, p. 803–815, 2012.

WAYNE, L. G. et al. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. **International Journal of Systematic and Evolutionary Microbiology**, v. 37, n. 4, p. 463–464, 1987.

ZIGA, E. D.; DRULEY, T.; BURNHAM, C. A. D. *Herbaspirillum* species bacteremia in a pediatric oncology patient. **Journal of Clinical Microbiology**, v. 48, n. 11, p. 4320–4321, 2010.