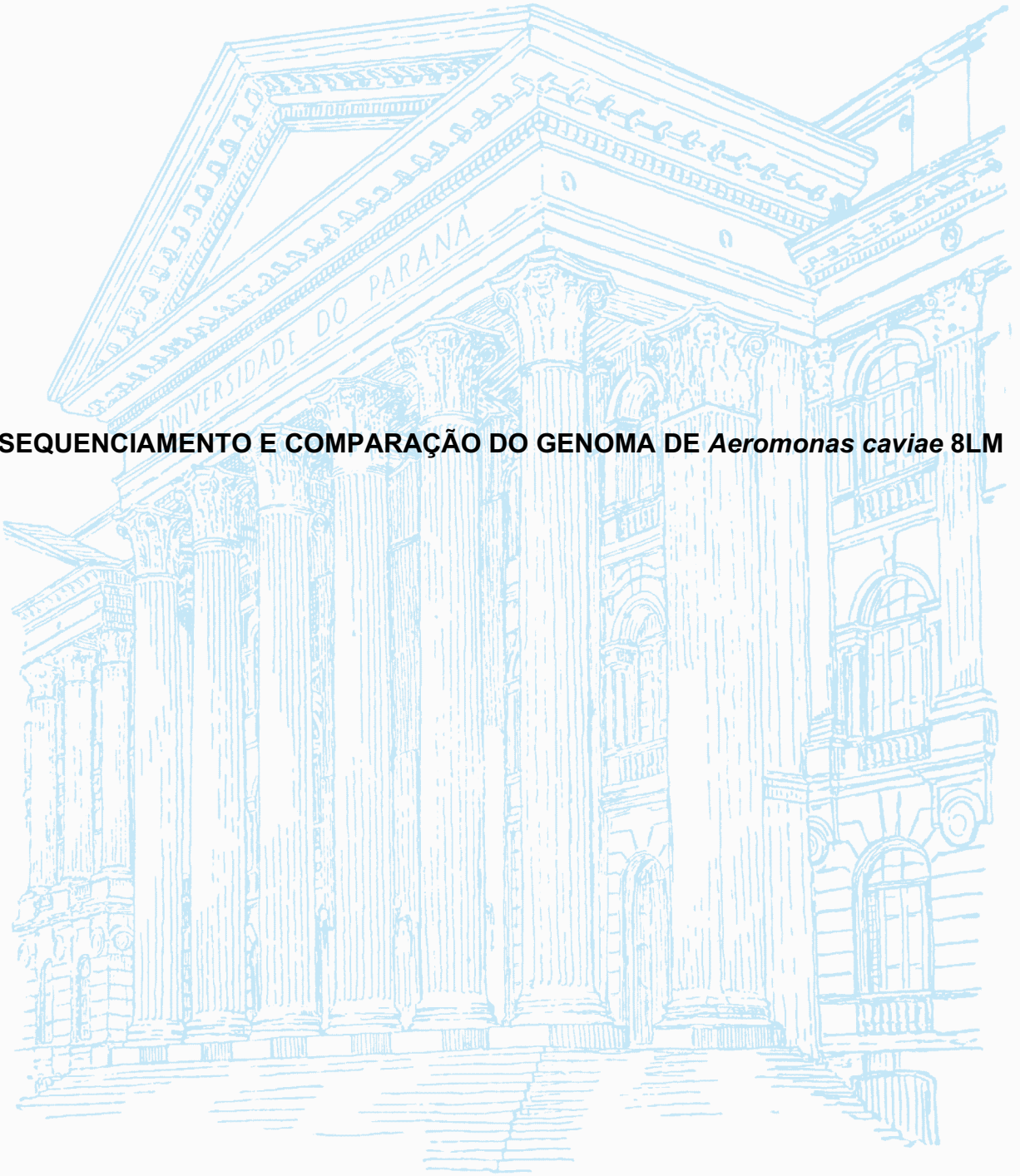


UNIVERSIDADE FEDERAL DO PARANÁ

BÁRBARA MORIEL

SEQUENCIAMENTO E COMPARAÇÃO DO GENOMA DE *Aeromonas caviae* 8LM



CURITIBA

2017

BÁRBARA MORIEL

SEQUENCIAMENTO E COMPARAÇÃO DO GENOMA DE *Aeromonas caviae* 8LM

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Ciências Farmacêuticas, no Curso de Pós-Graduação em Ciências Farmacêuticas, Setor de Ciências da Saúde, da Universidade Federal do Paraná.

Orientadores:

Profa. Dra. Cyntia Maria Telles Fadel-Picheth

Prof. Dr. Leonardo Magalhães Cruz

CURITIBA

2017

Moriel, Bárbara

Sequenciamento e comparação do genoma de *Aeromonas caviae* 8LM / Bárbara Moriel
– Curitiba, 2017.

148 f. ; il. (algumas color.) ; 30 cm

Orientadora: Professora Dra. Cyntia Maria Telles Fadel-Picheth

Coorientador: Professor Dr. Leonardo Magalhães Cruz

Tese (doutorado) – Programa de Pós-Graduação em Ciências Farmacêuticas, Setor de Ciências da Saúde. Universidade Federal do Paraná.

Inclui bibliografia

1. *Aeromonas*. 2. Genoma. 3. Genômica. I. Fadel-Picheth, Cyntia Maria Telles. II. Cruz, Leonardo Magalhães. III. Universidade Federal do Paraná. IV. Título.

CDD 616.014




TERMO DE APROVAÇÃO

BÁRBARA MORIEL

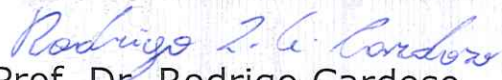
Título: **"Sequenciamento e comparação do genoma de *Aeromonas caviae* 8LM"**

Tese aprovada como requisito parcial para a obtenção de grau de Doutora, no Programa de Pós-Graduação em Ciências Farmacêuticas, da Universidade Federal do Paraná, área de concentração: Análises Clínicas.


Prof^a. Dr^a. Cyntia Maria Telles Fadel Picheth
Orientadora


Prof. Dr. Leonardo Magalhães Cruz
Co-orientador


Prof. Dr. Dieval Guizelini
Universidade Federal do Paraná


Prof. Dr. Rodrigo Cardoso
Universidade Federal do Paraná


Prof^a. Dr^a. Dayane Alberton
Universidade Federal do Paraná


Prof. Dr. Helisson Faoro
Instituto Carlos Chagas

Curitiba, 03 de julho de 2017.

AGRADECIMENTOS

Gostaria de agradecer os meus orientadores, Prof^ª. Cynthia e Prof. Leonardo da Universidade Federal do Paraná, pela oportunidade de trabalhar com esse projeto tão desafiador.

Ao Prof. Rupert Mutzel da Freie Universität Berlin e Prof. Torsten Semmler do Instituto Robert Koch em Berlin pela oportunidade de realizar o estágio de doutorado sanduíche.

Agradeço também o Programa de Pós-Graduação em Ciências Farmacêuticas e o Departamento de Bioquímica e Biologia Molecular da Universidade Federal do Paraná, onde foram realizadas as análises desse projeto.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes, pelo aporte financeiro durante o doutoramento.

Às minhas colegas de laboratório pela ajuda com as análises e as agradáveis conversas durante todos esses anos.

Aos Professores Dieval, Rodrigo e Helisson, do Departamento de Bioquímica e Biologia Molecular da Universidade Federal do Paraná, pela orientação e grande auxílio nas análises de bioinformática.

Ao Douglas por toda ajuda e paciência principalmente em explicar como lidar com as ferramentas de bioinformática.

À minha família e amigos pelo suporte durante todos esses anos.

“A ciência nunca resolve um problema sem criar pelo menos outros dez.”

(George Bernard Shaw)

RESUMO

Aeromonas são bacilos gram-negativos, encontrados em ambientes aquáticos, capazes de causar diversas infecções em humanos destacando-se a diarreia. Os objetivos deste trabalho foram sequenciar o genoma da estirpe *A. caviae* 8LM, isolada a partir de fezes diarreicas, e realizar análises de genômica comparativa. O sequenciamento foi realizado empregando os sistemas 454 GS Junior, Illumina MiSeq e PacBio. Nas análises de genômica comparativa foram empregadas diversas ferramentas de bioinformática para análise estrutural e do conteúdo gênico de estirpes de *Aeromonas*. O genoma de *A. caviae* 8LM apresenta 4,6Mb, conteúdo GC de 61,7%, dispondo de 4.101 CDS e 10 operons de rRNA distribuídos próximos da origem de replicação. Foram identificados diversos genes relacionados com virulência incluindo motilidade e adesão, toxinas, sistemas de secreção tipo II e VI, entre outros. As análises de genômica comparativa mostram a similaridade do genoma de *A. caviae* 8LM com outras estirpes da espécie. As análises realizadas indicaram divergências na identificação de algumas *Aeromonas* e outras, identificadas apenas ao nível de gênero puderam ser identificadas ao nível de espécie. O pan-genoma de *Aeromonas* é formado por 9.785 genes e o core genoma 701 genes. Foram detectadas regiões do genoma que podem conter possíveis marcadores moleculares para a espécie *A. caviae*. Os resultados deste trabalho trazem informações sobre o genoma e virulência de *A. caviae* 8LM, e sugerem a presença de possíveis marcadores moleculares com potencial para o desenvolvimento de testes diagnósticos.

Palavras-chave: *Aeromonas*. Genoma completo. Genômica comparativa.

ABSTRACT

Aeromonas are gram-negative bacilli, found in aquatic environments, are capable of causing several human infections, highlighting diarrhea. The proposes of this work were to sequence the genome of strain *A. caviae* 8LM, isolated from diarrheal feces, and carry out comparative genomics analyzes. Sequencing was performed applying the 454 GS Junior, Illumina MiSeq and PacBio systems. In comparative genomics analyzes, several bioinformatics tools were applied for structural and gene content analysis of *Aeromonas* strains. The genome of *A. caviae* 8LM presents 4.6Mb, GC content of 61.7%, 4,101 CDS and 10 distributed rRNA operons close to the origin of replication. Several genes related to virulence including motility and adhesion, toxins, type II and VI secretion systems, among others were identified. Comparative genomics analyzes exhibited the similarity of the genome of *A. caviae* 8LM with other strains of the species. The analyzes indicated differences in identification of some *Aeromonas* and others that were only identified at genus level could be identified at species level. The pan-genome of *Aeromonas* consists of 9,785 genes and the core genome 701 genes. Genomic regions containing possible molecular markers for *A. caviae* species were detected. The results of this work provide information about the genome and virulence of *A. caviae* 8LM, and suggest the presence of possible molecular markers with potential for the development of diagnostic tests.

Key-words: *Aeromonas*. Complete genome. Comparative genomics.

LISTA DE FIGURAS

FIGURA 1 – SISTEMAS DE SECREÇÃO EM BACTÉRIAS.....	27
FIGURA 2 – SEQUENCIA DE <i>dnaA box</i> EM <i>A. caviae</i> 8LM	49
FIGURA 3 – DISTRIBUIÇÃO DOS OPERONS DE rRNA EM <i>A. caviae</i> 8LM	52
FIGURA 4 – BLAST-ATLAS DE <i>Aeromonas</i> COM GENOMAS COMPLETOS	61
FIGURA 5 – PRESENÇA/AUSÊNCIA DE SEQUÊNCIAS NO PAN-GENOMA DE <i>Aeromonas</i> COM GENOMA COMPLETO.....	66
FIGURA 6 – NÚMERO DE SEQUENCIAS ÚNICAS NAS ESPÉCIES DE <i>Aeromonas</i>	67
FIGURA 7 – PRESENÇA/AUSÊNCIA DE GENES NO PAN-GENOMA DE <i>Aeromonas</i>	71
FIGURA 8 – ÁRVORE FILOGENÉTICA DOS GENES <i>HOUSEKEEPING</i>	74
FIGURA 9 – ÁRVORE FILOGENÉTICA DOS GENES DO CORE GENOMA	76
FIGURA 10 – ORGANIZAÇÃO GÊNICA DO FLAGELO LATERAL	81
FIGURA 11 – ORGANIZAÇÃO GÊNICA DO FLAGELO POLAR	82
FIGURA 12 – ORGANIZAÇÃO GÊNICA DO T6SS	83

LISTA DE GRÁFICOS

GRÁFICO 1 – NÚMERO DE GENOMAS DE <i>Aeromonas</i> DEPOSITADOS NO NCBI	46
GRÁFICO 2 – ANOTAÇÃO DO GENOMA DE <i>A. caviae</i> 8LM PELO RAST	48
GRÁFICO 3 – CURVAS Z PARA O GENOMA DE <i>A. caviae</i> 8LM	49
GRÁFICO 4 – <i>DOTPLOT</i> ENTRE <i>A. caviae</i> E <i>A. caviae</i> 8LM.....	57
GRÁFICO 5 – <i>DOTPLOT</i> ENTRE <i>A. hydrophila</i> E <i>A. caviae</i> 8LM.....	58
GRÁFICO 6 – <i>DOTPLOT</i> ENTRE AS DEMAIS <i>Aeromonas</i> E <i>A. caviae</i> 8LM.....	59
GRÁFICO 7 – NÚMERO DE GENES NO PAN-GENOMA DE <i>Aeromonas</i> COM GENOMAS COMPLETOS.....	65
GRÁFICO 8 – COMPRIMENTO DAS SEQUENCIAS DE AMINOÁCIDOS	73

LISTA DE TABELAS

TABELA 1 – NOMENCLATURA DOS GENES DO T6SS	29
TABELA 2 – TESTES BIOQUÍMICOS PARA IDENTIFICAÇÃO DE <i>A. caviae</i> 8LM.	31
TABELA 3 – GENES <i>HOUSEKEEPING</i> UTILIZADOS PARA A RECONSTRUÇÃO FILOGENÉTICA	41
TABELA 4 – DADOS DO SEQUENCIAMENTO DE <i>A. caviae</i> 8LM	44
TABELA 5 – GENOMAS DE <i>Aeromonas</i> spp. SEQUENCIADOS	47
TABELA 6 – OPERONS DE RNA RIBOSSOMAL EM <i>A. caviae</i> 8LM	50
TABELA 7 – MAIORES VALORES DE GGDC–DDH ENTRE <i>Aeromonas</i>	54
TABELA 8 – MAIORES VALORES DE ANI ENTRE <i>Aeromonas</i>	55
TABELA 9 – ILHAS GENÔMICAS DE <i>A. caviae</i> 8LM	62
TABELA 10 – NOVA IDENTIFICAÇÃO DAS ESTIRPES DE <i>Aeromonas</i>	77
TABELA 11 – FATORES DE VIRULÊNCIA EM <i>A. caviae</i> 8LM	79

LISTA DE ABREVIATURAS E SIGLAS

ANI	- Identidade média de nucleotídeos
ATP	- Adenosina trifosfato
BLAST	- <i>Basic Local Alignment Search Tool</i>
BLASTp	- <i>Protein Basic Local Alignment Search Tool</i>
cAMP	- Adenosina monofostato cíclico
CDS	- Sequência codificadora de ácido desoxirribonucleico
DDH	- Hibridização ácido desoxirribonucleico – ácido desoxirribonucleico
DNA	- Ácido desoxirribonucleico
GGDC	- <i>Genome-to-Genome Distance Calculator</i>
HSPs	- Segmentos de alta pontuação
MCL	- <i>Markov Cluster Algorithm</i>
MUMmer	- <i>Maximal Unique Matches</i>
NCBI	- <i>National Center for Biotechnology Information</i>
RAST	- <i>Rapid Annotation using Subsystem Technology</i>
RNA	- Ácido ribonucleico
rRNA	- Ácido ribonucleico ribossômico
T2SS	- Sistema de secreção tipo II
T3SS	- Sistema de secreção tipo III
T6SS	- Sistema de secreção tipo VI
tRNA	- Ácido ribonucleico transportador

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO DE LITERATURA.....	15
2.1	CARACTERÍSTICAS DE <i>AEROMONAS</i>	15
2.2	INFECÇÕES CAUSADAS POR <i>AEROMONAS</i>	16
2.2.1	Síndromes do trato gastrointestinal	16
2.2.2	Infecções de pele e tecidos moles.....	18
2.2.3	Septicemia	18
2.2.4	Infecções diversas	20
2.3	DIAGNÓSTICO LABORATORIAL	20
2.4	GENÔMICA	22
2.5	VIRULÊNCIA	24
3	OBJETIVOS	30
3.1	OBJETIVO GERAL.....	30
3.2	OBJETIVOS ESPECÍFICOS	30
4	MATERIAL E MÉTODOS.....	31
4.1	CARACTERIZAÇÃO DA AMOSTRA.....	31
4.2	CONDIÇÕES DE CULTIVO	31
4.3	EXTRAÇÃO DO DNA.....	32
4.4	SEQUENCIAMENTO.....	32
4.4.1	Plataforma 454 GS Jr.	33
4.4.2	Plataforma Illumina	33
4.4.3	Plataforma PacBio	33
4.5	MONTAGEM DOS CONTIGS	33
4.5.1	Plataforma 454 GS Junior	33
4.5.2	Plataforma Illumina MiSeq.....	34
4.5.3	Plataforma PacBio	34
4.6	ANOTAÇÃO DO GENOMA	34
4.6.1	Rapid Annotation using Subsystem Technology – RAST.....	34
4.6.2	Ori-Finder.....	34
4.6.3	RNAmmer	35
4.7	ORDENAÇÃO DOS CONTIGS E FECHAMENTO DE GAPS.....	35
4.7.1	Ordenação dos contigs.....	35

4.7.2	Fechamento de GAPS.....	36
4.7.3	Submissão das sequências.....	36
4.8	GENÔMICA COMPARATIVA.....	36
4.8.1	Similaridade entre genomas.....	37
4.8.1.1	Genome-to-Genome Distance Calculator – GGDC.....	37
4.8.1.2	Average Nucleotide Identity – ANI.....	37
4.8.2	Análise estrutural.....	38
4.8.3	Análise do conteúdo gênico.....	38
4.8.3.1	BLAST Ring Image Generator – BRIG.....	38
4.8.3.2	Ilhas genômicas.....	38
4.8.3.3	Agrupamento de genes.....	39
4.9	ANÁLISE FILOGENÉTICA.....	40
4.9.1	<i>Multilocus Sequence Analysis</i> – MLSA.....	40
4.9.2	Reconstrução filogenética baseada no <i>core</i> genoma.....	42
4.10	VIRULÊNCIA E RESISTÊNCIA.....	42
5	RESULTADOS E DISCUSSÃO.....	44
5.1	SEQUENCIAMENTO E MONTAGEM DOS CONTIGS.....	44
5.1.1	Genomas de <i>Aeromonas</i>	45
5.2	ANOTAÇÃO DO GENOMA.....	47
5.2.1	Rapid Annotation using Subsystem Technology – RAST.....	47
5.2.2	Ori-Finder.....	48
5.2.3	RNAmmer.....	50
5.3	GENÔMICA COMPARATIVA.....	53
5.3.1	Similaridade entre genomas.....	53
5.3.1.1	Genome-to-genome Distance Calculator – GGDC.....	53
5.3.1.2	Average Nucleotide Identity – ANI.....	55
5.3.2	Análise estrutural.....	56
5.4	ANÁLISE DO CONTEÚDO GÊNICO.....	60
5.4.1	BLAST Ring Image Generator – BRIG.....	60
5.4.2	Ilhas genômicas.....	61
5.4.3	Análise de pan-genoma.....	64
5.5	ANÁLISE FILOGENÉTICA.....	72
5.5.1	Multilocus Sequence Analysis – MLSA.....	72
5.5.2	Análise filogenética baseada no <i>core</i> genoma.....	75

5.6	VIRULÊNCIA E RESISTÊNCIA.....	78
5.6.1	Genes de virulência	78
5.6.2	Sistema de secreção tipo VI – T6SS.....	82
6	CONCLUSÃO.....	84
	ANEXO A – GENOMAS DE <i>AEROMONAS</i> DEPOSITADOS NO NCBI	109
	ANEXO C – SIMILARIDADE PELO GGDC–DDH.....	129
	ANEXO D – SIMILARIDADE PELO ANI	132
	ANEXO E – CORE GENOMA DE <i>AEROMONAS</i> (COMPLETOS).....	135
	ANEXO F – ÁRVORE FILOGENÉTICA DO CORE GENOMA	147

1 INTRODUÇÃO

Aeromonas são bacilos gram-negativos, anaeróbios facultativos, oxidase positiva e fermentam a glucose (MARTIN-CARNAHAN & JOSEPH, 2005). São amplamente distribuídas em ambientes aquáticos (KÜHN et al., 1997; KHAJANCHI et al., 2010; MORRIS & HORNEMAN, 2017), e algumas espécies são capazes de causar uma grande variedade de doenças em humanos. A manifestação mais comum entre as infecções causadas por *Aeromonas* em humanos é a diarreia (MORRIS & HORNEMAN, 2017).

As doenças diarreicas são a segunda maior causa de morte em todo o mundo. Estas infecções causam morbidade significativa e morte, particularmente em pessoas idosas e crianças abaixo de 5 anos de idade. Segundo o relatório de 2015 sobre os Níveis e Tendências em Mortalidade Infantil, divulgado pelo Fundo da Organização das Nações Unidas para a Infância (UNICEF) e outras agências colaboradoras, o número de mortes de crianças menores de 5 anos, devido à diarreia, caiu de 12,7 milhões por ano em 1990 para 5,9 milhões em 2015. Apesar desta redução, a diarreia infecciosa continua como um dos principais problemas de saúde pública em todo o mundo (KELLY, 2011).

A diarreia contribui para a má nutrição e para o retardamento do desenvolvimento físico e mental (O'LOUGHLIN et al., 1991). Embora a maioria dessas infecções se resolva após poucos dias, com ou sem intervenção medicamentosa, pode ocorrer progressão para quadros severos tais como bacteremia, meningite e síndrome urêmica-hemolítica deixando sequelas crônicas como insuficiência renal e problemas neurológicos (LINDSAY, 1997; SWAMINATHAN et al., 1999).

Os primeiros indícios sobre o papel etiológico de *Aeromonas* como causa de diarreia datam da década de 1960 (FARMER III, 2006). Mas estas bactérias emergiram como importantes enteropatógenos devido à suspeita de terem provocado surtos de doença diarreica (CHOPRA & HOUSTON, 1999; HOFER et al., 2006), e do seu isolamento de pessoas com diarreia do viajante (CHOPRA & HOUSTON, 1999; VILA et al., 2003).

No entanto, a maioria dos laboratórios clínicos não pesquisa *Aeromonas* na rotina laboratorial para a identificação de enteropatógenos (MORRIS & HORNEMAN, 2017), o que sugere que muitos casos de diarreia causados por estas bactérias não são diagnosticados.

Três espécies de *Aeromonas* são isoladas de casos de diarreia com maior frequência: *A. hydrophila*, *A. caviae* e *A. veronii* biovar sobria (KANNAN et al., 2001; AGUILERA-ARREOLA et al., 2007; von GRAEVENITZ, 2007). Entre estas, *A. hydrophila* é a espécie melhor estudada (ZHANG et al., 2005; YU et al., 2005; SESHADRI et al., 2006; GRIM et al., 2013, 2014; PANG et al., 2015).

Ainda há poucos estudos sobre *Aeromonas* realizados no Brasil, mas estes indicam que *A. caviae* é uma causa importante de diarreia no país (HOFER et al., 2006, GUERRA et al., 2007; SUREK et al., 2010; ASSIS et al., 2014).

Os avanços na tecnologia de sequenciamento de DNA e o desenvolvimento da bioinformática permitiram o sequenciamento em grande escala de genomas inteiros de patógenos e a sua comparação e análise detalhada, ampliando a capacidade para o estudo de agentes infecciosos (KWONG et al., 2015).

A genômica tem muitas aplicações potenciais na microbiologia médica, como o desenvolvimento de ferramentas para diagnóstico de agentes infecciosos, identificação de marcadores de virulência e de resistência a antimicrobianos e o desenvolvimento de meios de cultura específicos (FOURNIER et al., 2014).

Assim, o sequenciamento do genoma de *A. caviae* 8LM, isolada de fezes de uma criança com diarreia, pode revelar informações importantes sobre a biologia do patógeno, permitir o reconhecimento de marcadores de resistência a antimicrobianos, e auxiliar na compreensão da patogenicidade desse micro-organismo através da identificação de genes associados à virulência.

2 REVISÃO DE LITERATURA

2.1 CARACTERÍSTICAS DE *Aeromonas*

Aeromonas são bacilos gram-negativos retos, anaeróbios facultativos, quimio-organotróficos e não esporulam. Ocorrem isoladamente, ocasionalmente aos pares ou em cadeias curtas. Apresentam metabolismo fermentativo e respiratório com oxigênio como aceptor final de elétrons. Reduzem nitratos à nitritos mas não desnitrificam. Sais de amônio são utilizados como única fonte de nitrogênio. São catalase e oxidase positivos (MARTIN-CARNAHAN & JOSEPH, 2005).

Atualmente são reconhecidas 27 espécies de *Aeromonas* (BEAZ-HIDALGO et al., 2015), a maioria é mesofílica, mas existem espécies psicrófilas; geralmente são móveis, via flagelo polar, e algumas estirpes podem expressar flagelos laterais, e ainda algumas espécies são imóveis (MARTIN-CARNAHAN & JOSEPH, 2005).

Aeromonas são organismos primariamente aquáticos, isolados mais frequentemente de água doce (rios, lagos, águas subterrâneas, águas estuarinas) e em associação com animais aquáticos (MARTIN-CARNAHAN & JOSEPH, 2005; KHAJANCHI et al., 2010). Também podem ser encontradas em água potável, águas residuais e esgotos com vários níveis de tratamento (HOLMES et al., 1996; KÜHN et al., 1997; KHAJANCHI et al., 2010).

No Brasil, um dos estudos sobre a distribuição ambiental de *Aeromonas* foi realizado por FALCÃO e colaboradores (1998). A bactéria foi encontrada em 50% das amostras de água corrente, 33,4% das amostras de água de irrigação, e de 16,6% da água de reservatório na cidade de Araraquara – SP. A maioria dos isolados pertencia à espécie *A. hydrophila*.

Aeromonas também foram isoladas em água corrente na região de Londrina – PR, e os isolados foram identificados como *A. hydrophila*, *A. caviae* e *A. sobria*. As amostras de água analisadas foram coletadas em área não poluída, próxima a nascente (GIBOTTI et al., 2000). Portanto estes micro-organismos estão presentes em ambientes aquáticos de diferentes regiões geográficas.

Além disso, *Aeromonas* podem ser isoladas de animais domésticos, invertebrados, pássaros, insetos e do solo, sendo ubiqüitárias na biosfera microbiana (JANDA & ABBOTT, 2010). A presença de *Aeromonas* também foi detectada em diversos tipos de alimentos incluindo produtos lácteos, vegetais, carnes e aves,

mariscos e em peixes (BORREL et al., 1998; NEYTS et al., 2000; MCMAHON & WILSON, 2001).

A ampla distribuição de *Aeromonas* facilita seu contato com humanos e o eventual desenvolvimento de infecção (JANDA & ABBOTT, 2010).

2.2 INFECÇÕES CAUSADAS POR *Aeromonas*

Bactérias do gênero *Aeromonas* são importantes causas de doenças em peixes e animais peçonhentos, e algumas espécies também são responsáveis por várias infecções em humanos, tanto em indivíduos saudáveis como imunocomprometidos (JANDA & ABBOTT, 2010; CHUANG et al., 2011).

Três espécies – *A. hydrophila*, *A. caviae* e *A. veronii* bv. sobria – são responsáveis pela grande maioria ($\geq 85\%$) das infecções humanas causadas por bactérias deste gênero (JANDA & ABBOTT, 1998; PARKER & SHAW, 2011; TOMÁS, 2012; COLSTON et al., 2014).

As infecções clínicas causadas por *Aeromonas* em humanos são comumente classificadas em quatro categorias: síndromes do trato gastrointestinal, infecções em tecidos moles e lesões, septicemia e infecções diversas, que incluem doenças de menor frequência (JANDA & ABBOTT, 2010).

2.2.1 Síndromes do trato gastrointestinal

Entre as infecções humanas causadas por *Aeromonas* a gastroenterite é a mais frequente (FIGUEIRAS, 2005; PARKER & SHAW, 2011; ASSIS et al., 2014). É adquirida através da ingestão de água ou de alimentos contaminados com a bactéria (BUCHAMAN et al., 1984; NOJIMOTO et al., 1997; JANDA & ABBOTT, 2010; KHAJANCHI et al., 2010; TOMÁS, 2012). Ocorre mundialmente e acomete pessoas de todas as idades; a frequência varia de 0,09 a 13% como indicam diferentes estudos realizados em distintas regiões geográficas (ESSERS et al., 2000; JANDA & ABBOTT, 2010).

As manifestações clínicas decorrentes das infecções gastrointestinais causadas por *Aeromonas* variam e incluem diarreia secretora aguda, frequentemente acompanhada por vômito; diarreia disentérica aguda com sangue e muco; diarreia crônica durando mais de 10 dias; diarreia colérica com fezes “em água de arroz”; diarreia do viajante. Foram relatados dois casos associados com a síndrome urêmico-

hemolítica (MORRIS & HORNEMAN, 2017). Pode ainda haver complicações como a peritonite, colite e colangite (PARKER & SHAW, 2011).

Estudos têm demonstrado um aumento no isolamento dessas bactérias em amostras fecais durante os meses mais quentes do ano. A elevada concentração dessas bactérias nos ambientes aquáticos durante os meses quentes do ano pode levar a um aumento da exposição à *Aeromonas* e consequente desenvolvimento de infecção ou colonização por essas bactérias (JANDA & ABBOTT, 2010).

Essa mesma sazonalidade é observada em infecções extra-intestinais causadas por *Aeromonas*, como a septicemia, na qual 42 a 67% dos casos ocorrem durante o verão (KELLY et al., 1993; LLOPIS et al., 2004; TSAI et al., 2006).

O estudo de *Aeromonas* como patógeno entérico ainda é escasso no Brasil. NOJIMOTO e colaboradores (1997) observaram a presença de *Aeromonas* em 21,9% das fezes diarreicas de crianças na cidade de Goiânia – GO. Foram isoladas as seguintes espécies: *A. caviae* (7,7%), *A. salmonicida* subsp. *salmonicida* (6,6%), *A. sobria* (4,3%), *A. hydrophila* (2,2%) e *A. salmonicida* subsp. *achromogenes* (1,1%).

Em 2004 foi relatado um surto de diarreia ocorrido em São Bento do Una – PE, no qual foram registrados 2.170 casos da doença. Das 582 amostras analisadas 19,5% apresentavam *Aeromonas*, das quais 9,8% eram *A. caviae*, 3,9% *A. veronii* biovar *sobria*, 2,6% *A. veronii* biovar *veronii* e 3,2% de outras espécies (HOFER et al., 2006).

PEREIRA e colaboradores (2008) avaliaram 2.323 neonatos hospitalizados em unidades de terapia intensiva com sintomas de diarreia aguda na cidade do Rio de Janeiro durante o período de 1998 a 2006. Foram isoladas 56 estirpes de *Aeromonas*, sendo 42,8% *A. caviae*, 25% *A. media*, 10,7% *A. veronii* biogrupo *sobria*, 9% *A. hydrophila* e 5,3% *A. veronii* biogrupo *veronii*. Foi sugerido que *Aeromonas* seja considerada nas diretrizes de controle preventivo de infecções hospitalares em unidades de terapia intensiva neonatal.

No Paraná a prevalência de *Aeromonas* foi de 2,6% em uma amostragem de pessoas atendidas em laboratórios clínicos particulares. *A. caviae* foi a espécie predominante (SUREK et al., 2010). No entanto, em pacientes com diarreia atendidos pelo Sistema Único de Saúde, através do Laboratório Municipal de Saúde de Curitiba, a frequência de *Aeromonas* foi de 14,6%, sendo 10,5% de *A. caviae* (ASSIS et al., 2014).

2.2.2 Infecções de pele e tecidos moles

Pele e tecidos moles representam o segundo sítio anatômico do qual *Aeromonas* são mais frequentemente isoladas (JANDA & ABBOTT 1998; 2010). Estas infecções ocorrem frequentemente nas extremidades, após traumas aquáticos e podem variar de lesões tóxicas leves, como lesões postulares, a infecções graves com risco de vida (MORRIS & HORNEMAN, 2017) como mionecrose e fasciíte necrotizante (BEHERA et al., 2014; SOOD & NERURKAR, 2014).

Estas infecções são frequentemente decorrentes de traumas ocupacionais ou lesões decorrentes de atividades recreativas esportivas, tais como pesca e natação. O contato direto de pequenas lesões e abrasões com ambientes aquáticos contaminados com *Aeromonas* como riachos e lagos também pode resultar em infecção (JANDA e ABBOTT, 1998; LAI et al., 2007a). Isto pode ser demonstrado pelo fato de *Aeromonas* estar entre os isolados bacterianos mais comuns recuperados de ferimentos de vítimas do tsunami ocorrido no oceano Índico em 2004 e da inundação em Nova Orleans em 2005 (MORRIS e HORNEMAN, 2017).

Lesão tecidual mais acentuada pode resultar de traumas como mordidas de animais ou da introdução de corpos estranhos (terra, madeira ou metal) contendo *Aeromonas* nos tecidos mais profundos (JANDA & ABBOTT, 2010).

A infecção de feridas por *Aeromonas* através de picadas de cobra (MUKHOPADHYAY et al., 2008) e mordidas de vários animais é uma das rotas mais negligenciadas, no entanto a microbiota da orofaringe de répteis, principalmente de cobras, muitas vezes abriga espécies de *Aeromonas* (JANDA e ABBOTT, 2010).

Queimaduras por fogo ou eletricidade são muitas vezes controladas inicialmente pela imersão total ou parcial da área lesionada em água de torneira, bicas ou riachos contaminados por *Aeromonas* (KIENZLE et al., 2000). Tal ação pode resultar na colonização por *Aeromonas* ou infecção dos tecidos lesionados, resultando até em septicemia (KIENZLE et al., 2000; WILCOX et al.; 2000; CHIM & SONG, 2007; LAI et al., 2007b).

2.2.3 Septicemia

Três espécies de *Aeromonas*, *A. hydrophila*, *A. caviae* e *A. veronii* bv. sobria, são responsáveis por mais de 95% dos casos de septicemia causadas por *Aeromonas*

(JANDA & ABBOTT, 2010). No entanto, a espécie *A. dhakensis*, recentemente descrita (BEAZ-HIDALGO et al., 2013), está sendo relatada como causa importante de septicemia (MORINAGA et al., 2013).

A maioria dos casos de septicemia associada à *Aeromonas* é observada em pessoas imunocomprometidas. Neste grupo, os indivíduos com maior risco de desenvolver septicemia são os que apresentam doença crônica no fígado (TSAI et al., 2005; LAI et al., 2007a; WU et al., 2007), câncer e doenças hepatobiliares (WU et al., 2007; JANDA & ABBOTT 2010). As portas de entrada mais comuns de *Aeromonas* nesses pacientes são por meio do trato gastrointestinal, tecidos moles, intra-abdominal e o uso de catéteres de longa permanência contaminados.

Septicemia por *Aeromonas* também pode ocorrer em pacientes com histórico de evento traumático severo precedendo imediatamente a infecção. As principais portas de entrada para o micro-organismo são o trato respiratório e tecidos cutâneo e subcutâneo. Este grupo de indivíduos muitas vezes não apresenta condições pré-existentes que os levariam a desenvolver bacteremia causada por *Aeromonas* (JANDA & ABBOTT, 2010).

A septicemia causada por espécies de *Aeromonas* também pode ocorrer em pessoas saudias, sem fatores de risco aparentes, na ausência de traumas, e as portas de entrada do micro-organismo não são conhecidas. DWIVEDI e colaboradores (2008), descreveram casos de septicemia adquirida na comunidade em pessoas imunocompetentes, previamente saudáveis, sem história recente de diarreia. No entanto a mesma bactéria foi isolada das fezes dos doentes indicando que eram portadores assintomáticos de *Aeromonas* e sugerindo que o micro-organismo, a partir do intestino, teve acesso à corrente circulatória causando septicemia (DWIVEDI et al., 2008).

Por fim, outro grupo de pacientes com septicemia causada por *Aeromonas* inclui pessoas que se submeteram a cirurgia reconstrutiva. Nestes, a terapia medicinal com sanguessugas, as quais são frequentemente aplicadas em tecidos ou áreas de replante como resultado da cirurgia plástica e reconstrutiva a fim de aliviar o acúmulo sanguíneo, pode ter sido a porta de entrada de *Aeromonas* pelo tecido lesado (JANDA & ABBOTT, 2010).

2.2.4 Infecções diversas

Vários outros tipos de infecção podem ser causados por *Aeromonas* afetando diversos sítios anatômicos. Estas incluem artrite séptica (TSENG et al., 2013); infecções abdominais (TENA et al., 2007); do trato hepatobiliar (WU et al., 2007; LAMY et al., 2009); trato respiratório (MUKHOPADHYAY et al., 2008; LAMY et al., 2009; CHAO et al., 2013); peritonite (CHOI et al., 2008; LAMY et al., 2009); meningite (OUDERKIRK et al., 2004; BUKHARI et al., 2011; PAMPÍN et al., 2012); infecção do trato urinário (AL-BENWAN et al., 2007; TENA et al., 2007; LAMY et al., 2009); prostatite (HUANG et al., 2007); doença inflamatória pélvica (MUKHOPADHYAY et al., 2008); infecção ocular (LAMY et al., 2009).

2.3 DIAGNÓSTICO LABORATORIAL

Nas últimas décadas ocorreu uma grande expansão no número de novas espécies bacterianas. Este fato é devido em grande parte à relativa facilidade com que as sequências do gene 16S rRNA são determinadas.

Para o gênero *Aeromonas*, que em 1980 compreendia apenas quatro espécies – *A. hydrophila*, *A. punctata*, *A. salmonicida* e *A. sobria* – foram descritas sete novas espécies desde 2002 e atualmente existem 27 espécies validadas (JANDA & ABBOTT, 2010; COLSTON et al., 2014; BEAZ-HIDALGO et al., 2015).

A identificação de *Aeromonas* ao nível de espécie pode ser um desafio e pouquíssimos laboratórios clínicos são capazes de identificar as espécies clinicamente significativas além do nível de complexos (JANDA & ABBOTT, 2010), como os complexos *A. hydrophila* e *A. caviae*.

Aeromonas crescem bem em meios utilizados na rotina laboratorial como por exemplo, ágar sangue e ágar MacConkey. A hemólise em ágar sangue é variável. A maioria das espécies incluindo *A. hydrophila* e *A. veronii* apresenta hemólise total (β) e muitas *A. caviae* exibem hemólise parcial (α) embora também possam apresentar zonas estreitas de β hemólise. Em ágar MacConkey a maioria cresce como não fermentadores de lactose, mas muitos isolados de *A. caviae* fermentam a lactose (MARTIN-CARNAHAN & JOSEPH, 2005), como *Escherichia coli* e podem não ser distinguidos desta (MORRIS & HORNEMAN, 2017).

Aeromonas podem ser confundidas com bactérias de outros gêneros como *Vibrio* e *Plesiomonas* (ABBOTT et al., 2003; JANDA & ABBOTT, 2010). Testes específicos tais como a resistência ao agente vibriostático 2,4-diamino-6,7-diisopropil-pteridina (O129), habilidade de crescer em caldo nutriente com concentrações entre 0 a 3% de NaCl, e inabilidade de crescer em ágar TCBS são necessários, mas estes não são empregados na rotina dos laboratórios clínicos (JANDA & ABBOTT, 2010).

A identificação de *Aeromonas* por testes bioquímicos convencionais requer a utilização de um grande número de testes (ABBOTT et al., 2003; MARTIN-CARNAHAN & JOSEPH, 2005). Mesmo assim, nem sempre é possível chegar a identificação de uma espécie.

SUREK e colaboradores (2010) compararam a identificação de *Aeromonas* por provas bioquímicas convencionais e pelo kit comercial API 20E. Para todas as estirpes analisadas foi necessário realizar provas bioquímicas adicionais para distinguir entre *Vibrio fluvialis* e *Aeromonas* ou para identificar as estirpes ao nível de espécie. Este dado mostra que mesmo sistemas comerciais para identificação bacteriana podem apresentar dificuldades na identificação de *Aeromonas*.

A ferramenta molecular mais comumente empregada em laboratórios clínicos para a identificação bacteriana é o sequenciamento do gene 16S rRNA (JANDA & ABBOTT, 2007). Entretanto, esse gene não é considerado um bom marcador para o gênero *Aeromonas*, pois as espécies são estreitamente relacionadas, e níveis elevados de similaridade na sequência de 16S rRNA (97,8 a 100%) são observados. (MARTINEZ-MURCIA et al., 1992; MARTIN-CARNAHAN & JOSEPH, 2005). Outro fator é o elevado número de cópias do gene que varia de 8 a 11 nas diferentes estirpes (ROGER et al., 2012) e a sua heterogeneidade intragenômica, cuja frequência de polimorfismos varia de 0,06 a 1,5% (MORANDI et al., 2005; ALPERI et al., 2008). Assim, o sequenciamento de 16S rRNA define claramente um isolado apenas ao nível de gênero (MARTIN-CARNAHAN & JOSEPH, 2005).

Devido a esses fatores, o sequenciamento dos genes *gyrB* (subunidade beta da DNA girase), *rpoD* (fator sigma 70), ou *rpoB* (subunidade beta da RNA polimerase), tem sido utilizado para a separação das espécies de *Aeromonas*, já que permitem a diferenciação de grupos intimamente relacionados (SOLER et al., 2004; MARTIN-CARNAHAN & JOSEPH, 2005).

2.4 GENÔMICA

O sequenciamento de genomas tem contribuído enormemente para os estudos de taxonomia. Com o avanço tecnológico novas metodologias de sequenciamento foram desenvolvidas e o número de genomas bacterianos disponíveis nos bancos de dados públicos aumentou consideravelmente (van TONDER et al., 2014). Consequentemente, a utilização dessas sequências genéticas oferece novas oportunidades para a caracterização bacteriana e esta abordagem vem sendo utilizada para esclarecer posições taxonômicas duvidosas (COLSTON et al., 2014; BEAZ-HIDALGO et al., 2015).

Além disso, o aumento da capacidade computacional favoreceu os estudos de genômica comparativa envolvendo um número grande de sequências, bem como o desenvolvimento de metodologias analíticas fazendo com que as análises de genômica comparativa se tornem mais acessíveis à comunidade científica (van TONDER et al., 2014). Atualmente, a comparação de genomas inteiros representa uma conduta primordial na classificação de espécies bacterianas (COLSTON et al., 2014; BEAZ-HIDALGO et al., 2015).

Nas análises de genômica comparativa verifica-se que há genes comuns a todas as espécies de um determinado gênero, denomina-se este conjunto de genes como core-genoma. Os outros genes que são compartilhados por algumas espécies e aqueles que são específicos de uma determinada espécie fazem parte do chamado genoma acessório. O pan-genoma é o somatório do core-genoma e do genoma acessório (ALCARAZ, 2014). Dessa maneira, as espécies bacterianas podem ser definidas pelo seu pan-genoma (van TONDER et al., 2014).

Questões evolutivas também podem ser respondidas com a análise de core-genoma como, por exemplo, quais são os genes conservados ao longo da escala evolutiva, quais os valores de corte para similaridade. Além disso, a reconstrução filogenética utilizando a totalidade do genoma talvez esteja mais próxima da história evolutiva do organismo que a genealogia de genes individuais (LUCIANI et al., 2012).

Com os dados do pan-genoma, o padrão de presença e ausência de genes no genoma acessório das bactérias analisadas poderia dar indícios da perda e aquisição de genes por meio de transferência horizontal ao longo do processo evolutivo (RENO & HELD, 2009).

A relação de aquisição e perda de genes pode ser importante na determinação da estirpe, bem como na localização de espécies (filogeografia) e na dinâmica populacional. A determinação do pan-genoma é importante para a manutenção de um registro completo das capacidades metabólicas de um grupo de organismos. Assim como, as diferenças no genoma acessório de bactérias próximas podem dar indícios da adaptação ao ambiente e estilo de vida desses micro-organismos (RENO & HELD, 2009).

Na tentativa de melhor resolver a identificação de espécies bacterianas, foram desenvolvidas ferramentas para análise de similaridade entre genomas. A hibridização DNA-DNA (DDH) foi amplamente utilizada na determinação de espécies e por muito tempo foi o único método capaz de mensurar o grau de similaridade entre genomas (RICHTER & ROSSELLÓ-MÓRA, 2009).

Posteriormente, o GGDC-DDH foi criado como uma ferramenta *in silico* para atender as necessidades advindas dos avanços no sequenciamento de genomas. Nessa análise, o genoma em estudo é comparado *in silico* com um genoma de referência, através de sua sequência de nucleotídeos, e a distância intergenômica é calculada pela soma de todas as identidades encontradas nos segmentos de alta pontuação (HSPs) dividida pelo comprimento global de HSPs. Dessa maneira, os dados são convertidos em valores de similaridade análogas ao DDH para apoiar a decisão sobre o parentesco da nova estirpe com estirpes conhecidas (MEIER-KOLTHOFF, 2013; AUCH et al., 2010a, 2010b).

A abordagem da identidade nucleotídica média, do inglês *Average Nucleotide Identity*, (ANI) é semelhante a do GGDC-DDH, porém essa metodologia estima a identidade média de nucleotídeos entre dois genomas. O genoma em estudo é fragmentado *in silico* em segmentos de 1.020 nucleotídeos, os quais são alinhados utilizando o algoritmo do BLAST contra um genoma de referência a fim de calcular a identidade média dos nucleotídeos (GORIS et al., 2007).

Conforme estudos realizados por COLSTON e colaboradores (2014), essas ferramentas mostraram-se bastante promissoras na separação de espécies proximamente relacionadas, demonstrando que a genômica comparativa pode ser um bom caminho para a resolução taxonômica de *Aeromonas*.

2.5 VIRULÊNCIA

As bactérias do gênero *Aeromonas* podem produzir uma variedade de elementos denominados fatores de virulência, associados com a sua capacidade de causar doença. Estes fatores incluem características com variadas funções, e incluem desde adesinas e toxinas à sistemas de secreção (JANDA & ABBOTT, 2010; PARKER E SHAW, 2011).

A colonização do hospedeiro é uma etapa inicial do processo de infecção. Inicia-se pela adesão da bactéria às células do hospedeiro e está relacionada com a presença de estruturas como pilus (ou fímbrias) e adesinas.

As adesinas são estruturas que permitem a ligação da bactéria a receptores específicos nas células hospedeiras. Em *Aeromonas* diversas adesinas foram descritas e incluem as adesinas fimbriais, presentes na extremidade dos pili/fímbrias, e as localizadas na membrana externa da bactéria.

Pili são estruturas proteicas, filamentosas, formadas por subunidades de pilina. Diversas funções podem estar associadas com diferentes tipos de pili, tais como sítios receptores de bacteriófagos, transferências de material genético durante a conjugação, formação de biofilme, agregação celular, além de aderência e invasão de células hospedeiras. Foi observado que em *Aeromonas* pili do tipo IV desempenham função na adesão e colonização de células epiteliais (BARNETT & KIROV, 1999).

Ao menos dois pili tipo IV diferentes foram identificados em *Aeromonas* associadas com gastroenterite, o *bundle-forming pili* (Bfp) e Tap (BARNETT et al., 1997; CHOPRA et al., 2000; RASMUSSEN-IVEY et al., 2016). O pili Bfp está associado com a adesão às células intestinais (BARNETT & KIROV, 1999; KIROV et al., 1999; TOMÁS, 2012; RASMUSSEN-IVEY et al., 2016) e também é denominado *mannose-sensitive hemagglutinin* (Msh) em decorrência da homologia da sequência com o pili hemaglutinina sensível à manose de *Vibrio cholerae* (HADI et al., 2012; TOMÁS, 2012).

Já o pilus tipo Tap, apresenta homologia como pilus tipo IV de *Pseudomonas* e espécies patogênicas de *Neisseria* (BARNETT & KIROV, 1999; TOMÁS, 2012; RASMUSSEN-IVEY et al., 2016) e não parece ter um papel significativo na colonização intestinal por *Aeromonas*, diferentemente do observado para Bfp cuja remoção resulta em redução de cerca de 80% na adesão às células intestinais (KIROV et al. 1999, 2000).

Além de pili, outras estruturas associadas com adesão de *Aeromonas* são os flagelos. Dois tipos de flagelos foram descritos em bactérias desse gênero, o flagelo polar e os flagelos laterais. O flagelo polar é constitutivo e responsável pela motilidade em meios líquidos (*swimming*). Já o flagelo lateral, está presente somente em algumas estirpes e é expresso em meios sólidos, sendo responsável pela motilidade em superfície (*swarming*) (KIROV et al., 2002; 2004). A motilidade decorrente da expressão dos flagelos polar ou lateral auxilia a bactéria nos processos de adesão, colonização e invasão das células hospedeiras e na formação de biofilme (GAVÍN et al., 2002; KIROV et al., 2002; 2004; GRIM et al., 2014).

O papel dos flagelos polar e lateral na adesão de *Aeromonas* em linhagens de células intestinais Henle 407 e Caco-2 foi demonstrada por KIROV e colaboradores (2004). Estirpes sem flagelo polar praticamente não aderiram às células, enquanto a ausência dos flagelos laterais resultou na redução de 60% na adesão em comparação com a bactéria selvagem, indicando que ambos os flagelos são adesinas de enterócitos (KIROV et al., 2004).

Outros fatores associados com virulência em *Aeromonas* incluem a produção de uma série de toxinas e enzimas extracelulares como proteases, lipases, amilases, nucleases, entre outras. Essas enzimas podem contribuir para a sobrevivência dessas bactérias em diferentes nichos.

As proteases têm um papel na patogenicidade, promovendo invasão pelo dano direto de tecidos do hospedeiro, uma vez que são capazes de degradar proteínas presentes no tecido conectivo incluindo fibrinogênio, elastina e colágeno, bem como na ativação proteolítica de toxinas (JANDA e ABBOTT, 2010; TOMÁS, 2012).

Diversas exotoxinas já foram descritas em *Aeromonas*. Entre elas, a toxina RtxA, cuja atividade é contato dependente, sendo capaz de romper a actina do citoesqueleto de células HeLa resultando em alteração do fenótipo celular provocando arredondamento e induzindo apoptose (SUAREZ et al., 2012).

Em *Aeromonas* existem dois tipos de enterotoxinas, as citotóxicas e citotônicas. A enterotoxina citotóxica Act, causa degeneração das criptas e vilosidades intestinais e geralmente é encontrada em estirpes isoladas de pacientes com diarreia, apresentando atividades hemolítica, citotóxica e enterotóxica. As toxinas citotônicas, como a termo-lábil Alt e a termo-estável Ast, não produzem lesão epitelial e seu mecanismo de ação está associado ao aumento dos níveis de adenosina monofostato

cíclico (cAMP) e prostaglandinas no intestino, causando secreção de fluidos no hospedeiro (ALBERT et al., 2000; SHA et al., 2002; RASMUSSEN-IVEY et al., 2016).

Além das enterotoxinas citotóxicas com atividade hemolítica, *Aeromonas* podem produzir hemolisinas não enterotóxicas que são classificadas em dois grupos, α -hemolisinas e β -hemolisinas. As α -hemolisinas apresentam efeito citotóxico reversível e lise incompleta dos eritrócitos. Já as β -hemolisinas levam à formação de poros culminando na lise e destruição dos eritrócitos (TOMÁS, 2012).

Existem ao menos quatro tipos de hemolisinas em *Aeromonas*, a aerolisina A (AerA), a hemolisina termo-lábil extracelular (AHH1), AhyA e Asa1. Aquela prevalente em *Aeromonas hydrophila* é AHH1, entretanto as estirpes mais virulentas possuem tanto a AHH1 como AerA (WANG et al., 2003; RASMUSSEN-IVEY et al., 2016).

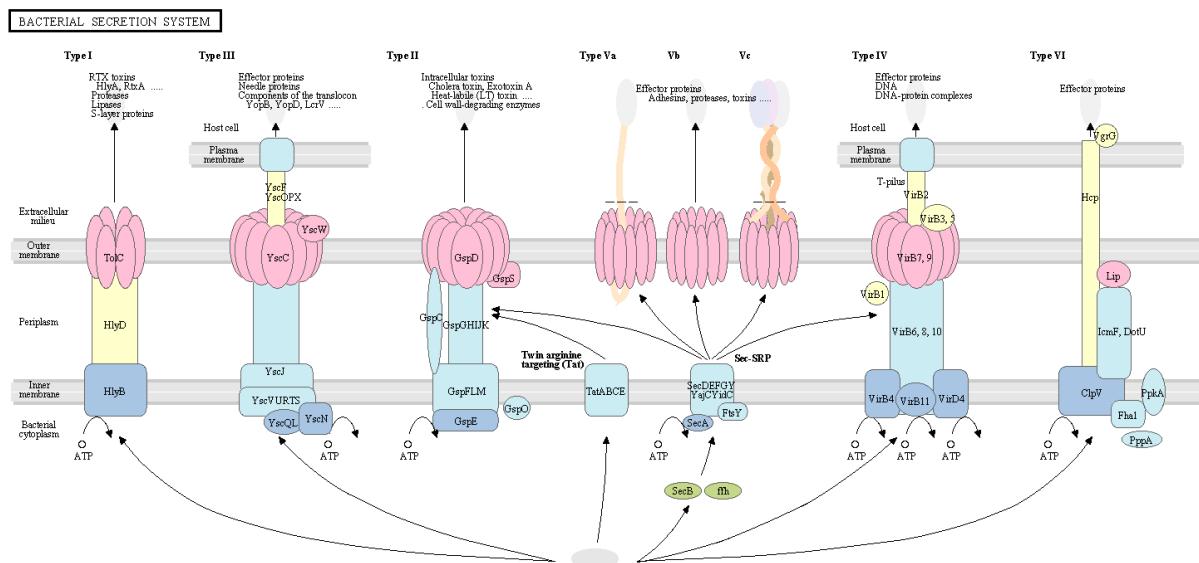
A secreção de proteínas é de fundamental importância no ciclo de vida dos micro-organismos, está envolvida na biogênese de organelas, aquisição de nutrientes, liberação de fatores de virulência e interação com outros organismos (TOMÁS, 2012; WANDERSMAN, 2013). É um processo complexo, principalmente para bactérias Gram-negativas, nas quais o processo de transporte ocorre a partir do seu local de síntese no citoplasma através do envelope celular (membrana interna, periplasma e membrana externa) (TOMÁS, 2012; WANDERSMAN, 2013).

Apesar do vasto número, diversidade e funções das proteínas que são secretadas, elas são translocadas utilizando um número limitado de mecanismos. Até o momento foram identificados seis sistemas de secreção em bactérias Gram-negativas, denominados sistemas de secreção tipo I – VI, mostrados na FIGURA 1.

Estes sistemas são estruturalmente distintos, variam em relação aos mecanismos de secreção dos respectivos substratos, que pode ocorrer em uma ou duas etapas, e ainda quanto ao destino do produto secretado – meio ambiente, ou injetado diretamente no citoplasma de células eucarióticas (TSENG, et al., 2009; HAYES et al., 2010).

Em *Aeromonas* três sistemas de secreção foram associados com virulência, os sistemas de secreção tipo II (ROSENZWEIG & CHOPRA, 2013), tipo III (VILCHES et al., 2004) e tipo VI (SUAREZ et al., 2010).

FIGURA 1 – SISTEMAS DE SECREÇÃO EM BACTÉRIAS



NOTA: Esquema da estrutura dos sistemas de secreção do tipo I a VI.

FONTE: Permissão autorizada pelo KEGG – Kyoto Encyclopedia of Genes and Genomes para Bacterial secretion system (map03070) (2017).

O sistema de secreção tipo II (T2SS) é utilizado por muitas bactérias Gram-negativas para exportar proteínas do citoplasma, já enoveladas, para a superfície celular ou para o meio externo (MICHEL & VOULHOUX, 2009; TAUSCHEK et al., 2002). Contém de 12 a 15 proteínas, as quais geram um poro que transpassa ambas as membranas, é amplamente conservado em bactérias Gram-negativas e tem sido estudado principalmente em patógenos de plantas, animais e humanos (MICHEL & VOULHOUX, 2009; RONDELET & CONDEMINÉ, 2013).

A função das proteínas secretadas pelo T2SS é extremamente diversa e incluem toxinas, fatores de virulência associados à superfície celular, citocromos e uma ampla gama de enzimas que hidrolizam macromoléculas como lipídeos, polissacarídeos. Estas proteínas podem estar envolvidas em vários processos biológicos como respiração, formação de biofilme, ligação a substrato e fixação bacteriana (RONDELET & CONDEMINÉ, 2013).

Em *Aeromonas* o T2SS está relacionado com a secreção extracelular de vários fatores de virulência como aerolisinas, amilases, DNAases e proteases (RASMUSSEN-IVEY et al., 2016). ROSENZWEIG e CHOPRA (2013) demonstraram que a enterotoxina Act é secretada por esse sistema.

O T3SS é um sistema estruturalmente complexo, formado por dois anéis que geram um canal contínuo através das membranas interna e externa da bactéria, incluindo a camada de peptidoglicano. Uma estrutura em agulha é associada com o anel presente na membrana externa e se projeta a partir da superfície bacteriana. As proteínas secretadas são denominadas “efetoras” e transportadas, através da estrutura, diretamente na célula hospedeira através de poros gerados pelas proteínas do T3SS. Muitas destas proteínas representam fatores de virulência que promovem alteração nas células hospedeiras (COBURN et al., 2007).

Esse sistema é formado por aproximadamente 20 proteínas e encontrado em várias bactérias gram-negativas, patogênicas e comensais, nas quais está associado a funções de adesão, colonização e invasão (SORG & CORNELIS, 2009).

Os efetores são na maioria das vezes distintos, com funções únicas adequadas à estratégia de virulência de um determinado patógeno. Entretanto, homólogos efetores também existem entre os diferentes tipos de bactérias que possuem esse sistema (COBURN et al., 2007).

Em *Aeromonas* esse sistema é encontrado principalmente em isolados clínicos (COBURN et al., 2007). Inicialmente descrito em *A. hydrophila*, o T3SS compreende 35 genes dispostos de maneira similar a encontrada em *Pseudomonas aeruginosa* (VILCHES et al., 2004). Mutações em genes que compõem o T3SS em *Aeromonas*, ocasionam redução da virulência em relação ao selvagem. Estudos realizados com estirpes desse gênero demonstram que este sistema é co-regulado pelo contato com a célula hospedeira (SHA et al., 2002).

O T6SS é codificado por um cluster de 13 genes conservados em diversos organismos (BOYER et al., 2009). É amplamente distribuído em bactérias Gram-negativas e sua função denota ser muito versátil, podendo ser utilizado para interagir com células eucarióticas ou bactérias competidoras (SHA et al., 2013).

Provavelmente desempenha um papel importante nas interações inter-bacterianas, afetando a estrutura das comunidades em vários nichos ambientais, na competição das bactérias patogênicas com rivais, facilitando, assim, a colonização de sítios de infecção nos quais a microbiota está presente. Além de poder atuar na interação direta de certos patógenos e simbioses com seus hospedeiros eucarióticos (CASCALES, 2008; SILVERMAN et al., 2012; DINIZ et al., 2015)

Em *Aeromonas* esse sistema foi inicialmente descrito em *A. hydrophila* e apresenta 18 genes conservados. A porção inicial do cluster é bastante variada com

a presença de genes acessórios (SUAREZ et al., 2008). Em algumas espécies o T6SS está relacionado com a formação de biofilme (SHA et al., 2013) e evasão do sistema imune (SUAREZ et al., 2010).

Como a nomenclatura dos genes do T6SS difere em diversos estudos adotamos a proposta por SHALOM e colaboradores (2007), como mostra a TABELA 1.

TABELA 1 – NOMENCLATURA DOS GENES DO T6SS

COG ID ^A	SHALOM ^B	RAST	<i>Aeromonas</i> ^C	<i>A. hydrophila</i> ATCC7966 ^D
COG3157	<i>hcp-2</i>	<i>hcp</i>	<i>hcp-2</i>	AHA_1826
COG3501	<i>vgrG-2</i>	<i>vgrG vgrG-3</i>	<i>vgrG-2</i>	AHA_1827
COG3516	<i>tssB</i>	<i>impB</i>	<i>impB</i>	AHA_1832
COG3517	<i>tssC</i>	<i>impC</i>	<i>impC</i>	AHA_1833
COG3518	<i>tssE</i>	<i>pvc109</i>	<i>pvc109</i>	AHA_1834
COG3519	<i>tssF</i>	<i>impG vasA</i>	<i>Vasa</i>	AHA_1835
COG3520	<i>tssG</i>	<i>impH vasB</i>	<i>vasB</i>	AHA_1836
COG3456	<i>fha</i>	<i>impI vasC</i>	<i>vasC</i>	AHA_1837
COG3521	<i>tssJ</i>		<i>vasD</i>	AHA_1838
COG3522	<i>tssK</i>	<i>impJ vasE</i>	<i>Vase</i>	AHA_1839
COG3455	<i>tssL</i>	<i>impK vasF</i> <i>opmA/motB</i>	<i>vasF</i>	AHA_1840
COG0542	<i>tssH</i>	<i>clpB</i>	<i>clpB</i>	AHA_1841
COG2204	<i>vasH</i>	sigma-54 dependent transcriptional regulator	<i>vasH</i>	AHA_1842
COG3515	<i>tssA(1)</i>		<i>impA</i>	AHA_1844
COG3523	<i>tssM</i>	<i>icmF</i>	<i>vasK</i>	AHA_1845
COG3515	<i>tssA(2)</i>	hypothetical protein		AHA_1846
COG4104	<i>PAAR</i>	Uropathogenic specific protein		AHA_1847
COG3501	<i>vgrG-3</i>	<i>vgrG vgrG-3</i>	<i>vgrG-3</i>	AHA_1848
COG3157	<i>hcp-1</i>	<i>hcp</i>	<i>hcp-1</i>	
COG3501	<i>vgrG-1</i>	<i>vgrG vgrG-3</i>	<i>vgrG-1</i>	

NOTA: A= o número do COG (BOYER et al. 2009). B= nomenclatura proposta por SHALOM et al. 2007. C= nomenclatura utilizada por Suarez et al. 2008. D= os respectivos loci em *A. hydrophila* subsp. *hydrophila* ATCC 7966 (SESHADRI et al. 2006, COULTHURST 2013).
 FONTE: a autora (2017).

3 OBJETIVOS

3.1 OBJETIVO GERAL

O objetivo geral deste trabalho foi sequenciar o genoma de *Aeromonas caviae* 8LM e realizar análises de genômica comparativa.

3.2 OBJETIVOS ESPECÍFICOS

I. Sequenciamento do genoma

- Sequenciar o genoma de *A. caviae* estirpe 8LM;
- Montar o genoma sequenciado;
- Anotar automaticamente os genes presentes no genoma de *A. caviae* estirpe 8LM;

II. Genômica comparativa

- Determinar a similaridade entre os genomas de *Aeromonas*;
- Comparar estruturalmente o genoma sequenciado com outras espécies de *Aeromonas*;
- Comparar o conteúdo gênico obtido com outras bactérias presentes no gênero;
- Estabelecer um pan-genoma utilizando as sequências genômicas de todas as espécies de *Aeromonas* depositadas no banco de dados do NCBI;
- Determinar o conjunto de genes representativos do core-genoma;
- Realizar análise filogenética das espécies de *Aeromonas*;
- Verificar a presença de genes associados com virulência em *Aeromonas*;

4 MATERIAL E MÉTODOS

4.1 CARACTERIZAÇÃO DA AMOSTRA

A bactéria *Aeromonas caviae* 8LM foi isolada a partir de cultura de fezes de uma criança de nove meses de idade, com diarreia e baixo peso. O exame parasitológico de fezes foi negativo, e a coprocultura negativa para *Salmonella*, *Shigella*, EPEC (*Escherichia coli* enteropatogênica) e EIEC (*Escherichia coli* enteroinvasiva).

A bactéria foi identificada por ASSIS (2014) através de testes bioquímicos segundo a metodologia proposta por ABBOTT e colaboradores (2003). Os resultados estão apresentados na TABELA 2. A estirpe *A. caviae* 8LM é mantida no Laboratório de Bacteriologia da Universidade Federal do Paraná em meio Cary–Blair a 4°C

TABELA 2 – TESTES BIOQUÍMICOS PARA IDENTIFICAÇÃO DE *A. caviae* 8LM

PROVA BIOQUÍMICA	RESULTADO	PROVA BIOQUÍMICA	RESULTADO
Agente vibriostático O129 (150µg)	Resistente	Indol	Positivo
Arabinose	Positivo	Inositol	Negativo
Arginina	Positivo	Lactose	Negativo
Cefalotina 30µg	Resistente	Lisina/ornitina	Negativo
Celobiose	Positivo	Manitol	Positivo
Citrato	Negativo	Manose	Negativo
Crescimento 0-3% NaCl (0-1-3) teste de concentração	Positivo	Motilidade	Positivo
Crescimento 6% NaCl	Negativo	Oxidase	Positivo
Esculina	Positivo	Rhamnose	Negativo
Fermentação da glucose	Positivo	Sacarose	Positivo
Gás glucose	Negativo	Sorbitol	Negativo
Hemólise em sangue de carneiro	Negativo	Ureia	Negativo
		VP	Negativo

FONTE: ASSIS et al., com modificações (2014).

4.2 CONDIÇÕES DE CULTIVO

Uma alíquota do estoque bacteriano mantido à 4°C, foi transferida para caldo BHI – *Brain Heart Infusion* – e incubada em estufa bacteriológica a 37°C por 24 horas. Após o crescimento as células foram utilizadas para a extração do DNA total da bactéria.

4.3 EXTRAÇÃO DO DNA

A extração do DNA de *A. caviae* 8LM foi realizada utilizando os kits “High Pure PCR template preparation kit®” – Roche e “MasterPure™ DNA Purification Kit for Blood Version II” – Epicentre Biotechnologies, ambos seguindo as recomendações do fabricante. A qualidade do DNA foi verificada por eletroforese em gel de agarose a 1%, e a quantificação foi realizada utilizando o NanoDrop (ThermoFisher Scientific).

4.4 SEQUENCIAMENTO

A metodologia de Sanger e suas modificações, baseadas no uso de terminadores dideoxynucleotídeostri-fosfatados acoplados a fluoróforos, foram amplamente utilizadas para o sequenciamento do DNA (ROGERS & VENTER, 2005).

Posteriormente, novas tecnologias de sequenciamento foram desenvolvidas e três delas foram utilizadas neste trabalho. O pirosequenciamento, tecnologia adotada na plataforma 454 (Roche), foi a primeira a ser comercializada (RONAGHI et al., 1998). É um método não-eletroforético baseado na emissão de bioluminescência gerada após a liberação de pirofosfato, quando da adição de desoxinucleotídeos à cadeia de DNA pela enzima DNA polimerase. Este é convertido em ATP, pela enzima ATP sulfúrilase, e utilizado pela enzima luciferase pela oxidação da luciferina produzindo um sinal de luz capturado pela câmera CCD (METZKER, 2010).

O sequenciamento na plataforma Illumina, amplamente utilizado, é baseado na síntese da molécula de DNA, por meio da terminação reversível cíclica. Este método utiliza a enzima DNA polimerase e nucleotídeos terminadores marcados com fluoróforos (CARVALHO & SILVA, 2010). Em um processo reversível e cíclico, ocorre a incorporação de nucleotídeos, sua clivagem e obtenção de imagens de fluorescência (METZKER, 2010).

A plataforma PacBio, usada mais recentemente, emprega uma metodologia de sequenciamento em tempo real (SCHADT et al., 2010). As sequências, produzidas por moléculas únicas de DNA polimerase, são obtidas através das imagens da incorporação contínua dos nucleotídeos marcados, que geram espectros de emissão de fluorescência distintos, durante a síntese de DNA (METZKER, 2010; RHOADS & AU, 2015).

4.4.1 Plataforma 454 GS Jr.

O DNA genômico de *A. caviae* 8LM foi extraído pelo “*High Pure PCR template preparation kit*®” (Roche) e o sequenciamento foi realizado pelo método de *whole genome shotgun* empregando a técnica de pirosequenciamento no sistema 454 GS Jr. Para isso, foram utilizados todos os kits e reagentes específicos para este tipo de sequenciamento e equipamento, seguindo todas as orientações preconizadas pelo fabricante.

4.4.2 Plataforma Illumina

O sequenciamento na plataforma Illumina MiSeq foi realizado no Departamento de Bioquímica da Universidade Federal do Paraná. O DNA foi extraído pelo kit “*High Pure PCR template preparation kit*®” (Roche) e para o sequenciamento foi utilizado o método de *pair-end* segundo as especificações descritas no protocolo do fabricante.

4.4.3 Plataforma PacBio

O DNA genômico da bactéria foi extraído pelo “*MasterPure™ DNA Purification Kit for Blood Version II*” (Epicentre Biotechnologies) e enviado para o sequenciamento na plataforma da PacBio. O preparo e envio do material genético foi realizado segundo as condições estabelecidas pela empresa.

O sequenciamento na plataforma PacBio foi realizado na Freie Universität Berlin e no Instituto Robert Koch, Berlim – Alemanha, como parte do projeto de doutorado sanduíche no período de setembro de 2015 a agosto de 2016.

4.5 MONTAGEM DOS CONTIGS

4.5.1 Plataforma 454 GS Junior

Os *reads* obtidos do sequenciamento na plataforma 454 GS Junior foram montados em contigs pelo *software* GS Data Analysis versão 2.7 empregando a ferramenta GS *de novo* Assembler.

4.5.2 Plataforma Illumina MiSeq

A montagem dos *reads* obtidos no sequenciamento da plataforma Illumina MiSeq em *contigs* foi realizada pelo *software* CLC Genomics Workbench versão 6.5.1.

4.5.3 Plataforma PacBio

No sequenciamento realizado na plataforma PacBio, os *reads* foram montados em *contigs* pelo *software* SMRT Analysis versão 2.3.0. Isto foi realizado durante o estágio de doutorado sanduíche na Alemanha.

4.6 ANOTAÇÃO DO GENOMA

4.6.1 Rapid Annotation using Subsystem Technology – RAST

A anotação do genoma de *A. caviae* 8LM foi realizada por meio do serviço *web* RAST (*Rapid Annotation using Subsystem Technology* – <http://rast.nmpdr.org>) usando os parâmetros padrão do programa, um serviço *online* automatizado para anotar genomas de Bacteria e Archaea.

O RAST identifica genes codificadores de proteínas, bem como de rRNA e tRNA, atribui funções para os genes, prevê quais subsistemas estão representados no genoma, utilizando esta informação para reconstruir uma rede metabólica e os resultados podem ser facilmente transferíveis para o usuário.

Além disso, o genoma anotado pode ser pesquisado em um ambiente que suporta a análise comparativa com os genomas anotados mantidos no ambiente SEED (AZIZ et al., 2008).

4.6.2 Ori-Finder

A origem de replicação do genoma de *A. caviae* 8LM foi inferida utilizando as configurações padrão do programa Ori-Finder (<http://tubic.tju.edu.cn/Ori-Finder/>), baseado em um método integrado que compreende a análise da assimetria da composição de base usando o método da curva Z, distribuição de *dnaA* boxes e a ocorrência de genes frequentemente próximos a oriC.

A curva Z é uma curva tridimensional que constitui uma representação da sequência de DNA, cujos componentes representam três distribuições independentes.

As componentes x_n , y_n e z_n exibem as distribuições de purina versus pirimidina (R vs. Y), amino versus ceto (M vs. K) e ligações fortes de hidrogênio versus ligações fracas de hidrogênio (S vs. W) ao longo da sequência de DNA, respectivamente.

Entre eles, os componentes x_n e y_n são denominados de curvas de disparidade RY e MK, respectivamente. As curvas de disparidade AT e GC são definidas por $\frac{(x_n+y_n)}{2}$ e $\frac{(x_n-y_n)}{2}$, que mostram o excesso de A sobre T e G sobre C ao longo da sequência de DNA, respectivamente.

As curvas de disparidade RY e MK, bem como as curvas de disparidade AT e GC, poderiam ser usadas para prever origens de replicação, uma vez que as curvas Z podem exibir as distribuições assimétricas de nucleotídeos em torno de oriC (GAO, 2015).

Os resultados foram exportados em um relatório HTML, que oferece visualizações em formatos gráficos e tabulares (GAO & ZHANG, 2008).

4.6.3 RNAmmer

As sequências de RNA ribossômico (rRNA) 5S, 16S e 23S foram preditas utilizando o programa RNAmmer 1.2 (<http://www.cbs.dtu.dk/services/RNAmmer/>), o qual utiliza a base de dados de RNA ribossomal europeia (WUYTS et al., 2004) e a de RNA ribossomal 5S (SZYMANSKI et al., 2002; LAGESEN, 2007).

As sequências de todos os rRNA foram alinhadas pelo Clustal Omega (SIEVERS et al., 2011; MCWILLIAN et al., 2013; LI et al., 2015) para verificação dos sítios variáveis.

4.7 ORDENAÇÃO DOS CONTIGS E FECHAMENTO DE GAPS

4.7.1 Ordenação dos contigs

Os contigs foram ordenados empregando a sequência de nucleotídeos da espécie referência *Aeromonas caviae* FDAARGOS 75. Os contigs foram utilizados para o alinhamento global de sequências contra a espécie referência usando o programa MUMmer versão 3.0 (*Maximal Unique Matches* – <http://mummer.sourceforge.net>).

Este programa é um sistema para o rápido alinhamento de genomas completos ou incompletos, baseado na estrutura de dados em *suffix tree*. É também um dos

sistemas mais rápidos e eficientes disponíveis para esta tarefa, permitindo sua utilização em sequências muito longas (KURTZ et al., 2004).

O sistema MUMmer é acondicionado como três módulos independentes: construção da *suffix tree* (árvore sufixo), classificação e extração do LIS (*longest increasing subsequence* – maior subsequência crescente) e na produção de alinhamentos Smith-Waterman para todas as regiões em MUM's (*maximal unique match* – correspondência única máxima) (DELCHER et al., 1999).

O novo sistema MUMmer, versão 3.0, inclui novos módulos gráficos para visualização das comparações de montagem, bem como a visualização de alinhamentos de espécies com parentesco mais distante (KURTZ et al., 2004).

Os contigs ordenados foram concatenados em uma única sequência utilizando o software TextWrangler (Bare Bones Software, Inc.©), unindo uns aos outros pela adição de um 'N', formando assim uma pseudomolécula.

4.7.2 Fechamento de GAPS

A pseudomolécula do genoma de *A. caviae* 8LM foi submetida à ferramenta FGAP (PIRO et al., 2014) utilizando uma biblioteca de contigs obtida das diversas montagens geradas pelo CLC Genomics Workbench versão 6.5.1 e pelo GS Data Analysis versão 2.7, além da sequência de nucleotídeos do genoma de referência *A. caviae* FDAARGOS 75.

4.7.3 Submissão das sequências

As sequências de nucleotídeos dos contigs foram depositadas no banco de dados do NCBI sob o número de acesso LAFH01.

4.8 GENÔMICA COMPARATIVA

Para a comparação genômica foram utilizados todos os genomas das espécies de *Aeromonas* depositadas no banco de dados no NCBI até o *release* 210 (outubro de 2015) (NCBI RESOURCE COORDINATORS, 2016), as estirpes podem ser observadas nas TABELAS A1 e A2 disponíveis no ANEXO A.

4.8.1 Similaridade entre genomas

4.8.1.1 Genome-to-Genome Distance Calculator – GGDC

A hibridização DNA-DNA (DDH) foi aplicada para obter uma estimativa da semelhança global entre dois genomas em estudo. Foi utilizado o GGDC (*Genome-to-Genome Distance Calculator* – <http://ggdc.dsmz.de/distcalc2.php>) versão 2.0.

O GGDC–DDH é um serviço *web* que calcula as distâncias intergenômicas e as converte em valores de similaridade análogas ao DDH para apoiar a decisão sobre o parentesco da nova estirpe com estirpes conhecidas (MEIER-KOLTHOFF, 2013; AUCH et al., 2010a, 2010b).

O GGDC–DDH compara o genoma em estudo com um genoma de referência e calcula uma distância intergenômica sob três fórmulas diferentes. A fórmula recomendada é a número 2, a qual soma todas as identidades encontradas em pares de segmento de alta pontuação (HSPs) e divide pelo comprimento total dessas HSPs. Essa fórmula também é independente do tamanho do genoma, podendo ser utilizada em genomas incompletos.

Estirpes que apresentaram porcentagens de similaridade acima de 70% demonstram que os genomas analisados pertencem à mesma espécie. O genoma de *A. caviae* 8LM foi comparado com todos os genomas de *Aeromonas* (ANEXO B).

4.8.1.2 Average Nucleotide Identity – ANI

Nessa metodologia o genoma em estudo é fragmentado *in silico* em segmentos de 1.020 nucleotídeos, os quais são alinhados utilizando o algoritmo do BLAST (*Basic Local Alignment Search Tool*) contra um genoma de referência a fim de calcular a identidade média dos nucleotídeos.

Valores superiores a 95% indicam que os genomas analisados pertencem à mesma espécie (GORIS et al., 2007).

O genoma de *A. caviae* 8LM foi analisado frente as outras espécies do gênero, disponíveis no ANEXO C, através do ANI Calculator (*Average Nucleotide Identity Calculator* – <http://enve-omics.ce.gatech.edu/ani/>) desenvolvido pelo laboratório Kostas (KONSTANTINIDIS & TIEDJE, 2005).

4.8.2 Análise estrutural

Para a comparação estrutural foram utilizadas as sequências de nucleotídeos em formato 'FASTA' do genoma de *A. caviae* 8LM e das *Aeromonas* com genoma completo foram submetidas ao *software* MUMmer versão 3.0, utilizando a ferramenta PROmer.

4.8.3 Análise do conteúdo gênico

4.8.3.1 BLAST Ring Image Generator – BRIG

A comparação do conteúdo gênico foi realizada através do programa BRIG (*BLAST Ring Image Generator*). Foram utilizadas nessa análise somente as estirpes de *Aeromonas* que apresentavam genoma completo.

As sequências de nucleotídeos em formato 'FASTA' do genoma sequenciado e das espécies de *Aeromonas* com genoma completo foram utilizadas para a produção de um *Blast Atlas*, por meio do algoritmo de alinhamento BLASTn (*Nucleotide Basic Local Alignment Search Tool*).

O BRIG é um aplicativo livre multi-plataforma que realiza comparações entre um grande número de genomas, por meio do algoritmo BLAST (ALIKHAN et al., 2011).

A imagem de saída mostra a semelhança entre uma sequência de referência central e outras sequências como um conjunto de anéis concêntricos, além de uma escala indicativa do percentual de identidade entre as sequências. É um programa de fácil acesso que não requer conhecimento computacional avançado (ALIKHAN et al., 2011).

4.8.3.2 Ilhas genômicas

A predição das ilhas genômicas em *A. caviae* 8LM foi realizada pelo *software* Gipsy versão 1.1.2, que considera os seguintes parâmetros na análise: (i) desvio da assinatura genômica, ou seja, diferenças no conteúdo de G+C e/ou no uso de codons; (ii) a presença de genes que codificam transposases, fatores de virulência, genes associados com metabolismo, resistência à antibióticos ou simbiose; (iii) presença de genes codificando RNA transportadores flanqueando a região; e (iv) ausência dos

genes em outros organismos do mesmo gênero ou espécies intimamente relacionadas (SOARES et al., 2016).

Nesta análise utilizou-se as sequências em formato *genbank*, foram aplicados os parâmetros padrão do programa e como referência a estirpe *A. caviae* FDAARGOS 75.

O *software* prevê quatro classes de ilhas genômicas classificadas em (i) ilhas de patogenicidade (PAIs), que carregam genes de virulência; (ii) ilhas metabólicas (MIs), que abrigam genes associados à biossíntese de metabólitos secundários; (iii) ilhas de resistência (RIs) que contêm genes codificadores de proteínas relacionadas com resistência à antibióticos; e (iv) ilhas simbióticas (ISs) que fornecem à bactéria um repertório genômico para sustentar uma relação simbiótica hospedeiro-bactéria (SOARES et al., 2016).

4.8.3.3 Agrupamento de genes

As análises de agrupamento gênico foram realizadas na Freie Universität Berlin e no Instituto Robert Koch, Berlim – Alemanha, como parte do projeto de doutorado sanduíche no período de setembro de 2015 a agosto de 2016.

Quando é realizada a comparação de um conjunto de genes de duas ou mais espécies pertencentes ao mesmo gênero, é possível determinar os genes comuns a todas as espécies, este conjunto de genes é denominado *core* genoma.

Os outros genes que são compartilhados com algumas espécies e aqueles que são específicos uma determinada espécie fazem parte do chamado genoma acessório. Já o pan-genoma é a soma de todos os genomas utilizados na análise comparativa.

Os genomas das 115 estirpes de *Aeromonas* depositadas no NCBI até o release 210, foram utilizados para a elaboração das análises de agrupamento de genes. Primeiro, foi realizado o agrupamento somente com os genomas completos de *Aeromonas* e posteriormente, uma análise com todos os genomas. Essas duas análises foram realizadas através do *software* Roary versão 3.6.1 (PAGE et al., 2015). Roary é um pipeline de alta velocidade que recebe montagens anotadas pelo Prokka (SEEMANN, 2014) e determina o pangenoma.

As sequências de nucleotídeos dos genomas foram anotadas pelo Prokka e posteriormente as proteínas preditas foram agrupadas pelo pipeline Roary de modo a investigar a variação genômica do conjunto de sequências.

O método baseia-se na filtragem e pré-agrupamento das proteínas mediante a comparação de todos contra todos usando o algoritmo BLASTp (*Protein Basic Local Alignment Search Tool*) e subsequente agrupamento com MCL (*Markov Cluster Algorithm*).

O critério utilizado para identificar o *core* genoma foi a presença dos genes em todos os isolados. Para definir o genoma acessório, os parâmetros utilizados foram: para o *soft* core genoma presença dos genes entre 95 a 99% dos isolados, para o *shell* genoma entre 15 a 95% dos isolados e para o *cloud* genoma em menos de 15% dos isolados (PAGE et al., 2015).

4.9 ANÁLISE FILOGENÉTICA

As análises filogenéticas foram efetuadas utilizando o software MEGA versão 7. É uma ferramenta integrada para a realização do alinhamento de sequências, inferir árvores filogenéticas, estimar tempos de divergência, extrair bases de dados online, estimar taxas de evolução molecular, inferir sequências ancestrais e testar hipóteses evolutivas (KUMAR et al., 2016).

Os alinhamentos múltiplos foram verificados no Gblocks Server (CASTRESANA, 2000; TALAVERA & CASTRESANA, 2007), o qual elimina posições mal alinhadas e regiões divergentes de um alinhamento de DNA ou proteína a fim de se tornar mais adequado para análise filogenética.

4.9.1 *Multilocus Sequence Analysis* – MLSA

Na análise de sequências multilocus foram utilizadas as sequências de aminoácidos em formato 'FASTA' de genes *housekeeping*, ou seja, mantenedores das funções celulares básicas, expressos em todos os organismos sob condições normais.

Os genes foram selecionados segundo estudos realizados com genes *housekeeping* em *Aeromonas* (SAAVEDRA et al., 2006; SESHADRI et al., 2006; MARTINEZ-MURCIA et al., 2011; BEAZ-HIDALGO et al., 2013; COLSTON et al., 2014; SANGLAS et al., 2017). A TABELA 3 apresenta os genes usados na análise.

TABELA 3 – GENES *HOUSEKEEPING* UTILIZADOS PARA A RECONSTRUÇÃO FILOGENÉTICA

SIGLA DO GENE	PROTEÍNA
<i>atpD</i>	Beta subunit of the membrane ATP synthase
<i>dnaG</i>	DNA primase
<i>dnaJ</i>	Molecular chaperone DnaJ
<i>dnaX</i>	DNA polymerase III subunits gamma and tau
<i>ftsZ</i>	Cell division protein
<i>glnA</i>	Glutamine synthetase
<i>gltA</i>	Citrate synthase
<i>groEL</i>	Heat-shock protein 60
<i>gyrB</i>	DNA gyrase beta subunit
<i>ileS</i>	Isoleucyl-tRNA synthetase
<i>infC</i>	Translation initiation factor IF-3
<i>mdh</i>	Malate dehydrogenase
<i>pyrG</i>	CTP synthase
<i>rplB</i>	Ribosomal protein L2
<i>rplC</i>	Ribosomal protein L3
<i>rplD</i>	Ribosomal protein L4
<i>rplE</i>	Ribosomal protein L5
<i>rplF</i>	Ribosomal protein L6
<i>rplM</i>	Ribosomal protein L13
<i>rplS</i>	Ribosomal protein L19
<i>rplT</i>	Ribosomal protein L20
<i>rpsE</i>	Ribosomal protein S5
<i>rpsI</i>	Ribosomal protein S9
<i>smpB</i>	SsrA-binding protein

FONTE: a autora (2017).

Inicialmente, foram obtidas as sequências de aminoácidos dos genes *housekeeping* no anotador automático RAST via BLASTp.

O alinhamento múltiplo das sequências foi realizado no *software* MEGA versão 7, empregando o algoritmo MUSCLE (EDGAR, 2004). Foram utilizadas as configurações padrões do programa para o alinhamento – score de -2.9 para abertura de gap, 0 para extensão de gap, 1.2 para multiplicador de hidrofobicidade e máximo de iterações 8.

Posteriormente, os alinhamentos foram submetidos ao Gblocks Server e as sequências resultantes concatenadas para construção da árvore filogenética.

A história evolutiva foi inferida usando o método de Máxima Verossimilhança baseado no modelo de matriz JTT (JONES et al., 1992). A árvore de consenso *bootstrap*, deduzida de 50 réplicas, foi utilizada para representar a história evolutiva dos taxa analisados (FELSENSTEIN, 1985).

Os ramos correspondentes às partições reproduzidos em menos de 50% das repetições de *bootstrap* foram recolhidos. A(s) árvore(s) inicial(is) para a pesquisa

heurística foram obtidas automaticamente aplicando os algoritmos *Neighbor-Joining* e BioNJ a uma matriz de distâncias *pairwise* estimadas usando um modelo JTT e selecionando a topologia com o valor elevado de verossimilhança logarítmica.

A análise envolveu 115 sequências de aminoácidos e todas as posições com gaps e dados em falta foram eliminadas. Análises evolutivas foram realizadas em MEGA 7 (KUMAR et al., 2016).

4.9.2 Reconstrução filogenética baseada no *core* genoma

A reconstrução filogenética baseada no *core* genoma foi realizada na Freie Universität Berlin e no Instituto Robert Koch, Berlim – Alemanha, como parte do estágio de doutorado sanduíche.

Os genes representativos do *core* genoma foram utilizados para a reconstrução filogenética.

A obtenção do *core* genoma, o alinhamento das sequências e a construção da árvore filogenética foram realizados utilizando o *software* Roary versão 3.6.1 (PAGE et al., 2015).

4.10 VIRULÊNCIA E RESISTÊNCIA

Como a anotação automática pode gerar erros foi necessária a verificação manual dos genes e sequências anotados para atestar a sua fidelidade.

Inicialmente, os genes de virulência relacionados aos sistemas de secreção, motilidade e adesão, resistência a antibióticos, *quorum-sensing*, toxina, aquisição de ferro e outros fatores foram verificados quanto a sua presença e ausência em *A. caviae* 8LM e foram obtidas as sequências de aminoácidos no anotador automático RAST (AZIZ et al., 2008).

Adicionalmente, foi realizada a anotação manual pelo alinhamento das sequências de aminoácidos por meio do algoritmo BLASTp, considerando como valores mínimos 95% de identidade e cobertura. Para isso, foram utilizadas como referências as espécies *Aeromonas hydrophila* 4AK4 (GAO et al., 2013), *Aeromonas hydrophila* AH-3 (CANALS et al., 2006), *Aeromonas hydrophila* SSU (SUAREZ et al., 2008), *Aeromonas hydrophila* subsp. *hydrophila* ATCC 7966 (SESHADRI et al., 2006) e *Aeromonas salmonicida* subsp. *salmonicida* A449 (REITH et al., 2008).

A presença do sistema de secreção tipo VI (T6SS) foi investigada quanto a sua presença em *A. caviae* 8LM. Para a comprovação e verificação das sequências, foi selecionada a estirpe *A. hydrophila* SSU como referência para a anotação manual, pois já existiam dados da literatura (SUAREZ et al., 2008, 2010a, 2010b; SHA et al., 2013; GRIM et al., 2014) mostrando experimentalmente a presença e funcionalidade do T6SS.

As sequências de proteínas do T6SS em *A. hydrophila* SSU está disponível no banco de dados do NCBI sob o número de acesso DQ667172. Na análise, seguimos a nomenclatura proposta por SHALOM e colaboradores (2007).

A verificação manual das sequências dos genes do T6SS foi realizada através da comparação frente a espécie referência utilizando o algoritmo de alinhamento local de sequências de proteínas BLASTp.

5 RESULTADOS E DISCUSSÃO

Este estudo relata a obtenção do genoma completo de *A. caviae* 8LM, sua comparação com genomas de outras estirpes de *Aeromonas*.

5.1 SEQUENCIAMENTO E MONTAGEM DOS CONTIGS

As leituras (*reads*) obtidas no sequenciamento do DNA realizado na plataforma 454 GS Jr. foram submetidas à ferramenta GS *de novo* assembler para a montagem dos contigs, enquanto que os *reads* obtidos na plataforma Illumina MiSeq foram analisados pelo *software* CLC Genomics Workbench versão 6.5.1.

No sequenciamento realizado na plataforma PacBio, os *reads* foram analisados com o *software* SMRT Analysis versão 2.3.0. Os resultados estão indicados na TABELA 4.

TABELA 4 – DADOS DO SEQUENCIAMENTO DE *A. caviae* 8LM

CARACTERÍSTICAS	454 GS Jr.	Illumina MiSeq	PacBio
Tamanho do genoma	4.374.623	4.468.794	4.560.001
Número de contigs	379	50	1
Reads	127.870	5.635.640	54.241
Reads mapeados	125.332	5.534.571	51.910
Reads não mapeados	2.538	101.069	2.331
Cobertura	10	385	113,07
GC%	61,7	61,8	61,7

FONTE: a autora (2017).

O sequenciamento pela plataforma PacBio resultou na obtenção de um único contig, no entanto, para fechar o genoma foi necessário associar os resultados das três tecnologias.

As montagens realizadas com as leituras obtidas com as plataformas de sequenciamento 454 GS Jr e Illumina MiSeq produziram, respectivamente, 379 e 50 contigs. Esses números são explicados pela presença de sequências de regiões repetidas em que os montadores não reconheceram como repetições e produziram sequências únicas, tais como os contigs formados apenas por uma cópia das sequências dos genes ribossomais.

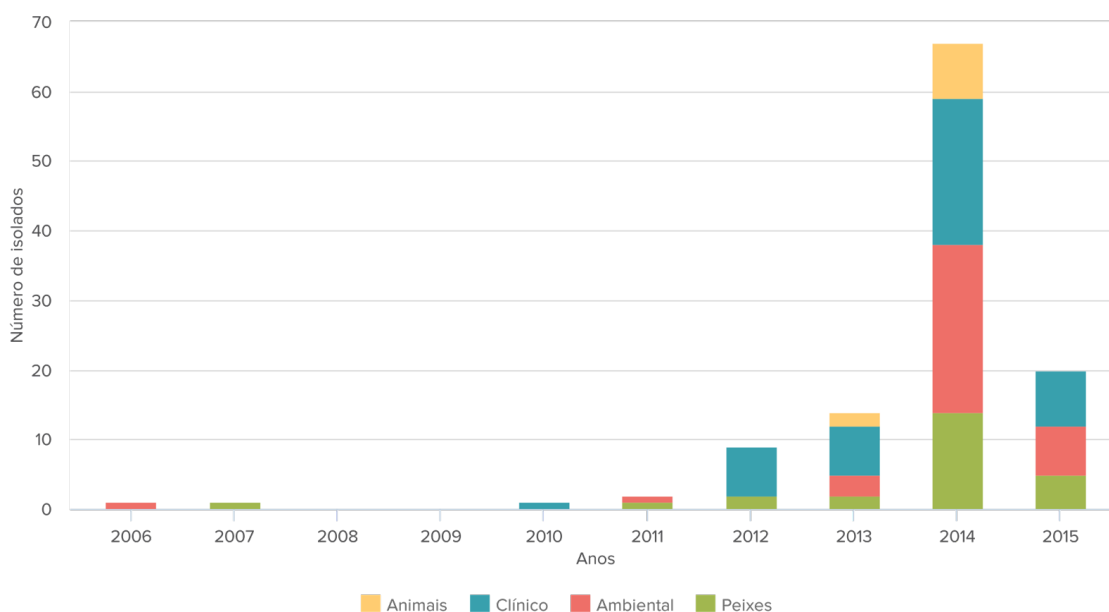
O sequenciamento realizado na plataforma PacBio resolveu os problemas de repetições, entretanto apresentava regiões de baixa qualidade e ausências de bases que puderam ser observadas na anotação através da detecção de *frameshifts*.

Ao mapear as leituras das duas primeiras corridas na montagem obtida dos dados sequenciados no PacBio, permitiu-se obter uma única sequência consenso que foi considerada a sequência completa do genoma de *A. caviae* 8LM.

5.1.1 Genomas de *Aeromonas*

A primeira estirpe de *Aeromonas* a ter seu genoma sequenciado foi *A. hydrophila* subsp. *hydrophila* ATCC 7966 em 2006 (SESHADRI et al., 2006), seguido pelo sequenciamento do genoma de *A. salmonicida* subsp. *salmonicida* A449 em 2008 (REITH et al., 2008).

Porém, a partir de 2012 houve um aumento considerável no número de genomas de *Aeromonas* sequenciados, culminando em 115 genomas até outubro de 2015, conforme release 210 do banco de dados do NCBI (NCBI RESOURCE COORDINATORS, 2016). O GRÁFICO 1 representa o número de genomas de *Aeromonas* sequenciados por ano e a fonte do isolado.

GRÁFICO 1 – NÚMERO DE GENOMAS DE *Aeromonas* DEPOSITADOS NO NCBI

NOTA: Quantidade de genomas de *Aeromonas* sequenciados por ano e fonte do isolado. No eixo x estão representados os anos, no eixo y, o número de isolados de *Aeromonas* disponíveis no banco de dados do NCBI até o release 210.

FONTE: a autora (2017).

Em relação à fonte do isolado bacteriano, no total, 44 são isolados clínicos, 36 ambientais, 25 de peixes e 10 de outros animais. A maioria dos genomas corresponde a isolados clínicos entre os quais predominam as espécies *A. hydrophila* (18), *A. caviae* (6) e *A. veronii* (5), refletindo sua importância como patógenos humanos, como citados em diversos estudos (JANDA & ABBOTT, 2010 PARKER & SHAW, 2011)

Em *Aeromonas* a espécie com o maior número de genomas sequenciados é *A. hydrophila*, representando 33% do total, e também com o maior número de genomas completos, como indicado na TABELA 5.

TABELA 5 – GENOMAS DE *Aeromonas* spp. SEQUENCIADOS

ESPÉCIES	NCBI	COMPLETOS	ESPÉCIES	NCBI	COMPLETOS
<i>Aeromonas allosaccharophila</i>	2	0	<i>Aeromonas lacus</i>	1	0
<i>Aeromonas aquatica</i>	1	0	<i>Aeromonas media</i>	4	1
<i>Aeromonas australiensis</i>	1	0	<i>Aeromonas molluscorum</i>	1	0
<i>Aeromonas bestiarum</i>	1	0	<i>Aeromonas piscicola</i>	1	0
<i>Aeromonas bivalvium</i>	1	0	<i>Aeromonas popofii</i>	1	0
<i>Aeromonas caviae</i>	11	2	<i>Aeromonas rivuli</i>	1	0
<i>Aeromonas dhakensis</i>	4	0	<i>Aeromonas salmonicida</i>	10	1
<i>Aeromonas diversa</i>	2	0	<i>Aeromonas sanarellii</i>	1	0
<i>Aeromonas encheleia</i>	1	0	<i>Aeromonas schubertii</i>	1	0
<i>Aeromonas enteropelogenes</i>	4	0	<i>Aeromonas simiae</i>	1	0
<i>Aeromonas eucrenophila</i>	1	0	<i>Aeromonas sobria</i>	1	0
<i>Aeromonas finlandensis</i>	1	0	<i>Aeromonas</i> sp.	6	0
<i>Aeromonas fluvialis</i>	1	0	<i>Aeromonas taiwanensis</i>	1	0
<i>Aeromonas hydrophila</i>	38	10	<i>Aeromonas tecta</i>	1	0
<i>Aeromonas jandaei</i>	3	0	<i>Aeromonas veronii</i>	12	1

FONTE: a autora (2017).

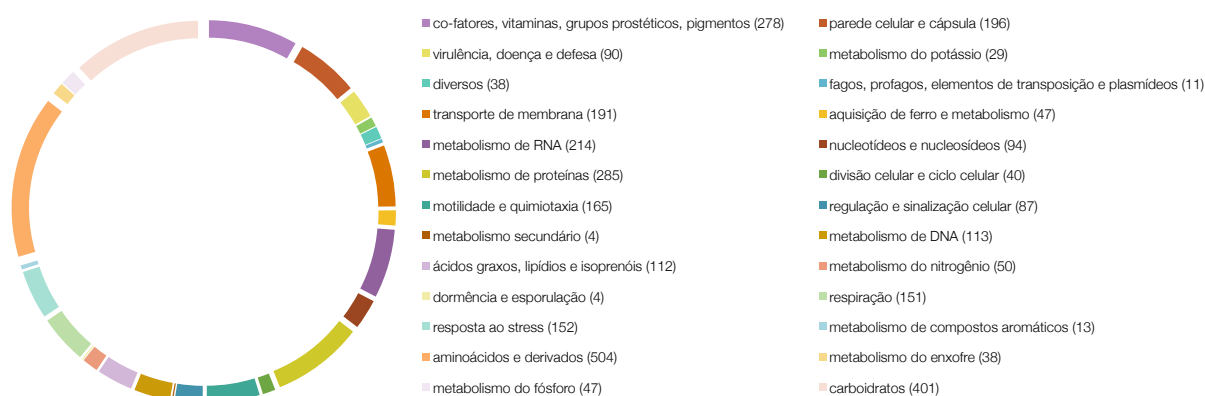
O genoma de *A. caviae* 8LM é o terceiro genoma completo da espécie, possui tamanho e conteúdo G+C (TABELA 4) semelhante a *A. caviae* FDAARGOS 72 e *A. caviae* FDAARGOS 75, dados disponíveis na TABELA A1 do ANEXO A.

5.2 ANOTAÇÃO DO GENOMA

5.2.1 Rapid Annotation using Subsystem Technology – RAST

O genoma de *A. caviae* 8LM foi anotado automaticamente pelo *web-service* RAST. O conteúdo de G+C do genoma é 61,7%, foram preditos 4.256 genes, 4.102 CDS (regiões codificadoras de proteína) e 154 RNA estruturais (RNA ribossômico e transportador).

Foram anotados 3.350 genes categorizados em diversos subsistemas, como indicado no GRÁFICO 2.

GRÁFICO 2 – ANOTAÇÃO DO GENOMA DE *A. caviae* 8LM PELO RAST

FONTE: a autora (2017).

Os genes responsáveis pelo metabolismo de carboidratos, aminoácidos e proteínas relacionados com os processos de obtenção de energia e manutenção celular, correspondem a mais de um quarto do conteúdo gênico.

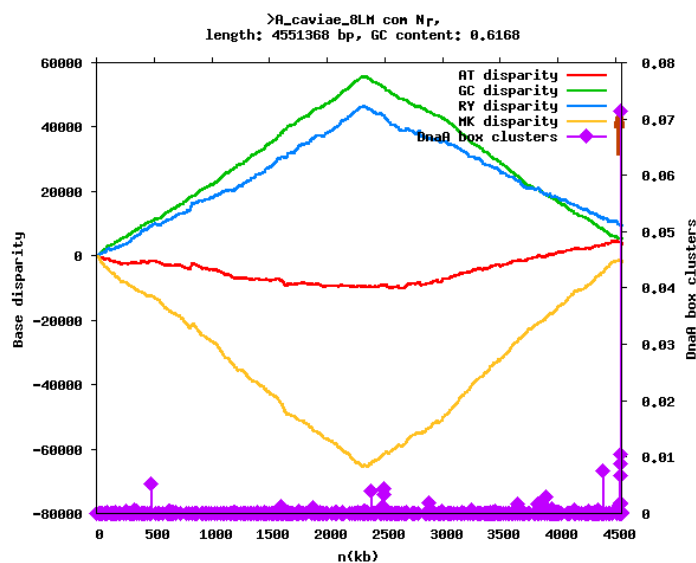
Entre os diversos genes anotados (GRÁFICO 2) destaca-se a presença daqueles relacionados com motilidade e quimiotaxia, resposta ao stress, virulência e resistência à antibióticos, bem como, genes pertencentes aos sistemas de secreção, alguns dos quais já foram associados com a capacidade do micro-organismo e causar doença.

5.2.2 Ori-Finder

A origem de duplicação (OriC) é uma região do cromossomo que contem de 100 a 1000 pb, na qual é iniciada o processo de duplicação do DNA. Apresenta um conteúdo de AT elevado em relação ao genoma (GAO & ZHANG, 2008; SERNOVA & GELFAND, 2008), além de possuir sítios de ligação para a proteína DnaA (proteína iniciadora da replicação) denominados *dnaA boxes* (GAO & ZHANG, 2008).

A OriC pode ser inferida *in silico* pela identificação dos múltiplos sítios de ligação de DnaA e pela diferença entre o conteúdo GC% (REITH et al., 2008). O software Ori-Finder foi utilizado para inferir a origem de duplicação do genoma de *A. caviae* 8LM.

Os resultados indicados no GRÁFICO 3, sugerem que a origem da duplicação está localizada entre os nucleotídeos 4535698 e 4536149, contendo 452 pb com conteúdo de AT de 54,42%.

GRÁFICO 3 – CURVAS Z PARA O GENOMA DE *A. caviae* 8LM

NOTA: Gráfico de curvas Z, distribuição das *dnaA boxes*, localização de genes indicadores putativos e predita região *oriC*. Os picos roxos com os diamantes indicam os conjuntos de *dnaA boxes*. A seta em vermelho representa a origem de replicação.

FONTE: a autora (2017).

Foram encontradas três possíveis *dnaA box* na origem de duplicação de *A. caviae* 8LM, utilizando como referência *Escherichia coli* cuja sequência é ttatccaca (GAO & ZHANG, 2008). As sequências estão representadas na FIGURA 2 e destacadas em maiúsculo e negrito.

FIGURA 2 – SEQUENCIA DE *dnaA box* EM *A. caviae* 8LM

```
acgggctgcccgtggttgctgcttcggttcctgcatgaccgccaccgacgacgatcacatcaa
attgttcatgggtattgcattggggatcaccttgatctgcacgcctgaagatctttgttgccc
tggctgaagggcggatcattctatctgtttagcatccggaggggaacgtctacttcaggat
cagtatggatcagggatcggttaaatagatcttaaggatctttagatcgatcttttattgga
tcttttattaggatccgcgacttcTGTTGATAAggcatttttgatcataaagatcattcggg
taaggatcgatcgctgctTGTGGAAAAgcttggatccggagggggcgaacgggggataacatg
ggtttTTATACACAgatacatggtcggcaagatctaatacatgaataactatagggtttgca
gagcttggtagcgaagTT
```

NOTA: Sequência de nucleotídeos da origem de duplicação do cromossomo de *A. caviae* 8LM, em negrito as possíveis *dna boxes*.

FONTE: a autora (2017).

Dentre os genomas completos sequenciados e publicados apenas o genoma de *A. salmonicida* subsp. *salmonicida* A449 (REITH et al., 2008) determina a

localização da origem de replicação. Portanto, os resultados obtidos para *A. caviae* 8LM foram comparados com esta estirpe.

Foi constatado que a suposta origem de duplicação em *A. caviae* 8LM é 102 pb maior que a de *A. salmonicida* subsp. *salmonicida* A449 e a região de término de replicação em ambas as estirpes se localiza em torno de 2,2 Mb.

5.2.3 RNAmmer

Os RNAs ribossômicos 5S, 16S e 23S foram preditos no RNAmmer 1.2 Server, e os resultados obtidos estão indicados na TABELA 6.

TABELA 6 – OPERONS DE RNA RIBOSSOMAL EM *A. caviae* 8LM

OPERON	rRNA	INÍCIO (pb)	FINAL (pb)	+/-
1	16S	74600	76132	+
	23S	76661	79547	+
	5S	79673	79786	+
2	16S	298597	300129	+
	23S	300608	303494	+
	5S	303576	303689	+
3	16S	379499	381031	+
	23S	381630	384516	+
	5S	384642	384755	+
4	16S	779719	781251	+
	23S	781782	784668	+
	5S	784779	784892	+
5	5S	3668605	3668718	-
	23S	3668800	3671686	-
	16S	3672165	3673697	-
6	5S	3827115	3827228	-
	23S	3827352	3830238	-
	16S	3830838	3832370	-
7	5S	4253279	4253392	-
	23S	4253503	4256389	-
	16S	4256920	4258452	-
8	5S	4379118	4379231	-
	5S	4379379	4379492	-
	23S	4379616	4382502	-
	16S	4382981	4384513	-
9	5S	4426457	4426570	-
	23S	4426696	4429582	-
	16S	4430113	4431645	-
10	5S	4504042	4504155	-
	23S	4504281	4507167	-
	16S	4507758	4509290	-

NOTA: Destacado em negrito a cópia adicional de rRNA 5S. operon= número do operon; rRNA= tipo de subunidade; início= posição de início; final= posição de término; +/- = sentido da sequência.
 FONTE: a autora (2017).

O genoma de *A. caviae* 8LM apresenta dez operons de RNA ribossômico e os genes rRNA 5S, 16S e 23S contêm respectivamente 114 pb, 1.533 pb e 2.887 pb. Foi observado que no operon 8 há uma cópia adicional do gene rRNA 5S.

A análise das sequências das estirpes de *Aeromonas* com genoma completo indicou que a quantidade de rRNAs variou de 29 a 33, distribuídos entre 9 e 10 operons, dados disponíveis na TABELA A1 do ANEXO A.

A maioria das estirpes de *Aeromonas* possuem 10 operons de rRNA, entretanto, *A. hydrophila* 4AK4 (GAO et al., 2013), *A. media* WS (CHAI et al., 2012) e *A. salmonicida* subsp. *salmonicida* A449 (REITH et al., 2008) apresentam 9 operons. Além disso, todas as estirpes, com genoma completo, apresentam uma cópia adicional do rRNA 5S (SESHADRI et al., 2006; LI et al., 2011; PRIDGEON et al., 2014; PANG et al., 2015; TEKEDAR et al., 2015, LIM et al., 2016). Por outro lado, a estirpe *A. hydrophila* AH10 possui uma cópia extra dos rRNA 16S e 23S. Já as estirpes *A. hydrophila* 4AK4 (GAO et al., 2013) e *A. media* WS (CHAI et al., 2012) apresentam, respectivamente, 1 e 2 cópias adicionais de rRNA 23S adicionais, porém esta última apresenta duas cópias extras, dados obtidos no RNAmmer 1.2 Server.

Para verificar a conservação dos rRNA 5S, 16S e 23S em *A. caviae* 8LM, as sequências de nucleotídeos foram alinhadas pelo Clustal Omega e encontram-se disponíveis no ANEXO B.

As sequências de rRNA 5S são idênticas, exceto a cópia adicional encontrada no operon 8. Esta apresenta o mesmo número de nucleotídeos que as demais sequências, porém possui substituições de nucleotídeos resultando em 96% de identidade.

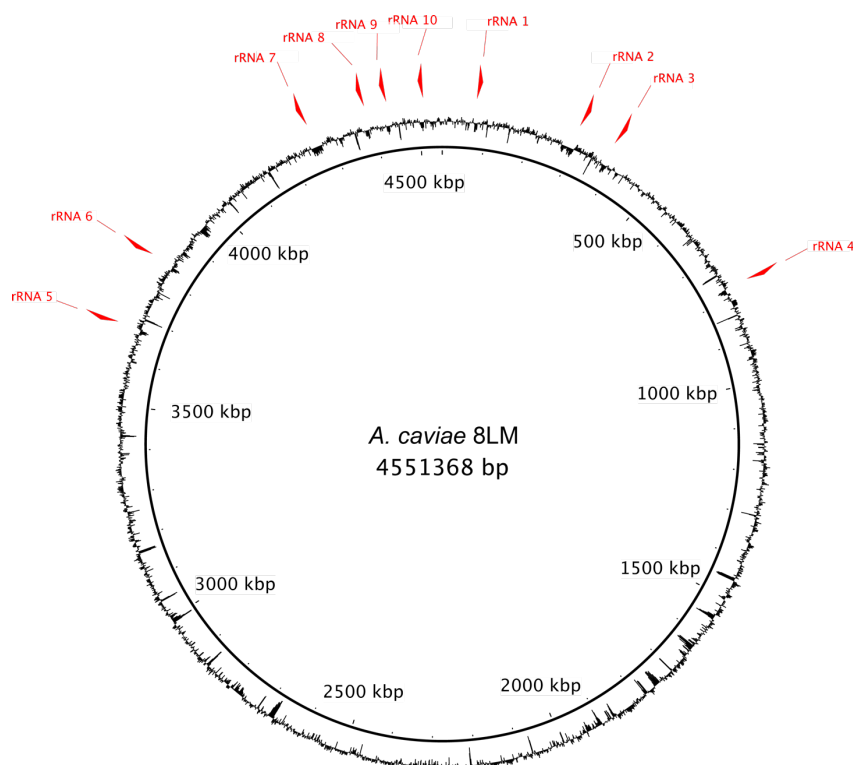
Das dez cópias de rRNA 23S encontradas no genoma de *A. caviae* 8LM, seis apresentam substituições de nucleotídeo, como mostrado no ANEXO B.

Em relação ao rRNA 16S, nenhuma das dez sequências é idêntica (ANEXO B), confirmando a heterogeneidade genética de rRNA 16S de *Aeromonas*, como relatado em outros estudos (MORANDI et al., 2005; ALPERI et al., 2008).

Conseqüentemente, a utilização da sequência de rRNA 16S como ferramenta para a identificação das bactérias do gênero *Aeromonas* não é apropriada, uma vez que dentro do próprio genoma bacteriano há heterogeneidade intragenômica, cuja frequência de polimorfismos é semelhante a empregada para distinção entre espécies (MORANDI et al., 2005; ALPERI et al., 2008; COLSTON et al., 2014).

Na FIGURA 3 está representada a distribuição dos operons de rRNA no genoma de *A. caviae* 8LM. Observa-se que eles estão localizados cerca de 800kbp em ambos os sentidos da origem de duplicação do cromossomo, a qual situa-se entre os operons 10 e 1.

FIGURA 3 – DISTRIBUIÇÃO DOS OPERONS DE rRNA EM *A. caviae* 8LM



NOTA: O círculo interno representa o genoma de *A. caviae* 8LM e o externo o conteúdo de GC%. As setas em vermelho indicam as posições e o sentido dos operons de rRNA.

FONTE: a autora (2017).

Dentre as estirpes de *Aeromonas* com genoma publicado apenas em *A. salmonicida* subsp. *salmonicida* A449 (REITH et al., 2008) é mostrada a distribuição desses operons, os quais estão dispostos ao redor da origem de replicação, resultado similar ao observado em *A. caviae* 8LM.

A presença de um elevado número de operons de rRNA no genoma de *Aeromonas* pode estar associado com a habilidade da bactéria em responder rapidamente às mudanças nas condições ambientais (SESHADRI et al., 2006).

5.3 GENÔMICA COMPARATIVA

5.3.1 Similaridade entre genomas

Para as análises realizadas a seguir foram utilizados os genomas completos e em formato *draft* das estirpes de *Aeromonas* disponíveis no banco de dados do NCBI até o release 210 (NCBI RESOURCE COORDINATORS, 2016).

A utilização de ferramentas de bioinformática como o GGDC–DDH e ANI são úteis para definir a taxonomia de procariotos (GORIS et al., 2007; MEYER-KOLTHOFF et al., 2013)

5.3.1.1 Genome-to-genome Distance Calculator – GGDC

O GGDC-DDH é baseado na hibridização DNA-DNA (DDH) *in silico*. O GGDC 2.0 (<http://ggdc.dsmz.de>) é uma versão atualizada e aprimorada com melhores modelos de previsão por meio do DDH, além de apresentar funcionalidades adicionais, como a estimativa de um intervalo de confiança.

Os resultados, mostrados de forma resumida na TABELA 7, indicam as estirpes de *Aeromonas* que apresentam os maiores valores de GGDC–DDH quando comparadas com *A. caviae* 8LM. No ANEXO C estão disponíveis os dados de todas as estirpes analisadas.

Pelos parâmetros estipulados (MEIER-KOLTHOFF et al., 2013) estirpes que pertencem à mesma espécie apresentam porcentagem de GGDC–DDH superior à 70%.

Observa-se na TABELA 7 que os valores de GGDC–DDH superiores a 70% foram encontrados para todas as estirpes desde *A. caviae* 429865 à *Aeromonas* sp. ZOR0002, com valores variando de 83,80% a 80,50%, respectivamente. Na sequência, os valores de GGDC–DDH caem drasticamente, a partir de *A. sanarellii* LMG 24682 com 40,10%, indicando baixa probabilidade de pertencer a mesma espécie que a estirpe 8LM.

TABELA 7 – MAIORES VALORES DE GGDC–DDH ENTRE *Aeromonas*

ORGANISMO	ESTIRPE	DDH	DIST.	DDH>70%
<i>Aeromonas caviae</i>	429865	83,80	0,02	93,11
<i>Aeromonas caviae</i>	A23	83,50	0,02	92,91
<i>Aeromonas caviae</i>	Ae398	82,60	0,02	92,41
<i>Aeromonas</i> sp.	HZM	82,40	0,02	92,28
<i>Aeromonas hydrophila</i>	BWH65	82,10	0,02	92,08
<i>Aeromonas caviae</i>	CECT 4221	82,00	0,02	92,02
<i>Aeromonas caviae</i>	CECT 838	81,80	0,02	91,88
<i>Aeromonas caviae</i>	FDAARGOS_75	81,80	0,02	91,89
<i>Aeromonas caviae</i>	FDAARGOS_76	81,70	0,02	91,80
<i>Aeromonas caviae</i>	L12	81,70	0,02	91,81
<i>Aeromonas caviae</i>	FDAARGOS_72	81,50	0,02	91,65
<i>Aeromonas caviae</i>	YL12	81,20	0,02	91,48
<i>Aeromonas</i> sp.	ZOR0002	80,50	0,02	90,96
<i>Aeromonas sanarellii</i>	LMG 24682	40,10	0,10	2,80
<i>Aeromonas taiwanensis</i>	LMG 24683	37,90	0,11	1,57
<i>Aeromonas media</i>	WS	35,00	0,12	0,67
<i>Aeromonas hydrophila</i>	4AK4	34,50	0,12	0,56
<i>Aeromonas media</i>	CECT 4232	34,80	0,12	0,62
<i>Aeromonas media</i>	ARB13	34,60	0,12	0,57
<i>Aeromonas media</i>	ARB20	34,60	0,12	0,57

NOTA: Maiores valores obtidos em relação à *A. caviae* 8LM. DDH= valor estimado. Dist.= distância evolutiva. DDH > 70%= Probabilidade de pertencem a mesma espécie.

FONTE: a autora (2017).

Entre *Aeromonas* com valores de GGDC–DDH superiores a 80%, 10 são *A. caviae*, o que confirma que a estirpe 8LM pertence a esta espécie.

No entanto, as bactérias identificadas apenas ao nível de gênero, *Aeromonas* sp. HZM e *Aeromonas* sp. ZOR0002, também apresentaram valores de GGDC–DDH superiores a 80%. Esses resultados indicam que ambas pertencem a espécie *A. caviae*, uma vez que apresentam probabilidade maior que 90% de pertencerem à mesma espécie que a estirpe em estudo.

Por outro lado, entre as estirpes com valores mais elevados de GGDC–DDH encontra-se *A. hydrophila* BWH65. Os resultados obtidos (TABELA 7) sugerem uma divergência na classificação dessa estirpe, uma vez que a probabilidade dessa bactéria pertencer a *A. caviae* é de 92,08%. Esses dados também mostram a dificuldade na identificação de *Aeromonas* ao nível de espécie por meio de provas bioquímicas.

Desta maneira, torna-se evidente que as análises genômicas de *Aeromonas* contribuem para a identificação taxonômica desse grupo.

5.3.1.2 Average Nucleotide Identity – ANI

A calculadora ANI, desenvolvida pelo laboratório Kostas (<http://enve-omics.ce.gatech.edu/ani/>), estima a identidade média de nucleotídeos entre dois genomas e valores superiores a 95% (GORIS et al., 2007) são critérios para considerar duas estirpes como membros de uma mesma espécie.

Os resultados mostrados na TABELA 8 mostram as estirpes de *Aeromonas* que apresentaram as maiores porcentagens de identidade quando comparadas com *A. caviae* 8LM. No ANEXO D estão disponíveis os dados de todas as estirpes investigadas.

TABELA 8 – MAIORES VALORES DE ANI ENTRE *Aeromonas*

ORGANISMO	ESTIRPE	ANI (%)
<i>Aeromonas caviae</i>	YL12	98,84
<i>Aeromonas caviae</i>	429865	98,12
<i>Aeromonas caviae</i>	A23	98,11
<i>Aeromonas</i> sp.	HZM	98,01
<i>Aeromonas caviae</i>	FDAARGOS_76	98,00
<i>Aeromonas caviae</i>	Ae398	97,99
<i>Aeromonas caviae</i>	CECT 4221	97,93
<i>Aeromonas hydrophila</i>	BWH65	97,90
<i>Aeromonas caviae</i>	L12	97,89
<i>Aeromonas caviae</i>	CECT 838	97,88
<i>Aeromonas</i> sp.	ZOR0002	97,87
<i>Aeromonas caviae</i>	FDAARGOS_75	97,85
<i>Aeromonas caviae</i>	FDAARGOS_72	97,83
<i>Aeromonas sanarellii</i>	LMG 24682	89,94
<i>Aeromonas taiwanensis</i>	LMG 24683	89,15
<i>Aeromonas media</i>	WS	88,07
<i>Aeromonas media</i>	CECT 4232	88,05
<i>Aeromonas media</i>	ARB20	87,85
<i>Aeromonas media</i>	ARB13	87,85
<i>Aeromonas hydrophila</i>	4AK4	87,74

NOTA: Maiores valores obtidos em relação à *A. caviae* 8LM. ANI= Average Nucleotide Identity.
 FONTE: a autora (2017).

Nessa análise, a porcentagem de identidade entre a estirpe 8LM e *A. caviae* é superior a 97%. Estes dados confirmam os dados do GGDC–DDH e a identificação da estirpe 8LM como *A. caviae*. Pelo ANI o genoma mais similar ao de *A. caviae* 8LM foi *A. caviae* YL12 com 98,84% de identidade, enquanto que na análise de GGDC–DDH foi *A. caviae* 429865, o que provavelmente se deve a diferenças na metodologia

de análise de cada software, porém os resultados em ambas as análises indicam que a estirpe *A. caviae* 8LM pertence à espécie *A. caviae*.

Observa-se também na TABELA 8, assim como na análise de GGDC–DDH, que os valores de ANI para *Aeromonas* sp. HZM, *Aeromonas* sp. ZOR0002 e *A. hydrophila* BWH65, sugerem que as estirpes pertencem a espécie *A. caviae*. Esses resultados confirmam o estudo de BEAZ-HIDALGO e colaboradores (2015), que realizaram análise de ANI para a classificação taxonômica de *Aeromonas* e também concluem que a estirpe *Aeromonas* sp. HZM pertence a espécie *A. caviae*.

Para as estirpes de *Aeromonas* pertencentes outras espécies (TABELA 8), o percentual de ANI foi inferior a 90%.

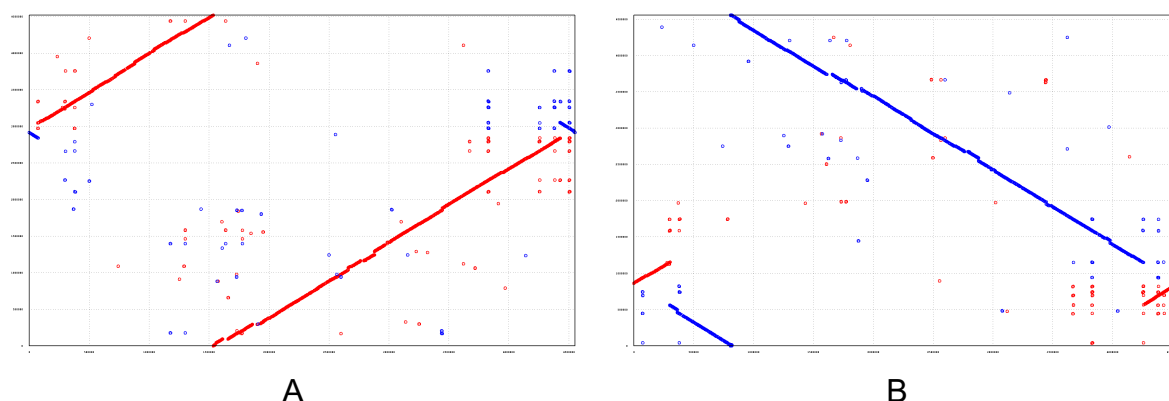
5.3.2 Análise estrutural

As análises realizadas com software MUMmer são referentes aos dados obtidos na montagem do genoma de *Aeromonas caviae* 8LM, a qual foi contrastada com as estirpes do gênero com genoma completo, respectivamente, *A. caviae* FDAARGOS 72, *A. caviae* FDAARGOS 75, *A. hydrophila* 4AK4, *A. hydrophila* AH10, *A. hydrophila* AL06-06, *A. hydrophila* AL09-71, *A. hydrophila* J-1, *A. hydrophila* ML09-119, *A. hydrophila* NJ-35, *A. hydrophila* pc104A, *A. hydrophila* YL17, *A. hydrophila* subsp. *hydrophila* ATCC7966, *A. media* WS, *A. salmonicida* subsp. *salmonicida* A449 e *A. veronii* B565.

Nos gráficos de *dotplot* gerados pelo software MUMmer, cada similaridade é representada por um ponto, e quando estes estão muito próximos acabam formando retas, ou seja, quanto mais pontos alinhados formando uma reta maior o grau de conservação das sequencias.

Os pontos representados em vermelho fora da “reta” indicam que a sequencia analisada é similar ao genoma de referência (eixo y), mas encontram-se em locais diferentes. Já os pontos em azul mostram que a sequencia é similar, mas encontra-se no sentido reverso à sequencia do genoma de referência.

Os gráficos *dotplot* a seguir mostram os resultados das comparações entre o genoma de *A. caviae* 8LM (eixo x) e os das estirpes de *Aeromonas* com genomas completos (eixo y).

GRÁFICO 4 – DOTPLOT ENTRE *A. caviae* E *A. caviae* 8LM

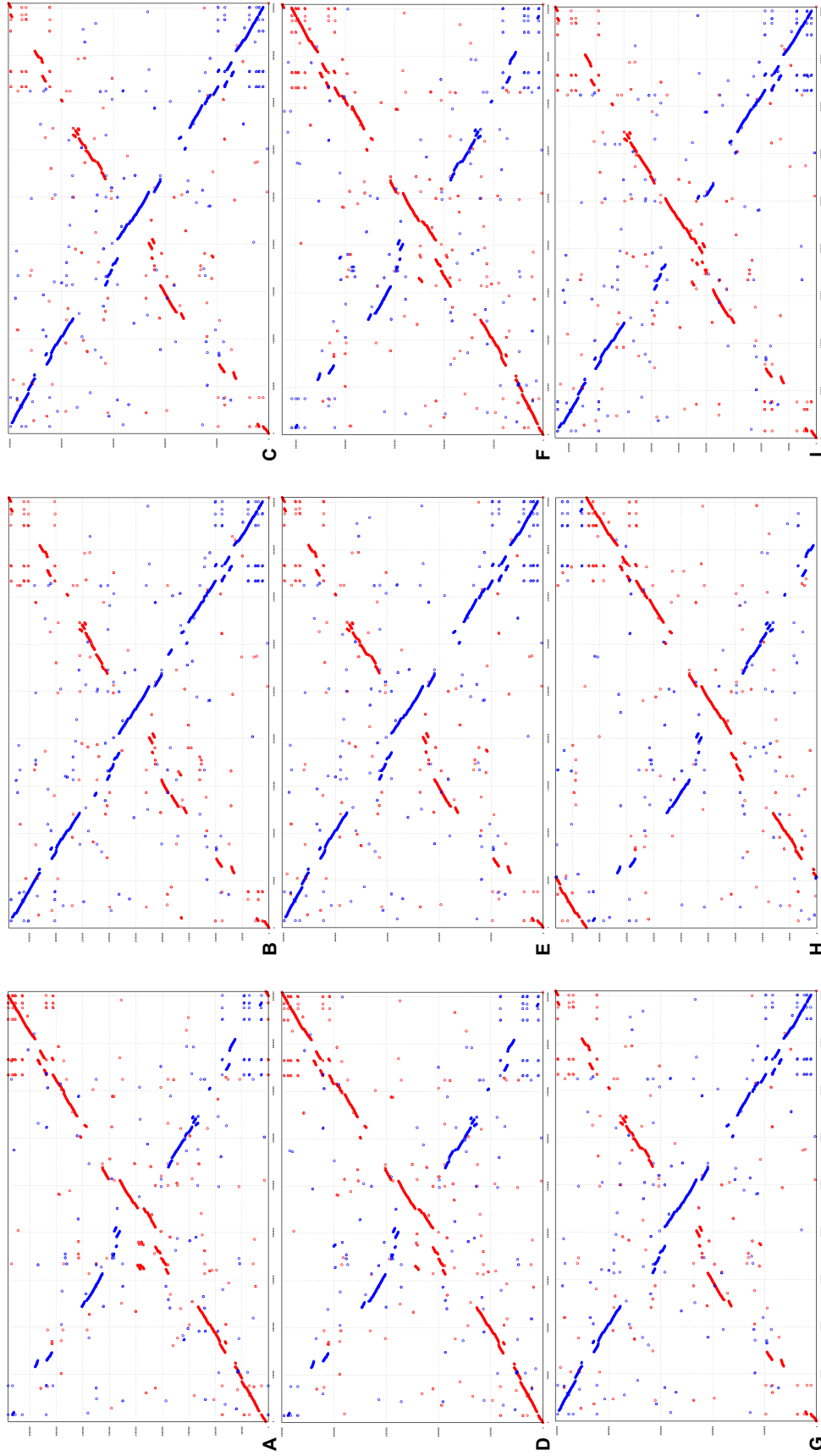
NOTA: A= *Aeromonas caviae* FDAARGOS 72. B= *Aeromonas caviae* FDAARGOS 75.
 FONTE: a autora (2017).

A estirpe *A. caviae* FDAARGOS 72 foi a que apresentou maior similaridade com o genoma de *A. caviae* 8LM (GRÁFICO 4A). Apesar do elevado grau de conservação, verifica-se que há regiões com sentido complementar reverso nas porções inicial e final dos genomas, bem como sequências que não se encontram na mesma posição no genoma de referência.

Também foi observada similaridade com o genoma da estirpe *A. caviae* FDAARGOS 75 (GRÁFICO 4B), no entanto, este encontra-se em ordem inversa.

O GRÁFICO 5 reúne os *dotplots* das estirpes de *A. hydrophila*. Nota-se que os genomas apresentam um grande número de sequências com sentido complementar reverso, em azul. Os gráficos reproduzem um “X”, portanto menos semelhantes são os genomas e menor o grau de conservação das sequências em relação à *A. caviae* 8LM.

GRÁFICO 5 – DOTPLOT ENTRE *A. hydrophila* E *A. caviae* 8LM

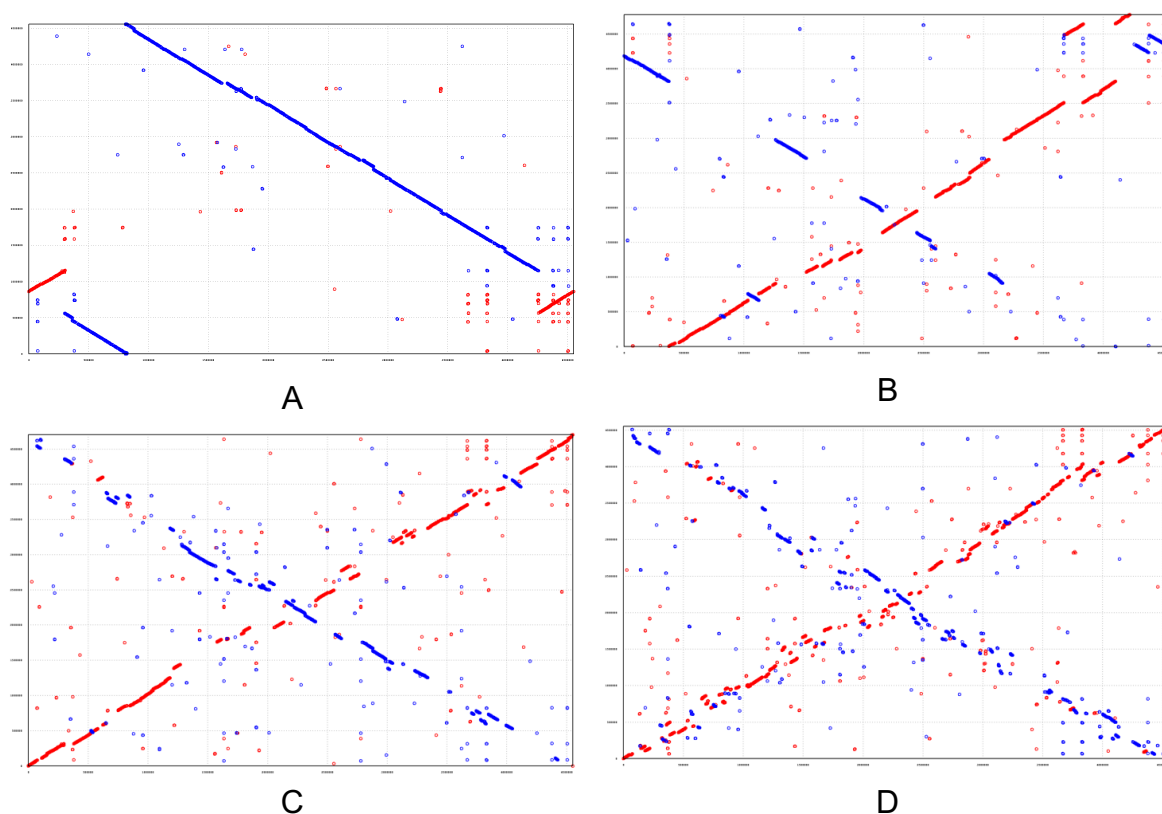


NOTA: A= *A. hydrophila* AH10. B= *A. hydrophila* AL06-06. C= *A. hydrophila* AL09-71. D= *A. hydrophila* J-1. E= *A. hydrophila* ML09-119. F= *A. hydrophila* pc104A. G= *A. hydrophila* subsp. *hydrophila* ATCC7966. H= *A. hydrophila* YL17. I= *A. hydrophila* subsp. *hydrophila* ATCC7966.
 FONTE: a autora (2017).

O GRÁFICO 6 representada os resultados da ordenação genômica para *A. hydrophila* 4AK4 e demais espécies de *Aeromonas*. Assim como nas estirpes de *A. hydrophila*, nota-se que os gráficos *dotplots* de *A. media* WS, *A. salmonicida* subsp. *salmonicida* A449 e *A. veronii* B565 formam um “X” e são menos semelhantes à *A. caviae* 8LM.

Porém, *A. hydrophila* 4AK4 é uma exceção e exibe maior semelhança com o genoma de *A. caviae* 8LM (GRÁFICO 6A), embora este genoma encontre-se em sentido complementar reverso, assim como a estirpe *A. caviae* FDAARGOS 75, GRÁFICO 4B.

GRÁFICO 7 – DOTPLOT ENTRE AS DEMAIS *Aeromonas* E *A. caviae* 8LM



NOTA: A= *A. hydrophila* 4AK4. B= *A. media* WS. C= *A. salmonicida* subsp. *salmonicida* A449. D= *A. veronii* B565.
 FONTE: a autora (2017).

Em suma, os resultados das análises de ordenação dos genomas indicam que *A. caviae* 8LM apresenta maior semelhança com as estirpes *A. caviae* FDAARGOS 72 e 75. A ordenação desses genomas é conservada, embora existam regiões com

sequências no sentido complementar reverso. Isto também foi observado para *A. hydrophila* 4AK4 o que esta em contraste com o que foi observado para outras estirpes de *A. hydrophila*.

Resultados de análises taxonômicas obtidos por BEAZ-HIDALGO e colaboradores (2015), sugerem que *A. hydrophila* 4AK4 não esta relacionada com outras estirpes de *A. hydrophila* e poderia pertencer a uma nova espécie de *Aeromonas*, o que poderia justificar os resultados na análise estrutural do genoma.

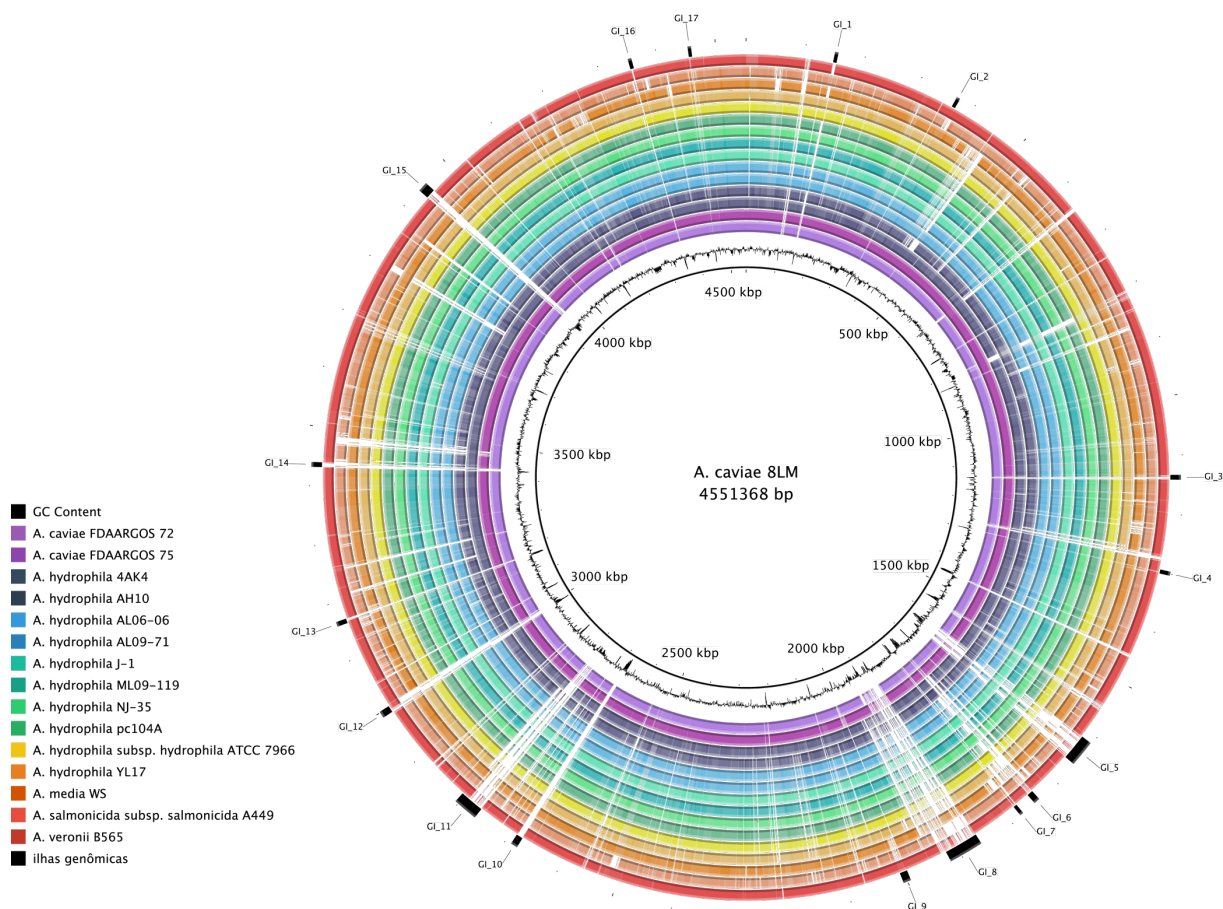
5.4 ANÁLISE DO CONTEÚDO GÊNICO

5.4.1 BLAST Ring Image Generator – BRIG

Os dados obtidos no BRIG estão relacionados a uma análise do conteúdo gênico, ou seja, presença ou ausência de um determinado gene. Esses dados são apresentados através de uma figura circular (*BLAST-Atlas*) e exibem o grau de semelhança entre os genomas analisados e a referência.

A FIGURA 4 representa os resultados do BRIG entre o genoma referência *A. caviae* 8LM e os genomas completos de *Aeromonas*.

Os círculos concêntricos (FIGURA 4) indicam que os genomas de *Aeromonas* são conservados, porém existem várias regiões onde formam-se lacunas na extensão do gráfico circular, as quais representam porções do genoma de *A. caviae* 8LM ausentes nas demais estirpes. Isto ocorre principalmente nas áreas em que o conteúdo de GC% é diferente do restante do genoma de *A. caviae* 8LM.

FIGURA 4 – BLAST-ATLAS DE *Aeromonas* COM GENOMAS COMPLETOS

NOTA: Ao centro, em preto, o genoma de *A. caviae* 8LM utilizado como referência e seu respectivo conteúdo de GC%. Os círculos em cores representam, outras estirpes de *Aeromonas*. Os segmentos externos destacados em preto mostram a distribuição das supostas ilhas genômicas encontradas no genoma de *A. caviae* 8LM.

FONTE: a autora (2017).

Essas regiões estão distribuídas de maneira desigual nos organismos estudados, o que torna cada espécie particular, com a capacidade de sobreviver em variados nichos e infectar diferentes hospedeiros.

5.4.2 Ilhas genômicas

As ilhas genômicas (IG) são regiões no genoma, que variam de 6 a 200 kb aproximadamente, as quais são adquiridas de outros organismos, por meio de transferência genética horizontal (HACKER & KAPER, 2000; HENSEL, 2004; SOARES et al., 2016).

Estas regiões são caracterizadas por um conjunto de fatores, entre eles o conteúdo de GC% diferente do restante do genoma, a presença de genes que codificam transposases e recombinases, bem como a existência de repetições diretas flaqueando a região da ilha genômica (HACHER & KAPER, 2000; SOARES et al., 2016).

As ilhas genômicas no genoma de *A. caviae* 8LM foram preditas pelo *software* Gipsy versão 1.1.2, utilizando como referência a estirpe *A. caviae* FDAARGOS 75, e posteriormente essas regiões foram anotadas automaticamente pelo RAST.

Dezessete IG foram preditas no genoma de *A. caviae* 8LM, como indicado na TABELA 9 e sua distribuição ao longo do genoma está representada na FIGURA 4.

TABELA 9 – ILHAS GENÔMICAS DE *A. caviae* 8LM

GI	INÍCIO	FINAL	TAMANHO	CDS	CLASSE
1	149521	155352	5831	6	Metabólica Resistência
2	366721	372673	5952	4	–
3	1133349	1141341	7992	4	–
4	1296019	1302819	6800	6	–
5	1613396	1658338	44942	45	–
6	1739919	1750133	10214	3	–
7	1775442	1781427	5985	4	–
8	1861619	1920313	58694	54	Patogenicidade
9	1994051	2007054	13003	9	Metabólica Simbiótica
10	2677921	2689314	11393	3	–
11	2763912	2805711	41799	25	–
12	2990110	3001471	11361	9	Patogenicidade Resistência Simbiótica
13	3160187	3168024	7837	5	Patogenicidade
14	3431430	3439544	8114	4	–
15	3935669	3952112	16443	8	–
16	4351419	4357491	6072	4	–
17	4453338	4458983	5645	5	Simbiótica

NOTA: GI= *genomic island*, ilha genômica. –= sem classificação, CDS= sequência de DNA codificante.
FONTE: a autora (2017).

Entre as supostas ilhas genômicas, apenas três foram classificadas com base no seu conteúdo gênico como ilhas de patogenicidade ou simbiótica. Para outras três, IG 1, 9 e 12, ocorreu uma sobreposição na classificação das ilhas genômicas e para as demais nenhuma classificação foi definida (TABELA 9).

No entanto, análises para a identificação de genes de virulência (item 5.6.2), indicaram a presença de um *cluster* gênico codificador do sistema de secreção tipo VI (T6SS), sobrepondo as IG 6 e 7, bem como a região compreendida entre elas.

Nessa região o conteúdo de G+C é 57,9%, diferente do restante do genoma que é de 61,7%. Além disso, essa região apresenta transposases, elementos móveis e RNA transportadores flanqueando os genes do T6SS. Estas características sugerem que este sistema está compreendido em uma ilha genômica.

O fato das GI 6 e 7 não serem identificadas como ilhas de patogenicidade, pode ser decorrente da utilização de *A. caviae* FDAARGOS 75 como referência. Esta estirpe também possui T6SS, e o Gipsy considera apenas genes ausentes no genoma de referência na busca pelas GI.

A GI 1 foi classificada como uma ilha metabólica e de resistência. A anotação pelo RAST indicou a presença de uma proteína da família glioxalase, a qual está associada com o metabolismo celular (INOUE & KIMURA, 1995). Também foi anotado nessa região a enzima aminoglicosídeo N3'-acetiltransferase, associada com resistência a gentamicina, netilmicina, tobramicina ou canamicina (DAVIES & O'CONNOR, 1978).

A GI 8 apresenta o maior número de CDSs e foi classificada como uma ilha de patogenicidade. Pela anotação realizada no RAST, estão presentes nesta região genes codificadores para várias proteínas de fago e elementos móveis, e para a DNA adenina metilase (EROVA et al., 2012). Além disso, há três proteínas relacionadas com o sistema de efluxo RND, o qual pode estar associado com o efluxo de diversos antibióticos e agentes quimioterapêuticos (NIKAIDO & TAKATSUKA, 2009), o que a classificaria como uma ilha de resistência.

A GI 9 foi categorizada como ilha metabólica e simbiótica. O anotador RAST indicou genes que codificam transportadores do tipo ABC para ferro, enzimas como hidrolases e aminotransferases, proteína reguladora do metabolismo e captação de aminoetilfosfonato. A presença de proteínas associadas ao transporte de ferro é importante para o processo de colonização (WYCKOFF et al., 2006; KREWULA & VOGEL, 2008).

A GI 12 apresenta nove CDSs anotadas pelo RAST, sendo uma proteína hipotética e as demais associadas com metabolismo incluindo síntese de polissacarídeos, biossíntese do ácido legionamínico, transferases e epimerase. Houve uma sobreposição na classificação dessa IG (TABELA 9), no entanto, a presença de genes associados com a biossíntese do ácido legionamínico, principal componente do lipopolissacarídeo (LPS) de *Legionella pneumophila* e outros patógenos humanos (MATTHIES et al., 2015), reforça sua classificação como ilha de patogenicidade.

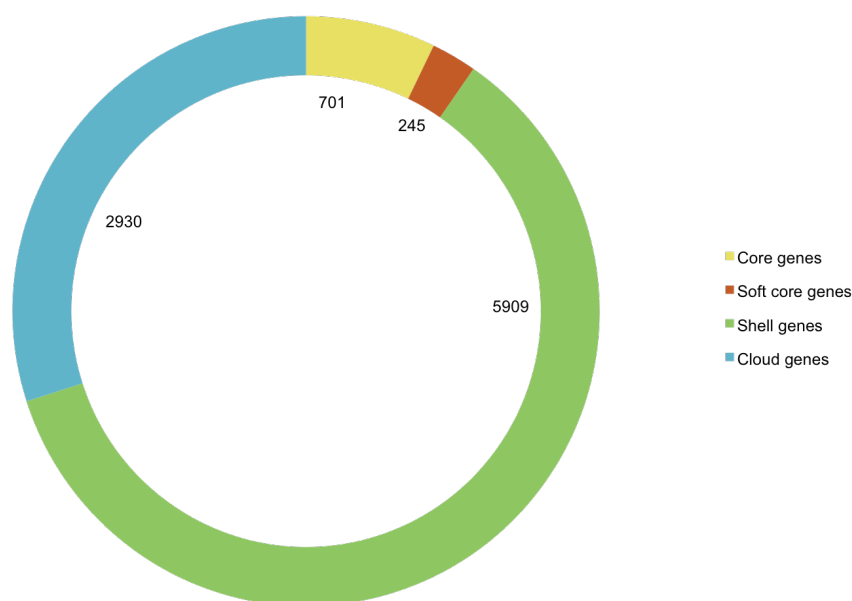
A GI 13 foi classificada como uma ilha de patogenicidade (TABELA 9), apresenta 5 CDSs anotadas pelo RAST, sendo duas proteínas hipotéticas, uma proteína desconhecida, uma proteína captadora de colina de alta afinidade e *5-methylcytosine-specific restriction related enzyme*. O sistema de aquisição de colina é ubiqüitário em bactérias, e a colina é um precursor da glicina betaína, a qual é um osmoprotetor e também poderia servir como fonte de carbono e nitrogênio. Também foi sugerido que esses compostos participem na interação entre eucariotos e procariotos (WARGO, 2013).

Por fim, a GI 17, classificada como simbiótica, possui cinco CDSs, sendo duas proteínas hipotéticas, uma proteína que contém domínio EAL, um precursor da proteína *ycgK* e uma proteína específica para o transporte de triptofano. O domínio EAL, também conhecido como domínio de função desconhecida 2 ou DUF2, apresenta atividade de fosfodiesterase (SCHMIDT et al, 2005; LINDENBERG et al., 2013), e transdução de sinal (GALPERIN et al., 2001; LINDENBERG et al., 2013).

5.4.3 Análise de pan-genoma

O *pipeline* Roary versão 3.6.1 (PAGE et al., 2015) foi utilizado para definir o conjunto do pan-genoma, *core* genoma e genoma acessório das espécies de *Aeromonas* depositadas no banco de dados do NCBI até o release 210 (NCBI RESOURCE COORDINATORS, 2016).

A análise requer a anotação dos genomas realizada pelo *software* Prokka (SEEMANN, 2014), assim apenas os genes anotados foram considerados para definir o pan-genoma. Primeiramente a análise foi realizada somente com as 16 estirpes de *Aeromonas* que apresentam genoma completo, constituindo um pan-genoma de 9.785 genes. Os resultados obtidos estão representados no GRÁFICO 7.

GRÁFICO 8 – NÚMERO DE GENES NO PAN-GENOMA DE *Aeromonas* COM GENOMAS COMPLETOS

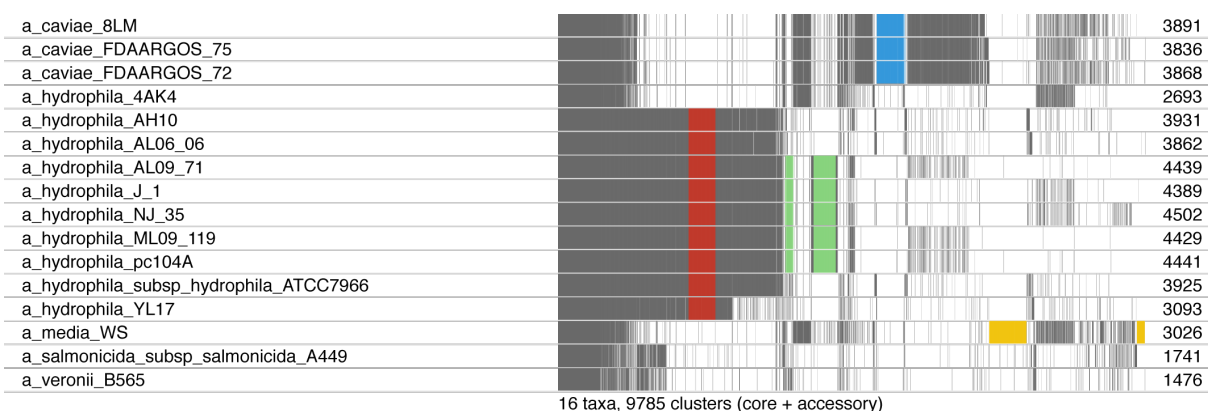
FONTE: a autora (2017).

O *core* genoma é formado por 701 genes, aproximadamente 7%, contendo basicamente genes essenciais associados com manutenção e divisão celular, metabolismo, motilidade e quimiotaxia. A relação desses genes e sua anotação estão disponíveis em ordem alfabética no ANEXO E.

O *soft core* genoma apresenta 245 genes presentes em 15 estirpes, o *shell* genoma compreende 5.909 genes encontrados entre 3 a 14 estirpes e o *cloud* genoma contém 2.930 genes identificados em 1 ou 2 estirpes. O conjunto desses grupos gênicos constitui o genoma acessório.

A seguir, os mesmos genomas foram comparados, também utilizando a *pipeline* Roary, para determinar a presença ou ausência de regiões gênicas nas diferentes estirpes, os resultados estão indicados na FIGURA 5.

FIGURA 5 – PRESENÇA/AUSÊNCIA DE SEQUÊNCIAS NO PAN-GENOMA DE *Aeromonas* COM GENOMA COMPLETO



NOTA: Em cinza está indicada a presença de *clusters* em uma ou mais espécies, em vermelho estão representadas regiões que contém possíveis marcadores moleculares para *A. hydrophila*, em verde apenas para um sub-conjunto de *A. hydrophila*, em azul para *A. caviae* e em amarelo para *A. media*. Os números a direita representam o número de clusters por estirpe.

FONTE: a autora (2017).

A FIGURA 5 representa o pangenoma de *Aeromonas*, verifica-se que existem regiões compartilhadas em determinadas espécies, destacadas em diferentes cores. Devido à resolução da imagem não é possível identificar com precisão essas regiões que representam potenciais marcadores moleculares.

Todas as estirpes de *A. hydrophila* possuem uma região destacada em vermelho na FIGURA 5. Entretanto, as estirpes *A. hydrophila* AL09-71, *A. hydrophila* ML09-119, *A. hydrophila* pc104A, *A. hydrophila* J-1 e *A. hydrophila* NJ-35 apresentam uma região adicional, destacada em verde. Uma vez que essas estirpes são todas isoladas de peixe, é possível que essa região contenha genes associados com virulência para esses animais.

Embora *A. hydrophila* YL17 possua a região realçada em vermelho, ela difere das outras *A. hydrophila* pela ausência de uma região presente nas demais estirpes (FIGURA 5) e poderia representar uma sub-espécie.

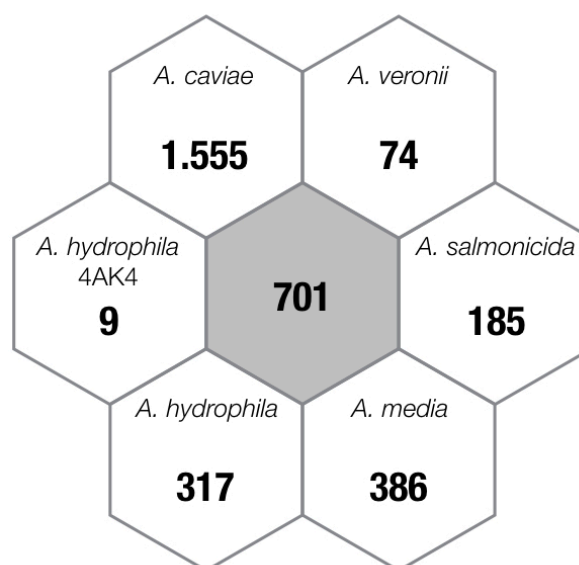
Essas informações indicam que há uma grande variabilidade genética entre as bactérias dessa espécie, o que pode estar relacionado com a sua habilidade em colonizar diversos ambientes e hospedeiros (ALBERT et al., 2000; JANDA & ABBOTT, 2010; KHAJANCHI et al., 2010).

Como também pode ser observado na FIGURA 5, *A. hydrophila* 4AK4 possui um perfil de presença e ausência de genes distinto de *A. hydrophila* e mais similar ao

de *A. media* WS. Porém esta estirpe não apresenta as regiões destacadas em amarelo que são características apenas de *A. media* WS. Este resultado reforça os dados encontrados por BEAZ-HIDALGO e colaboradores (2015), sugerindo que a estirpe *A. hydrophila* 4AK4 representa uma nova espécie.

A partir do genoma acessório foi determinado o número de sequencias únicas entre as espécies, como indicado na FIGURA 6. Devido às informações, as quais sugerem que *A. hydrophila* 4AK4 pode representar uma nova espécie, esta foi considerada separadamente nesta representação. Para a estirpe *A. caviae* 8LM foi observada a presença de 162 sequencias únicas.

FIGURA 6 – NÚMERO DE SEQUENCIAS ÚNICAS NAS ESPÉCIES DE *Aeromonas*



NOTA: O hexágono cinza ao centro indica o número de genes do core genoma e os demais o número de sequencias únicas em cada espécie.

FONTE: a autora (2017).

A FIGURA 6 indica que *A. caviae* é a espécie que contém o maior número de sequencias únicas, sugerindo uma maior diversidade.

Análise de eletroforese em campo pulsado (PFGE) de *Aeromonas*, utilizando a enzima XbaI como endonuclease de restrição, também sugere maior diversidade entre estirpes de *A. caviae*. DALLAGASSA (2016) verificou que entre 32 estirpes de *A. caviae* apenas duas apresentavam o mesmo perfil de PFGE, ao passo que entre

17 *A. hydrophila* cinco estirpes possuíam perfis indistinguíveis, e das 14 estirpes de *A. veronii* apenas sete exibiam perfil idêntico.

É possível que dentre esses genes alguns possam ser utilizados como marcadores moleculares para a identificação da espécie. Devido a semelhança fenotípica entre as espécies do gênero *Aeromonas* nem sempre é possível identificá-las com segurança através de provas bioquímicas convencionais e até mesmo por kits comerciais (JANDA & ABBOTT, 2010). Portanto, a identificação de um marcador molecular seria de grande valia para o diagnóstico de *A. caviae*.

Os resultados obtidos pelo Roary versão 3.6.1 indicaram a presença de 1.555 sequências únicas em *A. caviae*. Destes, 567 codificam proteínas hipotéticas. Entre as 988 sequências restantes aparentemente estão incluídos genes *housekeeping*, um achado inesperado, outros com funções associadas com metabolismo, motilidade, entre outras. Embora indicados como genes únicos de *A. caviae*, verificou-se através do algoritmo BLASTp que estes genes também estão presentes em outras espécies de *Aeromonas*.

Cinco desses genes foram analisados, entre eles alguns *housekeeping*. Foram obtidas as sequências de aminoácidos das proteínas codificadas pelos genes *damX*, *dnaA*, *dnaC*, *ftsA* e *pilM* para análise comparativa frente as bactérias que possuem genoma completo. As sequências foram alinhadas utilizando o algoritmo Clustal Omega.

O gene *damX* codifica uma proteína de divisão celular (ARENDIS et al., 2010). e a sequência de aminoácidos dessa proteína varia de 512 a 535 em *Aeromonas*. O alinhamento das sequências indicou identidade mínima de 72,89% entre as 6 espécies de *Aeromonas*, e de 98,46 a 100% entre as 3 estirpes de *A. caviae*. Foram identificadas 40 substituições de aminoácidos específicas para *A. caviae*. Além disso, a deleção de 18 aminoácidos nas posições 257-258, 293, 304 e 387a 390, estas também observadas em *A. media* e *A. hydrophila*. A região compreendida entre as posições 378 a 426 contém substituição de 18 aminoácidos, observados somente em *A. caviae*, além da deleção de 14 aminoácidos. Esta região constitui um potencial alvo para o desenvolvimento de testes diagnósticos para a identificação de *A. caviae*.

O gene *dnaC* foi identificado pelo Prokka, software utilizado pela *pipeline* Roary versão 3.6.1 para a anotação automática dos genomas, como *dnaA_2*. Quando a sequência foi analisada pelo algoritmo BLASTp, verificou-se que a sequência corresponde ao gene *dnaC*. Este gene codifica uma proteína carregadora de helicase

(MASAI & ARAI, 1988); em *Aeromonas* esta proteína contém 284 aminoácidos e identidade mínima de 81,78%. Foram identificadas, nas sequências de DnaC de *A. caviae*, 12 substituições de aminoácidos únicas para a espécie, todas localizadas na região N-terminal da proteína. Destas, 10 substituições se encontram entre as posições 8 a 53, que também poderia representar potencial marcador molecular para esta espécie.

A proteína PilM está relacionada com a expressão do pili tipo IV e motilidade (Boyd et al., 2008). Sua sequência compreende 353 aminoácidos, com identidade mínima de 72,80% em *Aeromonas*. É altamente conservado em *A. caviae*, com 99,4% de identidade. Foram encontradas 13 substituições de aminoácidos únicas para *A. caviae* distribuídas ao longo da proteína. Destas, 7 estão localizadas entre as posições 72 a 137, das quais 4 estão agrupadas entre 132 a 137. Esta também poderia representar um alvo para diagnóstico.

A proteína DnaA é ativadora do processo de replicação em bactérias (KONIECZNY et al., 1997; RIBER & LØBNER-OLESEN, 2005) e possui 456 aminoácidos em *Aeromonas*. É altamente conservada entre as espécies de *Aeromonas* analisadas, com identidade mínima de 96,93%. Nas sequências de *A. caviae* foram identificadas 4 substituições de aminoácidos, únicas para esta espécie; 3 delas encontram-se entre as posições 102 e 111.

O gene *ftsA* codifica uma proteína essencial de divisão celular (PICHOFF & LUTKENHAUS, 2002) e possui 419 aminoácidos em *Aeromonas*. A partir do alinhamento observou-se que esse gene é altamente conservado nas estirpes avaliadas com identidade mínima de 99,05%. Foram encontradas diferenças na sequência de aminoácidos em apenas seis posições, mas nenhuma delas é específica para *A. caviae*. Portanto, não seria adequado para uso como marcador.

Portanto, entre as sequências identificadas pelo software Roary como únicas para *A. caviae*, muitas estão presentes em outras espécies. No entanto, mesmo assim há potenciais marcadores para espécie entre essas moléculas, que merecem estudo mais aprofundado.

Devido ao baixo número de estirpes com genomas completos, foi realizada outra análise de pan-genoma incluindo os genomas em formato *draft*, compreendendo 115 estirpes de 30 espécies de *Aeromonas*.

Verificou-se que o pan-genoma apresenta 37.418 genes, destes 65 genes fazem parte do *core* genoma e 37.353 genes do genoma acessório. O número de

genes representativos do core genoma foi menor que 1%, esse valor reduzido pode estar relacionado a quantidade de genomas analisados, uma vez que a maioria deles se encontra na forma de contigs, mas também devido a utilização de espécies não proximamente relacionadas filogeneticamente.

No entanto, a representação gráfica da matriz de presença e ausência dos genes no pan-genoma, indicada na FIGURA 7, reforça os achados da análise realizada com genomas completos e confirma a presença de uma região específica presente apenas em *A. caviae*.

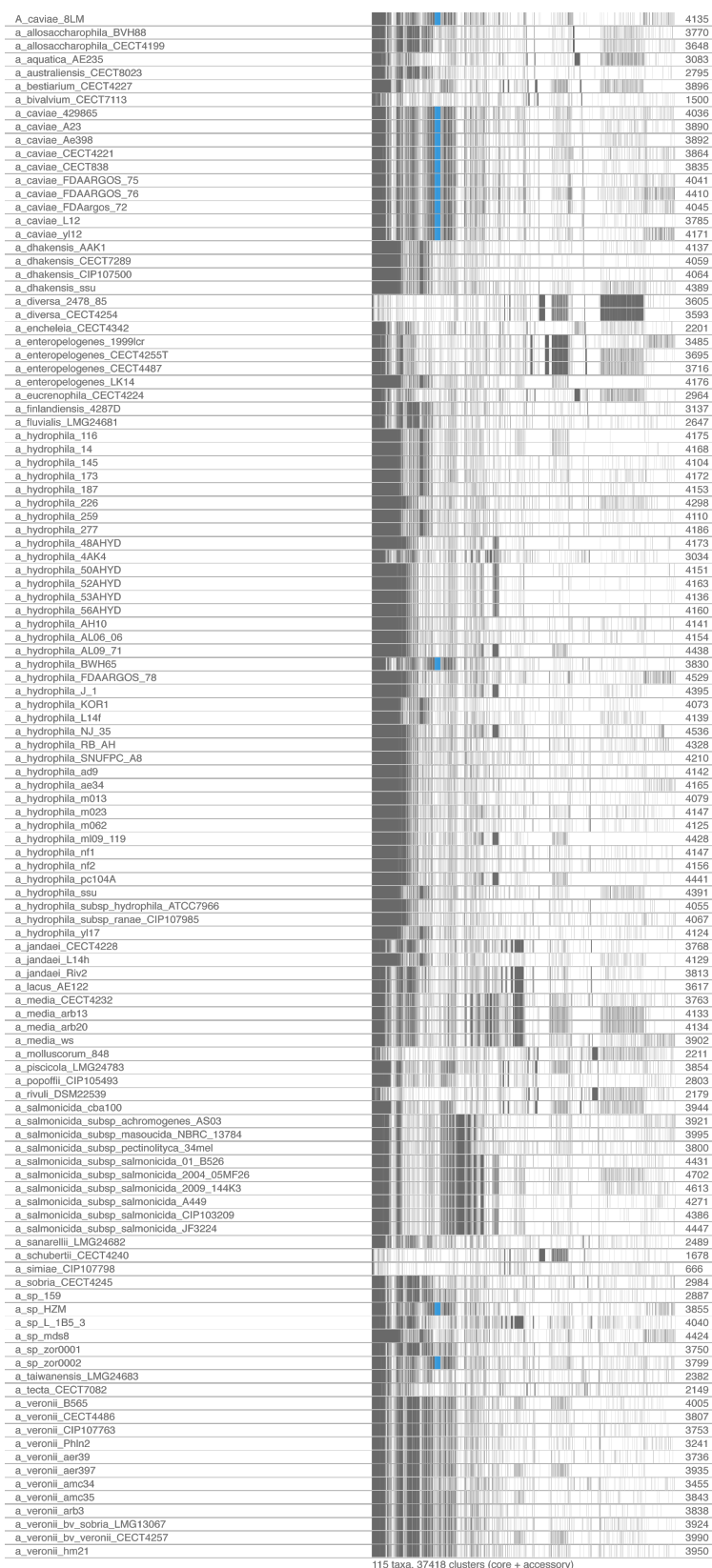
O número crescente de genomas de *Aeromonas* depositados no banco de dados do NCBI proporcionou a realização de estudos de genômica comparativa (GRIM et al., 2013; 2014; PANG et al., 2015; GHATAK et al., 2016). Os estudos publicados com pan-genoma de *Aeromonas* são escassos. GRIM e colaboradores (2013) determinaram o core genoma de *A. hydrophila*, usando apenas três estirpes. As estirpes analisadas compartilharam 3.896 genes, entre eles genes associados aos sistemas de transporte de proteína (Sec, Tat, T1SS e T6SS), flagelo polar, pilus tipo IV (tap e msh) e a enterotoxina citotóxica *at*.

Outro estudo realizado por GRIM e colaboradores (2014) incluiu 7 estirpes de *A. hydrophila*. O core genoma totalizou 3.617 CDSs. Os genes relacionados ao flagelo polar e pilus tipo IV continuaram presentes no core genoma, porém a enterotoxina citotóxica *act* não pertencia mais a este conjunto, o que sugere, segundo o autor, a transferência vertical desse gene em *Aeromonas*.

Em outro estudo GHATAK e colaboradores (2016), determinaram um pan-genoma de 14.150 genes e um core genoma de 2.679 para 46 genomas compreendendo as espécies *A. caviae*, *A. hydrophila* e *A. veronii*.

No presente trabalho o número de genes encontrados no core genoma considerando as 115 estirpes e as 16 com genoma completo, 65 e 701 respectivamente, foram diferentes dos encontrados em outros estudos (GRIM et al., 2013; 2014; PANG et al., 2015; GHATAK et al., 2016).

Essas disparidades podem estar relacionadas com a quantidade de espécies distintas utilizadas, quanto mais distantes as espécies, maiores serão as diferenças nos genomas, bem como o uso de genomas completos e *draft*, reduzindo o número de genes presentes no core genoma, além da metodologia empregada nas análises.

FIGURA 7 – PRESENÇA/AUSÊNCIA DE GENES NO PAN-GENOMA DE *Aeromonas*

NOTA: Em azul os possíveis marcadores moleculares para *A. caviae*.
 FONTE: a autora (2017).

5.5 ANÁLISE FILOGENÉTICA

Para a avaliação filogenética de *Aeromonas* foram utilizadas duas abordagens, uma empregando os genes *housekeeping* e outra com os genes do *core* genoma.

5.5.1 Multilocus Sequence Analysis – MLSA

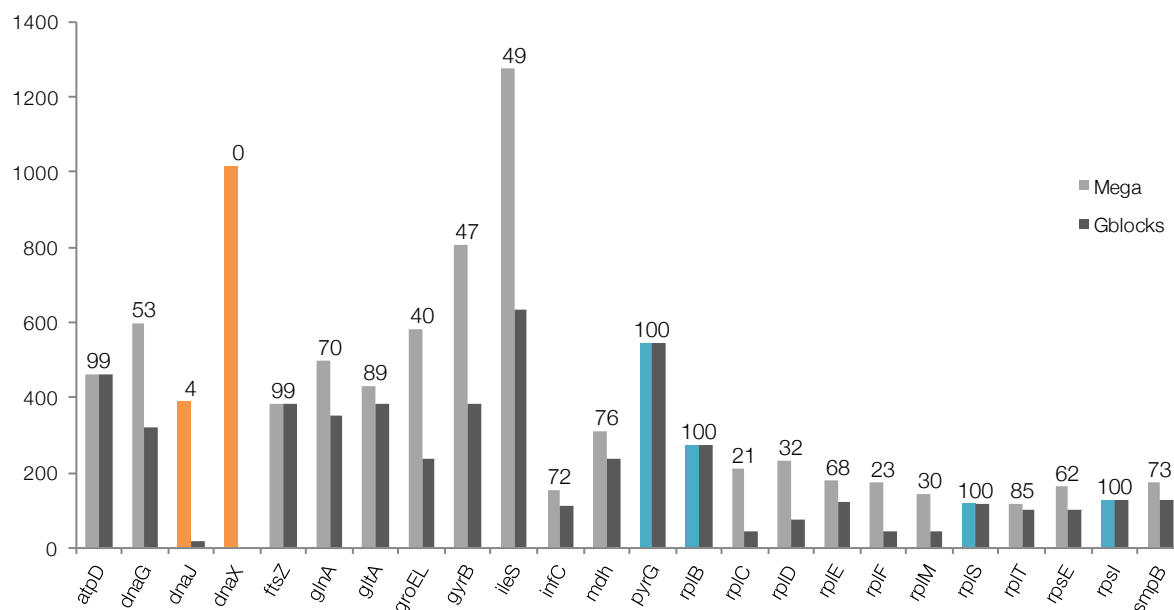
Inicialmente, foram obtidas as sequências de aminoácidos dos produtos dos genes *housekeeping* *atpD*, *dnaG*, *dnaJ*, *dnaX*, *ftsZ*, *glnA*, *gltA*, *groEL*, *gyrB*, *ileS*, *infC*, *mdh*, *pyrG*, *rplB*, *rplC*, *rplD*, *rplE*, *rplF*, *rplM*, *rplS*, *rplT*, *rpsE*, *rpsI* e *smpB* (SAAVEDRA et al., 2006; SESHADRI et al., 2006; MARTINEZ-MURCIA et al., 2011; BEAZ-HIDALGO et al., 2015; COLSTON et al., 2014; SANGLAS et al., 2017) no anotador automático RAST via BLASTp.

As sequências dos genes foram alinhadas separadamente no software MEGA versão 7 utilizando o algoritmo Muscle, no qual foram utilizadas as configurações padrões do programa para o alinhamento. Posteriormente, os alinhamentos foram submetidos ao software Gblocks Server e as sequências resultantes foram concatenadas para construção da árvore filogenética.

Como o software Gblocks elimina posições mal alinhadas e regiões divergentes de um alinhamento de DNA ou proteína a fim de se tornar mais adequado para análise filogenética, as sequências de aminoácidos alinhadas no software MEGA 7 pelo algoritmo Muscle sofreram alteração.

O GRÁFICO 8 mostra a relação gráfica entre a quantidade de aminoácidos nas sequências dos genes *housekeeping* antes e depois do processamento pelo Gblocks.

GRÁFICO 9 – COMPRIMENTO DAS SEQUÊNCIAS DE AMINOÁCIDOS



NOTA: No eixo x os genes *housekeeping*, no eixo y a quantidade de aminoácidos na sequência. Em laranja, as sequências dos genes menos preservados após o processamento pelo Gblocks, em azul, as sequências mais conservadas. Os números acima das colunas indicam a conservação das sequências em porcentagem.

FONTE: a autora (2017).

Posteriormente, a análise filogenética foi inferida usando o método de Máxima verossimilhança baseado no modelo em matriz JTT (JONES et al., 1992). A árvore de consenso *bootstrap*, inferida a partir de 50 repetições, representa a história evolutiva dos taxa analisados (FIGURA 8) (FELSENSTEIN, 1985).

As árvores iniciais para a pesquisa heurística foram obtidas automaticamente aplicando os algoritmos Neighbor-Joining e BioNJ a uma matriz de distâncias *pairwise* estimadas usando um modelo JTT e selecionando a topologia com valor de verossimilhança logarítmica superior.

FIGURA 8 – ÁRVORE FILOGENÉTICA DOS GENES *HOUSEKEEPING*



NOTA: A estirpe em estudo está destacada com uma esfera preta.

FONTE: a autora (2017).

A análise envolveu seqüências de aminoácidos de 115 genomas. Todas as posições com lacunas e dados em falta foram eliminadas. Foram analisadas 5237 posições no conjunto final de dados.

A partir da árvore filogenética (FIGURA 9) nota-se que as estirpes de *A. caviae* formam um ramo muito bem delimitado, separando-se das demais espécies, e mais

próximas filogeneticamente de *A. taiwanensis* e *A. sanarellii*. Resultados similares foram encontrados por BEAZ-HIDALGO e colaboradores (2015) e SANGLAS e colaboradores (2017), demonstrando que o uso de genes *housekeeping* é uma opção para a separação de espécies de *A. caviae*.

5.5.2 Análise filogenética baseada no *core* genoma

Nessa análise foram utilizados os genes representativos do *core* genoma – obtidos por meio da *pipeline* Roary empregada na análise de pan-genoma descrita acima no item 5.4.3 – para relacionar evolutivamente as estirpes analisadas em uma árvore binária usando FastTree (PRICE et al., 2010).

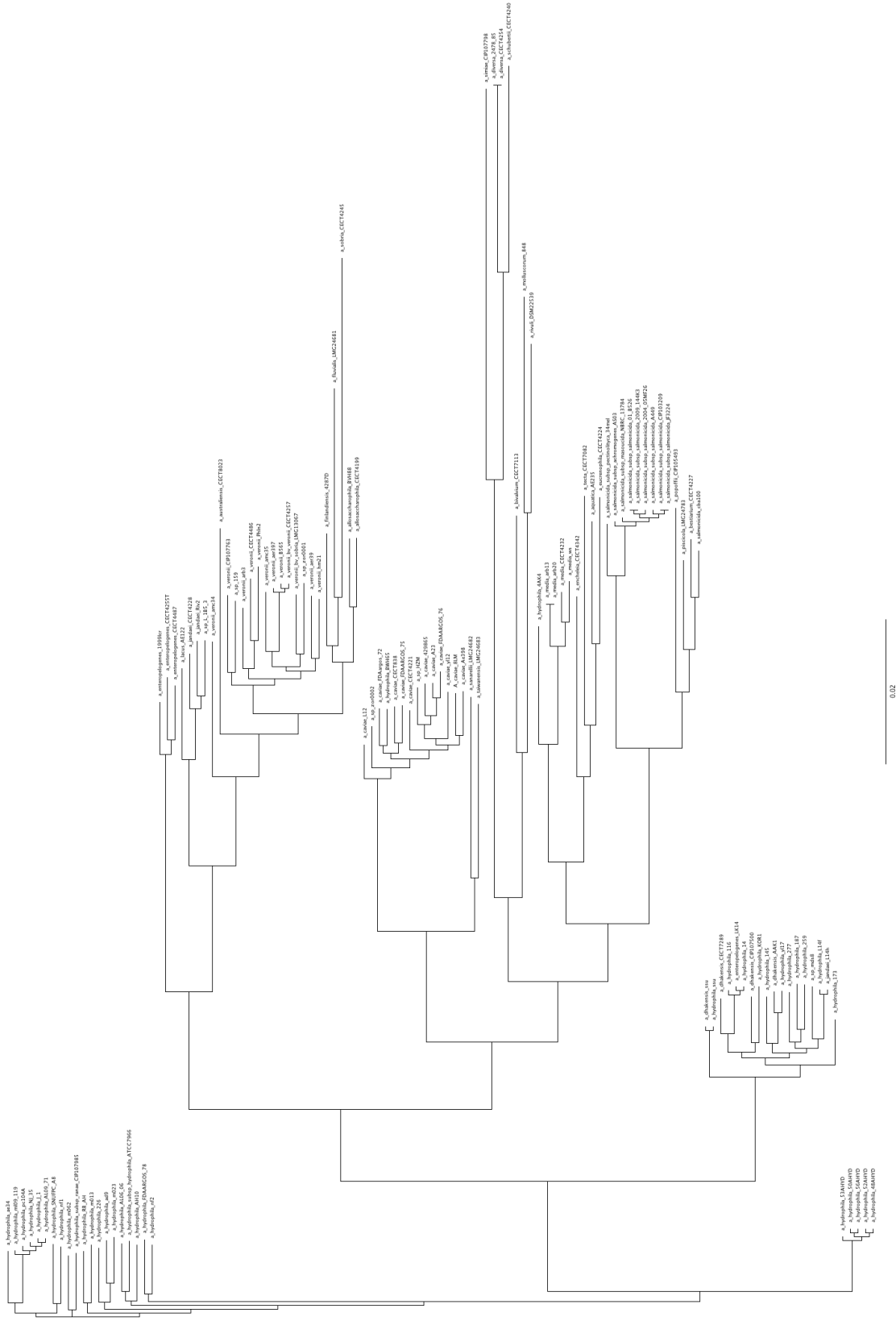
O alinhamento das sequências de nucleotídeos dos genes foi realizado por meio do algoritmo PRANK (LÖYTYNOJA, 2014), que emprega um alinhamento de codons, os resultados obtidos podem ser observados na FIGURA 9, no ANEXO F encontra-se a figura ampliada.

As estirpes de *A. caviae* formam um ramo muito bem delimitado, sendo as espécies *A. taiwanensis* e *A. sanarellii* as mais próximas filogeneticamente, o mesmo padrão foi encontrado nas outras análises filogenéticas empregando os genes *housekeeping*.

A árvore filogenética baseada nos genes do *core* genoma separa melhor as estirpes de *A. hydrophila* que a obtida com a utilização dos genes *housekeeping*, uma vez que o número de genes nessa análise é maior, 65 contra 24, sendo possível a melhor separação das espécies estudadas.

A partir dos resultados das duas abordagens filogenéticas, aliados aos obtidos nas análises de similaridade de genomas (item 5.3.1), constatou-se uma divergência na identificação da estirpe *A. hydrophila* BWH65, a qual é mais próxima de *A. caviae* que *A. hydrophila*.

FIGURA 9 – ÁRVORE FILOGENÉTICA DOS GENES DO CORE GENOMA



NOTA: Escala da distância intergênica entre as estirpes de *Aeromonas*
FONTE: a autora (2017).

Além disso, as estirpes *Aeromonas* sp. ZOR0002 e *Aeromonas* sp. HZM identificadas apenas no nível de gênero puderam agora ser classificadas como pertencentes à espécie *A. caviae*, pois agrupam-se filogeneticamente com essa espécie, confirmando os resultados também encontrados por BEAZ-HIDALGO e colaboradores (2015) e SANGLAS e colaboradores (2017).

Também foram observadas divergências na identificação de outras *Aeromonas* indicadas na TABELA 10. A maioria delas foi previamente identificada como *A. hydrophila*, porém os resultados de similaridade de genomas e das análises filogenéticas apresentados nesse trabalho indicam que efetivamente pertencem à espécie *A. dhakensis*.

TABELA 10 – POSSÍVEL NOVA IDENTIFICAÇÃO DAS ESTIRPES DE *Aeromonas*

REFERÊNCIA NCBI	NOVA CLASSIFICAÇÃO
<i>A. hydrophila</i> BWH65	<i>A. caviae</i>
<i>Aeromonas</i> sp. HZM	<i>A. caviae</i>
<i>Aeromonas</i> sp. ZOR0001	<i>A. caviae</i>
<i>A. enteropelogenes</i> LK14	<i>A. dhakensis</i>
<i>A. hydrophila</i> 14	<i>A. dhakensis</i>
<i>A. hydrophila</i> 116	<i>A. dhakensis</i>
<i>A. hydrophila</i> 145	<i>A. dhakensis</i>
<i>A. hydrophila</i> 173	<i>A. dhakensis</i>
<i>A. hydrophila</i> 187	<i>A. dhakensis</i>
<i>A. hydrophila</i> 259	<i>A. dhakensis</i>
<i>A. hydrophila</i> 277	<i>A. dhakensis</i>
<i>A. hydrophila</i> KOR1	<i>A. dhakensis</i>
<i>A. hydrophila</i> L14f	<i>A. dhakensis</i>
<i>A. hydrophila</i> SSU	<i>A. dhakensis</i>
<i>A. hydrophila</i> YL17	<i>A. dhakensis</i>
<i>A. jandaei</i> L14h	<i>A. dhakensis</i>
<i>Aeromonas</i> sp. 159	<i>A. veronii</i>
<i>A. hydrophila</i> 4AK4	Nova espécie
<i>A. veronii</i> AMC34	Nova espécie

FONTE: a autora (2017).

Os dados das análises filogenéticas indicam que *A. hydrophila* 4AK4 forma um ramo único e mais próximo de *A. media* que de *A. hydrophila*. Assim como a estirpe *A. veronii* AMC34, a qual também forma um ramo separado de *A. veronii*, reforçando os dados encontrados por BEAZ-HIDALGO e colaboradores (2015), os quais sugerem que se tratam de novas espécies.

5.6 VIRULÊNCIA E RESISTÊNCIA

5.6.1 Genes de virulência

A TABELA 11 indica os genes associados com virulência e resistência à antibióticos detectados no genoma de *A. caviae* 8LM.

Entre os genes cujos produtos atuam nas etapas de adesão e colonização, essenciais para o desenvolvimento de infecção, foram encontradas adesinas e pili do tipo IV Msh e Tap, amplamente difundidos entre *Aeromonas* (BOYD et al., 2008; TOMÁS, 2012; GRIM et al., 2013; GRIM et al., 2014). A produção de pilus tipo IV Msh foi associada com diarreia (BARNETT et al., 1997; CHOPRA et al., 2000; GRIM et al., 2014; RASMUSSEN-IVEY et al., 2016).

Também foram identificados genes codificadores de flagelos polar e lateral, que são estruturas importantes para a motilidade de *Aeromonas* (KIROV et al., 2004; GRIM et al., 2013; 2014) e possivelmente auxiliam a bactéria a se deslocar pela camada de muco intestinal. Adicionalmente, foi mostrado que os flagelos de *Aeromonas* atuam como adesinas facilitando a ligação a células intestinais e também na formação de biofilme (KIROV et al., 2002; 2004 GRIM et al., 2014).

A habilidade de *A. caviae* 8LM aderir a linhagens de células intestinais foi confirmada em nosso laboratório por SUREK (2014), onde verificou-se que esta estirpe aderiu a células epiteliais de adenocarcinoma de cólon (HRT-18) mostrando padrão agregativo. Além disso, a expressão dos flagelos polares foi inferida a partir da motilidade tipo *swarming* observada para esta estirpe em estudo realizado por DALLAGASSA (2016).

Também foram identificados genes associados com a produção de sideróforos, os quais permitem a captação de ferro, elemento essencial para a sobrevivência da bactéria no hospedeiro (WYCKOFF et al., 2006; KREWULAK & VOGEL, 2008).

Adicionalmente, foram identificados genes relacionados com a hemolisina termoestável e hemolisina III, como mostrado na TABELA 11. A atividade hemolítica de *A. caviae* 8LM frente a eritrócitos humanos foi confirmada por WOLF (2012), indicando que uma ou ambas as hemolisinas são expressas pela bactéria.

TABELA 11 – FATORES DE VIRULÊNCIA EM *A. caviae* 8LM

FUNÇÃO	NOME	LOCUS	
Secreção	T2SS	AHA_3785–3786 AHA_0568–0579	
	T6SS	AHA_1826–1848 AHA_1118–1119	
Motilidade e adesão	Flagelo polar	AHA_1364–1390 AHA_1698–1703 AHA_2824–2826 AHA_2832–2847 AHA_0660 AHA_4179	
	Flagelo lateral	ASA_0346–0386	
	Pilus tipo IV Msh	AHA_0383–0399	
	Pilus tipo IV Tap	AHA_3868–3871 AHA_3190–3194 AHA_3665–3666 AHA_2739 B565_3462–3469 AHA_2681 AHA_1757	
	Pseudopilina Tpp	AHA_0686–0696	
	Fimbrilina	AHA_0062	
	Adesina	AHA_3491 AHA_2697	
	Resistência à antibióticos	Cloranfenicol acetiltransferase	AHA_0037 AHA_3656
		Fosmidomicina	AHA_0056
		β -lactamase classe C	AHA_3135
		β -lactamase classe D	AHA_4258
	<i>Quorum sensing</i>	AI-1	AHA_0556–0557
		AI-2 (LuxS)	AHA_0700
Sistema de dois componentes <i>QseBC</i>		AHA_3223–3222	
Toxinas	Hemolisina termoestável	AHA_3217	
	Hemolisina III	AHA_3493	
	Toxina RtxA	AHA_1359	
	Enterotoxina citotônica termolábil (Alt)	AAA96668	
Enzimas extracelulares	Elastase	AHA_0851	
	Protease (EprA1)	AHA_2712–2713	
	Fosfolipase A1	AHA_0104	
	Hialuronidase	AHA_1474	
Aquisição de ferro	Colagenase	AHA_0517	
	Síntese de sideroforo	AHA_2473–2479	
	Regulador da captação de ferro <i>Fur</i>	AHA_1530	
	GidA	AHA_4273	
Outros fatores	Enolase	AHA_0821 AHA_0994	
	Serine protease	KMK92093	
	DNA adenina metiltransferase	AHA_3186	
	Invasina	AHA_1066	
	Cold-shock exoribonuclease <i>VacB</i>	AHA_0702	

NOTA: AHA = *Aeromonas hydrophila* subsp. *hydrophila* ATCC 7966, KMK92093 = *Aeromonas enteropelogenes* LK14, AAA96668 = *Aeromonas hydrophila* SSU, ASA = *Aeromonas salmonicida* subsp. *salmonicida*, B565 = *Aeromonas veronii* B565.

FONTE: a autora (2017).

Outras toxinas no genoma de *A. caviae* incluem a toxina RtxA, que promove atividade citotóxica (LINHARTOVÁ et al., 2010) e a enterotoxina Alt, associada com a

elevação dos níveis de AMP cíclico (cAMP) e prostaglandina, levando à secreção de fluidos no intestino (ALBERT et al., 2000).

Possivelmente a presença de pili Msh e a produção da enterotoxina Alt sejam os principais mecanismos de virulência de *A. caviae* 8LM diretamente associados com diarreia.

Além disso, há no genoma da bactéria genes codificadores para diversas enzimas extracelulares como elastase, proteases, hialuronidases, entre outras, que podem auxiliar a bactéria a obter nutrientes, bem como invadir tecidos do hospedeiro.

A. caviae 8LM também contém um gene codificador de enolase, a qual é uma enzima glicolítica e apresenta atividade “moonlight” sendo expressa na superfície bacteriana com funções de proteína *heat-shock* e reguladora da transcrição (RASMUSSEN-IVEY et al., 2016).

Além disso, no genoma de *A. caviae* 8LM há genes codificadores para os autoindutores AI-1 e 2, bem como o sistema de dois componentes QseBC. Estes fazem parte do sistema de sinalização *quorum sensing* que atuam na regulação da virulência bacteriana influenciando diversas características tais como a motilidade, atividade hemolítica, formação de biofilmes, secreção de alguns fatores de virulência, entre outros (KHAJANCHI et al., 2012; RASMUSSEN-IVEY et al., 2016).

Dois sistemas de secreção, T2SS e T6SS, associados com virulência, foram identificados no genoma de 8LM. O T2SS é amplamente conservado em *Aeromonas* e essencial para o transporte de vários fatores de virulência como aerolisinas, fosfolipases, proteases e Dnases. Essas proteínas estão associadas com destruição tecidual (PANG et al., 2015; RASMUSSEN-IVEY et al., 2016) contribuindo para a obtenção de nutrientes e invasão do hospedeiro (ABBOTT e JANDA, 2010).

O T6SS em *Aeromonas* está relacionado com virulência através do transporte de proteínas efetoras diretamente do citoplasma bacteriano para o citoplasma da célula hospedeira levando a processos citotóxicos, apoptose e evasão do sistema imune (SUAREZ et al., 2008, 2010a, 2010b; SHA et al., 2013; GRIM et al., 2014).

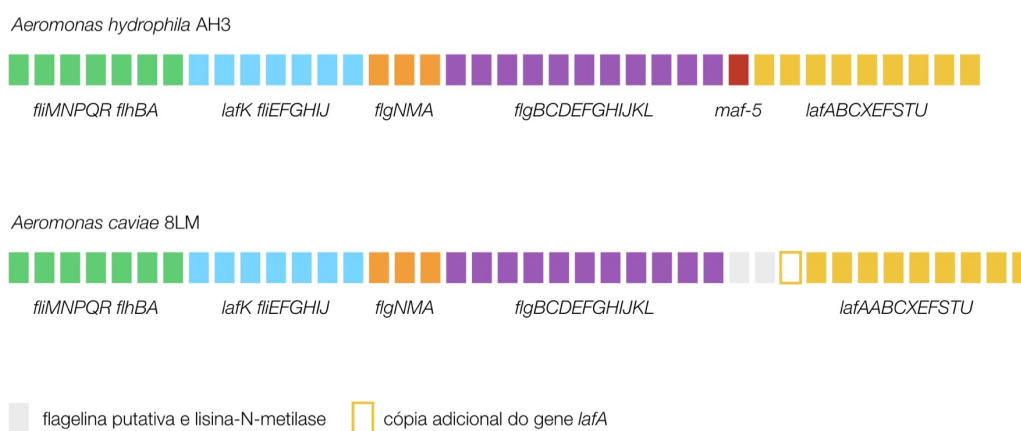
Esses dados sugerem que *A. caviae* 8LM apresenta um repertório de genes associados com virulência que lhe permitem causar diarreia e potencial para provocar outras patologias.

Em relação a resistência a antibióticos, foram encontrados genes que codificam para β -lactamases classe C e D, os quais estão relacionadas com a resistência a ampicilina, cefalotina e cefazolina observada em *A. caviae* 8LM por ASSIS (2014). No

entanto, no mesmo estudo não há dados laboratoriais disponíveis em relação ao cloranfenicol e à fosmidomicina (ASSIS, 2014).

A disposição dos genes dos flagelos lateral e polar foram comparadas com a estirpe *A. hydrophila* AH3 (TOMÁS, 2012), os resultados estão indicados nas FIGURAS 10 e 11, respectivamente.

FIGURA 11 – ORGANIZAÇÃO GÊNICA DO FLAGELO LATERAL

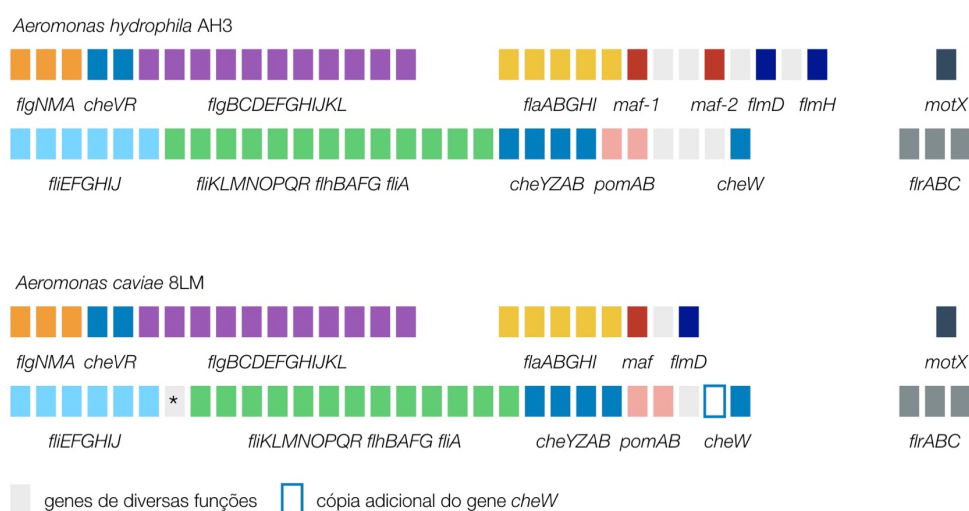


FONTE: a autora (2017).

Em *A. hydrophila* AH-3 o flagelo lateral é codificado por 38 genes agrupados em um único *cluster* gênico (CANALS et al., 2006; TOMÁS, 2012). A organização dos genes que codificam o flagelo lateral (FIGURA 10) nas duas espécies é bastante semelhante. Porém *A. caviae* 8LM não apresenta o gene *maf-5* envolvido na glicosilação do flagelo, nesta região contém uma flagelina putativa e uma lisina-N-metilase; além disso, há uma cópia extra do gene *lafA*, em relação a estirpe referência.

Enquanto que o flagelo polar é codificado por 55 genes distribuídos em cinco regiões ao longo do genoma, como mostrado na FIGURA 11 (CANALS et al., 2006; TOMÁS, 2012).

FIGURA 12 – ORGANIZAÇÃO GÊNICA DO FLAGELO POLAR



FONTE: a autora (2017).

A organização gênica de *A. caviae* 8LM exibe diferenças em relação a referência, a medida que *A. hydrophila* AH3 possui os genes *maf-1* e *maf-2*, *A. caviae* 8LM apresenta apenas uma cópia do gene *maf*. Neste mesmo agrupamento de genes, podemos notar também a deleção de cinco genes. Além disso, observa-se a inserção de um gene, destacado com um asterisco, entre os genes *fliJ* e *fliK*, bem como, duas deleções após o gene *pomB* e uma duplicação do gene *cheW*.

5.6.2 Sistema de secreção tipo VI – T6SS

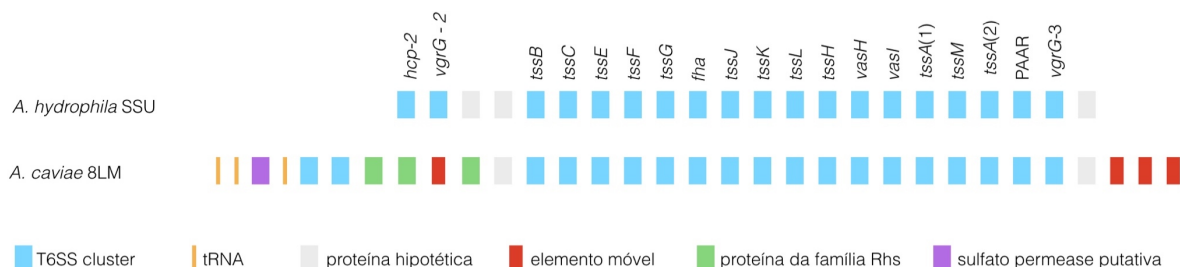
A presença de genes codificadores para o T6SS foi verificado em *A. caviae* 8LM pela anotação automática do RAST. Porém alguns genes não foram identificados pelo anotador, como o *hcp* e *impK/vasF*, *ompA/motB*, uma proteína hipotética conservada e uma proteína uropatogênica. Além disso, as proteínas de referência *vgrG-1*, *vgrG-2* e *vgrG-3* (SUAREZ et al., 2008; COULTHURST, 2013) foram distribuídas entre *vgrG* e *vgrG-3* no RAST.

Esse sistema é codificado por 13 genes conservados presentes em todos os organismos, além de genes acessórios adicionais (SUAREZ et al., 2008; BOYER et al., 2009; COULTHURST, 2013).

A disposição dos genes do T6SS em *A. caviae* 8LM e *A. hydrophila* SSU pode ser observada na FIGURA 12. Verifica-se que a estirpe *A. caviae* 8LM possui todo o

cluster gênico do T6SS e segue a mesma ordenação gênica, de *tssB* (*impB*) à *vgrG-3*, que a referência *A. hydrophila* SSU.

FIGURA 13 – ORGANIZAÇÃO GÊNICA DO T6SS



FONTES: a autora (2017).

No entanto, a região 5' do cluster é variável, com a presença de genes da família Rhs, elementos móveis e integrases. Esta organização gênica também foi observada para a maioria dos genomas de *Aeromonas* analisados por GRIM e colaboradores (2014), compreendendo 16 estirpes das espécies *A. hydrophila*, *A. dhakensis*, *A. salmonicida* e *A. jandaei*.

Assim como a referência, *A. caviae* 8LM possui um único gene *hcp* (*hcp-2*) e dois genes *vgrG* (*vgrG-2* e *vgrG-3*) dentro do *cluster* do T6SS, como mostrado na FIGURA 12.

Porém, num *locus* cromossômico separado foram encontrados um segundo gene *hcp* (*hcp-1*) e uma terceira *vgrG* (*vgrG-1*), esses mesmo genes também foram encontrados fora do cluster do T6SS na referência *A. hydrophila* SSU (SUAREZ et al., 2008) e em *A. hydrophila* ATCC 7966 (GRIM et al., 2014).

Em *A. caviae* 8LM, o T6SS está codificado na região localizada entre as possíveis ilhas genômicas 6 e 7 (item 5.3.2). A presença de RNAs transportadores e elementos móveis flanqueando o *cluster*, sugerem que esta região foi adquirida por transferência horizontal.

6 CONCLUSÃO

Com o sequenciamento do genoma de *Aeromonas caviae* 8LM foi possível obter o genoma completo que contém 4,6Mb, conteúdo GC de 61,7%, 4.256 genes, 4.101CDS e 154 RNA estruturais. A estirpe em estudo apresenta 10 operons de rRNA distribuídos próximos da origem de replicação.

O genoma da estirpe em estudo é similar ao de outras espécies do gênero, as diferenças ocorrem principalmente nas regiões onde o conteúdo de GC% de *A. caviae* 8LM é diferente do restante do seu genoma.

Foram preditas 17 ilhas genômicas em *A. caviae* 8LM, destas apenas seis foram classificadas como ilhas metabólicas, patogênicas, de resistência ou simbióticas. As ilhas 6 e 7 que não foram classificadas, compreendem o *cluster* do T6SS, associado com virulência, representando uma região única que pode ser classificada como ilha de patogenicidade.

O pan-genoma de *Aeromonas* com genoma completo possui 9.785, dos quais 701 constituem o *core* genoma. A análise do pan-genoma possibilitou detectar regiões genômicas contendo possíveis marcadores moleculares para a espécie *A. caviae*.

No genoma de *A. caviae* 8LM estão presentes genes associados com virulência incluindo sistemas de secreção tipo II e VI, motilidade e adesão, resistência a antibióticos, *quorum-sensing*, diversas toxinas, aquisição de ferro, entre outros fatores.

A organização dos genes para os flagelos lateral e polar em *A. caviae* 8LM é semelhante a encontrada em *A. hydrophila* AH3, porém há algumas deleções, inserções e duplicações gênicas no *cluster* desses dois tipos de flagelos.

O *cluster* do T6SS é conservado em *A. caviae* 8LM, porém há variação na região 5' do *cluster*, com a presença de genes da família Rhs, proteínas transmembranares, elementos móveis, tranposases e integrases.

As análises filogenéticas e de similaridade confirmaram a identificação laboratorial de *A. caviae* 8LM, bem como indicaram algumas divergências na classificação de *Aeromonas*.

REFERÊNCIAS

ABBOTT, S. L.; CHEUNG, W. K. W.; JANDA J. M. The Genus *Aeromonas*: Biochemical Characteristics, Atypical Reactions, and Phenotypic Identification Schemes. **Journal of Clinical Microbiology**. p. 2348–2357, 2003.

AGUILERA-ARREOLA, M. G.; HERNÁNDEZ-RODRÍGUEZ, C.; ZÚÑIGA, G.; FIGUERAS, M. J.; GARDUÑO, R. A.; CASTRO-ESCARPULLI, G. Virulence potential and genetic diversity of *Aeromonas caviae*, *Aeromonas veronii*, and *Aeromonas hydrophila* clinical isolates from Mexico and Spain: a comparative study. **Canadian Journal of Microbiology**. v. 53, n. 7, p. 877–887, 2007.

AL-BENWAN, K.; ABBOTT, S.; JANDA, J. M.; ALBERT, M. J. Cystitis caused by *Aeromonas caviae*. **Journal of Clinical Microbiology**. v. 45, p. 2348–2350, 2007.

ALBERT, M. J.; ANSARUZZAMAN, M.; TALUKDER, K. A.; CHOPRA, A. K.; KUHN, I.; RAHMAN, M.; FARUQUE, A. S. G.; ISLAM, M. S.; SACK, R. B.; MOLLBY, R. Prevalence of enterotoxin genes in *Aeromonas* spp. isolated from children with diarrhea, healthy controls, and the environment. **Journal of Clinical Microbiology**. v. 38, n. 10, p. 3785–3790, 2000.

ALCAIDE, E.; BLASCO, M. D.; ESTEVE, C. Mechanisms of quinolone resistance in *Aeromonas* species isolated from humans, water and eels. **Research in Microbiology**. v. 161, p. 40–45, 2010.

ALCARAZ L. D. Pan-genomics: Unmasking the gene diversity hidden in the bacteria species. **PeerJ PrePrints**. v. 2, n. 113, 2014.

ALIKHAN, N.-F.; PETTY, N. K.; BEN ZAKOUR, N. L.; BEATSON, S. A. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. **BMC Genomics**. v. 12, n. 402, 2011.

ALPERI, A.; FIGUEIRAS, M. J.; INZA, I.; MARTÍNEZ-MURCIA, A. J. Analysis of 16S rRNA gene mutations in a subset of *Aeromonas* strains and their impact in species delineation. **International Microbiology**. v. 11, p. 185–194, 2008.

ARENDS, S. J. R.; WILLIAMS, K.; SCOTT, R. J.; ROLONG, S.; POPHAM, D.; WEISS, D. S. Discovery and Characterization of Three New *Escherichia coli* Septal

Ring Proteins That Contain a SPOR Domain: DamX, DedD, and RlpA. **Journal of Bacteriology**. v. 192, n. 1, p. 242–255, 2010.

ASSIS, F. E. A.; WOLF, S.; SUREK, M.; DE TONI, F.; SOUZA, E. M.; PEDROSA, F. O.; FARAH, S. M. S. S.; PICHETH, G.; FADEL-PICHETH, C. M. T. Impact of *Aeromonas* and diarrheagenic *Escherichia coli* screening in patients with diarrhea in Paraná, southern Brazil. **Journal of Infection in Developing Countries**. v. 8, n. 12, p.1609–1614, 2014.

(a) AUCH, A.F.; KLENK, H.-P.; GÖKER, M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. **Standards in Genomic Sciences**. v. 2, p. 142–148, 2010.

(b) AUCH, A.F.; VON JAN, M.; KLENK, H.-P.; GÖKER, M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. **Standards in Genomic Sciences**. v. 2, p. 117–134, 2010.

AZIZ, R. K.; BARTELS, D.; BEST, A. A.; DEJONGH, M.; DISZ, T.; EDWARDS, R. A.; FORMSMA, K.; GERDES, S.; GLASS, E. M.; KUBAL, M.; MEYER, F.; OLSEN, G. J.; OLSON, R.; OSTERMAN, A. L.; OVERBEEK, R. A.; MCNEIL, L. K.; PAARMANN, D.; PACZIAN, T.; PARRELLO, B.; PUSCH, G. D.; REICH, C.; STEVENS, R.; VASSIEVA, O.; VONSTEIN, V.; WILKE, A.; ZAGNITKO, O. The RAST Server: rapid annotations using subsystems technology. **BMC Genomics**. v. 9, n. 75, 2008.

BARNETT, T. C.; KIROV, S. M.; STROM, M. S.; SANDERSON, K. *Aeromonas* spp. possess at least two distinct type IV pilus families. **Microbial Pathogenesis**. v. 13, n. 4, p. 241–247, 1997.

BARNETT, T. C.; KIROV, S. M. The type IV *Aeromonas* pilus (Tap) gene cluster is widely conserved in *Aeromonas* species. **Microbial Pathogenesis**. v.26, p. 77–84, 1999.

BEATSON, S. A.; DE LUNA, M. G.; BACHMANN, N. L.; ALIKHAN, N.-F.; HANKS, K. R.; SULLIVAN, M. J.; WEE, B. A.; FREITAS-ALMEIDA, A. C.; DOS SANTOS, P. A.; DE MELO, J. T. B.; SQUIRE, D. J. P.; CUNNINGHAM, A. F.; FITZGERALD, J. R.; HENDERSON, I. R. Genome Sequence of the Emerging Pathogen *Aeromonas caviae*. **Journal of Bacteriology**. v. 193, n. 5, p. 1286–1287, 2011.

BEAZ-HIDALGO, R.; MARTÍNEZ-MURCIA, A.; FIGUERAS, M. J. Corrigendum to "Reclassification of *Aeromonas hydrophila* subsp. *dhakensis* Huys et al. 2002 and *Aeromonas aquariorum* Martínez-Murcia et al. 2008 as *Aeromonas dhakensis* sp. nov. comb nov. and emendation of the species *Aeromonas hydrophila*." [Systematic and Applied Microbiology. v. 36, p. 171–176, 2013.] **Systematic and Applied Microbiology**. v. 37, p. 543, 2014.

BEAZ-HIDALGO, R.; HOSSAIN, M. J.; LILES, M. R.; FIGUERAS, M. -J. Strategies to Avoid Wrongly Labelled Genomes Using as Example the Detected Wrong Taxonomic Affiliation for *Aeromonas* Genomes in the GenBank Database. **PLoS ONE**. v. 10, n. 1, e. 0115813, 2015.

BEHERA, B.; KUMAR, M. T.; NARASARAJU, K.; RANJAN, P. N.; MURALI, M. R.; NIMMALA, P. Fatal spontaneous *Aeromonas hydrophila* myonecrosis and sepsis without antecedent trauma. **JMM Case Reports**. p. 1–5, 2014.

BINGLE, L. E.; BAILEY, C. M.; PALLEN, M. J. Type VI secretion: a beginner's guide. **Current opinion in microbiology**. v. 11, p. 3–8, 2008.

BORREL, N.; FIGUEIRAS, M. J.; GUARRO, J. Phenotypic identification of *Aeromonas* genomospecies from clinical and environmental sources. **Canadian Journal of Microbiology**. v. 44, p. 103–108, 1998.

BOYD, J. M.; DACANAY, A.; KNICKLE, L. C.; TOUHAMI, A.; BROWN, L. L.; JERICO, M. H.; JOHNSON, S. C.; REITH, M. Contribution of Type IV Pili to the Virulence of *Aeromonas salmonicida* subsp. *salmonicida* in Atlantic Salmon (*Salmo salar* L.). **Infection and Immunity**. v. 76, n. 4, p. 1445–1455, 2008.

BOYER, F.; FICHANT, G.; BERTHOD, J.; VANDENBROUCK, Y.; ATTREE, I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? **BMC Genomics**. v. 10, n. 104, 2009.

BUCHAMAN, R. L. The "New" pathogens; and update of selected examples. **Association Food Drug Official Chemistry Bulletin**. v. 48, p. 142–155, 1984.

BUKHARI, E. E.; ALRABIAHET, A. A. A review of clinically suspected sepsis and meningitis in infants under 90 days old in a tertiary care center in Saudi Arabia. **Journal of Microbiology and Infectious Diseases**. v. 1, n. 2, p. 47–52, 2011.

CANALS, R.; RAMIREZ, S.; VILCHES, S.; HORSBURGH, G.; SHAW, J. G.; TOMÁS, J. M.; MERINO, S. Polar Flagellum Biogenesis in *Aeromonas hydrophila*. **Journal of Bacteriology**. v. 188, n. 2, p. 542–555, 2006.

CARVALHO, M. C. da C. G.; da SILVA, D. C. G. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciência Rural**. v. 40, n. 3, p. 735–744, 2010.

CASADEVALL, A.; PIROFSKI, L.-A. Host-pathogen interactions: basic concepts of microbial commensalism, colonization, infection, and disease. **Infection and Immunity**. v. 68, n. 12, p. 6511–6518, 2000.

CASCALES, C. The type VI secretion toolkit. **EMBO reports**. v. 9, p. 735–741, 2008.

CASTRESANA, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. **Molecular Biology and Evolution**. v. 17, p. 540–552, 2000.

CHAI, B.; WANG, H.; CHEN, X. Draft genome sequence of high-melanin-yielding *Aeromonas media* strain WS. **Journal of Bacteriology**. v. 194, n. 23, p. 6693–6694, 2012.

CHAN, K.-G.; CHIN, P.-S.; TEE, K. K.; CHANG, C.-Y.; YIN, W.-F.; SHENG, K.-Y. Draft genome sequence of *Aeromonas caviae* strain L12, a *quorum-sensing* strain isolated from a freshwater lake in Malaysia. **Genome Announcements**. v. 3, n. 2, e00079-15, 2015.

CHAO, C. M.; LAI, C. C.; TANG, H. J.; KO, W. C.; HSUEH, P.-R. Biliary tract infections caused by *Aeromonas* species. **European Journal of Clinical Microbiology & Infectious Diseases**. v. 32, n. 2, p. 245–251, 2013.

CHIM, H.; SONG, C. *Aeromonas* infection in critically ill burn patients. **Burns**. v. 33, p. 756–759, 2007.

CHOI, J. -P.; LEE, S. -O.; KWON, H. -H.; KWAK, Y. G.; CHOI, S. -H.; LIM, S. K.; KIM, M. N.; JEONG, J. -Y.; CHOI, S. -H.; WOO, J. H.; KIM, Y. S. Clinical significance of spontaneous *Aeromonas* bacterial peritonitis in cirrhotic patients: a matched case-control study. **Clinical Infectious Diseases**. v. 47, p. 66–72, 2008.

CHOPRA, A. K.; XU, X.-J.; RIBARDO, D.; GONZALEZ, M.; KUHL, K.; PETERSON, J. W.; HOUSTON C. W. The cytotoxic enterotoxin of *Aeromonas hydrophila* induces proinflammatory cytokine production and activates arachidonic acid metabolism in macrophages. **Infection and Immunity**. v. 68, n. 5, p. 2808–2818, 2000.

CHOPRA, A. K. Characterization of waterborne *Aeromonas* species for their virulence potential Copyright © by Awwa Research Foundation, 2008.

CHOPRA, A. K.; HOUSTON C. W. Enterotoxins in *Aeromonas*-associated gastroenteritis. **Microbes and Infection**. v. 1, n. 13, p. 1129–1137, 1999.

CHUANG, H. C.; HO, Y. H.; LAY, C. J.; WANG, L. S.; TSAI, Y. S.; TSAI, C. C. Different clinical characteristics among *Aeromonas hydrophila*, *Aeromonas veronii* biovar *sobria* and *Aeromonas caviae* monomicrobial bacteremia. **Journal of Korean Medical Science**. v. 26, p. 1415–1420, 2011.

COBURN, B.; SEKIROV, I.; FINLAY, B. B. Type III secretion systems and disease. **Clinical Microbiology Reviews**. v.20, n. 4, p. 535–49, 2007.

COLSTON, S. M.; FULLMER, M.; BEKA, L.; LAMY, B.; GOGARTEN, J. P.; GRAF, J. Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. **mBio**. v. 5, p. 6, e.02136–14, 2014.

COULTHURST, S. J. The Type VI secretion system e a widespread and versatile cell targeting system. **Research in Microbiology**. v. 164, p. 640–654, 2013.

DAVIES, J.; O'CONNOR, S. Enzymatic Modification of Aminoglycoside Antibiotics:3-N- Acetyltransferase with Broad Specificity that Determines Resistance to the Novel Aminoglycoside Apramycin. **Antimicrobial Agents and Chemotherapy**. v. 14, n. 1, p. 69–72, 1978.

DELAMARE, A. P. L.; COSTA, S. O. P.; DA SILVEIRA, M. M.; ECHEVERRIGARAY, S. Growth of *Aeromonas* species on increasing concentrations of sodium chloride. **Letters in Applied Microbiology**. v. 30, p. 45–60, 2000.

DELAMARE, A. P. L.; Dalcin, T.; Müller, G.; da Costa, S. O. P.; Echeverrigaray, S. The effect of organic osmoprotectors on *Aeromonas trota* and *A. hydrophila* grown under high sodium chloride concentrations. **Brazilian Journal of Microbiology**. v. 34, n. 1, p. 128–130, 2003.

DELCHER, A. L.; KASIF, S.; FLEISCHMANN, R. D.; PETERSON, J.; WHITE, O. SALZBERG, S. L. Alignment of whole genomes. **Nucleic Acids Research**. v. 27, n.11, p. 2369–2376, 1999.

DINIZ, J. A.; LIU, Y.-C.; COULTHURST, S. J. Molecular weaponry: diverse effectors delivered by the Type VI secretion system **Cell Microbiology**. v. 17, n. 12, p. 1742–1751, 2015.

DWIVEDI et al., 2008 – DWIVEDI, M.; MISHRA, A.; PRASAD, A.; AZIM, A.; SINGH, R. K.; BARONIA, A. K.; PRASAD, K. N.; DWIVEDI, U. N. *Aeromonas caviae* septicemia in immunocompetent gastrointestinal carriers. **The Brazilian Journal of Infectious Diseases**. v. 12, n. 6, p. 547–548, 2008.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**. v. 32, n. 5, 2004.

ESSERS, B.; BURNENS, A. P.; LANFRANCHINI, F. M.; SOMARUGA, S. G.; von VIGIER, R. O.; SCHAAD, U. B.; AEBI, C.; BIANCHETTI, M. G. Acute community-acquired diarrhea requiring hospital admission in Swiss children. **Clinical Infectious Diseases**. v. 31, n. 1, p. 192–196, 2000.

EROVA, T. E.; KOSYKH, V. G.; SHA, J.; CHOPRA, A. K. DNA adenine methyltransferase (Dam) controls the expression of the cytotoxic enterotoxin (act) gene of *Aeromonas hydrophila* via tRNA modifying enzyme-glucose-inhibited division protein (GidA). **Gene**. v. 498, p. 280–287, 2012.

FARMER III, J. J.; M. J. ARDUINO; F. W. HICKMAN-BRENNER. The Genera *Aeromonas* and *Plesiomonas*, The Prokaryotes. 1991–2005, Springer-Verlag New York.

FALCÃO, D. P.; LUSTRI, W. R.; BAUAB, T. M. Incidence of non-01 *Vibrio cholerae* and *Aeromonas* spp. in fresh water in Araraquara, Brazil. **Current Microbiology**. v. 37, p. 28–31, 1998.

FELSENSTEIN J. Confidence limits on phylogenies: An approach using the bootstrap. **Evolution**. v. 39, p. 783–791, 1985.

FIGUEIRAS, M. J. Clinical relevance of *Aeromonas* sM503. **Reviews in Medical Microbiology**. v. 16, p. 145–153, 2005.

FOSSE, T. Chapter 42: *Aeromonas*, *Vibrio* and *Plesiomonas*. In: COURVALIN, P.; LECLERCQ, R.; RICE, L. B. **Antibiogram**. 1ª Edição. Paris – França: Eska Publishing, ASM Press, 2010. 511–515.

FOURNIER, P.-E.; DUBOURG, G.; RAOULT, D. Clinical detection and characterization of bacterial pathogens in the genomics era. **Genome Medicine**. v. 6, n. 114, 2014.

GAO, X.; JIAN, J.; LI, W.-J.; YANG, Y.-C.; SHEN, X.-W.; SUN, Z.-R.; WU, Q.; CHEN, G.-Q. Genomic study of polyhydroxyalkanoates producing *Aeromonas hydrophila* 4AK4. **Applied Microbiology and Biotechnology**. v. 97, n. 20, p. 9099–9109, 2013.

GAO, F. Bacteria may have multiple replication origins. **Frontiers in Microbiology**. v. 6, n. 324, 2015.

GAO, F.; ZHANG, C.-T. Ori-Finder: a web-based system for finding oriCs in unannotated bacterial genomes. **BMC Bioinformatics**. v. 9, n. 79, 2008.

GHATAK, S.; BLOM, J.; DAS, S.; SANJUKTA, R.; PURO, K.; MAWLONG, M.; SHAKUNTALA, I.; SEN, A.; GOESMANN, A.; KUMAR, A.; NGACHAN, S. V. Pan-genome analysis of *Aeromonas hydrophila*, *Aeromonas veronii* and *Aeromonas caviae* indicates phylogenomic diversity and greater pathogenic potential for *Aeromonas hydrophila*. **Antonie van Leeuwenhoek**. v. 109, p. 945–956, 2016.

GAVÍN, R.; RABAAN, A. A.; MERINO, S.; TOMÁS, J. M.; GRYLLOS, I.; SHAW, J. G. Lateral flagella of *Aeromonas* species are essential for epithelial cell adherence and biofilm formation. **Molecular Microbiology**. v. 43, n. 2, p. 383–397, 2002.

GAVÍN, R.; MERINO, S.; ALTARRIBA, M.; CANALS, R.; SHAW, J. G.; TOMÁS, J. M. Lateral flagella are required for increased cell adherence, invasion and biofilm formation by *Aeromonas* spp. **FEMS Microbiology Letters**. v. 224, p. 77–83, 2003

GIBOTTI, A.; SARIDAKIS, H. O.; PELAYO, J. S.; TAGLIARI, K. C.; FALCÃO, D P. Prevalence and virulence properties of *Vibrio cholerae* non-O1, *Aeromonas* spp. and *Plesiomonas shigelloides* isolated from Cambé Stream (State of Paraná, Brazil). **Journal of Applied Microbiology**. v. 89, p. 70–75, 2000.

GORIS, J.; KONSTANTINIDIS, K. T.; KLAPPENBACH, J. A.; COENYE, T.; VANDAMME, P.; TIEDJE, J. M. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. **International Journal of Systematic and Evolutionary Microbiology**. v. 57, p. 81–91, 2007.

Von GRAEVENITZ, A. The role of *Aeromonas* in diarrhea: a review. **Infection**. v. 35, p. 59–64 2007.

GRIM, C. J.; KOZLOVA, E. V.; SHA, J.; FITTS, E. C.; VAN LIER, C. J.; KIRTLEY, M. L.; JOSEPH, S. J.; READ, T. D.; BURD, E. M.; TALL, B. D.; JOSEPH, S. W.; HORNEMAN, A. J.; CHOPRA, A. K.; SHAK, J. R. Characterization of *Aeromonas hydrophila* wound pathotypes by comparative genomic and functional analyses of virulence genes. **mBio**. v. 4, n. 2, e00064-13, 2013.

GRIM, C. J.; KOZLOVA, E. V.; PONNUSAMY, D.; FITTS, E. C.; SHA, J.; KIRTLEY, M. L.; VAN LIER, C. J.; TINNER, B. L.; EROVA, T. E.; JOSEPH, S. J.; READ, T. D.; SHAK, J. R.; JOSEPH, S. W.; SINGLETARY, E.; FELLAND, T.; BAZE, W. B.; HORNEMAN, A. J.; CHOPRA, A. K. Functional genomic characterization of virulence factors from necrotizing fasciitis-causing strains of *Aeromonas hydrophila*. **Applied and Environmental Microbiology**. v. 80, n. 14, p. 4162–4183, 2014.

GUERRA, I. M. F.; FADANELLI, R.; FIGUEIRÓ, M.; SCHREINER, F.; DELAMARE, A. P. L.; WOLLHEIM, C.; COSTA, S. O. P.; ECHEVERRIGARAY, S. *Aeromonas* associated diarrhoeal disease in South Brazil: prevalence, virulence factors and antimicrobial resistance. **Brazilian Journal of Microbiology**. v. 38, p. 638-643, 2007.

HACKER, J.; KAPER, J. B. Pathogenicity islands and the evolution of microbes. **Annual Review of Microbiology**. v. 54, p. 641–79, 2000.

HADI, N.; YANG, Q.; BARNETT, T. C.; TABELI, S. M. B.; KIROV, S. M.; SHAW, J. G. Bundle-Forming Pilus Locus of *Aeromonas veronii* bv. *Sobria*. **Infection and Immunity**. v. 80, n. 4, p. 1351-1360, 2012.

HAN, J. E.; KIM, J. H.; CHERESCA JR., C. H.; SHIN, S. P.; JUN, J. W.; CHAI, J. Y.; HAN, S. Y.; PARK, S. C. First description of the qnrS-like (qnrS5) gene and analysis of quinolone resistance-determining regions in motile *Aeromonas* spp. from diseased fish and water. **Research in Microbiology**. v. 163, 73e79, 2012.

HAYES, C. S.; AOKI, S. K.; LOW, D. A. Bacterial contact-dependent delivery systems. **Annual Review of Genetics**. v. 44, p. 71–90, 2010.

HENGSTLER, K. A.; HAMMANN, R.; FAHR A.-M. Evaluation of BBL CHROMagar Orientation Medium for Detection and Presumptive Identification of Urinary Tract Pathogens. **Journal of Clinical Microbiology**. v. 35, n. 11, p. 2773–2777, 1997.

HENSEL, M. Evolution of pathogenicity islands of *Salmonella enterica*. **International Journal of Medical Microbiology**. v. 294, n. 2–3, p. 95–102, 2004.

HOFER, E.; dos REIS, C. M. F.; THEOPHILO, G. N. D.; CAVALCANTI, V. O.; de LIMA, N. V.; HENRIQUES, M. F. C. M. Envolvimento de *Aeromonas* em surto de doença diarreica aguda em São Bendo do Una, Pernambuco. **Revista da Sociedade Brasileira de Medicina Tropical**. v. 39, n. 2, p. 217–220, 2006.

HOLMES, P.; NICCOLLS, L. M.; SARTORY, D. P. The ecology of mesophilic *Aeromonas* in the aquatic environment, p. 127-150. In B. Austin, M. Altwegg, P. J. Gosling, and S. Joseph (ed.), *The genus Aeromonas*. John Wiley & Sons Ltd., West Sussex, England, 1996.

HUANG, H. -C.; YU, W. -L.; HUAN, K. -H.; CHENG, K. -C.; CHUANG, Y. -C. *Aeromonas sobria* prostatitis and septic shock in a healthy man with chronic alcoholic consumption. **Japanese Journal of Infectious Diseases**. v. 60, p. 400-401, 2007.

IGBINOSA, I. H.; IGUMBOR, E. U.; AGHDASI, F.; TOM, M.; OKOH, A. I. Emerging *Aeromonas* Species Infections and Their Significance in Public Health. **The Scientific World Journal**. v. 2012, Article ID 625023, p. 1–13, 2012.

INOUE, Y.; KIMURA, A. Methylglyoxal and regulation of its metabolism in microorganisms. **Advances in Microbial Physiology**. v. 37, p. 177-227, 1995.

JANDA, J. M.; ABBOTT, S. L. Evolving concepts regarding the genus *Aeromonas*: an expanding panorama of species, disease presentation, and unanswered questions. **Clinical Infectious Diseases**. v. 27, p. 332–344, 1998.

JANDA, J. M.; ABBOTT, S. L. The Genus *Aeromonas*: Taxonomy, Pathogenicity, and Infection. **Clinical Microbiology Reviews**. v. 23, n. 1, p. 35–73, 2010.

JONES, D. T.; TAYLOR, W. R.; THORNTON, J. M. The rapid generation of mutation data matrices from protein sequences. **Computer Applications in the Biosciences**. v. 8, n. 3, p. 275–282, 1992.

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.

KANNAN, S.; CHATTOPADHYAY, U. K.; PAL, D.; SHIMADA, T.; TAKEDA, Y.; BHATTACHARYA, S. K.; ANANTHANARAYANAN, P. H. Isolation and identification of *Aeromonas* from patients with acute diarrhoea in Kolkata, India. **Indian Journal of Medicinal Microbiology**. v. 19, n. 4, p. 190–192, 2001.

KAWECKI, T. J.; LENSKI, R. E.; EBERT, D.; HOLLIS, B.; OLIVIERI, I.; WHITLOCK, M. C. Experimental evolution. **Trends in Ecology and Evolution**. v. 27, n. 10, p. 547–560, 2012.

KELLY, K. A.; KOEHLER, J. M.; ASHDOWN, L. R. Spectrum of extraintestinal disease due to *Aeromonas* species in tropical Queensland, Australia. **Clinical Infectious Diseases**. v. 16, p. 574–579, 1993.

KELLY, P. Infectious diarrhea. **Medicine**. v. 39, n. 4, 201–206, 2011.

KHAJANCHI, B. K.; FADL, A. A.; BORCHARDT, M. A.; BERG, R. L.; HORNEMAN, A. J.; STEMPER, M. E.; JOSEPH, S. W.; MOYER, N. P.; SHA, J.; CHOPRA, A. K. Distribution of virulence factors and molecular fingerprinting of *Aeromonas* species isolates from water and clinical samples: suggestive evidence of water-to-human transmission. **Applied and Environmental Microbiology**. v. 76, n. 7, 2313–2325, 2010.

KHAJANCHI, B. K.; KOZLOVA, E. V.; SHA, J.; POPOV, V. L.; CHOPRA, A. K. The two-component QseBC signalling system regulates *in vitro* and *in vivo* virulence of *Aeromonas hydrophila*. **Microbiology**. v. 158, p. 259–271, 2012.

KIENZLE, N.; MULLER, M.; PEGG, S. *Aeromonas* wound infections in burns. **Burns**. v. 26, p. 478–482, 2000.

KIROV, S. M.; JACOBS, I.; HAYWARD, L. J.; HAPIN, R. H. “Electron microscopic examination of factors influencing the expression of filamentous surface structures on clinical and environmental isolates of *Aeromonas veronii* biotype sobria,” **Microbiology and Immunology**. v. 39, n. 5, p. 329–338, 1995.

KIROV, S. M.; O'DONOVAN, L. A.; SANDERSON, K. Functional Characterization of Type IV Pili Expressed on Diarrhea-Associated Isolates of *Aeromonas* species. **Infection and Immunity**. v. 67, n. 10, p. 5447–5454, 1999.

KIROV, S. M.; BARNETT, T. C.; PEPE, C. M.; STROM, M. S.; ALBERT, M. J. Investigation of the Role of Type IV *Aeromonas* Pilus (Tap) in the Pathogenesis of *Aeromonas* Gastrointestinal Infection. **Infection and immunity**. v. 68, n. 7, p. 4040–4048, 2000.

KIROV, S. M.; TASSELL, B. C.; SEMMLER, A. B. T.; O'DONOVAN, L. A.; RABAAN, A. A.; SHAW, J. G. Lateral Flagella and Swarming Motility in *Aeromonas* Species. **Journal of Bacteriology**. v. 184, n. 2, p. 547–555, 2002.

KIROV, S. M.; CASTRISIOS, M.; SHAW, J. G. *Aeromonas* Flagella (Polar and Lateral) Are Enterocyte Adhesins That Contribute to Biofilm Formation on Surfaces. **Infection and Immunity**. v. 72, n. 4, p. 1939–1945, 2004.

KONIECZNY, I.; DORAN, K. S.; HELINSKI, D. R.; BLASINA, A. Role of TrfA and DnaA Proteins in Origin Opening during Initiation of DNA Replication of the Broad Host Range Plasmid RK2. **The Journal of Biological Chemistry**. v. 272, n. 32, p. 20173–20178, 1997.

KONSTANTINIDIS, K. T. & TIEDJE, J. M. Genomic insights that advance the species definition for prokaryotes. **PNAS**. v. 102, n. 7, p. 2567–2572, 2005.

KREWULAK, K. D.; VOGEL, H. J. Structural biology of bacterial iron uptake. **Biochimica et Biophysica Acta**. n. 1778, p. 1781–1804, 2008.

KÜHN, I.; ALLESTAM, G.; HUYS, G.; JANSSEN, P.; KERSTERS, K; KROVACEK, K.; STENSTRÖM, T. A. Diversity, persistence and virulence of *Aeromonas* strains isolated from drinking water distribution systems in Sweden. **Applied and Environmental Microbiology**. v. 63, n. 7, p. 2708–2715, 1997.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. **Molecular Biology and Evolution**. v. 33, n. 7, p. 1870–1874, 2016.

KURTZ, S.; PHILLIPPY, A.; DELCHER, A. L.; SMOOT, M.; SHUMWAY, M.; ANTONESCU, C.; SALZBERG, S. L. Versatile and open software for comparing large genomes. **Genome Biology**. v. 5, n. 2, p. R12, 2004.

KWONG, J. C.; MCCALLUM, N.; SINTCHENKO, V.; HOWDEN, B. P. Whole genome sequencing in clinical and public health microbiology. **Pathology**. v. 47, n. 3, p. 199–210, 2015.

LAGESEN, K.; HALLIN, P.; RØDLAND, E. A.; STÆRFELDT, H.-H.; ROGNES, T.; USSERY, D. W. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. **Nucleic Acids Research**. v. 35, n. 9, p. 3100–3108, 2007.

(a) LAI, C. -C.; DING, L. -W.; HSUEH, P. -R. Wound infection and septic shock due to *Aeromonas trota* in a patient with liver disease. **Clinical Infectious Diseases**. v. 44, p. 1523–1524, 2007.

(b) LAI, C. -C.; SHIAO, C. -C.; LU, G. -D.; DING, L. -W. *Aeromonas hydrophila* and *Aeromonas sobria* bacteremia: rare pathogens of infection in a burn patient. **Burns**. v. 33, p. 255–257, 2007.

LAMY, B.; KODJO, A.; THE colBVH STUDY GROUP; LAURENT, F. Prospective nationwide study of *Aeromonas* infections in France. **Journal of Clinical Microbiology**. v. 47, p. 1234–1237, 2009.

LI, Y.; LIU, Y.; ZHOU, Z.; HUANG, H.; REN, Y.; ZHANG, Y.; LI, G.; ZHOU, Z.; WANG, L. Complete Genome Sequence of *Aeromonas veronii* Strain B565. **Journal of Bacteriology**. v. 193, n. 13, p. 3389–3390, 2011.

LI, W.; COWLEY, A.; ULUDAG, M.; GUR, T.; MCWILLIAM, H.; SQUIZZATO, S.; PARK, Y. M.; BUSO, N.; LOPEZ, R. The EMBL-EBI bioinformatics web and programmatic tools framework. **Nucleic Acids Research**. v. 43(Web Server issue), W580–W584, 2015.

LIM, Y.-L.; ROBERTS, R. J.; EE, R.; YIN, W.-F.; CHAN, K.-G. Complete genome sequence and methylome analysis of *Aeromonas hydrophila* Strain YL17, isolated from a compost pile. **Genome Announcements**. v. 4, n. 2, e00060–16, 2016.

LINDENBERG, S.; KLAUCK, G.; PESAVENTO, C.; KLAUCK, E. HENGGE, R. The EAL domain protein YciR acts as a trigger enzyme in a c-di-GMP signalling cascade in *E. coli* biofilm control. **The EMBO Journal**. v. 32, p. 2001–2014, 2013.

LINDSAY, J. A. Chronic Sequelae of Foodborne Disease. **Emerging Infectious Diseases**. v. 3, n. 4, p. 443-452, 1997.

LINHARTOVÁ, I.; BUMBA, L.; MAŠÍN, J.; BASLER, M.; OSIČKA, R.; KAMANOVÁ, J.; PROCHÁZKOVÁ, K.; ADKINS, I.; HEJNOVÁ-HOLUBOVÁ, J.; SADÍLKOVÁ, L.; MOROVÁ, J.; ŠEBO, P. RTX proteins: a highly diverse family secreted by a common mechanism. **FEMS Microbiology Reviews**. v. 34, p. 1076–1112, 2010.

LLOPIS, F.; GRAU, I.; TUBAU, F.; CISNAL, M.; PALLARES, R. Epidemiological and clinical characteristics of bacteraemia caused by *Aeromonas* spp. as compared with *Escherichia coli* and *Pseudomonas aeruginosa*. **Scandinavian Journal of Infectious Diseases**. v. 36, p. 335-341, 2004.

LÖYTYNOJA, A. Phylogeny-aware alignment with PRANK. **Methods in Molecular Biology**. v. 1079, p. 155-70, 2014.

LOWRY, R.; BALBOA, S.; PARKER, J. L.; SHAW, J. G. *Aeromonas* flagella and colonisation mechanisms. **Advances in Microbial Physiology**. v. 65, p. 203–256, 2014.

LUCIANI, F.; BULL, R. A.; LLOYD, A. R. Next generation deep sequencing and vaccine design: today and tomorrow. **Trends in biotechnology**. v. 30, n. 9, p. 443–452, 2012.

MARLIÈRE, P.; PATROUX, J.; DÖRING, V.; HERDEWIJN, P.; TRICOT, S.; CRUVEILLER, S.; BOUZON, M.; MUTZEL, R. Chemical Evolution of a Bacterium's Genome. **Angewandte Chemie International Edition**. v. 123, p. 1–7, 2011.

MARTIN-CARNAHAN, A. & JOSEPH, S. W. *Aeromonas*. In: *Bergey's Manual of Systematic Bacteriology, Second Edition*, Brenner, Krieg, Staley, Garrity (Eds), Williams and Wilkins, New York 2005. Vol 2.

MARTINEZ-MURCIA, A. J.; MONERA, A.; SAAVEDRA, M. J.; ONCINA, R.; LOPEZ-ALVAREZ, M.; LARA, E.; FIGUERAS, M. J. Multilocus phylogenetic analysis of the genus *Aeromonas*. **Systematic and Applied Microbiology**. 34:189–199, 2011.

MARTINS, M. L.; MIYAZAKI, D. M. Y.; MOURIÑO, J. L. P. *Aeromonas caviae* durante surto de mortalidade em tilápia do nilo e suplementação com vitamina C na dieta. **Boletim do Instituto de Pesca**. v. 34, n. 4, p. 585-590, 2008.

MASAI, H.; ARAI, K. Operon Structure of *dnaT* and *dnaC* Genes Essential for Normal and Stable DNA Replication of *Escherichia coli* Chromosome. **The Journal of Biological Chemistry**. v. 263, n. 29, p. 15083-15093, 1988.

MATTHIES, S.; STALLFORTH, P.; SEEBERGER, P. H. Total Synthesis of Legionaminic Acid as Basis for Serological Studies. **Journal of the American Chemical Society**. v. 137, p. 2848–2851, 2015.

MCMAHON, M. A. S. & WILSON, I. G. The occurrence of enteric pathogens and *Aeromonas* species in organic vegetables. **International Journal of Food Microbiology**, v. 70, p. 155–162, 2001.

MCWILLIAM, H.; LI, W.; ULUDAG, M.; SQUIZZATO, S.; PARK, Y. M.; BUSO, N.; COWLEY, A. P.; LOPEZ, R. Analysis Tool Web Services from the EMBL-EBI. **Nucleic Acids Research**. v. 41, W597–W600, 2013.

MEIER-KOLTHOFF, J. P.; AUCH, A. F.; KLENK, H. -P.; GÖKER, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. **BMC Bioinformatics**. v. 14, n. 60, 2013.

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

MERINO, S.; RUBIRES, X.; KNØCHEL, S.; TOMÁS, J. M. Emerging pathogens: *Aeromonas* spp. **International Journal of Food Microbiology**. v. 28, n. 2, p. 157–168, 1995.

MICHEL, G. P. F. & VOULHOUX, R. Type II Secretory System (T2SS) in Gram-negative Bacteria: a Molecular Nanomachine for Secretion of Sec and Tat-dependent Extracellular Proteins. *Bacterial Secreted Proteins: Secretory Mechanisms and Role in Pathogenesis*. Caister Academic Press: Norfolk, UK. 2009.

MORANDI, A.; ZHAXYBAYEVA, O.; GOGARTEN, J. P.; GRAF, J. Evolutionary and diagnostic implications of intragenomic heterogeneity in the 16S rRNA gene in *Aeromonas* strains. **Journal of Bacteriology**. v. 187, p. 6561–6564, 2005.

MORINAGA, Y.; YANAGIHARA, K.; EUGENIN, F. L.; BEAZ-HIDALGO, R.; KOHNO, S.; FIGUERAS SALVAT M. J. Identification error of *Aeromonas aquariorum*: A causative agent of septicemia. **Diagnostic Microbiology and Infectious Disease**. v. 76, n. 1, p. 106–109, 2013.

MORRIS, J. G. & HORNEMAN, A. *Aeromonas* infections – UpToDate Disponível em < <https://www.uptodate.com/contents/aeromonas-infections>>. Acesso em 1 maio 2017.

MUKHOPADHYAY, C.; CHAWLA, K.; SHARMA, Y.; BAIRY, I. Emerging extra-intestinal infections with *Aeromonas hydrophila* in coastal region of southern Karnataka. **Journal of Postgraduate Medicine**. v. 54, p. 199–202, 2008.

NCBI RESOURCE COORDINATORS. Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**. v. 44, 2016.

NEYTS, K.; HUYS, G.; UYTENDAELE, M.; SWINGS, J.; DEBEVERE, J. Incidence and identification of mesophilic *Aeromonas* spp. from retail foods. **Letters in Applied Microbiology**. v. 31, p. 359–363, 2000.

NIKAIDO, H.; TAKATSUKA, Y. Mechanisms of RND Multidrug Efflux Pumps. **Biochimica et Biophysica Acta**. v. 1794, n. 5, p. 769–781, 2009.

NOJIMOTO, I. T. I.; BEZANA, C. S. C.; CARMO, C.; VALADÃO, L. M.; CARRIJO, K.M. Prevalência de *Aeromonas* spp em fezes diarréicas de crianças menores de 5 anos de idade na cidade de Goiânia, Goiás, no biênio 1995-1996. **Revista da Sociedade Brasileira de Medicina Tropical**. v. 30, p.385–388, 1997.

O'LOUGHLIN, E. V.; SCOTT, R. B.; GALL, D. G. Pathophysiology of infectious diarrhea: changes in intestinal structure and function. **Journal of Pediatric Gastroenterology and Nutrition**. v. 12, n. 1, p. 05-20, 1991.

OUDERKIRK, J.P.; BEKHOR, D.; TURETT, G.S.; MURALI, R. *Aeromonas* meningitis complicating medicinal leech therapy. **Clinical Infectious Diseases**. v.38, p.36–37, 2004.

PAGE, A. J.; CUMMINS, C. A.; HUNT, M.; WONG, V. K.; REUTER, D.; HOLDEN, M. T. G.; FOOKES, M.; FALUSH, D.; KEANE, J. A.; PARKHILL, J. Roary: Rapid large-scale prokaryote pan genome analysis. **Bioinformatics**. v. 31, n. 22, p. 3691–3693, 2015.

PAMPÍN, F.; BOUB, G.; GALEIRASA, R.; FREIREA, D.; BOUZAA, M. T.; ZÚÑIGA, M. DEL C. *Aeromonas* y meningitis: una presentación infrecuente. **Neurocirugia**. v. 23, n. 5, p. 200–202, 2012.

PANG, M.; JIANG, J.; XIE, X.; WU, Y.; DONG, Y.; KWOK, A. H. Y.; ZHANG, W.; YAO, H.; LU, C.; LEUNG, F. C.; LIU, Y. Novel insights into the pathogenicity of epidemic *Aeromonas hydrophila* ST251 clones from comparative genomics. **Scientific Reports**. v. 5, 09833, 2015.

PARKER, J. L. & SHAW J. G. *Aeromonas* spp. *Clinical Microbiology and Disease*. **Journal of Infection**. p. 1–10. 2010.

PEREIRA, C. S.; AMORIM, S. D.; SANTOS, A. F. M.; dos REIS, C M. F.; THEOPHILO, G. N. D.; RODRIGUES, D. P. Caracterização de *Aeromonas* spp isoladas de neonatos hospitalizados. **Revista da Sociedade Brasileira de Medicina Tropical**. v. 41, n. 2, p. 179–182, 2008.

PICHOFF, S.; LUTKENHAUS, J. Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. **The EMBO Journal**. v. 21, n. 4, p. 685–693, 2002.

PIRO, V. C.; FAORO, H.; WEISS, V. A.; STEFFENS, M. B. R.; PEDROSA, F. O.; SOUZA, E. M.; RAITTZ, R. T. FGAP: an automated gap closing tool. **BMC Research Notes**. v. 7, n. 371, 2014.

PRICE, M. N.; DEHAL, P. S.; ARKIN, A. P. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. **PLoS ONE**. v. 5, n. 3, e9490, 2010.

PRIDGEON, J. W.; ZHANG, D.; ZHANG, L. Complete genome sequence of a moderately virulent *Aeromonas hydrophila* strain, pc104A, isolated from soil of a catfish pond in west Alabama. **Genome Announcements**. v. 2, n. 3, e00554-14, 2014.

PROFT, T. & BAKER, E. N. Pili in Gram-negative and Gram-positive bacteria – structure, assembly and their role in disease. **Cellular and Molecular Life Sciences**. v. 66, n. 4, p. 613–635, 2009.

RASMUSSEN-IVEY, C. R.; FIGUERAS, M. J.; MCGAREY, D.; LILES, M. R. Virulence Factors of *Aeromonas hydrophila*: In the Wake of Reclassification. **Frontiers in Microbiology**. v. 7, n. 1337, 2016.

REITH, M. E.; SINGH, R. K.; CURTIS, B.; BOYD, J. M.; BOUEVITCH, A.; KIMBALL, J.; MUNHOLLAND, J.; MURPHY, C.; SARTY, D.; WILLIAMS, J.; NASH, J. H. E.; JOHNSON, S. C.; BROWN, L. L. The genome of *Aeromonas salmonicida* subsp. *salmonicida* A449: insights into the evolution of a fish pathogen. **BMC Genomics**. v. 9, n. 427, 2008.

RENO, M. & HELD, N. Biogeography of the *Sulfolobus islandicus* pan-genome. **PNAS**. v. 106, n. 21, 2009.

RHOADS, A. & AU, K. F. PacBio Sequencing and Its Applications. **Genomics Proteomics Bioinformatics**. v. 13, p. 278–289, 2015.

RIBER, L.; LØBNER-OLESEN, A. Coordinated Replication and Sequestration of *oriC* and *dnaA* Are Required for Maintaining Controlled Once-per-Cell-Cycle Initiation in *Escherichia coli*. **Journal of Bacteriology**. v. 187, n. 16, p. 5605–5613, 2005.

RICHAUD, C.; MENGIN-LECREULX, D.; POCHET, S.; JOHNSON, E. J.; COHEN, G. N.; MARLIÈRE, P. **The Journal of Biological Chemistry**. v. 268, p. 26827–26835, 1993.

RICHTER, M.; ROSSELLÓ-MÓRA, R. Shifting the genomic gold standard for the prokaryotic species definition. **PNAS**. v. 106, n. 45, p. 19126–19131, 2009.

RHOADS, A.; AU, K. F. PacBio Sequencing and Its Applications. **Genomics Proteomics Bioinformatics**. v. 13, p. 278–289, 2015.

ROGER, F.; LAMY, B.; JUMAS-BILAK, E.; KODJO, A; THE COLBVH STUDY GROUP, *ET AL.* Ribosomal Multi-Operon Diversity: An Original Perspective on the Genus *Aeromonas*. **PLoS ONE**. v. 7, n. 9, e. 46268, 2012.

ROGERS, Y.-H.; VENTER, C. Massively parallel sequencing. **Nature**. v. 437, p. 326–327, 2005.

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

RONDELET, A.; CONDEMINÉ, G. Type II secretion: the substrates that won't go away. **Research in Microbiology**. v. 164, p. 556-561, 2013.

ROSENZWEIG, J. A.; CHOPRA, A. K. Modulation of host immune defenses by *Aeromonas* and *Yersinia* species: convergence on toxins secreted by various secretion systems. **Frontiers in Cellular and Infection Microbiology**. v. 3, n. 70, 2013.

SAAVEDRA, M. J.; FIGUERAS, M. J.; MARTÍNEZ-MURCIA, A. J. Updated phylogeny of the genus *Aeromonas*. **International Journal of Systematic and Evolutionary Microbiology**. v. 56, p. 2481–2487, 2006.

SANGLAS, A.; ALBARRAL, V.; FARFÁN, M.; LORÉN, J. G.; FUSTÉ, M. C. Evolutionary Roots and Diversification of the Genus *Aeromonas*. **Frontiers in Microbiology**. v. 8, n. 127, 2017.

SCHADT, E. E.; TURNER, S.; KASARSKIS, A. A window into third-generation sequencing. **Human Molecular Genetics**. v. 19, n. R2, R227-R240, 2010.

SCHMIDT, A. J.; RYJENKOV, D. A.; GOMELSKY, M. The Ubiquitous Protein Domain EAL Is a Cyclic Diguanylate-Specific Phosphodiesterase: Enzymatically Active and Inactive EAL Domains. **Journal of Bacteriology**. v. 187, v. 14, p. 4774–4781, 2005.

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

SERNOVA, N. V.; GELFAND, M. S. Identification of replication origins in prokaryotic genomes. **Briefings in Bioinformatics**. v. 9, n. 5, p. 376–391, 2008.

SESHADRI, R.; JOSEPH, S. W.; CHOPRA, A. K.; SHA, J.; SHAW, J.; GRAF, J.; HAFT, D.; WU, M.; REN, Q.; ROISOVITZ, M. J.; MADUPU, R.; TALLON, L.; KIM, M.; JIN, S.; VUONG, H.; STINE, O. C.; ALI, A.; HORNEMAN, A. J.; HEIDELBERG, J. F. Genome Sequence of *Aeromonas hydrophila* ATCC 7966T: Jack of All Trades. **Journal of Bacteriology**. v. 188, n. 23, p. 8272–8282, 2006.

SHA, J.; KOZLOVA, E. V.; CHOPRA, A. K. Role of Various Enterotoxins in *Aeromonas hydrophila*-Induced Gastroenteritis: Generation of Enterotoxin Gene-Deficient Mutants and Evaluation of Their Enterotoxic Activity. **Infection and Immunity**. v. 70, n. 4, p. 1924–1935, 2002.

SHA, J.; ROSENZWEIG, J. A.; KOZLOVA, E. V.; WANG, S.; EROVA, T. E.; KIRTLEY, M. L.; VAN LIER, C. J.; CHOPRA, A. K. Evaluation of the roles played by Hcp and VgrG type 6 secretion system effectors in *Aeromonas hydrophila* SSU pathogenesis. **Microbiology**. v. 159, p. 1120–1135, 2013.

SHAKIR, Z.; KHAN, S.; SUNG, K.; KHARE, S.; KHAN, A.; STEELE, R.; NAWAZ, M. Molecular characterization of fluoroquinolone-resistant *Aeromonas* spp. isolated from imported shrimp. **Applied and Environmental Microbiology**. v. 78, n. 22, p. 8137–8141, 2012.

SHALOM, G.; SHAW, J. G.; THOMAS, M. S. *In vivo* expression technology identifies a type VI secretion system locus in *Burkholderia pseudomallei* that is induced upon invasion of macrophages. **Microbiology**. v. 153, p. 2689–2699, 2007.

SIEVERS, F.; WILM, A.; DINEEN, D.; GIBSON, T. J.; KARPLUS, K.; LI, W.; LOPEZ, R.; MCWILLIAM, H.; REMMERT, M.; SÖDING, J.; THOMPSON, J. D.; HIGGINS, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**. v. 7, n. 539, 2011.

SILVERMAN, J. M.; BRUNET, Y. R.; CASCALES, E.; MOUGOUS, J. D. Structure and Regulation of the Type VI Secretion System. **Annual Review of Microbiology**. v. 66, p. 453–472, 2012.

SOARES, S. C.; GEYIK, H.; RAMOS, R. T. J.; DE SÁ, P. H. C. G.; BARBOSA, E. G. V.; BAUMBACH, J.; FIGUEIREDO, H. C. P.; MIYOSHI, A.; TAUCH, A.; SILVA, A.; AZEVEDO, V GIPSy: Genomic island prediction software. **Journal of Biotechnology**. v. 232, p. 2–11, 2016.

SOLER, L.; YÁÑES, M. A.; CHACON, M. R.; AGUILERA-ARREOLA, M. G.; CATALÁN, V.; FIGUERAS, M. J.; MARTÍNEZ-MURCIA, A. J. Phylogenetic analysis of the genus *Aeromonas* based on two housekeeping genes. **International Journal of Systematic and Evolutionary Microbiology**. v. 54, p. 1511–1519, 2004.

SOOD, S. & NERURKAR, V. Fatal Necrotizing Soft Tissue Infection by *Aeromonas hydrophila*. **Journal of Clinical and Diagnostic Research**. v. 8, n. 4, p. 6–7, 2014.

SORGE, I.; CORNELIS, G. R. The Type III Secretion System. Bacterial Secreted Proteins: Secretory Mechanisms and Role in Pathogenesis. Caister Academic Press: Norfolk, UK. 2009.

SUAREZ, G.; SIERRA, J. C.; SHA, J.; WANG, S.; EROVA, T. E.; FADL, A. A.; FOLTZ, S. M.; HORNEMAN, A. J.; CHOPRA, A. K. Molecular characterization of a functional type VI secretion system from a clinical isolate of *Aeromonas hydrophila*. **Microbial Pathogenesis**. v. 44, p. 344–361, 2008.

(a) SUAREZ, G.; SIERRA, J. C.; EROVA, T. E.; SHA, J.; HORNEMAN, A. J.; CHOPRA, A. K. A Type VI Secretion system effector protein, VgrG1, from *Aeromonas hydrophila* That induces host cell toxicity by ADP ribosylation of actin. **Journal of Bacteriology**. v. 192, n. 1, p. 155–168, 2010.

(b) SUAREZ, G.; SIERRA, J. C.; KIRTLEY, M. L.; CHOPRA, A. K. Role of Hcp, a type 6 secretion system effector, of *Aeromonas hydrophila* in modulating activation of host immune cells. **Microbiology**. v. 156, n. 12, p. 3678–3688, 2010.

SUAREZ, G.; KHAJANCHI, B. K.; SIERRA, J. C.; EROVA, T. E.; SHA, J.; CHOPRA, A. K. Actin cross-linking domain of *Aeromonas hydrophila* repeat in toxin A (RtxA) induces host cell rounding and apoptosis. **Gene**. v. 506, n. 2, p. 369–376, 2012.

SUREK, M.; VIZZOTTO, B. S.; SOUZA, E. M.; PEDROSA, F. O.; DALLAGASSA, C. B.; FARAH, S. M. S. S.; FADEL-PICHETH, C. M. T. Identification and antimicrobial susceptibility of *Aeromonas* spp. isolated from stool samples of Brazilian subjects with diarrhoea and healthy controls. **Journal of Medical Microbiology**. v. 59, p. 373–374, 2010.

SWAMINATHAN, B. et al. In: Murray, P.R. et al. Manual of Clinical Microbiology. 7.ed. p.174–190, ASM Press, 1999.

SZYMANSKI, M.; BARCISZEWSKA, M. Z.; ERDMANN, V. A.; BARCISZEWSKI, J. 5S Ribosomal RNA Database. **Nucleic Acids Research**. v. 30, 2002.

TALAVERA, G.; CASTRESANA, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. **Systematic Biology**. v. 56, p. 564–577, 2007.

TAUSCHEK, M.; GORRELL, R. J.; STRUGNELL, R. A.; ROBINS-BROWNE, R. M. Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of *Escherichia coli*. **PNAS**. v. 99, n. 10, p. 7066–7071, 2002.

TEKEDAR, H. C.; KARSI, A.; AKGUL, A.; KALINDAMAR, S.; WALDBIESER, G. C.; SONSTEGARD, T.; SCHROEDER, S. G.; LAWRENCE, M. L. Complete Genome Sequence of Fish Pathogen *Aeromonas hydrophila* AL06-06. **Genome Announcements**. v. 3, n. 2, e00368–15, 2015.

TENA, D.; GONZALEZ-PRAETORIUS, A.; GUMENO, C.; PEREZ-POMATA, MT.; BISQUERT, J. Infeccion extraintestinal por *Aeromonas* spp.: revision de 38 casos. **Enfermedades Infecciosas y Microbiología Clínica**. v. 25, n. 4, p. 235–241, 2007.

TOMÁS, J. M. The Main *Aeromonas* Pathogenic Factors. **ISRN Microbiology**. v. 2012, Article ID 256261, 22 pages, 2012.

van TONDER, A. J.; MISTRY, S.; BRAY, J. E.; HILL, D. M. C.; CODY, A. J.; FARMER, C. L.; KLUGMAN, K. P.; VON GOTTBURG, A.; BENTLEY, S. D.; PARKHILL, J.; JOLLEY, K. A.; MAIDEN, M. C. J.; BRUEGGEMANN, A. B. Defining the Estimated Core Genome of Bacterial Populations Using a Bayesian Decision Model. **PLOS Computational Biology**. v. 10, n. 8, e. 1003788, 2014.

TSAI, M.-S.; KUO, C.-Y.; WANG, M.-C.; WU, H.-C.; CHIEN, C.-C.; LIU, J.-W. Clinical features and risk factors for mortality in *Aeromonas* bacteremic adults with hematologic malignancies. **Journal of Microbiology, Immunology and Infection**. v. 39, p. 150–154, 2006.

TSENG, T.-T.; TYLER, B. M.; SETUBAL, J. C. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. **BMC Microbiology**. v. 9, n.1, S2, 2009.

TSENG et al., 2013 – Tsenga, C.-C.; Wub, C.-D.; Chengb, S.-C.; Linb, W.-T.; Chana, H.-T.; Chen, P.-Y. Septic arthritis caused by *Aeromonas hydrophila*. **Formosan Journal of Musculoskeletal Disorders**. v. 4, n. 2, p. 53–55, 2013.

VILA, J.; RUIZ, J.; GALLARDO, J.; VARGAS, M.; SOLER, L.; FIGUERAS, M.J.; GASCÓN, J. *Aeromonas* spp. and traveler's diarrhea: clinical features and antimicrobial resistance. **Emerging Infectious Diseases**. v. 9, p. 552-555, 2003.

VILCHES, S.; URGELL, C.; MERINO, S.; CHACÓN, M. R.; SOLER, L.; CASTRO-ESCARPULLI, G.; FIGUERAS, M. J.; TOMÁS, J. M. Complete Type III Secretion System of a Mesophilic *Aeromonas hydrophila* Strain. **Applied and Environmental Microbiology**. v. 70, n. 11, p. 6914–6919, 2004.

WANDERSMAN, C. Concluding remarks on the special issue dedicated to bacterial secretion systems: function and structural biology. **Research in Microbiology**. v. 164, p. 683–687, 2013.

WANG, G.; CLARK, C. G.; LIU, C.; PUCKNELL, C.; MUNRO, C. K.; KRUK, T. M. A. C.; CALDEIRA, R.; WOODWARD, D. L.; RODGERS, F. G. Detection and Characterization of the Hemolysin Genes in *Aeromonas hydrophila* and *Aeromonas sobria* by Multiplex PCR. **Journal of Clinical Microbiology**. v. 41, n. 3, p. 1048–1054, 2003.

WARGO, M. J. Homeostasis and Catabolism of Choline and Glycine Betaine: Lessons from *Pseudomonas aeruginosa*. **Applied and Environmental Microbiology**. v. 79, n. 7, p. 2112–2120, 2013.

WILCOX, R. A.; CHIN, G. K.; SEGASOTHY, M. *Aeromonas hydrophila* infection secondary to an electrical burn. **Medical Journal of Australia**. v. 173, p. 219-220, 2000.

WOLF, S. Características de virulência em estirpes de *Aeromonas* spp. Dissertação de Mestrado. Programa de Pós-Graduação em Ciências Farmacêuticas. Universidade Federal do Paraná. 75 p., 2011.

WU, C.-J.; WU, J.-J.; YAN, J.-J.; LEE, H.-C.; LEE, N.-Y.; CHANG, C.-M.; SHIH, H.-I.; WU, H.-M.; WANG, L.-R.; KO, W.-C. Clinical significance and distribution of putative virulence markers of 116 consecutive clinical *Aeromonas* isolates in southern Taiwan. **Journal of Infection**. v. 54, p. 151e158, 2007.

WUYTS, J.; PERRIERE, G.; VAN DE PEER, Y. The European ribosomal RNA database. **Nucleic Acids Research**. v. 32, Database issue, D101–D103, 2004.

WYCKOFF, E. E.; MEY, A. R.; LEIMBACH, A.; FISHER, C. F.; PAYNE, S. M. Characterization of Ferric and Ferrous Iron Transport Systems in *Vibrio cholerae*. **Journal of Bacteriology**. v. 188, n. 18, p. 6515–6523, 2006.

XU, X.-J.; FERGUSON, M. R.; POPOV, V. L.; HOUSTON, C. W.; PETERSON, J. W.; CHOPRA, A. K. Role of a cytotoxic enterotoxin in *Aeromonas*-mediated infections: development of transposon and isogenic mutants. **Infection and Immunity**. v. 66, n. 8, p. 3501–3509, 1998.

YÁÑEZ, M. A.; CATALÁN, V.; APRÁIZ, D.; FIGUERAS, M. J.; MARTÍNEZ-MURCIA, A. J. Phylogenetic analysis of members of the genus *Aeromonas* based on *gyrB* gene sequences. **International Journal of Systematic and Evolutionary Microbiology**. v. 53, p. 875–883, 2003.

YU, H. B.; ZHANG, Y. L.; LAU, Y. L.; YAO, F.; VILCHES, S.; MERINO, S.; TOMAS, J. M.; HOWARD, S. P.; LEUNG, K. Y. Identification and Characterization of Putative Virulence Genes and Gene Clusters in *Aeromonas hydrophila* PPD134/91. **Applied and Environmental Microbiology**. v. 71, n. 8, p. 4469–4477, 2005.

ZHANG, C. T.; GAO, F.; ZHANG, R. Segmentation algorithm for DNA sequences. **Physical Review**. v. 72, e. 041917, p. 1-6, 2005.

ANEXO A – GENOMAS DE *AEROMONAS* DEPOSITADOS NO NCBI

ANEXO A – GENOMAS DE AEROMONAS DEPOSITADOS NO CBNCI

TABELA A1 – GENOMAS COMPLETOS

Especie	Entrada	Foto	Tipos de fonte	Tecnologia de sequenciamento	Tamanho (Mb)	Sexo	Scfoffords	GC%	Cobertura	NSO	rRNA	SRNA	Outras rRNA	Proteínas	Genes	Pseudogenes	Biofilme	BioProjct	Assembly	Cromossomo	INSDC	Fluenciômicos	INSDC	WGS	Data de sequenciamento	Data de identificação			
A. caviae	A. caviae	FDAR002_72	stool	clinical	PacBio	4.81762	Contig	1	61.70	21.9	4,017,618	31	102	4	2042	4077	276	SAMN0204520	PRJNA227021	GCA_000870773.1				JTB001	2014/11/14	2015/02/27			
	A. caviae	FDAR002_76	stool	clinical	PacBio	4.65521	Contig	1	61.60	16.63	4,655,013	31	103	1	2674	4162	307	SAMN0204527	PRJNA227021	GCA_000870776.1				JTB001	2014/11/14	2015/02/27			
	A. hydrophila	4H41	environmental	fish	454 / Illumina	4.52769	Completed	1	62.00	74.4	4,527,690	29	104	1	3977	4074	23	SAMN0204150	PRJNA120524	GCA_000512163.1		CP066761.1			2013/10/20	2015/07/31			
	A. hydrophila	grass carp	stool	environment	fish	4.93027	Completed	1	61.10	41.00	4,926,965	33	126	1	4197	4074	23	SAMN0241555	PRJNA276509	GCA_000903645.1					2015/03/26	2015/07/31			
	A. hydrophila	AL06_06	gill/skin	fish	illumina	4.90159	Completed	1	61.37	48.88	4,901,593	31	112	1	4270	4447	33	SAMN0206653	PRJNA270827	GCA_000943015.1	NZ_CP010447.1			CP100447.1 NZ_CP010448.1 NZ_CP010449.1 NZ_CP010450.1	CP100448.1 CP100449.1 CP100450.1		2015/03/04	2015/07/19	
	A. hydrophila	AL06_71	catfish	fish	Illumina	5.02386	Completed	1	60.80	47.9	5,023,861	31	111	1	4306	4066	56	SAMN0272600	PRJNA272037	GCA_000931751.1				CP075861.1			2014/04/18	2015/07/31	
A. hydrophila	ML019_19	catfish	fish	454 GS FLX	5.02045	Completed	1	60.80	259	5,020,451	31	110	1	4274	4462	65	SAMN0205811	PRJNA227042	GCA_000819523.1				CP066863.1			2015/07/16	2015/07/31		
	ML019_19	catfish	fish	454 GS FLX Illumina	5.02045	Completed	1	60.80	30.7	5,020,300	31	111	1	4210	4501	47	SAMN0205811	PRJNA227042	GCA_000819523.1				CP066863.1			2015/05/24	2015/07/31		
	A. hydrophila	ye10A4	catfish pond	fish	Illumina	5.02263	Completed	1	60.80	46.4	5,022,630	31	111	1	4238	4241	68	SAMN0271403	PRJNA270203	GCA_000918663.1				CP068701.1			2015/04/03	2015/07/31	
	A. hydrophila	ye10A4	catfish pond	fish	Illumina	5.02263	Completed	1	60.80	46.4	5,022,630	31	111	1	4238	4241	65	SAMN0270409	PRJNA270203	GCA_000905955.1				CP067576.1			2015/04/02	2015/07/31	
	A. hydrophila	YL17	compost	environment	PacBio RSE	4.83627	Completed	1	61.80	174.34	4,836,271	31	125	1	4183	4481	41	SAMN0271348	PRJNA234373	GCA_000913075.1				CP076151.1			2014/04/02	2015/07/31	
	A. hydrophila subsp. hydrophila	ATCC 7966	compost	environment	PacBio RSE	4.74445	Completed	1	61.50	4.744	4,744,448	31	128	1	4119	4284	7	SAMN0200452	PRJNA268617	GCA_000814851.1				CP004625.1			2014/06/16	2014/12/22	
	A. media	A. media	WS	water sample	environment	ABI 3730XL / Roche 454 GS FLX / PacBio RSE II	4.76843	Completed	2	60.60	210	4,768,430	30	126	1	4385	4658	28	SAMN0247129	PRJNA170164	GCA_000287151.3				CP075671.1 NZ_CP007058.1	CP075661.1		2014/04/22	2015/07/31
		A. salmonicida subsp. salmonicida	A. salmonicida subsp. salmonicida	AA49	brown trout	fish	pac	5.04054	Completed	6	58.17	6	5,040,541	31	109	9	3981	4320	193	SAMN0200364	PRJNA268631	GCA_000196356.1	NZ_CP000481.1			NC_004925.1 NC_004946.1 NC_004923.1 NC_004924.1 NC_004926.1	AY010364.1 AY010365.1 AY010366.1 AY010367.1		2002/04/06
	A. veronii	A. veronii	B565	aquaculture pond sediment	environment	454 GS FLX / Illumina	4.55178	Completed	1	58.70	1805	4,551,783	31	101	1	3650	4100	17	SAMN0200340	PRJNA268623	GCA_000204151.1	NZ_CP15424.1			CP00267.1			2011/04/13	2015/07/30

TABELA A2 – DRAFT GENOMAS

Especie	Entrada	Foto	Tipos de fonte	Tecnologia de sequenciamento	Tamanho (Mb)	Sexo	Scfoffords	GC%	Cobertura	NSO	rRNA	SRNA	Outras rRNA	Proteínas	Genes	Pseudogenes	Biofilme	BioProjct	Assembly	Cromossomo	INSDC	Fluenciômicos	INSDC	WGS	Data de sequenciamento	Data de identificação					
A. alboscarchophila	A. alboscarchophila	BH-88	human blood	clinical	4.70589	Contig	131	58.40	72.44	4.705.89				1110	4230		SAMEA2752403	PRJEB1343	GCA_000822341.1							CCDB01	2014/11/24	2015/09/16			
	A. alboscarchophila	CECT 4169	soil	clinical	4.63639	Contig	30	58.40	114.54	4.636.39				4236	4276		SAMEA2752403	PRJEB1343	GCA_000819285.1							CCDB01	2014/11/24	2015/09/16			
A. aquatica	A. aquatica	AE235	reservoir lake	environment	Illumina MiSeq	4.58203	Contig	171	61.20	25	67,531	4	80	2	3963	4163	115	SAMN0200375	PRJNA24116	GCA_000794655.1						JR0010	2014/10/08	2014/10/14			
	A. aquatica	CECT 8023	irrigation water	environment	Illumina	4.11003	Contig	113	58.10	96.005	6	80	1	3070	3735	75	SAMEA2752426	PRJEB1021	GCA_000819725.1						CCDB01	2014/11/24	2015/08/05				
A. balearum	A. balearum	CECT 4227	infected fish	fish	4.60082	Contig	41	60.60	135.742	4	83	1	423	4153	42	SAMEA2752425	PRJEB1022	GCA_000819745.1						CCDB01	2014/11/24	2015/08/16					
	A. balearum	CECT 7113	cockles (Cardium sp.)	animals	4.43039	Contig	63	62.20	70.548	4	86	1	3694	3910	127	SAMEA2752424	PRJEB1022	GCA_000819765.1						CCDB01	2014/11/24	2015/08/16					
A. caviae	A. caviae	42665	urine	clinical	Illumina MiSeq	4.70029	Contig	4	61.00	100	1,256,313	31	114	1	4230	4272	96	SAMN0200367	PRJNA269298	GCA_000720765.1						LF001	2015/08/19	2015/08/25			
	A. caviae	BL8	human kg wound	clinical	Illumina MiSeq	4.67418	Contig	143	61.80	285	64,403	18	87	1	3921	4142	134	SAMN0200367	PRJNA269298	GCA_000720765.1						LF001	2015/08/19	2015/08/25			
	A. caviae	A53	supernatant	environment	Illumina MiSeq	4.52123	Contig	143	61.00	100	64,403	18	87	1	3921	4142	134	SAMN0200367	PRJNA269298	GCA_000720765.1						LF001	2015/07/21	2015/08/21			
	A. caviae	AC638	stool	clinical	454 GS-FLX/Illumina	4.43022	Contig	149	61.40	76.364	6	72	1	3876	4040	83	SAMEA2727404	PRJEA51877	GCA_000828825.1						CA001	2011/10/23	2015/08/16				
	A. caviae	CECT 838	epitoxic of young (parrot fish)	animals	Illumina	4.47414	Contig	111	61.70	17,441	5	98	1	3860	4047	114	SAMEA2752423	PRJEB1024	GCA_000819785.1						CCDB01	2014/11/24	2015/08/16				
	A. caviae	CECT 4221	used oil emulsions	environment	illumina	4.57592	Contig	332	61.10	30.787	16	87	1	3800	4233	239	SAMEA2752420	PRJEB1046	GCA_000820262.2						CCDB01	2014/11/24	2015/08/20				
	A. caviae	FDAR002_76	stool	clinical	PacBio	4.87779	Contig	10	61.00	19.15	3,550,007	31	124	2	3656	4149	704	SAMN0200458	PRJNA231221	GCA_000793695.1						JTB001	2014/11/14	2015/04/20			
	A. caviae	L12	water sample	clinical	Illumina HiSeq	4.37672	Contig	91	61.70	120	125,255	3	90	1	3866	3965	45	SAMN0206338	PRJNA270276	GCA_000821732.1						JW001	2015/01/02	2015/08/17			
	A. caviae	YL12	compost	environment	PacBio	4.84757	Contig	30	61.50	176.74	723,300	35	141	1	3907	4457	283	SAMN0200764	PRJNA242225	GCA_000821855.1						JW001	2014/07/02	2015/08/17			
	A. chalthensis	A. chalthensis	AAK1	blood	clinical	454 GS junior	4.76789	Contig	37	61.80	20.61	405,457	1	87	1	4179	4288	80	SAMN0200816	PRJEB070	GCA_000315106.1						NC_010041.1	AP012431.1	BALD1	2013/02/17	2015/09/16
		A. chalthensis	CECT 7289	aquarium of ornamental fish	clinical	illumina	4.69224	Contig	78	61.90	111,283	12	88	1	4074	4294	119	SAMEA2752427	PRJEB1020	GCA_000819705.1						CCDB01	2014/11/24	2015/08/16			
		A. chalthensis	CP 107500	bacon of children with diarrhea	clinical	illumina	4.71188	Contig	73	61.80	166,885	5	94	1	4099	4288	89	SAMEA2752400	PRJEB1048	GCA_000820305.1						CCDB01	2014/11/24	2015/08/16			
		A. chalthensis	BSU	stool	clinical	454	4.87787	Contig	271	61.60	11	36,733	5	84	1	4116	4077	371	SAMN0200748	PRJNA237007	GCA_000798065.1						JDM001	2014/08/16	2015/08/17		
A. divera	A. divera	CDC 2478.85	kg wound	clinical	454	4.03642	Contig	104	61.70	15	64,780	2	68	1	3521	3877	175	SAMN0247138	PRJNA183611	GCA_000838784.51						AP001	2013/04/19	2015/08/14			
	A. divera	CECT 4254	human kg wound	clinical	454	4.02641	Contig	37	61.60	140,915	5	95	1	3500	3778	188	SAMEA2752422	PRJEB1026	GCA_000819605.1						CCDB01	2014/11/24	2015/08/16				
	A. divera	CECT 4362	healthy outpatient urine	clinical	illumina	4.72523	Contig	35	62.00	170,359	10	94	1	3919	4068	98	SAMN0200369	PRJEB1027	GCA_000819625.1						CCDB01	2014/11/24	2015/08/16				
A. enteropneustes	A. enteropneustes	198809	reservoir water	clinical	454 GS junior	4.45548	Contig	133	60.10	15.1	3,382,347	7	78	3	3189	3823	303	SAMN0272206	PRJNA242526	GCA_000838755.1						CCDB01	2014/02/09	2014/05/05			
	A. enteropneustes	CECT 4265T	human stool	clinical	illumina	4.33725	Contig	27	60.10	224,321	6	88	1	3758	3803	60	SAMEA2752405	PRJEB1043	GCA_000820205.1						CCDB01	2014/11/24	2015/08/16				
	A. enteropneustes	CECT 4487	human stool	clinical	illumina	4.47474	Contig	49	59.70	119,983	3	86	1	3866	4060	84	SAMEA2752420	PRJEB1028	GCA_000819845.1						CCDB01	2014/11/24	2015/08/16				
	A. enterop																														

ANEXO B – ALINHAMENTO DAS SEQUÊNCIAS DE rRNA

5S rRNA

Percent Identity Matrix - created by Clustal2.1

1: rRNA_A_caviae_8LM_4504042-4504155_DIR-	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
2: rRNA_A_caviae_8LM_3668605-3668718_DIR-	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
3: rRNA_A_caviae_8LM_384642-384755_DIR+	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
4: rRNA_A_caviae_8LM_4426457-4426570_DIR-	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
5: rRNA_A_caviae_8LM_4253279-4253392_DIR-	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
6: rRNA_A_caviae_8LM_303576-303689_DIR+	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
7: rRNA_A_caviae_8LM_4379379-4379492_DIR-	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
8: rRNA_A_caviae_8LM_784779-784892_DIR+	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
9: rRNA_A_caviae_8LM_79673-79786_DIR+	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
10: rRNA_A_caviae_8LM_3827115-3827228_DIR-	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.61
11: rRNA_A_caviae_8LM_4379118-4379231_DIR-	95.61	95.61	95.61	95.61	95.61	95.61	95.61	95.61	95.61	95.61	95.61	100.00

CLUSTAL O(1.2.4) multiple sequence alignment

```

rRNA_A_caviae_8LM_4504042-4504155_DIR-      TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_3668605-3668718_DIR-      TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_384642-384755_DIR+        TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_4426457-4426570_DIR-      TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_4253279-4253392_DIR-      TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_303576-303689_DIR+        TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_4379379-4379492_DIR-      TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_784779-784892_DIR+        TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_79673-79786_DIR+          TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_3827115-3827228_DIR-      TGGCGGCAATAGCGCAGTGGAAACCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCTG
rRNA_A_caviae_8LM_4379118-4379231_DIR-      TGACGGCAATAGCGCCGTGGAAACCACCTGATCCCATGCCGAACCTCAGAAGTGAAACGCGG
* * * * *

```

```

rRNA_A_caviae_8LM_4504042-4504155_DIR-      TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_3668605-3668718_DIR-      TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_384642-384755_DIR+        TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_4426457-4426570_DIR-      TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_4253279-4253392_DIR-      TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_303576-303689_DIR+        TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_4379379-4379492_DIR-      TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_784779-784892_DIR+        TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_79673-79786_DIR+          TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_3827115-3827228_DIR-      TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGCCAGGCA
rRNA_A_caviae_8LM_4379118-4379231_DIR-      TAGCGCCGATGGTAGTGTGGCATTTCGCCATGCGAGAGTAGGACACTGTCAGGCA
* * * * *

```

16S rRNA

Percent Identity Matrix - created by Clustal2.1

1: rRNA_A_caviae_8LM_4382981-4384513_DIR-	100.00	99.80	99.74	99.80	99.80	99.87	99.87	99.87	99.87	99.87
2: rRNA_A_caviae_8LM_379499-381031_DIR+	99.80	100.00	99.93	99.87	99.74	99.80	99.80	99.80	99.80	99.80
3: rRNA_A_caviae_8LM_74600-76132_DIR+	99.74	99.93	100.00	99.93	99.80	99.87	99.87	99.87	99.87	99.87
4: rRNA_A_caviae_8LM_4507758-4509290_DIR-	99.80	99.87	99.93	100.00	99.87	99.93	99.93	99.93	99.93	99.93
5: rRNA_A_caviae_8LM_4430113-4431645_DIR-	99.80	99.74	99.80	99.87	100.00	99.93	99.93	99.93	99.93	99.93
6: rRNA_A_caviae_8LM_3830838-3832370_DIR-	99.87	99.80	99.87	99.93	99.93	100.00	100.00	100.00	100.00	100.00
7: rRNA_A_caviae_8LM_298597-300129_DIR+	99.87	99.80	99.87	99.93	99.93	100.00	100.00	100.00	100.00	100.00
8: rRNA_A_caviae_8LM_4256920-4258452_DIR-	99.87	99.80	99.87	99.93	99.93	100.00	100.00	100.00	100.00	100.00
9: rRNA_A_caviae_8LM_779719-781251_DIR+	99.87	99.80	99.87	99.93	99.93	100.00	100.00	100.00	100.00	100.00
10: rRNA_A_caviae_8LM_3672165-3673697_DIR-	99.87	99.80	99.87	99.93	99.93	100.00	100.00	100.00	100.00	100.00

CLUSTAL O(1.2.4) multiple sequence alignment

```

rRNA_A_caviae_8LM_4382981-4384513_DIR-      AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_379499-381031_DIR+        AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_74600-76132_DIR+          AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_4507758-4509290_DIR-      AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_4430113-4431645_DIR-      AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_3830838-3832370_DIR-      AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_298597-300129_DIR+        AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_4256920-4258452_DIR-      AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_779719-781251_DIR+        AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
rRNA_A_caviae_8LM_3672165-3673697_DIR-      AGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGC
*****

```

```

rRNA_A_caviae_8LM_4382981-4384513_DIR-      GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_379499-381031_DIR+        GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_74600-76132_DIR+          GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_4507758-4509290_DIR-      GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_4430113-4431645_DIR-      GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_3830838-3832370_DIR-      GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_298597-300129_DIR+        GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_4256920-4258452_DIR-      GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_779719-781251_DIR+        GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
rRNA_A_caviae_8LM_3672165-3673697_DIR-      GGCAGCGGGAAGTAGCTTGCTACTTTTGCCGGCGAGCGCGGACGGGTGAGTAATGCCT
*****

```

```

rRNA_A_caviae_8LM_4382981-4384513_DIR-      GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATGCGCC
rRNA_A_caviae_8LM_379499-381031_DIR+        GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATGCGCC
rRNA_A_caviae_8LM_74600-76132_DIR+          GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
rRNA_A_caviae_8LM_4507758-4509290_DIR-      GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
rRNA_A_caviae_8LM_4430113-4431645_DIR-      GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
rRNA_A_caviae_8LM_3830838-3832370_DIR-      GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
rRNA_A_caviae_8LM_298597-300129_DIR+        GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
rRNA_A_caviae_8LM_4256920-4258452_DIR-      GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
rRNA_A_caviae_8LM_779719-781251_DIR+        GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
rRNA_A_caviae_8LM_3672165-3673697_DIR-      GGGAAATGCCCAGTCGAGGGGGATAACAGTTGGAAACGACTGCTAATACCGCATACGCC
*****

```

```

rRNA_A_caviae_8LM_4382981-4384513_DIR-      CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_379499-381031_DIR+      CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_74600-76132_DIR+        CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_4507758-4509290_DIR-    CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_4430113-4431645_DIR-    CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_3830838-3832370_DIR-    CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_298597-300129_DIR+      CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_4256920-4258452_DIR-    CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_779719-781251_DIR+     CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
rRNA_A_caviae_8LM_3672165-3673697_DIR-    CTACGGGGGAAAGCAGGGGACCTTCGGGCCCTTGCGCGATTGGATATGCCAGGTGGGATT
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_379499-381031_DIR+     AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_74600-76132_DIR+       AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_4507758-4509290_DIR-    AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_4430113-4431645_DIR-    AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_3830838-3832370_DIR-    AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_298597-300129_DIR+     AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_4256920-4258452_DIR-    AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_779719-781251_DIR+     AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
rRNA_A_caviae_8LM_3672165-3673697_DIR-    AGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATG
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_379499-381031_DIR+     ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_74600-76132_DIR+       ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_4507758-4509290_DIR-    ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_4430113-4431645_DIR-    ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_3830838-3832370_DIR-    ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_298597-300129_DIR+     ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_4256920-4258452_DIR-    ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_779719-781251_DIR+     ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
rRNA_A_caviae_8LM_3672165-3673697_DIR-    ATCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAAT
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    ATTGCACAATGGGGGAAACCTTGATGCAGCCATACCGCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_379499-381031_DIR+     ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_74600-76132_DIR+       ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_4507758-4509290_DIR-    ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_4430113-4431645_DIR-    ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_3830838-3832370_DIR-    ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_298597-300129_DIR+     ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_4256920-4258452_DIR-    ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_779719-781251_DIR+     ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
rRNA_A_caviae_8LM_3672165-3673697_DIR-    ATTGCACAATGGGGGAAACCTTGATGCAGCCATGCCCGTGTGTGAAGAAGGCCCTTCGGG
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_379499-381031_DIR+     TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_74600-76132_DIR+       TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_4507758-4509290_DIR-    TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_4430113-4431645_DIR-    TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_3830838-3832370_DIR-    TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_298597-300129_DIR+     TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_4256920-4258452_DIR-    TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_779719-781251_DIR+     TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
rRNA_A_caviae_8LM_3672165-3673697_DIR-    TTGTAAAGCACTTTCAGCGAGGAGGAAAGGTCAGTAGCTAATAATCTGCTGGCTGTGACGT
*****

```

```

rRNA_A_caviae_8LM_4382981-4384513_DIR-      TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_379499-381031_DIR+        TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_74600-76132_DIR+          TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_4507758-4509290_DIR-      TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_4430113-4431645_DIR-      TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_3830838-3832370_DIR-      TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_298597-300129_DIR+        TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_4256920-4258452_DIR-      TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_779719-781251_DIR+        TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
rRNA_A_caviae_8LM_3672165-3673697_DIR-      TACTCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGC
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-      AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_379499-381031_DIR+        AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_74600-76132_DIR+          AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_4507758-4509290_DIR-      AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_4430113-4431645_DIR-      AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_3830838-3832370_DIR-      AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_298597-300129_DIR+        AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_4256920-4258452_DIR-      AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_779719-781251_DIR+        AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
rRNA_A_caviae_8LM_3672165-3673697_DIR-      AAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGCGGTTGGATAAGTTAGATGT
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-      GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_379499-381031_DIR+        GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_74600-76132_DIR+          GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_4507758-4509290_DIR-      GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_4430113-4431645_DIR-      GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_3830838-3832370_DIR-      GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_298597-300129_DIR+        GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_4256920-4258452_DIR-      GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_779719-781251_DIR+        GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
rRNA_A_caviae_8LM_3672165-3673697_DIR-      GAAAGCCCCGGGCTCAACCTGGGAATTGCATTTAAAAGTGTCCAGCTAGAGTCTTGTAGA
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-      GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_379499-381031_DIR+        GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_74600-76132_DIR+          GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_4507758-4509290_DIR-      GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_4430113-4431645_DIR-      GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_3830838-3832370_DIR-      GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_298597-300129_DIR+        GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_4256920-4258452_DIR-      GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_779719-781251_DIR+        GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
rRNA_A_caviae_8LM_3672165-3673697_DIR-      GGGGGGTAGAATTCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGG
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-      CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_379499-381031_DIR+        CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_74600-76132_DIR+          CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_4507758-4509290_DIR-      CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_4430113-4431645_DIR-      CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_3830838-3832370_DIR-      CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_298597-300129_DIR+        CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_4256920-4258452_DIR-      CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_779719-781251_DIR+        CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
rRNA_A_caviae_8LM_3672165-3673697_DIR-      CGAAGGCGGCCCCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGAGCAAAACAG
*****

```



```

rRNA_A_caviae_8LM_4382981-4384513_DIR-      TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_379499-381031_DIR+      TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_74600-76132_DIR+        TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_4507758-4509290_DIR-    TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_4430113-4431645_DIR-    TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_3830838-3832370_DIR-    TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_298597-300129_DIR+      TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_4256920-4258452_DIR-    TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_779719-781251_DIR+     TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
rRNA_A_caviae_8LM_3672165-3673697_DIR-    TGGGTTAAGTCCCACAACGAGCGCAACCCCTGTCCTTTGTTGCCAGCACGTAATGGTGGG
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_379499-381031_DIR+     AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_74600-76132_DIR+       AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_4507758-4509290_DIR-    AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_4430113-4431645_DIR-    AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_3830838-3832370_DIR-    AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_298597-300129_DIR+      AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_4256920-4258452_DIR-    AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_779719-781251_DIR+     AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
rRNA_A_caviae_8LM_3672165-3673697_DIR-    AACTCAAGGGAGACTGCCGGTGATAAACCCGAGGAAGTGGGGATGACGTCAAGTCATCA
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_379499-381031_DIR+     TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_74600-76132_DIR+       TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_4507758-4509290_DIR-    TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_4430113-4431645_DIR-    TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_3830838-3832370_DIR-    TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_298597-300129_DIR+      TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_4256920-4258452_DIR-    TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_779719-781251_DIR+     TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
rRNA_A_caviae_8LM_3672165-3673697_DIR-    TGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGCGCGTACAGAGGGCTGCAAGCT
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_379499-381031_DIR+     AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_74600-76132_DIR+       AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_4507758-4509290_DIR-    AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_4430113-4431645_DIR-    AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_3830838-3832370_DIR-    AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_298597-300129_DIR+      AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_4256920-4258452_DIR-    AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_779719-781251_DIR+     AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
rRNA_A_caviae_8LM_3672165-3673697_DIR-    AGCGATAGTGAGCGAATCCCAAAAAGCGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGA
*****

rRNA_A_caviae_8LM_4382981-4384513_DIR-    CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_379499-381031_DIR+     CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_74600-76132_DIR+       CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_4507758-4509290_DIR-    CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_4430113-4431645_DIR-    CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_3830838-3832370_DIR-    CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_298597-300129_DIR+      CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_4256920-4258452_DIR-    CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_779719-781251_DIR+     CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
rRNA_A_caviae_8LM_3672165-3673697_DIR-    CTCCATGAAGTCGGAATCGCTAGTAATCGCAAATCAGAATGTTGCGGTGAATACGTTCCC
*****

```


23S rRNA

Percent Identity Matrix - created by Clustal2.1

1: rRNA_A_caviae_8LM_3827352-3830238_DIR-	100.00	99.65	99.58	99.69	99.69	99.69	99.69	99.72	99.72	99.72
2: rRNA_A_caviae_8LM_4504281-4507167_DIR-	99.65	100.00	99.86	99.83	99.83	99.83	99.83	99.86	99.86	99.86
3: rRNA_A_caviae_8LM_3668800-3671686_DIR-	99.58	99.86	100.00	99.90	99.90	99.90	99.90	99.86	99.86	99.86
4: rRNA_A_caviae_8LM_4426696-4429582_DIR-	99.69	99.83	99.90	100.00	100.00	100.00	100.00	99.97	99.97	99.97
5: rRNA_A_caviae_8LM_4379616-4382502_DIR-	99.69	99.83	99.90	100.00	100.00	100.00	100.00	99.97	99.97	99.97
6: rRNA_A_caviae_8LM_4253503-4256389_DIR-	99.69	99.83	99.90	100.00	100.00	100.00	100.00	99.97	99.97	99.97
7: rRNA_A_caviae_8LM_781782-784668_DIR+	99.69	99.83	99.90	100.00	100.00	100.00	100.00	99.97	99.97	99.97
8: rRNA_A_caviae_8LM_381630-384516_DIR+	99.72	99.86	99.86	99.97	99.97	99.97	99.97	100.00	100.00	100.00
9: rRNA_A_caviae_8LM_300608-303494_DIR+	99.72	99.86	99.86	99.97	99.97	99.97	99.97	100.00	100.00	100.00
10: rRNA_A_caviae_8LM_76661-79547_DIR+	99.72	99.86	99.86	99.97	99.97	99.97	99.97	100.00	100.00	100.00

CLUSTAL O(1.2.4) multiple sequence alignment

```

rRNA_A_caviae_8LM_3827352-3830238_DIR-      TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_4504281-4507167_DIR-      TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_3668800-3671686_DIR-      TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_4426696-4429582_DIR-      TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_4379616-4382502_DIR-      TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_4253503-4256389_DIR-      TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_781782-784668_DIR+        TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_381630-384516_DIR+        TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_300608-303494_DIR+        TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
rRNA_A_caviae_8LM_76661-79547_DIR+          TTAAGTGACTAAGCGTACATGGTGGATGCCTTGGCAGTCAGAGGCGATGAAGGACGTACT
*****

```

```

rRNA_A_caviae_8LM_3827352-3830238_DIR-      AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_4504281-4507167_DIR-      AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_3668800-3671686_DIR-      AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_4426696-4429582_DIR-      AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_4379616-4382502_DIR-      AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_4253503-4256389_DIR-      AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_781782-784668_DIR+        AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_381630-384516_DIR+        AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_300608-303494_DIR+        AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
rRNA_A_caviae_8LM_76661-79547_DIR+          AACCTGCGATAAGCCGTGAGCAGTCGGTAAGAGACATTACGACTCACGGATTTCCGAATG
*****

```

```

rRNA_A_caviae_8LM_3827352-3830238_DIR-      GGGAAACCCACTTAGCATAAGCTAGGTATCGTGCCTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_4504281-4507167_DIR-      GGGAAACCCACTTAGCATAAGCTAGGTATCGTGCCTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_3668800-3671686_DIR-      GGGAAACCCACTTAGCATAAGCTAGGTATCGTTACGTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_4426696-4429582_DIR-      GGGAAACCCACTTAGCATAAGCTAGGTATCGTTACGTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_4379616-4382502_DIR-      GGGAAACCCACTTAGCATAAGCTAGGTATCGTTACGTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_4253503-4256389_DIR-      GGGAAACCCACTTAGCATAAGCTAGGTATCGTTACGTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_781782-784668_DIR+        GGGAAACCCACTTAGCATAAGCTAGGTATCGTTACGTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_381630-384516_DIR+        GGGAAACCCACTTAGCATAAGCTAGGTATCGTGCCTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_300608-303494_DIR+        GGGAAACCCACTTAGCATAAGCTAGGTATCGTGCCTGAATACATAGCGTAGCGAGGCGA
rRNA_A_caviae_8LM_76661-79547_DIR+          GGGAAACCCACTTAGCATAAGCTAGGTATCGTGCCTGAATACATAGCGTAGCGAGGCGA
*****

```

```

rRNA_A_caviae_8LM_3827352-3830238_DIR- ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_4504281-4507167_DIR- ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_3668800-3671686_DIR- ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_4426696-4429582_DIR- ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_4379616-4382502_DIR- ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_4253503-4256389_DIR- ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_781782-784668_DIR+ ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_381630-384516_DIR+ ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_300608-303494_DIR+ ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
rRNA_A_caviae_8LM_76661-79547_DIR+ ACCGGGAGAAGTAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCTCA
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_4504281-4507167_DIR- GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_3668800-3671686_DIR- GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_4426696-4429582_DIR- GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_4379616-4382502_DIR- GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_4253503-4256389_DIR- GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_781782-784668_DIR+ GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_381630-384516_DIR+ GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_300608-303494_DIR+ GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
rRNA_A_caviae_8LM_76661-79547_DIR+ GTAGCGCGAGCGAACCGGGATTAGCCCTTAAGCATCTTGGAAAGTTAGTGGAAACGGTCTT
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_4504281-4507167_DIR- GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_3668800-3671686_DIR- GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_4426696-4429582_DIR- GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_4379616-4382502_DIR- GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_4253503-4256389_DIR- GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_781782-784668_DIR+ GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_381630-384516_DIR+ GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_300608-303494_DIR+ GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
rRNA_A_caviae_8LM_76661-79547_DIR+ GGAAAGGCCGCGGATACAGGGTGATAGCCCGTACACGAAAACAACCTTGATGTGAAATC
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_4504281-4507167_DIR- GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_3668800-3671686_DIR- GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_4426696-4429582_DIR- GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_4379616-4382502_DIR- GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_4253503-4256389_DIR- GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_781782-784668_DIR+ GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_381630-384516_DIR+ GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_300608-303494_DIR+ GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
rRNA_A_caviae_8LM_76661-79547_DIR+ GAGTAGGGCGGGACACGTGACATCCTGTCTGAATATGGGGGGACCATCCTCCAAGGCTAA
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_4504281-4507167_DIR- ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_3668800-3671686_DIR- ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_4426696-4429582_DIR- ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_4379616-4382502_DIR- ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_4253503-4256389_DIR- ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_781782-784668_DIR+ ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_381630-384516_DIR+ ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_300608-303494_DIR+ ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
rRNA_A_caviae_8LM_76661-79547_DIR+ ATACTCCTGACTGACCGATAGTGAACCAAGTACCCTGAGGGAAAGCGAAAAGAACCCTG
*****

```

```

rRNA_A_caviae_8LM_3827352-3830238_DIR-    TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_4504281-4507167_DIR-    TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_3668800-3671686_DIR-    TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_4426696-4429582_DIR-    TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_4379616-4382502_DIR-    TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_4253503-4256389_DIR-    TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_781782-784668_DIR+      TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_381630-384516_DIR+      TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_300608-303494_DIR+      TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
rRNA_A_caviae_8LM_76661-79547_DIR+       TGAGGGGAGTGAAATAGAACCTGAAACCGTGTACGTACAAGCAGTGGGAGCCCTTCGGGG
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_4504281-4507167_DIR-    TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_3668800-3671686_DIR-    TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_4426696-4429582_DIR-    TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_4379616-4382502_DIR-    TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_4253503-4256389_DIR-    TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_781782-784668_DIR+      TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_381630-384516_DIR+      TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_300608-303494_DIR+      TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
rRNA_A_caviae_8LM_76661-79547_DIR+       TGACTGCGTACCTTTTGTATAATGGGTCAGCGACTTACATTTTGTAGCGAGGTTAACCGT
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_4504281-4507167_DIR-    ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_3668800-3671686_DIR-    ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_4426696-4429582_DIR-    ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_4379616-4382502_DIR-    ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_4253503-4256389_DIR-    ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_781782-784668_DIR+      ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_381630-384516_DIR+      ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_300608-303494_DIR+      ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
rRNA_A_caviae_8LM_76661-79547_DIR+       ATAGGGGAGCCGTAGGGAAACCGAGTCTTAACCTGGGCGTCTAGTTGCAAGGTGTAGACCC
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_4504281-4507167_DIR-    GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_3668800-3671686_DIR-    GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_4426696-4429582_DIR-    GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_4379616-4382502_DIR-    GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_4253503-4256389_DIR-    GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_781782-784668_DIR+      GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_381630-384516_DIR+      GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_300608-303494_DIR+      GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
rRNA_A_caviae_8LM_76661-79547_DIR+       GAAACCGGTTGATCTAGCCATGGGCAGGTTGAAGGTTGAGTAACATCAACTGGAGGACCG
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_4504281-4507167_DIR-    AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_3668800-3671686_DIR-    AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_4426696-4429582_DIR-    AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_4379616-4382502_DIR-    AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_4253503-4256389_DIR-    AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_781782-784668_DIR+      AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_381630-384516_DIR+      AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_300608-303494_DIR+      AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
rRNA_A_caviae_8LM_76661-79547_DIR+       AACCCACTAACGTTGCAAAGTTAGGGGATGACCTGTGGCTGGGGTGAAAGGCCAATCAA
*****

```

```

rRNA_A_caviae_8LM_3827352-3830238_DIR-    ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_4504281-4507167_DIR-    ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_3668800-3671686_DIR-    ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_4426696-4429582_DIR-    ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_4379616-4382502_DIR-    ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_4253503-4256389_DIR-    ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_781782-784668_DIR+      ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_381630-384516_DIR+      ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_300608-303494_DIR+      ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
rRNA_A_caviae_8LM_76661-79547_DIR+       ACTCGGAGATAGCTGGTTCTCCCCGAAAGCTATTTAGGTAGCGCCTCGGACGAATACTAC
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_4504281-4507167_DIR-    TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_3668800-3671686_DIR-    TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_4426696-4429582_DIR-    TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_4379616-4382502_DIR-    TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_4253503-4256389_DIR-    TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_781782-784668_DIR+      TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_381630-384516_DIR+      TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_300608-303494_DIR+      TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
rRNA_A_caviae_8LM_76661-79547_DIR+       TGGGGGTAGAGCACTGTTTGGGCTAGGGGGTCATCCCCTACTTACCAACCCCATGCAAACCT
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    CCGAATACCAGTAAGTACTATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_4504281-4507167_DIR-    CCGAATACCAGTAAGTACTATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_3668800-3671686_DIR-    CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_4426696-4429582_DIR-    CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_4379616-4382502_DIR-    CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_4253503-4256389_DIR-    CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_781782-784668_DIR+      CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_381630-384516_DIR+      CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_300608-303494_DIR+      CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
rRNA_A_caviae_8LM_76661-79547_DIR+       CCGAATACCAGTAAGTAATATCCGGGAGACACACGGCGGGTGCTAACGTCCCGTCGTGAAG
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_4504281-4507167_DIR-    AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_3668800-3671686_DIR-    AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_4426696-4429582_DIR-    AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_4379616-4382502_DIR-    AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_4253503-4256389_DIR-    AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_781782-784668_DIR+      AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_381630-384516_DIR+      AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_300608-303494_DIR+      AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
rRNA_A_caviae_8LM_76661-79547_DIR+       AGGGAACAACCCAGACCCCGGCTAAGGTCCCAAAGTTCCTGGTTAAGTGGGAAACGATG
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-    TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_4504281-4507167_DIR-    TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_3668800-3671686_DIR-    TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_4426696-4429582_DIR-    TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_4379616-4382502_DIR-    TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_4253503-4256389_DIR-    TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_781782-784668_DIR+      TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_381630-384516_DIR+      TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_300608-303494_DIR+      TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
rRNA_A_caviae_8LM_76661-79547_DIR+       TGGGAAGGCTCAGACAGCTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAAAGCGT
*****

```



```

rRNA_A_caviae_8LM_3827352-3830238_DIR-      TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_4504281-4507167_DIR-      TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_3668800-3671686_DIR-      TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_4426696-4429582_DIR-      TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_4379616-4382502_DIR-      TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_4253503-4256389_DIR-      TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_781782-784668_DIR+        TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_381630-384516_DIR+        TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_300608-303494_DIR+        TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
rRNA_A_caviae_8LM_76661-79547_DIR+         TCCTGCACGACTTGTAAATTCGCATGGGGGGACGGAGAAGGCTAGGTGGGCCAGGCACGG
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-      TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_4504281-4507167_DIR-      TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_3668800-3671686_DIR-      TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_4426696-4429582_DIR-      TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_4379616-4382502_DIR-      TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_4253503-4256389_DIR-      TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_781782-784668_DIR+        TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_381630-384516_DIR+        TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_300608-303494_DIR+        TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
rRNA_A_caviae_8LM_76661-79547_DIR+         TTGTCCCTGGTGAAAGTGCCTAGGTGGTGTTCCTAGGCAAAATCCGGAGACACAACACTGAG
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-      ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_4504281-4507167_DIR-      ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_3668800-3671686_DIR-      ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_4426696-4429582_DIR-      ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_4379616-4382502_DIR-      ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_4253503-4256389_DIR-      ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_781782-784668_DIR+        ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_381630-384516_DIR+        ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_300608-303494_DIR+        ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
rRNA_A_caviae_8LM_76661-79547_DIR+         ACACGAGACGAACGCACCACGGTGCCTGCGGAAGCCATTGATGCCTGCTTCCAGGAAAAGCCT
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-      CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_4504281-4507167_DIR-      CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_3668800-3671686_DIR-      CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_4426696-4429582_DIR-      CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_4379616-4382502_DIR-      CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_4253503-4256389_DIR-      CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_781782-784668_DIR+        CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_381630-384516_DIR+        CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_300608-303494_DIR+        CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
rRNA_A_caviae_8LM_76661-79547_DIR+         CTAAGCTTCAGATTACAAGTCATCGTACCCCAAACCGACACAGGTGGTTCGGGTAGAGAAT
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR-      ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_4504281-4507167_DIR-      ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_3668800-3671686_DIR-      ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_4426696-4429582_DIR-      ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_4379616-4382502_DIR-      ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_4253503-4256389_DIR-      ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_781782-784668_DIR+        ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_381630-384516_DIR+        ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_300608-303494_DIR+        ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
rRNA_A_caviae_8LM_76661-79547_DIR+         ACCAAGGCGCTTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATAGAACCCTAACTTCGG
*****

```



```
rRNA_A_caviae_8LM_3827352-3830238_DIR- ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_4504281-4507167_DIR- ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_3668800-3671686_DIR- ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_4426696-4429582_DIR- ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_4379616-4382502_DIR- ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_4253503-4256389_DIR- ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_781782-784668_DIR+ ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_381630-384516_DIR+ ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_300608-303494_DIR+ ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
rRNA_A_caviae_8LM_76661-79547_DIR+ ATCTGGTCGGACATCAGGAGGTTAGTGCAATGGCATAAAGCCAGCTTAAGTGCAGACGG
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_4504281-4507167_DIR- ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_3668800-3671686_DIR- ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_4426696-4429582_DIR- ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_4379616-4382502_DIR- ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_4253503-4256389_DIR- ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_781782-784668_DIR+ ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_381630-384516_DIR+ ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_300608-303494_DIR+ ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
rRNA_A_caviae_8LM_76661-79547_DIR+ ACAGGTCGAGCAGGTACGAAAGTAGGTCATAGTGATCCGGTGGTCTGAATGGAAGGGCC
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_4504281-4507167_DIR- ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_3668800-3671686_DIR- ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_4426696-4429582_DIR- ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_4379616-4382502_DIR- ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_4253503-4256389_DIR- ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_781782-784668_DIR+ ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_381630-384516_DIR+ ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_300608-303494_DIR+ ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
rRNA_A_caviae_8LM_76661-79547_DIR+ ATCGCTCAACGGATAAAAAGGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCAT
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_4504281-4507167_DIR- ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_3668800-3671686_DIR- ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_4426696-4429582_DIR- ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_4379616-4382502_DIR- ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_4253503-4256389_DIR- ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_781782-784668_DIR+ ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_381630-384516_DIR+ ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_300608-303494_DIR+ ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
rRNA_A_caviae_8LM_76661-79547_DIR+ ATCGACGGCGGTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTCGGT
*****

rRNA_A_caviae_8LM_3827352-3830238_DIR- CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_4504281-4507167_DIR- CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_3668800-3671686_DIR- CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_4426696-4429582_DIR- CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_4379616-4382502_DIR- CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_4253503-4256389_DIR- CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_781782-784668_DIR+ CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_381630-384516_DIR+ CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_300608-303494_DIR+ CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
rRNA_A_caviae_8LM_76661-79547_DIR+ CCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTCAGAACGTCGTG
*****
```


rRNA_A_caviae_8LM_3827352-3830238_DIR-	TAACCAT
rRNA_A_caviae_8LM_4504281-4507167_DIR-	TAACCAT
rRNA_A_caviae_8LM_3668800-3671686_DIR-	TAACCAT
rRNA_A_caviae_8LM_4426696-4429582_DIR-	TAACCAT
rRNA_A_caviae_8LM_4379616-4382502_DIR-	TAACCAT
rRNA_A_caviae_8LM_4253503-4256389_DIR-	TAACCAT
rRNA_A_caviae_8LM_781782-784668_DIR+	TAACCAT
rRNA_A_caviae_8LM_381630-384516_DIR+	TAACCAT
rRNA_A_caviae_8LM_300608-303494_DIR+	TAACCAT
rRNA_A_caviae_8LM_76661-79547_DIR+	TAACCAT

ANEXO C – SIMILARIDADE PELO GGDC–DDH

Organismo	Estirpe	DDH	Dist.	DDH > 70%	Dif. G+C
<i>Aeromonas allosaccharophila</i>	BVH88	27,90	0,15	0,04	3,07
<i>Aeromonas allosaccharophila</i>	CECT 4199	28,20	0,15	0,05	3,29
<i>Aeromonas aquatica</i>	AE235	30,20	0,14	0,12	0,53
<i>Aeromonas australiensis</i>	CECT 8023	27,60	0,16	0,04	3,56
<i>Aeromonas bestiarum</i>	CECT 4227	29,60	0,14	0,09	1,05
<i>Aeromonas bivalvium</i>	CECT 7113	28,10	0,15	0,04	0,61
<i>Aeromonas caviae</i>	429865	83,80	0,02	93,11	0,64
<i>Aeromonas caviae</i>	A23	83,50	0,02	92,91	0,41
<i>Aeromonas caviae</i>	Ae398	82,60	0,02	92,41	0,25
<i>Aeromonas caviae</i>	CECT 838	81,80	0,02	91,88	0,08
<i>Aeromonas caviae</i>	CECT 4221	82,00	0,02	92,02	0,60
<i>Aeromonas caviae</i>	FDAARGOS_72	81,50	0,02	91,65	0,00
<i>Aeromonas caviae</i>	FDAARGOS_75	81,80	0,02	91,89	0,01
<i>Aeromonas caviae</i>	FDAARGOS_76	81,70	0,02	91,80	0,64
<i>Aeromonas caviae</i>	L12	81,70	0,02	91,81	0,02
<i>Aeromonas caviae</i>	YL12	81,20	0,02	91,48	0,23
<i>Aeromonas dhakensis</i>	AAK1	32,20	0,13	0,25	0,07
<i>Aeromonas dhakensis</i>	CECT 7289	32,20	0,13	0,25	0,27
<i>Aeromonas dhakensis</i>	CIP 107500	32,20	0,13	0,25	0,10
<i>Aeromonas dhakensis</i>	SSU	32,40	0,13	0,26	0,09
<i>Aeromonas diversa</i>	CDC 2478-85	23,00	0,19	0,00	0,01
<i>Aeromonas diversa</i>	CECT 4254	23,20	0,19	0,00	0,03
<i>Aeromonas encheleia</i>	CECT 4342	30,30	0,14	0,12	0,28
<i>Aeromonas enteropelogenes</i>	1999lcr	29,40	0,15	0,08	1,51
<i>Aeromonas enteropelogenes</i>	CECT 4255T	28,50	0,15	0,05	1,61
<i>Aeromonas enteropelogenes</i>	CECT 4487	28,40	0,15	0,05	1,98
<i>Aeromonas enteropelogenes</i>	LK14	32,30	0,13	0,26	0,28
<i>Aeromonas eucrenophila</i>	CECT 4224	30,00	0,14	0,10	0,51
<i>Aeromonas finlandiensis</i>	4287D	28,10	0,15	0,04	3,11
<i>Aeromonas fluvialis</i>	LMG 24681	27,90	0,15	0,04	3,35
<i>Aeromonas hydrophila</i>	14	32,30	0,13	0,26	0,28
<i>Aeromonas hydrophila</i>	116	32,30	0,13	0,26	0,29
<i>Aeromonas hydrophila</i>	145	32,20	0,13	0,25	0,23
<i>Aeromonas hydrophila</i>	173	32,20	0,13	0,25	0,03
<i>Aeromonas hydrophila</i>	187	32,30	0,13	0,26	0,05
<i>Aeromonas hydrophila</i>	226	32,10	0,13	0,24	0,81
<i>Aeromonas hydrophila</i>	259	32,30	0,13	0,26	0,03
<i>Aeromonas hydrophila</i>	277	32,20	0,13	0,25	0,04
<i>Aeromonas hydrophila</i>	48_AHYD	32,30	0,13	0,26	0,09
<i>Aeromonas hydrophila</i>	4AK4	34,50	0,12	0,56	0,33
<i>Aeromonas hydrophila</i>	50_AHYD	0,00	0,00	0,00	0,05
<i>Aeromonas hydrophila</i>	52_AHYD	32,30	0,13	0,26	0,07
<i>Aeromonas hydrophila</i>	53_AHYD	32,30	0,13	0,26	0,05
<i>Aeromonas hydrophila</i>	56_AHYD	0,00	0,00	0,00	0,07
<i>Aeromonas hydrophila</i>	AD9	32,30	0,13	0,26	0,36
<i>Aeromonas hydrophila</i>	Ae34	32,20	0,13	0,25	0,10
<i>Aeromonas hydrophila</i>	AH10	32,60	0,13	0,29	0,53
<i>Aeromonas hydrophila</i>	AL06-06	32,30	0,13	0,26	0,30
<i>Aeromonas hydrophila</i>	AL09-71	32,30	0,13	0,26	0,86
<i>Aeromonas hydrophila</i>	BWH65	82,10	0,02	92,08	0,11

Organismo	Estirpe	DDH	Dist.	DDH > 70%	Dif. G+C
<i>Aeromonas hydrophila</i>	FDAARGOS_78	32,30	0,13	0,26	0,72
<i>Aeromonas hydrophila</i>	J-1	32,50	0,13	0,28	0,79
<i>Aeromonas hydrophila</i>	KOR1	32,50	0,13	0,28	0,14
<i>Aeromonas hydrophila</i>	L14f	32,20	0,13	0,25	0,03
<i>Aeromonas hydrophila</i>	M013	32,20	0,13	0,25	0,64
<i>Aeromonas hydrophila</i>	M023	32,10	0,13	0,24	0,74
<i>Aeromonas hydrophila</i>	M062	32,10	0,13	0,25	0,54
<i>Aeromonas hydrophila</i>	ML09119	32,30	0,13	0,26	0,87
<i>Aeromonas hydrophila</i>	NF1	32,00	0,13	0,24	0,61
<i>Aeromonas hydrophila</i>	NF2	32,10	0,13	0,25	0,41
<i>Aeromonas hydrophila</i>	NJ-35	32,40	0,13	0,27	1,17
<i>Aeromonas hydrophila</i>	pc104A	32,30	0,13	0,26	0,86
<i>Aeromonas hydrophila</i>	RB-AH	32,30	0,13	0,26	0,86
<i>Aeromonas hydrophila</i>	SNUFPC-A8	32,00	0,13	0,23	0,86
<i>Aeromonas hydrophila</i>	SSU	32,40	0,13	0,26	0,13
<i>Aeromonas hydrophila</i>	YL17	32,60	0,13	0,28	0,06
<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i>	ATCC 7966	32,40	0,13	0,27	0,13
<i>Aeromonas hydrophila</i> subsp. <i>ranae</i>	CIP107985	32,10	0,13	0,24	0,13
<i>Aeromonas jandaei</i>	CECT 4228	28,00	0,15	0,04	2,59
<i>Aeromonas jandaei</i>	L14h	32,20	0,13	0,25	0,02
<i>Aeromonas jandaei</i>	Riv2	28,00	0,15	0,04	2,64
<i>Aeromonas lacus</i>	AE122	27,80	0,15	0,04	2,72
<i>Aeromonas media</i>	ARB13	34,60	0,12	0,57	0,67
<i>Aeromonas media</i>	ARB20	34,60	0,12	0,57	0,71
<i>Aeromonas media</i>	CECT 4232	34,80	0,12	0,62	0,54
<i>Aeromonas media</i>	WS	35,00	0,12	0,67	0,94
<i>Aeromonas molluscorum</i>	848	24,70	0,18	0,01	2,50
<i>Aeromonas piscicola</i>	LMG 24783	28,90	0,15	0,07	2,47
<i>Aeromonas popoffii</i>	CIP 105493	28,30	0,15	0,05	3,07
<i>Aeromonas rivuli</i>	DSM 22539	25,10	0,17	0,01	1,66
<i>Aeromonas salmonicida</i>	CBA100	29,80	0,14	0,10	1,13
<i>Aeromonas salmonicida</i> subsp. <i>achromogenes</i>	AS03	27,90	0,15	0,04	3,00
<i>Aeromonas salmonicida</i> subsp. <i>masoucida</i>	NBRC 13784	27,90	0,15	0,04	2,88
<i>Aeromonas salmonicida</i> subsp. <i>pectinolytica</i>	34mel	28,10	0,15	0,05	3,20
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	01-B526	27,80	0,16	0,04	3,34
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	2004-05MF26	27,80	0,15	0,04	3,43
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	2009-144K3	27,90	0,15	0,04	3,33
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	A449	28,10	0,15	0,04	3,17
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	CIP 103209	27,80	0,15	0,04	3,17
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	JF3224	27,90	0,15	0,04	3,24
<i>Aeromonas sanarellii</i>	LMG 24682	40,10	0,10	2,80	1,45
<i>Aeromonas schubertii</i>	CECT 4240	23,10	0,19	0,00	0,01
<i>Aeromonas simiae</i>	CIP 107798	23,10	0,19	0,00	0,32
<i>Aeromonas sobria</i>	CECT 4245	27,00	0,16	0,03	4,06
<i>Aeromonas</i> sp.	159	28,50	0,15	0,05	3,49
<i>Aeromonas</i> sp.	312M	28,30	0,15	0,05	3,01
<i>Aeromonas</i> sp.	HZM	82,40	0,02	92,28	0,02
<i>Aeromonas</i> sp.	L_1B5_3	28,00	0,15	0,04	2,99
<i>Aeromonas</i> sp.	MDS8	32,40	0,13	0,26	0,12
<i>Aeromonas</i> sp.	ZOR0001	28,50	0,15	0,06	2,99
<i>Aeromonas</i> sp.	ZOR0002	80,50	0,02	90,96	0,47
<i>Aeromonas taiwanensis</i>	LMG 24683	37,90	0,11	1,57	1,14
<i>Aeromonas tecta</i>	CECT 7082	28,90	0,15	0,06	1,57

Organismo	Estirpe	DDH	Dist.	DDH > 70%	Dif. G+C
<i>Aeromonas veronii</i>	AER39	28,50	0,15	0,05	2,78
<i>Aeromonas veronii</i>	AER397	28,40	0,15	0,05	2,76
<i>Aeromonas veronii</i>	AMC34	28,20	0,15	0,05	3,19
<i>Aeromonas veronii</i>	AMC35	28,60	0,15	0,06	3,10
<i>Aeromonas veronii</i>	ARB3	28,20	0,15	0,05	2,91
<i>Aeromonas veronii</i>	B565	28,60	0,15	0,06	2,96
<i>Aeromonas veronii</i>	CECT 4486	28,20	0,15	0,05	3,25
<i>Aeromonas veronii</i>	CIP 107763	28,40	0,15	0,05	2,81
<i>Aeromonas veronii</i>	Hm21	28,10	0,15	0,05	2,74
<i>Aeromonas veronii</i>	Phln2	28,30	0,15	0,05	2,84
<i>Aeromonas veronii</i> bv. sobria	LMG 13067	28,30	0,15	0,05	2,95
<i>Aeromonas veronii</i> bv. veronii	CECT 4257	29,80	0,14	0,10	2,89

NOTA: DDH = valor estimado de DDH. Dist. = distância intergenômica. DDH>70% = propabilidade de pertencerem à mesma espécie. Dif G+C = diferença no conteúdo G+C.

ANEXO D – SIMILARIDADE PELO ANI

Organismo	Estirpe	ANI	DP
<i>Aeromonas allosaccharophila</i>	BVH88	85,09	5,10
<i>Aeromonas allosaccharophila</i>	CECT 4199	85,11	5,11
<i>Aeromonas aquatica</i>	AE235	85,91	4,57
<i>Aeromonas australiensis</i>	CECT 8023	84,93	5,16
<i>Aeromonas bestiarum</i>	CECT 4227	85,77	4,52
<i>Aeromonas bivalvium</i>	CECT 7113	85,77	5,69
<i>Aeromonas caviae</i>	429865	98,12	1,78
<i>Aeromonas caviae</i>	A23	98,11	1,91
<i>Aeromonas caviae</i>	Ae398	97,99	1,86
<i>Aeromonas caviae</i>	CECT 838	97,88	1,81
<i>Aeromonas caviae</i>	CECT 4221	97,93	2,03
<i>Aeromonas caviae</i>	FDAARGOS_72	97,83	1,88
<i>Aeromonas caviae</i>	FDAARGOS_75	97,85	1,85
<i>Aeromonas caviae</i>	FDAARGOS_76	98,00	2,21
<i>Aeromonas caviae</i>	L12	97,89	1,80
<i>Aeromonas caviae</i>	YL12	98,84	1,99
<i>Aeromonas dhakensis</i>	AAK1	87,37	5,11
<i>Aeromonas dhakensis</i>	CECT 7289	87,37	5,07
<i>Aeromonas dhakensis</i>	CIP 107500	87,39	5,11
<i>Aeromonas dhakensis</i>	SSU	87,54	5,10
<i>Aeromonas diversa</i>	CDC 2478-85	81,93	5,48
<i>Aeromonas diversa</i>	CECT 4254	81,94	5,56
<i>Aeromonas encheleia</i>	CECT 4342	85,91	4,50
<i>Aeromonas enteropelogenes</i>	1999lcr	85,46	5,29
<i>Aeromonas enteropelogenes</i>	CECT 4255T	85,44	5,18
<i>Aeromonas enteropelogenes</i>	CECT 4487	85,43	5,21
<i>Aeromonas enteropelogenes</i>	LK14	87,40	5,13
<i>Aeromonas eucrenophila</i>	CECT 4224	85,83	4,62
<i>Aeromonas finlandiensis</i>	4287D	85,10	5,08
<i>Aeromonas fluvialis</i>	LMG 24681	84,89	5,01
<i>Aeromonas hydrophila</i>	14	87,44	5,09
<i>Aeromonas hydrophila</i>	116	87,48	5,08
<i>Aeromonas hydrophila</i>	145	87,43	5,09
<i>Aeromonas hydrophila</i>	173	87,43	5,10
<i>Aeromonas hydrophila</i>	187	87,52	5,13
<i>Aeromonas hydrophila</i>	226	87,24	5,13
<i>Aeromonas hydrophila</i>	259	87,40	5,09
<i>Aeromonas hydrophila</i>	277	87,48	5,11
<i>Aeromonas hydrophila</i>	48_AHYD	87,35	5,23
<i>Aeromonas hydrophila</i>	4AK4	87,74	4,41
<i>Aeromonas hydrophila</i>	50_AHYD	87,39	5,20
<i>Aeromonas hydrophila</i>	52_AHYD	87,42	5,20
<i>Aeromonas hydrophila</i>	53_AHYD	87,29	5,25
<i>Aeromonas hydrophila</i>	56_AHYD	87,37	5,22
<i>Aeromonas hydrophila</i>	AD9	87,41	5,21
<i>Aeromonas hydrophila</i>	Ae34	87,26	5,31
<i>Aeromonas hydrophila</i>	AH10	87,49	5,29
<i>Aeromonas hydrophila</i>	AL06-06	87,34	5,27
<i>Aeromonas hydrophila</i>	AL09-71	87,46	5,19
<i>Aeromonas hydrophila</i>	BWH65	97,90	1,80
<i>Aeromonas hydrophila</i>	FDAARGOS_78	87,46	5,23
<i>Aeromonas hydrophila</i>	J-1	87,45	5,26

<i>Aeromonas hydrophila</i>	KOR1	87,47	5,22
<i>Aeromonas hydrophila</i>	L14f	87,40	5,13
<i>Aeromonas hydrophila</i>	M013	87,39	5,16
<i>Aeromonas hydrophila</i>	M023	87,27	5,18
<i>Aeromonas hydrophila</i>	M062	87,38	5,17
<i>Aeromonas hydrophila</i>	ML09119	87,36	5,24
<i>Aeromonas hydrophila</i>	NF1	87,31	5,20
<i>Aeromonas hydrophila</i>	NF2	87,43	5,22
<i>Aeromonas hydrophila</i>	NJ-35	87,45	5,22
<i>Aeromonas hydrophila</i>	pc104A	87,43	5,18
<i>Aeromonas hydrophila</i>	RB-AH	87,36	5,20
<i>Aeromonas hydrophila</i>	SNUFPC-A8	87,30	5,24
<i>Aeromonas hydrophila</i>	SSU	87,53	5,17
<i>Aeromonas hydrophila</i>	YL17	87,54	5,19
<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i>	ATCC 7966	87,47	5,26
<i>Aeromonas hydrophila</i> subsp. <i>ranae</i>	CIP107985	87,28	5,10
<i>Aeromonas jandaei</i>	CECT 4228	85,30	5,26
<i>Aeromonas jandaei</i>	L14h	87,44	5,11
<i>Aeromonas jandaei</i>	Riv2	85,20	5,33
<i>Aeromonas lacus</i>	AE122	85,14	5,20
<i>Aeromonas media</i>	ARB13	87,85	4,48
<i>Aeromonas media</i>	ARB20	87,85	4,45
<i>Aeromonas media</i>	CECT 4232	88,05	4,43
<i>Aeromonas media</i>	WS	88,07	4,67
<i>Aeromonas molluscorum</i>	848	82,92	5,02
<i>Aeromonas piscicola</i>	LMG 24783	85,43	4,56
<i>Aeromonas popoffii</i>	CIP 105493	84,94	4,53
<i>Aeromonas rivuli</i>	DSM 22539	83,21	5,17
<i>Aeromonas salmonicida</i>	CBA100	85,80	4,63
<i>Aeromonas salmonicida</i> subsp. <i>achromogenes</i>	AS03	84,60	4,59
<i>Aeromonas salmonicida</i> subsp. <i>masoucida</i>	NBRC 13784	84,64	4,62
<i>Aeromonas salmonicida</i> subsp. <i>pectinolytica</i>	34mel	84,76	4,54
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	01-B526	84,62	4,58
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	2004-05MF26	84,59	4,57
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	2009-144K3	84,61	4,62
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	A449	84,70	4,68
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	CIP 103209	84,60	4,60
<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	JF3224	84,63	4,59
<i>Aeromonas sanarellii</i>	LMG 24682	89,94	4,29
<i>Aeromonas schubertii</i>	CECT 4240	81,92	5,46
<i>Aeromonas simiae</i>	CIP 107798	81,81	5,37
<i>Aeromonas sobria</i>	CECT 4245	84,31	4,76
<i>Aeromonas</i> sp.	159	85,36	5,23
<i>Aeromonas</i> sp.	312M	85,14	5,20
<i>Aeromonas</i> sp.	HZM	98,01	2,13
<i>Aeromonas</i> sp.	L_1B5_3	85,21	5,33
<i>Aeromonas</i> sp.	MDS8	87,50	5,13
<i>Aeromonas</i> sp.	ZOR0001	85,33	5,15
<i>Aeromonas</i> sp.	ZOR0002	97,87	2,05
<i>Aeromonas taiwanensis</i>	LMG 24683	89,15	4,40
<i>Aeromonas tecta</i>	CECT 7082	85,23	4,56
<i>Aeromonas veronii</i>	AER39	85,19	5,29
<i>Aeromonas veronii</i>	AER397	85,27	5,25
<i>Aeromonas veronii</i>	AMC34	85,33	5,25
<i>Aeromonas veronii</i>	AMC35	85,36	5,33
<i>Aeromonas veronii</i>	ARB3	85,20	5,22

<i>Aeromonas veronii</i>	B565	85,42	5,32
<i>Aeromonas veronii</i>	CECT 4486	85,10	5,24
<i>Aeromonas veronii</i>	CIP 107763	85,16	5,15
<i>Aeromonas veronii</i>	Hm21	85,16	5,23
<i>Aeromonas veronii</i>	PhIn2	85,45	5,26
<i>Aeromonas veronii</i> bv. <i>sobria</i>	LMG 13067	85,35	5,15
<i>Aeromonas veronii</i> bv. <i>veronii</i>	CECT 4257	85,26	5,18

NOTA: ANI = average nucleotide identify. DP = desvio padrão.

ANEXO E – CORE GENOMA DE *AEROMONAS* (COMPLETOS)

Gene – Anotação

accC – Biotin carboxylase
 accD – Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
 aceA – Isocitrate lyase
 aceE – Pyruvate dehydrogenase E1 component
 aceF – Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex
 ackA_1 – Acetate kinase
 acnB – Aconitate hydratase B
 acpP – Acyl carrier protein
 add2 – Aminodeoxyfualosine deaminase
 adeP – Adenine permease AdeP
 adk – Adenylate kinase
 aguA – Agmatine deiminase
 ahpC – Alkyl hydroperoxide reductase subunit C
 ahpF – NADH dehydrogenase
 alaA – Glutamate-pyruvate aminotransferase AlaA
 alr – Alanine racemase, biosynthetic
 alsT_1 – Amino-acid carrier protein AlsT
 alx – Inner membrane protein alx
 ansA – L-asparaginase 1
 ansB – L-asparaginase 2
 appA_2 – Oligopeptide-binding protein AppA
 apt – Adenine phosphoribosyltransferase
 arcA_1 – Aerobic respiration control protein ArcA
 arcA_2 – Arginine deiminase
 arcC1 – Carbamate kinase 1
 argA – Amino-acid acetyltransferase
 argB – Acetylglutamate kinase
 argC – N-acetyl-gamma-glutamyl-phosphate reductase
 argF_3 – Ornithine carbamoyltransferase
 argG – Argininosuccinate synthase
 argH – Argininosuccinate lyase
 argR_2 – Arginine repressor
 argS – Arginine--tRNA ligase
 argT – Lysine/arginine/ornithine-binding periplasmic protein
 aroA – 3-phosphoshikimate 1-carboxyvinyltransferase
 aroF – Phospho-2-dehydro-3-deoxyheptonate aldolase, Tyr-sensitive
 aroG – Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive
 aroH_1 – Phospho-2-dehydro-3-deoxyheptonate aldolase, Trp-sensitive
 aroK – Shikimate kinase 1
 aroQ – 3-dehydroquinone dehydratase
 artI – Putative ABC transporter arginine-binding protein 2
 artJ_2 – ABC transporter arginine-binding protein 1
 artP – Arginine transport ATP-binding protein ArtP
 artQ – Arginine ABC transporter permease protein ArtQ
 asd – Aspartate-semialdehyde dehydrogenase
 asd2 – Aspartate-semialdehyde dehydrogenase 2
 asnA – Aspartate--ammonia ligase
 asnB – Asparagine synthetase B [glutamine-hydrolyzing]
 asnC – Regulatory protein AsnC
 aspA_1 – Aspartate ammonia-lyase
 aspC – Aspartate aminotransferase
 aspS – Aspartate--tRNA ligase
 aspT_1 – Aspartate/alanine antiporter
 aspT_2 – Aspartate/alanine antiporter
 astA – Arginine N-succinyltransferase

atpA – ATP synthase subunit alpha
 atpC – ATP synthase epsilon chain
 atpD – ATP synthase subunit beta
 atpG – ATP synthase gamma chain
 avtA – Valine--pyruvate aminotransferase
 bauA_1 – Beta-alanine--pyruvate aminotransferase
 betS – Glycine betaine/proline betaine transporter BetS
 can – Carbonic anhydrase 2
 carA – Carbamoyl-phosphate synthase small chain
 carB – Carbamoyl-phosphate synthase large chain
 ccmC – Heme exporter protein C
 ccmE – Cytochrome c-type biogenesis protein CcmE
 ccoN1 – Cbb3-type cytochrome c oxidase subunit CcoN1
 cfxP – Phosphoribulokinase, plasmid
 cheR2 – Chemotaxis protein methyltransferase Cher2
 cheV_2 – Chemotaxis protein CheV
 cheW_2 – Chemotaxis protein CheW
 clpA – ATP-dependent Clp protease ATP-binding subunit ClpA
 clpB – Chaperone protein ClpB
 clpS – ATP-dependent Clp protease adapter protein ClpS
 clpX – ATP-dependent Clp protease ATP-binding subunit ClpX
 cmoM – tRNA 5-carboxymethoxyuridine methyltransferase
 coaA – Pantothenate kinase
 coaBC – Coenzyme A biosynthesis bifunctional protein CoaBC
 comM – Competence protein ComM
 corC_2 – Magnesium and cobalt efflux protein CorC
 crp – cAMP-activated global transcriptional regulator CRP
 crr – Glucose-specific phosphotransferase enzyme IIA component
 cspD – Cold shock-like protein CspD
 cspV_1 – Cold shock protein CspV
 cspV_2 – Cold shock protein CspV
 csrA – Carbon storage regulator
 cvpA – Colicin V production protein
 cysD – Sulfate adenylyltransferase subunit 2
 cysE – Serine acetyltransferase
 cysH – Phosphoadenosine phosphosulfate reductase
 cysI – Sulfite reductase [NADPH] hemoprotein beta-component
 cysK – Cysteine synthase A
 cysN – Sulfate adenylyltransferase subunit 1
 cysS – Cysteine--tRNA ligase
 dapA_1 – 4-hydroxy-tetrahydrodipicolinate synthase
 dapD – 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
 dapF – Diaminopimelate epimerase
 dcd – Deoxycytidine triphosphate deaminase
 dcuB – Anaerobic C4-dicarboxylate transporter DcuB
 ddpD – Oligopeptide transport ATP-binding protein OppD
 deoA – Thymidine phosphorylase
 deoB – Phosphopentomutase
 deoC – Deoxyribose-phosphate aldolase
 deoD – Purine nucleoside phosphorylase DeoD-type
 der – GTPase Der
 dhaK – PEP-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit DhaK
 dhaL – PEP-dependent dihydroxyacetone kinase, ADP-binding subunit DhaL
 diaA – DnaA initiator-associating protein DiaA
 dksA_2 – RNA polymerase-binding transcription factor DksA
 dksA_3 – RNA polymerase-binding transcription factor DksA
 dmlR_1 – HTH-type transcriptional regulator DmlR
 dmlR_10 – HTH-type transcriptional regulator DmlR
 dnaA – Chromosomal replication initiator protein DnaA
 dnaB – Replicative DNA helicase

dnaE – DNA polymerase III subunit alpha
 dnaG – DNA primase
 dnaJ_1 – Chaperone protein DnaJ
 dnaK_1 – Chaperone protein DnaK
 dnaN – DNA polymerase III subunit beta
 dpp5 – Dipeptidyl-peptidase 5
 dppB – Dipeptide transport system permease protein DppB
 dppC_1 – Dipeptide transport system permease protein DppC
 dps – DNA protection during starvation protein
 dus – putative tRNA-dihydrouridine synthase
 eda – KHG/KDPG aldolase
 efp – Elongation factor P
 eno – Enolase
 epd – D-erythrose-4-phosphate dehydrogenase
 epsE_2 – Type II secretion system protein E
 epsE_3 – Type II secretion system protein E
 era – GTPase Era
 erpA – Iron-sulfur cluster insertion protein ErpA
 etfA – Electron transfer flavoprotein subunit alpha
 ettA – Energy-dependent translational throttle protein EttA
 exbB_1 – Biopolymer transport protein ExbB
 exbD_2 – Biopolymer transport protein ExbD
 exbD_3 – Biopolymer transport protein ExbD
 fabA – 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase
 fabB – 3-oxoacyl-[acyl-carrier-protein] synthase 1
 fabD – Malonyl CoA-acyl carrier protein transacylase
 fabF_1 – 3-oxoacyl-[acyl-carrier-protein] synthase 2
 fabG_4 – 3-oxoacyl-[acyl-carrier-protein] reductase FabG
 fabH – 3-oxoacyl-[acyl-carrier-protein] synthase 3
 fabV_2 – Enoyl-[acyl-carrier-protein] reductase [NADH] 1
 fabZ – 3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ
 fadA – 3-ketoacyl-CoA thiolase
 fadB – Fatty acid oxidation complex subunit alpha
 fadD – Long-chain-fatty-acid--CoA ligase
 fadE_1 – Acyl-coenzyme A dehydrogenase
 fadI – 3-ketoacyl-CoA thiolase
 fbp – Fructose-1,6-bisphosphatase class 1
 fdx – 2Fe-2S ferredoxin
 ffh – Signal recognition particle protein
 fkIB – FKBP-type 22 kDa peptidyl-prolyl cis-trans isomerase
 fkpA_1 – putative FKBP-type peptidyl-prolyl cis-trans isomerase FkpA
 fldA – Flavodoxin-1
 flgC – Flagellar basal-body rod protein FlgC
 fnr – Fumarate and nitrate reduction regulatory protein
 frdC – Fumarate reductase subunit C
 frdD – Fumarate reductase subunit D
 fre – NAD(P)H-flavin reductase
 frmA – S-(hydroxymethyl)glutathione dehydrogenase
 frr – Ribosome-recycling factor
 ftnA – putative bacterial non-heme ferritin
 ftsA_1 – Cell division protein FtsA
 ftsE – Cell division ATP-binding protein FtsE
 ftsH – ATP-dependent zinc metalloprotease FtsH
 ftsI – Peptidoglycan synthase FtsI
 ftsX – Cell division protein FtsX
 ftsY – Signal recognition particle receptor FtsY
 ftsZ – Cell division protein FtsZ
 fur – Ferric uptake regulation protein
 fusA – Elongation factor G
 futA1 – Iron uptake protein A1

gapA – Glyceraldehyde-3-phosphate dehydrogenase A
gcvA_3 – Glycine cleavage system transcriptional activator
gcvH – Glycine cleavage system H protein
gcvP – Glycine dehydrogenase (decarboxylating)
gdhB – NAD-specific glutamate dehydrogenase
glgC_1 – Glucose-1-phosphate adenylyltransferase
glmM – Phosphoglucosamine mutase
glmS – Glutamine--fructose-6-phosphate aminotransferase [isomerizing]
glnA – Glutamine synthetase
glnB – Nitrogen regulatory protein P-II
glnG – Nitrogen regulation protein NR(I)
glnK – Nitrogen regulatory protein P-II 2
glnL – Nitrogen regulation protein NR(II)
glnS – Glutamine--tRNA ligase
glpC – Anaerobic glycerol-3-phosphate dehydrogenase subunit C
glpF – Glycerol uptake facilitator protein
glpK – Glycerol kinase
glpR – Glycerol-3-phosphate regulon repressor
glpX – Fructose-1,6-bisphosphatase 1 class 2
glsA2 – Glutaminase 2
gltA – Citrate synthase
gltB – Glutamate synthase [NADPH] large chain
gltX – Glutamate--tRNA ligase
glyA – Serine hydroxymethyltransferase
glyQ – Glycine--tRNA ligase alpha subunit
glyS – Glycine--tRNA ligase beta subunit
gmhA – Phosphoheptose isomerase
gntP – High-affinity gluconate transporter
gor – Glutathione reductase
gpml – 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
gpsA – Glycerol-3-phosphate dehydrogenase [NAD(P)+]
gpt – Xanthine phosphoribosyltransferase
grcA – Autonomous glycy radical cofactor
greA – Transcription elongation factor GreA
groL – 60 kDa chaperonin
groS – 10 kDa chaperonin
group_1024 – hypothetical protein
group_1025 – hypothetical protein
group_1059 – hypothetical protein
group_1081 – hypothetical protein
group_1111 – Aspartate kinase
group_1132 – hypothetical protein
group_1134 – PhoH-like protein
group_115 – hypothetical protein
group_1229 – hypothetical protein
group_1290 – Purine nucleoside phosphoramidase
group_1317 – hypothetical protein
group_1406 – hypothetical protein
group_1523 – hypothetical protein
group_1568 – hypothetical protein
group_1576 – hypothetical protein
group_1593 – hypothetical protein
group_1606 – hypothetical protein
group_1708 – hypothetical protein
group_1802 – hypothetical protein
group_1819 – hypothetical protein
group_1828 – hypothetical protein
group_1916 – hypothetical protein
group_1959 – hypothetical protein
group_1964 – RNA-binding protein

group_1975 – hypothetical protein
group_2339 – putative acyl-CoA thioester hydrolase
group_249 – hypothetical protein
group_2508 – hypothetical protein
group_2515 – hypothetical protein
group_2577 – hypothetical protein
group_2636 – hypothetical protein
group_2677 – hypothetical protein
group_2896 – hypothetical protein
group_293 – hypothetical protein
group_2972 – hypothetical protein
group_3014 – hypothetical protein
group_3030 – hypothetical protein
group_305 – hypothetical protein
group_306 – hypothetical protein
group_348 – hypothetical protein
group_355 – hypothetical protein
group_3714 – hypothetical protein
group_3836 – hypothetical protein
group_394 – hypothetical protein
group_3988 – hypothetical protein
group_4087 – Glycine cleavage system transcriptional activator
group_4108 – hypothetical protein
group_530 – hypothetical protein
group_536 – hypothetical protein
group_652 – hypothetical protein
group_690 – hypothetical protein
group_710 – hypothetical protein
group_735 – hypothetical protein
group_748 – hypothetical protein
group_754 – hypothetical protein
group_755 – hypothetical protein
group_759 – hypothetical protein
group_771 – hypothetical protein
group_790 – hypothetical protein
group_805 – hypothetical protein
group_808 – hypothetical protein
group_810 – hypothetical protein
group_821 – hypothetical protein
group_867 – hypothetical protein
group_887 – hypothetical protein
group_900 – hypothetical protein
group_910 – hypothetical protein
group_9781 – hypothetical protein
gshB – Glutathione synthetase
gsiA_2 – Glutathione import ATP-binding protein GsiA
gsiC_1 – Glutathione transport system permease protein GsiC
guaA – GMP synthase [glutamine-hydrolyzing]
guaB_2 – Inosine-5'-monophosphate dehydrogenase
gyrB – DNA gyrase subunit B
hda – DnaA regulatory inactivator Hda
hemE – Uroporphyrinogen decarboxylase
hemL – Glutamate-1-semialdehyde 2,1-aminomutase
hexR – HTH-type transcriptional regulator HexR
hflC – Modulator of FtsH protease HflC
hflD – High frequency lysogenization protein HflD
hflK_2 – Modulator of FtsH protease HflK
hflX – GTPase HflX
hfq_1 – RNA-binding protein Hfq
hisG – ATP phosphoribosyltransferase

hisM – Histidine transport system permease protein HisM
 hisQ – Histidine transport system permease protein HisQ
 hisS – Histidine--tRNA ligase
 hldD – ADP-L-glycero-D-manno-heptose-6-epimerase
 hldE – Bifunctional protein HldE
 hpf – Ribosome hibernation promoting factor
 hpt – Hypoxanthine phosphoribosyltransferase
 hslU – ATP-dependent protease ATPase subunit HslU
 hslV – ATP-dependent protease subunit HslV
 htpX – Protease HtpX
 hupA – DNA-binding protein HU-alpha
 hutH – Histidine ammonia-lyase
 hutU – Urocanate hydratase
 ibaG – Acid stress protein IbaG
 ibpA_1 – Small heat shock protein IbpA
 icd_2 – Isocitrate dehydrogenase [NADP]
 iga – IgA-specific serine endopeptidase autotransporter
 ihfA – Integration host factor subunit alpha
 ihfB – Integration host factor subunit beta
 ileS – Isoleucine--tRNA ligase
 ilvA – L-threonine dehydratase biosynthetic IlvA
 ilvC – Ketol-acid reductoisomerase
 ilvD – Dihydroxy-acid dehydratase
 ilvH – Acetolactate synthase isozyme 3 small subunit
 ilvI – Acetolactate synthase isozyme 3 large subunit
 infA – Translation initiation factor IF-1
 infB – Translation initiation factor IF-2
 iscA – Iron-binding protein IscA
 iscR – HTH-type transcriptional regulator IscR
 iscS – Cysteine desulfurase IscS
 iscU – Iron-sulfur cluster assembly scaffold protein IscU
 ispB – Octaprenyl-diphosphate synthase
 ispF – 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
 ispG – 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)
 ispH – 4-hydroxy-3-methylbut-2-enyl diphosphate reductase
 kdsA – 2-dehydro-3-deoxyphosphooctonate aldolase
 ktrA – Ktr system potassium uptake protein A
 ldc – Lysine/ornithine decarboxylase
 lepA – Elongation factor 4
 leuA_2 – 2-isopropylmalate synthase
 leuB – 3-isopropylmalate dehydrogenase
 leuC – 3-isopropylmalate dehydratase large subunit
 leuS – Leucine--tRNA ligase
 lexA_1 – LexA repressor
 lipA – Lipoyl synthase
 livF – High-affinity branched-chain amino acid transport ATP-binding protein LivF
 livH – High-affinity branched-chain amino acid transport system permease protein LivH
 livJ – Leu/Ile/Val/Thr-binding protein
 lon_2 – Lon protease
 lpdA – Dihydrolipoyl dehydrogenase
 lpoB_2 – Penicillin-binding protein activator LpoB
 lptB_1 – Lipopolysaccharide export system ATP-binding protein LptB
 lpxA – Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
 lpxC – UDP-3-O-acyl-N-acetylglucosamine deacetylase
 lpxD – UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase
 lrp_1 – Leucine-responsive regulatory protein
 lrp_2 – Leucine-responsive regulatory protein
 lrp_3 – Leucine-responsive regulatory protein
 luxS – S-ribosylhomocysteine lyase
 lysA – Diaminopimelate decarboxylase

lysN – 2-aminoadipate transaminase
 lysO – Lysine exporter LysO
 maeA_2 – NAD-dependent malic enzyme
 malE – Maltose-binding periplasmic protein
 malG – Maltose transport system permease protein MalG
 malK – Maltose/maltodextrin import ATP-binding protein MalK
 malP_1 – Maltodextrin phosphorylase
 malZ – Maltodextrin glucosidase
 map – Methionine aminopeptidase
 mazG – Nucleoside triphosphate pyrophosphohydrolase
 mdh_1 – Malate dehydrogenase
 mdlY – Mandelamide hydrolase
 menB – 1,4-dihydroxy-2-naphthoyl-CoA synthase
 metB – Cystathionine gamma-synthase
 metE – 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase
 metH – Methionine synthase
 metJ – Met repressor
 metK – S-adenosylmethionine synthase
 metL – Bifunctional aspartokinase/homoserine dehydrogenase 2
 mgIA – Galactose/methyl galactoside import ATP-binding protein MglA
 minD – Septum site-determining protein MinD
 mlaE – putative phospholipid ABC transporter permease protein MlaE
 mlaF – putative phospholipid import ATP-binding protein MlaF
 mnmA – tRNA-specific 2-thiouridylase MnmA
 mnmE – tRNA modification GTPase MnmE
 mnmG – tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG
 mntH – Divalent metal cation transporter MntH
 mpl – UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptandioate ligase
 mraY – Phospho-N-acetylmuramoyl-pentapeptide-transferase
 mrdB – Rod shape-determining protein RodA
 mreB – Rod shape-determining protein MreB
 mreC – Cell shape-determining protein MreC
 mscS – Small-conductance mechanosensitive channel
 mtnN – 5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase
 mukB – Chromosome partition protein MukB
 mukE – Chromosome partition protein MukE
 mukF – Chromosome partition protein MukF
 murA – UDP-N-acetylglucosamine 1-carboxyvinyltransferase
 murC – UDP-N-acetylmuramate--L-alanine ligase
 murD – UDP-N-acetylmuramoylalanine--D-glutamate ligase
 mutS – DNA mismatch repair protein MutS
 nadB – L-aspartate oxidase
 nadK – NAD kinase
 nagC_1 – N-acetylglucosamine repressor
 nagC_2 – N-acetylglucosamine repressor
 nagD – Ribonucleotide monophosphatase NagD
 narL – Nitrate/nitrite response regulator protein NarL
 nasD_2 – Nitrite reductase [NAD(P)H]
 ndk – Nucleoside diphosphate kinase
 nfuA – Fe/S biogenesis protein NfuA
 nhaR_2 – Transcriptional activator protein NhaR
 nirC – Nitrite transporter NirC
 nqrB – Na(+)-translocating NADH-quinone reductase subunit B
 nqrD – Na(+)-translocating NADH-quinone reductase subunit D
 nqrE_1 – Na(+)-translocating NADH-quinone reductase subunit E
 nqrF – Na(+)-translocating NADH-quinone reductase subunit F
 nrdA – Ribonucleoside-diphosphate reductase 1 subunit alpha
 nrdB – Ribonucleoside-diphosphate reductase 1 subunit beta
 nrdD – Anaerobic ribonucleoside-triphosphate reductase
 nrfA – Cytochrome c-552

nsrR_1 – HTH-type transcriptional repressor NsrR
nth – Endonuclease III
nuoB – NADH-quinone oxidoreductase subunit B
nuoH – NADH-quinone oxidoreductase subunit H
nuoK – NADH-quinone oxidoreductase subunit K
nupC – Nucleoside permease NupC
nupX_1 – Putative nucleoside permease NupX
nusB – N utilization substance protein B
obg – GTPase Obg
ompR_2 – Transcriptional regulatory protein OmpR
ompR_3 – Transcriptional regulatory protein OmpR
oppB – Oligopeptide transport system permease protein OppB
oppD_3 – Oligopeptide transport ATP-binding protein OppD
oppF_1 – Oligopeptide transport ATP-binding protein OppF
oppF_2 – Oligopeptide transport ATP-binding protein OppF
orn – Oligoribonuclease
pal – Peptidoglycan-associated lipoprotein
parB – putative chromosome-partitioning protein ParB
parC – DNA topoisomerase 4 subunit A
parE – DNA topoisomerase 4 subunit B
pckA – Phosphoenolpyruvate carboxykinase [ATP]
pcm – Protein-L-isoaspartate O-methyltransferase
pdhR – Pyruvate dehydrogenase complex repressor
pdxJ – Pyridoxine 5'-phosphate synthase
pepA_1 – Cytosol aminopeptidase
petA – Ubiquinol-cytochrome c reductase iron-sulfur subunit
petB – Cytochrome b
pfkA – ATP-dependent 6-phosphofructokinase isozyme 1
pflB – Formate acetyltransferase 1
pgk – Phosphoglycerate kinase
pgm – Phosphoglucomutase
pgsA – CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase
pheA – P-protein
pheS – Phenylalanine--tRNA ligase alpha subunit
pheT – Phenylalanine--tRNA ligase beta subunit
phoB – Phosphate regulon transcriptional regulatory protein PhoB
phoU – Phosphate-specific transport system accessory protein PhoU
pilT_2 – Twitching mobility protein
plsB – Glycerol-3-phosphate acyltransferase
plsX – Phosphate acyltransferase
pmbA_1 – Metalloprotease PmbA
pncB – Nicotinate phosphoribosyltransferase
pnp – Polyribonucleotide nucleotidyltransferase
pntA – NAD(P) transhydrogenase subunit alpha
pntB – NAD(P) transhydrogenase subunit beta
potA_4 – Spermidine/putrescine import ATP-binding protein PotA
potH – Putrescine transport system permease protein PotH
ppa – Inorganic pyrophosphatase
ppiB_1 – Peptidyl-prolyl cis-trans isomerase B
ppk – Polyphosphate kinase
prc – Tail-specific protease
prfB – Peptide chain release factor 2
prfC – Peptide chain release factor 3
priA – Primosomal protein N'
prmA – Ribosomal protein L11 methyltransferase
prmB – 50S ribosomal protein L3 glutamine methyltransferase
proB – Glutamate 5-kinase
proS – Proline--tRNA ligase
prs – Ribose-phosphate pyrophosphokinase
pspB_1 – Phage shock protein B

pstA – Phosphate transport system permease protein PstA
 pstB – Phosphate import ATP-binding protein PstB
 pstS – Phosphate-binding protein PstS
 psuG – Pseudouridine-5'-phosphate glycosidase
 pta – Phosphate acetyltransferase
 pth – Peptidyl-tRNA hydrolase
 ptsG_1 – PTS system glucose-specific EIICB component
 ptsH – Phosphocarrier protein HPr
 ptsI_2 – Phosphoenolpyruvate-protein phosphotransferase
 ptsN – Nitrogen regulatory protein
 ptsO – Phosphocarrier protein NPr
 purB – Adenylosuccinate lyase
 purF – Amidophosphoribosyltransferase
 purL – Phosphoribosylformylglycinamide synthase
 purM – Phosphoribosylformylglycinamide cyclo-ligase
 purU – Formyltetrahydrofolate deformylase
 puuA – Gamma-glutamylputrescine synthetase PuuA
 puuC – Aldehyde dehydrogenase PuuC
 pykA – Pyruvate kinase II
 pyrB – Aspartate carbamoyltransferase catalytic chain
 pyrE – Orotate phosphoribosyltransferase
 pyrG – CTP synthase
 pyrH – Uridylate kinase
 qseB_2 – Transcriptional regulatory protein QseB
 queA – S-adenosylmethionine:tRNA ribosyltransferase-isomerase
 radA – DNA repair protein RadA
 rapA_1 – RNA polymerase-associated protein RapA
 rarA – Replication-associated recombination protein A
 rbgA – Ribosome biogenesis GTPase A
 rbsB – Ribose import binding protein RbsB
 rbsC_2 – Ribose import permease protein RbsC
 rbsR – Ribose operon repressor
 rdgB – dITP/XTP pyrophosphatase
 rdgC – Recombination-associated protein RdgC
 recA – Protein RecA
 recF_1 – DNA replication and repair protein RecF
 recJ – Single-stranded-DNA-specific exonuclease RecJ
 recO – DNA repair protein RecO
 recQ_1 – ATP-dependent DNA helicase RecQ
 recR – Recombination protein RecR
 relA – GTP pyrophosphokinase
 rep – ATP-dependent DNA helicase Rep
 rfbD_2 – UDP-galactopyranose mutase
 rhIE_2 – ATP-dependent RNA helicase RhIE
 ribB – 3,4-dihydroxy-2-butanone 4-phosphate synthase
 ribBA – Riboflavin biosynthesis protein RibBA
 ribF – Riboflavin biosynthesis protein RibF
 ribN_1 – Riboflavin transporter
 rimK_1 – Ribosomal protein S6--L-glutamate ligase
 rimP – Ribosome maturation factor RimP
 rimH – Ribosomal RNA large subunit methyltransferase H
 rimJ – Ribosomal RNA large subunit methyltransferase J
 rimL – Ribosomal RNA large subunit methyltransferase K/L
 rluA_1 – Ribosomal large subunit pseudouridine synthase A
 rluC_1 – Ribosomal large subunit pseudouridine synthase C
 rluD – Ribosomal large subunit pseudouridine synthase D
 rnc – Ribonuclease 3
 rng – Ribonuclease G
 rnpA – Ribonuclease P protein component
 rnn – Ribonuclease R

rnt – Ribonuclease T
rpe – Ribulose-phosphate 3-epimerase
rph – Ribonuclease PH
rplA – 50S ribosomal protein L1
rplB – 50S ribosomal protein L2
rplC – 50S ribosomal protein L3
rplD – 50S ribosomal protein L4
rplE – 50S ribosomal protein L5
rplF – 50S ribosomal protein L6
rplI – 50S ribosomal protein L9
rplJ – 50S ribosomal protein L10
rplK – 50S ribosomal protein L11
rplL – 50S ribosomal protein L7/L12
rplM – 50S ribosomal protein L13
rplN – 50S ribosomal protein L14
rplO – 50S ribosomal protein L15
rplP – 50S ribosomal protein L16
rplQ – 50S ribosomal protein L17
rplS – 50S ribosomal protein L19
rplT – 50S ribosomal protein L20
rplU – 50S ribosomal protein L21
rplV – 50S ribosomal protein L22
rplW – 50S ribosomal protein L23
rpmB – 50S ribosomal protein L28
rpmC – 50S ribosomal protein L29
rpmD – 50S ribosomal protein L30
rpmE – 50S ribosomal protein L31
rpmF – 50S ribosomal protein L32
rpmG – 50S ribosomal protein L33
rpmH – 50S ribosomal protein L34
rpmI – 50S ribosomal protein L35
rpoA – DNA-directed RNA polymerase subunit alpha
rpoB – DNA-directed RNA polymerase subunit beta
rpoC – DNA-directed RNA polymerase subunit beta'
rpoD – RNA polymerase sigma factor RpoD
rpoE_1 – ECF RNA polymerase sigma-E factor
rpoH – RNA polymerase sigma factor RpoH
rpoS – RNA polymerase sigma factor RpoS
rpoZ – DNA-directed RNA polymerase subunit omega
rppH_1 – RNA pyrophosphohydrolase
rpsA_2 – 30S ribosomal protein S1
rpsB – 30S ribosomal protein S2
rpsC – 30S ribosomal protein S3
rpsD – 30S ribosomal protein S4
rpsE – 30S ribosomal protein S5
rpsF – 30S ribosomal protein S6
rpsG – 30S ribosomal protein S7
rpsH – 30S ribosomal protein S8
rpsI – 30S ribosomal protein S9
rpsJ – 30S ribosomal protein S10
rpsK – 30S ribosomal protein S11
rpsL – 30S ribosomal protein S12
rpsM – 30S ribosomal protein S13
rpsN – 30S ribosomal protein S14
rpsO – 30S ribosomal protein S15
rpsP – 30S ribosomal protein S16
rpsQ – 30S ribosomal protein S17
rpsR – 30S ribosomal protein S18
rpsS – 30S ribosomal protein S19
rpsT – 30S ribosomal protein S20

rpsU – 30S ribosomal protein S21
rraA – Regulator of ribonuclease activity A
rsmA – Ribosomal RNA small subunit methyltransferase A
rsmB_1 – Ribosomal RNA small subunit methyltransferase B
rsmE – Ribosomal RNA small subunit methyltransferase E
rsmG – Ribosomal RNA small subunit methyltransferase G
rsmH – Ribosomal RNA small subunit methyltransferase H
rsmI – Ribosomal RNA small subunit methyltransferase I
sarZ – HTH-type transcriptional regulator MhqR
satP – Succinate-acetate/proton symporter SatP
sdaA_1 – L-serine dehydratase 1
sdcS_1 – Sodium-dependent dicarboxylate transporter SdcS
sdhA – Succinate dehydrogenase flavoprotein subunit
sdhB – Succinate dehydrogenase iron-sulfur subunit
secA – Protein translocase subunit SecA
secB – Protein-export protein SecB
secE – Protein translocase subunit SecE
secY – Protein translocase subunit SecY
serA – D-3-phosphoglycerate dehydrogenase
serC – Phosphoserine aminotransferase
serS – Serine--tRNA ligase
skp – Chaperone protein Skp
slmA – Nucleoid occlusion factor SlmA
slyD – FKBP-type peptidyl-prolyl cis-trans isomerase SlyD
smg – Protein Smg
sodB_2 – Superoxide dismutase [Fe]
soj – Sporulation initiation inhibitor protein Soj
speA – Biosynthetic arginine decarboxylase
spoT – Bifunctional (p)ppGpp synthase/hydrolase SpoT
srlR_3 – Glucitol operon repressor
srmB – ATP-dependent RNA helicase SrmB
sspA – Stringent starvation protein A
sspB – Stringent starvation protein B
sucA – 2-oxoglutarate dehydrogenase E1 component
sucC – Succinyl-CoA ligase [ADP-forming] subunit beta
sucD – Succinyl-CoA ligase [ADP-forming] subunit alpha
suhB – Inositol-1-monophosphatase
talB – Transaldolase B
tatD – 3'-5' ssDNA/RNA exonuclease TatD
tcdA – tRNA threonylcarbamoyladenine dehydratase
tcyC_2 – L-cystine import ATP-binding protein TcyC
tdh – L-threonine 3-dehydrogenase
tdk – Thymidine kinase
thrS – Threonine--tRNA ligase
thyA – Thymidylate synthase
tig – Trigger factor
tktA – Transketolase 1
tldD_1 – Metalloprotease TldD
tmpC – Membrane lipoprotein TmpC
topA_2 – DNA topoisomerase 1
tpiA_1 – Triosephosphate isomerase
treB – PTS system trehalose-specific EIIBC component
trkA – Trk system potassium uptake protein TrkA
trkI – Trk system potassium uptake protein TrkI
trmB – tRNA (guanine-N(7)-)-methyltransferase
trmD – tRNA (guanine-N(1)-)-methyltransferase
trpA – Tryptophan synthase alpha chain
trpB – Tryptophan synthase beta chain
trpG_1 – Anthranilate synthase component 2
trpS – Tryptophan--tRNA ligase

truB – tRNA pseudouridine synthase B
trxA_1 – Thioredoxin-1
tsaD – tRNA N6-adenosine threonylcarbamoyltransferase
tsf – Elongation factor Ts
ttcA – tRNA 2-thiocytidine biosynthesis protein TtcA
tusA_2 – Sulfurtransferase TusA
typA – GTP-binding protein TypA/BipA
tyrB – Aromatic-amino-acid aminotransferase
tyrS_2 – Tyrosine--tRNA ligase
ubiB – putative protein kinase UbiB
ubiD – 3-octaprenyl-4-hydroxybenzoate carboxy-lyase
ubiE – Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE
udk – Uridine kinase
udp_2 – Uridine phosphorylase
upp – Uracil phosphoribosyltransferase
uvrB – UvrABC system protein B
uvrC – UvrABC system protein C
uvrD – DNA helicase II
valS – Valine--tRNA ligase
xanP – Xanthine permease XanP
xerD_1 – Tyrosine recombinase XerD
yadH – Inner membrane transport permease YadH
ybaK – Cys-tRNA(Pro)/Cys-tRNA(Cys) deacylase YbaK
ybiT – putative ABC transporter ATP-binding protein YbiT
ycaO – Ribosomal protein S12 methylthiotransferase accessory factor YcaO
ycca – Modulator of FtsH protease Ycca
ychF – Ribosome-binding ATPase YchF
yciK – putative oxidoreductase YciK
ydcP_1 – putative protease YdcP
ydcV_2 – Inner membrane ABC transporter permease protein YdcV
ydgI_1 – Putative arginine/ornithine antiporter
yebC – putative transcriptional regulatory protein YebC
ygdH – LOG family protein YgdH
yhbU – putative protease YhbU
yhhQ – Inner membrane protein YhhQ
yhjX – putative MFS-type transporter YhjX
yidC – Membrane protein insertase YidC
yjiA – putative GTP-binding protein YjiA
yofA – HTH-type transcriptional regulator YofA
ypdA – Sensor histidine kinase YpdA
yqgF – Putative pre-16S rRNA nuclease
yxeM – putative amino-acid-binding protein YxeM
yxeN – putative amino-acid permease protein YxeN
zapB – Cell division protein ZapB
znuB – High-affinity zinc uptake system membrane protein ZnuB
znuC – Zinc import ATP-binding protein ZnuC

ANEXO F – ÁRVORE FILOGENÉTICA DO CORE GENOMA

