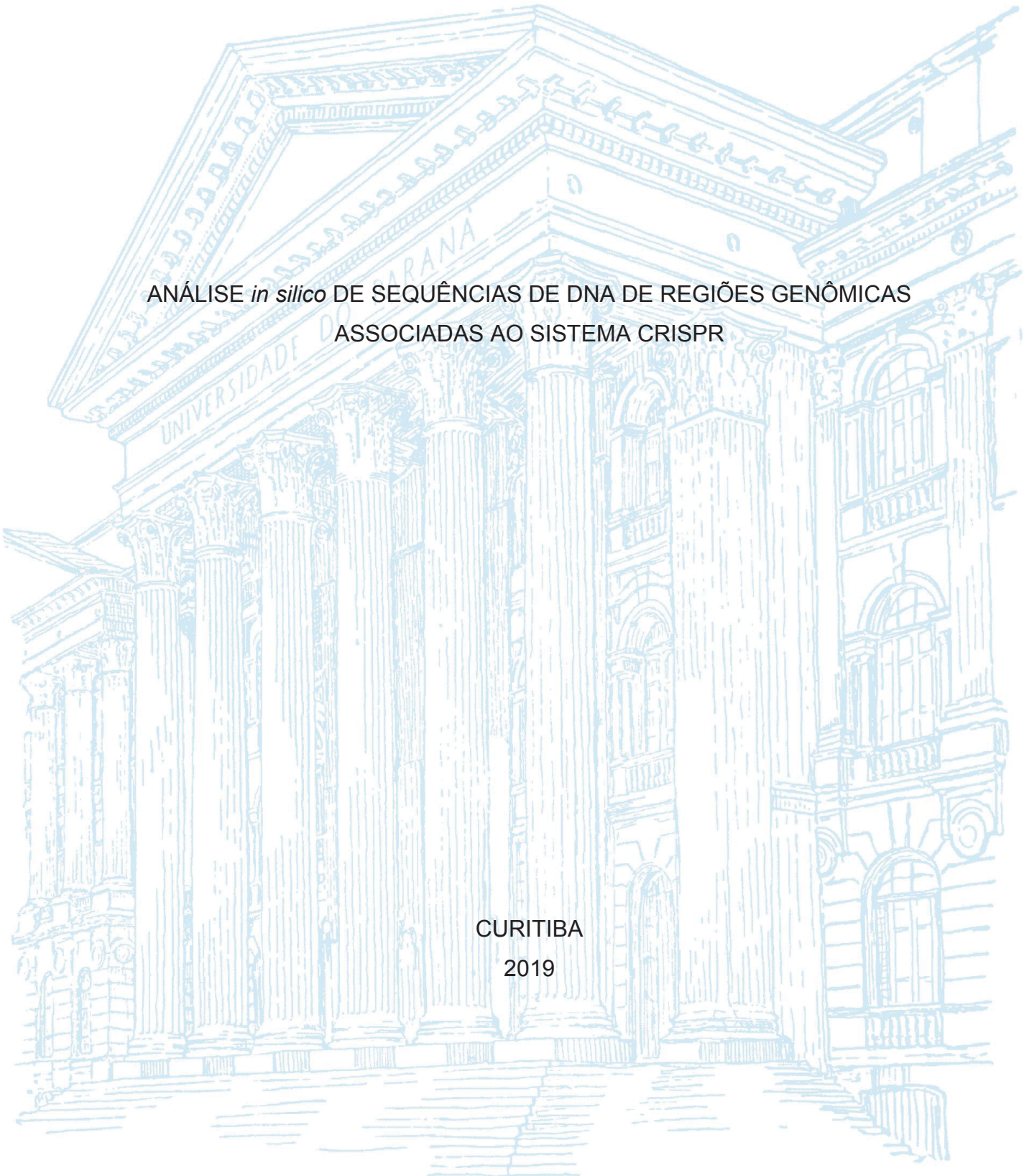
UNIVERSIDADE FEDERAL DO PARANÁ

HELLEN CRISTINE MACHADO



ANÁLISE *in silico* DE SEQUÊNCIAS DE DNA DE REGIÕES GENÔMICAS ASSOCIADAS AO SISTEMA CRISPR

CURITIBA

2019

HELLEN CRISTINE MACHADO

ANÁLISE *in silico* DE SEQUÊNCIAS DE DNA DE REGIÕES GENÔMICAS ASSOCIADAS AO SISTEMA CRISPR

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Bioinformática, Setor de Educação Profissional e Tecnológica, da Universidade Federal do Paraná como requisito parcial para a obtenção do grau de Mestre em Bioinformática.

Orientador: Prof. Dr. Dieval Guizelini

CURITIBA

2019

FICHA CATALOGRÁFICA

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

Machado, Hellen Cristine
M149 Análise *in silico* de sequências de DNA de regiões genômicas associadas ao sistema CRISPR / Hellen Cristine Machado. - Curitiba, 2019.
80 p.: il.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Educação Profissional e Tecnológica, Curso de Pós-Graduação em Bioinformática, 2019.
Orientador: Dieval Guizelini

1. Sequência de nucleotídios. 2. Algoritmos genéticos. 3. Bioinformática. I. Guizelini, Dieval. II. Título. III. Universidade Federal do Paraná.

CDD 575.113

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

TERMO DE APROVAÇÃO

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

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

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em BIOINFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **HELLEN CRISTINE MACHADO** intitulada: **"Análise *in silico* de sequências de DNA de regiões genômicas associadas ao sistema CRISPR"**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 25 de fevereiro de 2019.

Dr. Dieval Guizelini
Presidente/Programa de Pós-graduação em Bioinformática – UFPR

Drª. Saloê Bispo Poubel
Avaliadora Externa/Laboratório de Parasitologia Molecular - UFPR

Dr. Roberto Tadeu Raittz
Avaliador Interno/Programa de Pós-graduação em Bioinformática – UFPR

Dedico este trabalho a Deus, criador
de todos os mistérios da vida, sem
os quais a ciência não teria sentido.

AGRADECIMENTOS

A Deus:

Por ser meu refúgio, minha fortaleza e meu guia. Por me mostrar que não estou sozinha e segurar minha mão mesmo nos momentos onde pareço perder minha fé.

Aos meus pais, Vera e Aristides:

Por sempre me mostrarem o caminho do bem, do amor e do respeito. Por sempre me incentivarem a correr atrás dos meus sonhos acima de tudo. Por serem meu exemplo de construção de família. Foi tudo por vocês, sempre!

Ao meu amor, Du:

Por ser o amor da minha vida. Meu melhor amigo, minha melhor companhia, o dono do melhor abraço e dos meus melhores sorrisos. Se estou aqui hoje é porque você trilhou esse caminho comigo, me incentivando, apoiando, enxugando minhas lágrimas e me puxando pra cima quando o desânimo bateu. Sem você eu não conseguiria. Obrigada! Por me amar, me acompanhar, aguentar meus “surtos” e, óbvio, ser o responsável por me apresentar a doce área da Bioinformática. Eu te amo sempre, pra sempre e sempre mais!

Aos meus irmãos, Lili e Helder:

Obrigada, Lili, por ter sido meu espelho durante a minha infância, por me mostrar que mesmo nos erros podemos encontrar o melhor caminho; por ser meu ombro amigo e minha conselheira mesmo a 75km de distância. Helder, obrigada por, mesmo sem saber, ser um motivador da minha vida acadêmica, principalmente mostrando que nunca é tarde pra recomeçar. E, obrigada por junto da Michele, ter trazido os bens mais preciosos da nossa família, o nosso polaco lindo e a nossa princesa que está chegando.

Aos meus pequenos, Murilo, Anthony, Melissa e Isabela:

Por serem meus potinhos de alegria. Por trazerem felicidade, paz e amor para o meu coração, às vezes angustiado e temeroso. Que um dia vocês cresçam e se orgulhem da tia/dinda de vocês.

Às famílias Marangoni e Tieppo:

Que sempre me acolheram com muito amor e se tornaram minha família.

Às minhas meninas, Jéssica e Vanessa:

Por, mesmo com a distância, estarem sempre presentes em pensamento e coração. Obrigada pela torcida, pelas risadas e momentos de descontração. Mesmo raros nesse período, eles foram essenciais!

Aos presentes da bioinfo, Mariane, Sheyla e Aniele:

Vocês são o que de melhor a bioinfo me trouxe. Obrigada pelas sessões de terapia em grupo, pelas infinitas xícaras de café, pelos abraços apertados, pelos conselhos acadêmicos e da vida, e pela paciência em me ensinar coisas novas. Vocês são incríveis!

Ao meu orientador, Prof. Dr. Dieval Guizelini:

Por ter aceitado o desafio de me orientar, pela dedicação e envolvimento no projeto, e por me incentivar a “quebrar a resistência” e sair da zona de conforto.

Aos professores do PPG em Bioinformática:

Por todo ensinamento passado nesses dois anos.

À secretaria do programa, em especial à Suzana:

Por sempre estar pronta para intervir, resolver e solucionar todos os tipos de problemas e dúvidas, sempre com um carinho enorme e um abraço apertado.

À CAPES e órgãos de fomento:

Pelo auxílio financeiro.

~ It is our choices that show what we truly are, far more than our abilities. ~

Albus Percival Wulfric Brian Dumbledore
Harry Potter and the Chamber of Secrets

RESUMO

As repetições palindrômicas curtas, interespaçadas e regularmente agrupadas – CRISPR – formam um sistema de imunidade adquirida em bactérias e arqueas. O CRISPR é um dos sistemas mais estudados na última década, especialmente como ferramenta de edição gênica, devido à sua capacidade de gerar *indels* em sequências alvo. Entretanto, as características nucleotídicas da região, a origem das sequências estruturais básicas e sua relação com outras estruturas conhecidas ainda são pouco descritas. Por isso, aqui nós mostramos uma análise exploratória *in silico* de sequências genômicas de regiões CRISPR em procariotos. Sequências de regiões CRISPR foram obtidas de diferentes bases de dados e foram agrupadas com a ferramenta RAFTS³G com critério de 50% de identidade. Os *clusters* formados foram confrontados com bases de dados públicas para predição de funções e estruturas biológicas. Os resultados indicam relação entre as sequências de repetição direta (DR) e outras estruturas, e há evidências de transferência horizontal de genes entre os domínios Bacteria, Archaea e Eukarya. As 7.081 sequências DR de bactérias agrupadas constituem 1.547 *clusters,* que compartilham 50% de identidade. Os maiores *clusters* são compostos por 1.001 (14%) e 140 sequências (2%), porém há baixa diversidade intracluster visto que esses grupos apresentam 30 e 32 sequências distintas, respectivamente. Já a análise de predição funcional indica que há grande similaridade entre sequências DR e estruturas conhecidas, como RNAs e alguns MGEs. Alinhamentos de sequências indicam a transferência horizontal de arranjos CRISPR entre *Bradyrhizobium sp. BTAi 1* e a espécie de trigo selvagem *Triticum urartu.* Já as sequências de espaçadores CRISPR produziram muitos agrupamentos, todos com poucos membros e baixa similaridade com os elementos genéticos móveis conhecidos, demonstrando que a origem dos espaçadores precisa ser esclarecida. Nosso estudo demonstra que os componentes principais do arranjo CRISPR – DR e espaçadores – estão relacionados com diferentes estruturas funcionais conhecidas. A abordagem desse trabalho produziu diversos grupos que precisam ainda ser analisados, no intuito de ampliar o conhecimento do arranjo CRISPR. Também, a origem das sequências DR e dos espaçadores não foi revelada; além disso, contrapondo o que é descrito na literatura, observamos que os genes *Cas1* e *Cas2* não são universais, e detectamos a presença de CRISPR em eucarioto, visto que até o momento a estrutura era descrita unicamente em procariotos.

Palavras-chave: CRISPR. Sequências de Repetição Direta. Espaçadores. Agrupamento.

ABSTRACT

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems form an acquired immunity system that are widespread in bacteria and archaea. CRISPR are one of largely studied systems in last decade, especially as genome editing tool due to its ability to generate indels in target sequences. However, nucleotide characteristics of this region, the origin of basic structural sequences and its relations with well-known structures are poorly described. Therefore, here we show an *in silico* exploratory analysis of genomic sequences of CRISPR regions in prokaryotes. CRISPR sequences were collected from different databases and clustered by RAFTS³G tool with 50% of identity. Resulting clusters were matched against public databases in order to predict biological functions and structures. Results indicate a relationship between direct repeat sequences (DR) and other structures, and we found evidences of horizontal gene transfer between the Bacteria, Archaea and Eukarya domains. The 7.081 clustered DR sequences formed 1.547 clusters, which share 50% identity. The largest clusters are composed by 1.001 (14%) and 140 sequences (2%), but have low intracluster diversity, with 30 and 32 distinct sequences, respectively. The functional prediction analysis suggests high similarity between DR sequences and well-known structures, such as RNAs and some MGEs. Sequence alignments indicate horizontal transfer of CRISPR arrays from *Bradyrhizobium sp. BTAi 1* to the wild wheat specie *Triticum urartu*. CRISPR spacers sequences resulted in a large number of clusters, all with few members and low similarity to known mobile genetic elements, indicating that the origin of the spacers needs to be elucidated. This study demonstrates that the main components of CRISPR array - DR and spacers - are closely related with well-known functional structures. The approach used in this research produced several clusters that still need to be analyzed in order to increase CRISPR arrays understanding. In addition, the origin of DR and spacers sequences was not found out; furthermore, in contrast with the literature, we observed that the Cas1 and Cas2 genes are not universal, and we detected the presence of CRISPR in eukaryote, whereas the structure was described only in prokaryotes.

Keywords: CRISPR. Direct Repeat Sequences. Spacers. Clustering.

LISTA DE FIGURAS

LISTA DE QUADROS

LISTA DE SIGLAS

ACLAME	Classification of Mobile Genetic Elements
Cas	CRISPR Associated
Cascade	CRISPR-Associated Complex for Antiviral Defense
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
CSV	Comma-Separated Values
DNA	Ácido Desoxirribonucleico
DR	Direct Repeat
dsDNA	Double-Stranded DNA
ENA	European Nucleotide Archive
HGT	Horizontal Gene Transfer
IG	Ilha Genômica
MGE	Mobile Genetic Element
mRNA	RNA Mensageiro
NCBI	National Center for Biotechnology Information
ncRNA	RNA Não-Codificante
PAM	Motivo Adjacente ao Protospaçador
pb	Par de Base
pré-crRNA	Pré-CRISPR RNA
RAFTS³G	Rapid Alignment Free Tool for Sequences Similarity Search to Groups
RNA	Ácido Ribonucleico
seqcons	Sequence Conservation
SQL	Structured Query Language
TALEN	Transcription Activator-Like Effector Nucleases
TM	Temperatura de Melting
UFPR	Universidade Federal do Paraná

SUMÁRIO

1 INTRODUÇÃO

CRISPR (do inglês, *Clustered Regularly Interspaced Short Palindromic Repeats*) são repetições palindrômicas curtas, interespaçadas e regularmente agrupadas que funcionam como um sistema imune adaptativo em procariotos (arqueas e bactérias) (BULT et al., 1996; HORVATH; BARRANGOU, 2010; ISHINO et al., 1987; SOREK; KUNIN; HUGENHOLTZ, 2008). Os organismos que apresentam essa estrutura são capazes de capturar trechos de sequências de elementos genéticos móveis (MGE – do inglês, *mobile genetic elements*) e incorporar à sua sequência como um espaçador (MOJICA et al., 2005; POURCEL; SALVIGNOL; VERGNAUD, 2006). Esses espaçadores são separados entre si por sequências repetitivas parcialmente palindrômicas, chamadas sequências DR (do inglês, *direct repeats*), e são os responsáveis pelo reconhecimento da sequência caso o mesmo elemento genético móvel tente novamente infectar aquele organismo (GRISSA; VERGNAUD; POURCEL, 2007; HORVATH; BARRANGOU, 2010; JANSEN et al., 2002; SOREK; KUNIN; HUGENHOLTZ, 2008).

Além da separação dos espaçadores e da formação de grampos durante o reconhecimento do alvo, não se tem muitas outras informações a respeito da importância e função das sequências DR (GRISSA; VERGNAUD; POURCEL, 2007; HORVATH; BARRANGOU, 2010; JANSEN et al., 2002; SOREK; KUNIN; HUGENHOLTZ, 2008), e diversas questões estão em aberto em relação a essas sequências. Entre elas, destaca-se: é o próprio organismo o formador da estrutura ou ela tem origem em uma transferência horizontal? Existe relação com regiões regulatórias ou regiões de reconhecimento?

O conhecimento sobre a estrutura e função do sistema CRISPR justifica a curiosidade sobre a origem da composição de suas sequências e de regiões de inserção dentro dos organismos procarióticos. Assim, buscou-se explorar as regiões de repetição do sistema CRISPR e compará-las com outras estruturas já bem descritas como diferentes tipos de RNA, para tentar inferir alguma nova função ou concordância entre as sequências DR e estruturas funcionalmente conhecidas.

2 REVISÃO BIBLIOGRÁFICA

2.1 CLUSTERED REGULARLY INTERSPACED SHORT PALINDROMIC REPEATS (CRISPR)

Em 1987, Ishino e colaboradores encontraram um padrão diferente de repetições dentro da sequência de *Escherichia coli K-12* durante ensaios sobre o gene responsável pela conversão da isoenzima da fosfatase alcalina. Esse padrão apresentava 14 repetições parcialmente palindrômicas de 29 pares de base (pb) interespaçadas por sequências não repetitivas de cerca de 33pb; mais tarde, essa estrutura recebeu o nome de CRISPR (ISHINO et al., 1987; JANSEN et al., 2002). O sistema CRISPR trata-se de repetições palindrômicas curtas, interespaçadas e regularmente agrupadas que se comportam como um sistema de imunidade adquirida em bactérias e arqueas, sendo encontrado em cerca de 40% e 90% dos genomas desses organismos, respectivamente. Quanto ao número de arranjos CRISPR vistos em um mesmo organismo, segundo a literatura o maior número de arranjos descritos é oriundo da arquea *Methanocaldococcus jannaschii* que possui 18 loci CRISPR distintos (BULT et al., 1996; HORVATH; BARRANGOU, 2010; ISHINO et al., 1987; SOREK; KUNIN; HUGENHOLTZ, 2008). Entretanto, a ferramenta CRISPRCasdb traz a bactéria *Moorea producens JHB* como portadora de 121 arranjos CRISPR, porém, apenas 4 deles apresentam grau de evidência maior que 1, um indicador de confiança de CRISPR dado pela ferramenta (COUVIN et al., 2018).

O arranjo CRISPR é caracterizado por várias repetições diretas, chamadas de sequências DR, que têm tamanho variável entre 23 e 55pb, separadas entre si por sequências de espaçadores que são derivadas do material genético exógeno, variando seu tamanho entre 21 e 72pb e geralmente estão adjacentes aos genes associados a CRISPR (*Cas – CRISPR-associated genes*), além de conter também uma sequência líder (GRISSA; VERGNAUD; POURCEL, 2007; HORVATH; BARRANGOU, 2010; JANSEN et al., 2002; SOREK; KUNIN; HUGENHOLTZ, 2008).

2.1.1 Composição da região CRISPR

As sequências DR têm tamanho variável entre 23 e 55pb e variam em quantidade entre 2 e 374 dentro do arranjo CRISPR (MARRAFFINI; SONTHEIMER, 2010). Essas sequências são bastante conservadas dentro de um mesmo lócus CRISPR, mas variam entre as diferentes espécies com relação ao número de

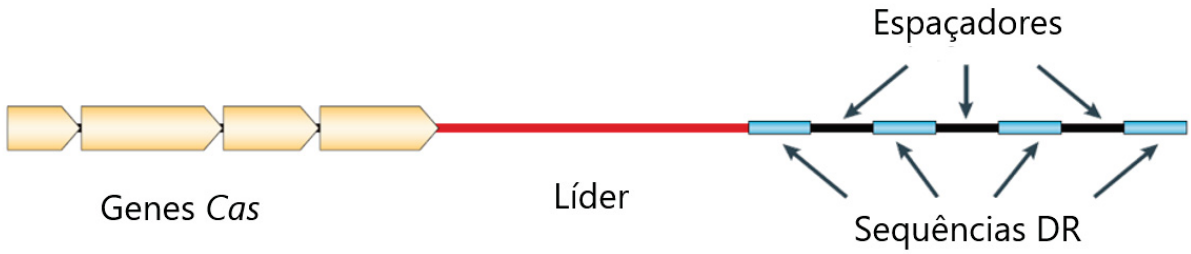
repetições, ao tamanho e ao padrão de sequência. Essa conservação vista nas sequências DR contribui para que sejam parcialmente palindrômicas (cerca de 7pb), potencializando a possibilidade de formar estruturas secundárias estáveis e conservadas, como a formação de *stem-loop* ou *hairpin*, importante para o reconhecimento e ação do CRISPR (GRISSA; VERGNAUD; POURCEL, 2007; HORVATH; BARRANGOU, 2010; JANSEN et al., 2002; SOREK; KUNIN; HUGENHOLTZ, 2008). Algumas sequências DR possuem no seu terminal 3' uma sequência conservada de GAAA (G/C), a qual acredita-se ter atuação como sítio de ligação para uma ou mais proteínas Cas (KUNIN; SOREK; HUGENHOLTZ, 2007).

Já os espaçadores são sequências com tamanho variável entre 21 e 72pb, vindas de elementos genéticos móveis exógenos como bacteriófagos, profagos, plasmídeos e transposons, responsáveis por infectar bactérias e/ou arqueas (MOJICA et al., 2005; POURCEL; SALVIGNOL; VERGNAUD, 2006). Diferentemente das sequências DR, os espaçadores podem ter diferentes comprimentos de sequência dentro de um mesmo CRISPR, e espaçadores idênticos nunca são encontrados no mesmo arranjo, mas podem ser encontrados em CRISPRs diferentes (GE et al., 2016; LILLESTOL et al., 2006). Curiosamente, já foram descritos casos de espaçadores com tamanho consideravelmente maior que 72pb, como ocorre na bactéria *Clostridium novyi NT,* que possui um espaçador de 857pb segundo o banco de dados CRISPRdb (GRISSA; VERGNAUD; POURCEL, 2007). É importante dizer que os elementos genéticos móveis são incapazes de infectar cepas que apresentam sequências espaçadoras homólogas no CRISPR; isso mostra que os espaçadores são os responsáveis por gerar a imunidade adquirida em bactérias e arqueas (MOJICA et al., 2005; POURCEL; SALVIGNOL; VERGNAUD, 2006).

O sistema CRISPR inclui, também, uma sequência líder de até 550pb, que está localizada na extremidade 5' do lócus CRISPR, adjacente aos genes *Cas* e é caracterizada por ser comumente rica em A e T, não conservada entre as espécies, e acredita-se que ela atua como uma região de reconhecimento para a adição de novos espaçadores (BARRANGOU et al., 2007; JANSEN et al., 2002; POURCEL; SALVIGNOL; VERGNAUD, 2006; SOREK; KUNIN; HUGENHOLTZ, 2008). Os novos espaçadores são, então, adicionados na extremidade 5' ao lado da sequência líder, fazendo com que os espaçadores mais antigos se tornem relativamente comuns entre os isolados de uma determinada espécie, enquanto que os novos são menos comuns

(POURCEL; SALVIGNOL; VERGNAUD, 2006). Na FIGURA 1 é possível ver a estrutura padrão de um lócus CRISPR.

FIGURA 1: LÓCUS TÍPICO DE CRISPR



FONTE: Adaptado de Sorek (2008).

NOTA: A figura acima ilustra um lócus CRISPR típico com seus 4 componentes estruturais principais: sequências DR, espaçadores, sequência líder e genes *Cas*. As sequências DR, representadas como retângulos azuis, têm tamanho variável de 23 a 55pb, são conservadas dentro do arranjo e parcialmente palindrômicas, e supõe-se que a origem e inserção de uma nova DR no arranjo seja uma ação dependente de *Cas1* e *Cas2*. Já os espaçadores, representados por retângulos pretos, têm um tamanho variável de 21 a 72pb, mas há dados na literatura de espaçadores com mais de 800pb. Eles têm origem de sequências de elementos genéticos móveis, por isso são os responsáveis por gerar a imunidade. Já a sequência líder, em vermelho, tem tamanho de até 550pb e fica localizada na extremidade 5' do arranjo, funcionando como um sítio de reconhecimento para inserção de novos espaçadores. Os genes *Cas*, representados em amarelo, são responsáveis por codificar uma ampla família de proteínas e se organizam como operons. Os genes *Cas* presentes no arranjo CRISPR direcionam a ação de degradação e determinam a classe à qual aquele CRISPR pertence: a classe 1 é comandada por *Cascade*, um complexo multiproteico de 4 a 7 proteínas visto em aproximadamente 90% dos arranjos; a classe 2 utiliza uma proteína única na degradação, como Cas9, por exemplo. A classe 2 é o menos comum, vista em apenas 10% dos arranjos CRISPR conhecidos.

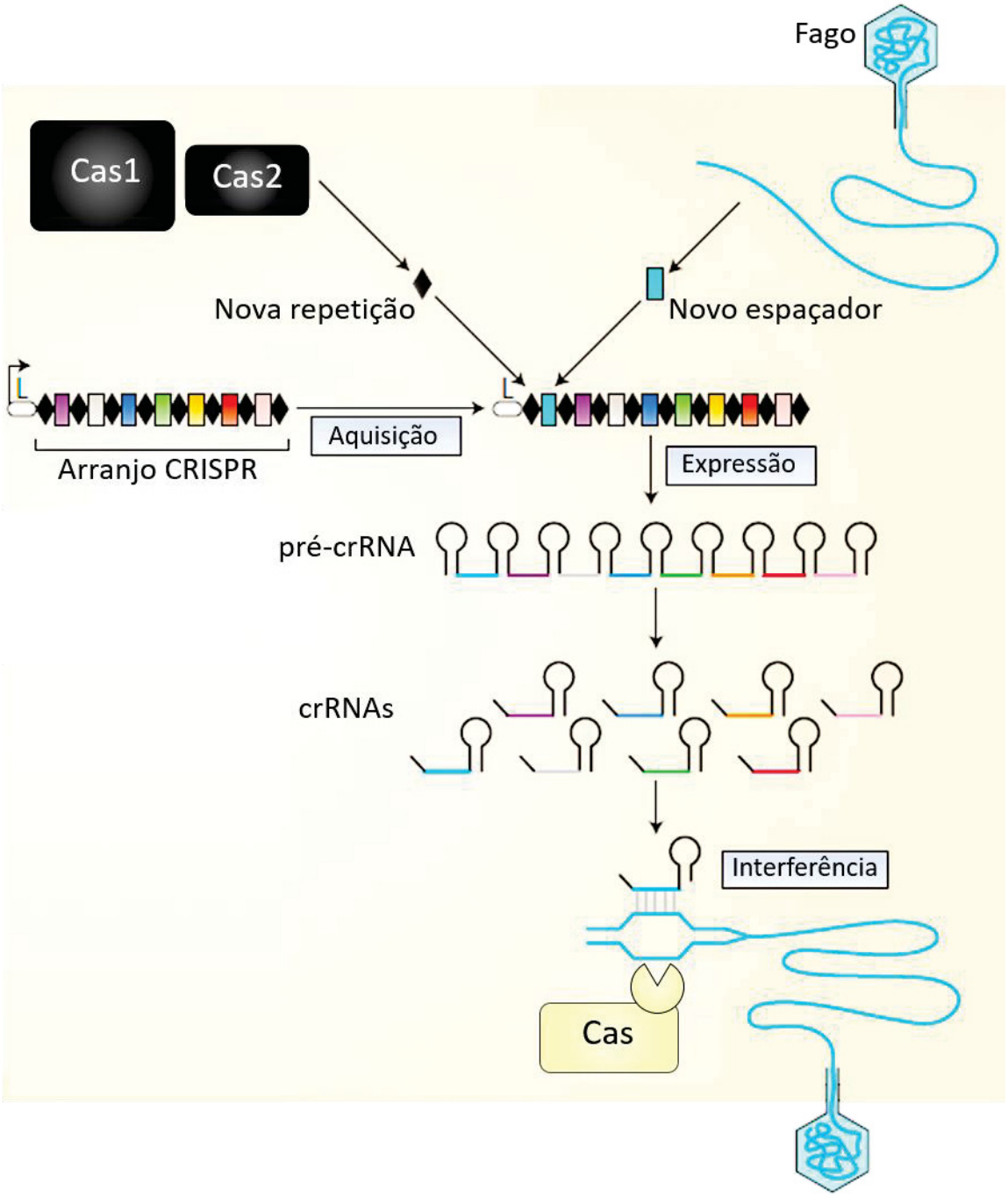
2.1.2 Mecanismo de ação de CRISPR

A ação de imunidade gerada por CRISPR funciona contra a invasão de material genético exógeno vindo de bacteriófagos, plasmídeos e outros elementos móveis. Com a contaminação, uma pequena quantidade de material genético do invasor é incorporada ao genoma do hospedeiro como um novo espaçador de CRISPR (BARRANGOU et al., 2007; BISWAS et al., 2013; LILLESTOL et al., 2006; MOJICA et al., 2005; POURCEL; SALVIGNOL; VERGNAUD, 2006). A cada nova infecção, a bactéria ou arquea adquire um novo espaçador e cria sua resistência mediada por CRISPR (BARRANGOU et al., 2007).

O exato mecanismo utilizado pelo sistema CRISPR ainda não é inteiramente esclarecido, mas já foi visto que um espaçador é transcrito em uma única etapa e processado em pequenos RNAs que correspondem a um espaçador completo flanqueado por partes de repetições, o que facilitaria a formação de estruturas em forma de *hairpin* (TANG et al., 2002, 2005). O RNA do sistema CRISPR, também

chamado de crRNA (ou CRISPR-RNA), é derivado de um precursor longo, ou pré-crRNA, e contém um espaçador que é incorporado junto às proteínas Cas. Esse conjunto é o responsável por destruir o material genético invasor (BISWAS et al., 2013). A estrutura e modo de ação do sistema CRISPR estão ilustradas na FIGURA 2.

FIGURA 2: ESTRUTURA E AÇÃO DE CRISPR



FONTE: Adaptado de Barrangou & Horvath (2017).

NOTA: A figura acima demonstra esquematicamente a estrutura do arranjo CRISPR e a forma como é gerada a resposta imune a partir dele. Na fase de adaptação, quando um MGE entra em contato com o organismo, uma pequena região da sua sequência é incorporada como um novo espaçador (retângulo azul) na matriz CRISPR ao lado da sequência líder (indicada por L), e então uma nova sequência DR (losango preto) é inserida em uma ação dependente de *Cas1* e *Cas2*. Na fase de expressão, o arranjo

CRISPR é transcrito em um RNA longo chamado de pré-crRNA que é processado em pequenos RNAs chamados de crRNAs, cada um contendo um espaçador único ligado a uma sequência DR responsável pela conformação de grampo. Na fase de interferência, quando aquele organismo entra em contato novamente com um MGE conhecido, os crRNAs maduros são responsáveis pelo reconhecimento da sequência por complementariedade e assim guiam a proteína Cas efetora (em amarelo) para que ela efetue a clivagem do contaminante e assim gere a imunidade mediada por CRISPR. O modo de ação das proteínas Cas será visto na Seção 2.2.

Há indicações de que o sistema CRISPR/Cas seja transferido horizontalmente através do auxílio de plasmídeos, megaplasmídeos e prófagos. Essa transferência horizontal de genes (HGT – do inglês, *horizontal gene transfer*) enfatiza o poder de defesa contra material genético exógeno e torna todo o conjunto CRISPR/Cas bastante desejável (HORVATH; BARRANGOU, 2010).

Estudos referentes à resposta dos fagos sobre a resistência gerada por CRISPR demonstraram que os fagos haviam mudado sua sequência por mutação, fazendo com que ela não fosse mais idêntica aos espaçadores adquiridos pelos organismos. Além disso, foi visto que a sequência motivo de reconhecimento, ou motivo adjacente ao protospaçador (PAM – do inglês, *protospacer adjacent motif*), que se encontra à jusante do terminal 3', também havia mudado, fortalecendo a ideia de que esse motivo é necessário para o reconhecimento e funcionamento do CRISPR. Os fagos conseguem fugir da imunidade gerada por CRISPR também através da deleção da sequência alvo. Esses achados demonstram que o sistema CRISPR impõe certa pressão seletiva sobre os fagos e acaba se tornando algo altamente relacionado com a sua evolução (DEVEAU et al., 2008; HORVATH; BARRANGOU, 2010).

Além da função de imunidade adquirida, sugerem-se outras funções para o sistema CRISPR como rearranjo cromossômico, regulação da expressão gênica de genes vizinhos, funcionamento como alvo de proteínas de ligação e ação sobre o reparo do DNA, pois muito genes *Cas* contêm domínios responsáveis pela manipulação do DNA (HORVATH; BARRANGOU, 2010; SOREK; KUNIN; HUGENHOLTZ, 2008). Devido a isso, o sistema CRISPR tem sido amplamente utilizado na edição gênica, assim como os já reconhecidos métodos de *zinc-fingers* (dedos de zinco), e enzimas de restrição conhecidas como TALEN (do inglês, *transcription activator-like effector nucleases*). No caso do sistema CRISPR acompanhado por proteínas Cas, a edição gênica feita por Cas9 é capaz de parear mais facilmente com o DNA alvo, tornando o método mais específico e eficiente na grande maioria dos casos (RAN et al., 2013).

O uso do sistema CRISPR como uma tecnologia de edição gênica pode ter aplicação em vários campos, como estudos ecológicos, epidemiológicos, industriais, entre outros. Em conjunto com as descobertas através da Bioinformática e dos métodos aplicados através da Engenharia Genética, é possível utilizar essa tecnologia para impedir elementos indesejáveis, usando como marcador de resistência a antibióticos, por exemplo, e também para limitar a disseminação de elementos genéticos móveis (HORVATH; BARRANGOU, 2010).

2.2 PROTEÍNAS CAS

Os genes *Cas* codificam uma ampla família de proteínas que se caracterizam como nucleases, helicases, polimerases e proteínas de ligação (HORVATH; BARRANGOU, 2010). Esses genes são encontrados tanto em arqueas quanto em bactérias que apresentam o sistema CRISPR, estando localizados adjacentes ao lócus CRISPR e organizados como operons (HORVATH; BARRANGOU, 2010; KARIMI et al., 2018).

Uma extensa pesquisa sobre os genes *Cas* foi realizada por Haft e colaboradores (2005), onde foram descritas 45 famílias de proteínas, sendo 4 delas (chamadas de Cas1 a Cas4) estritamente relacionadas com o arranjo CRISPR, sempre ocorrendo próximas ao agrupamento de repetições (HAFT et al., 2005). Em contrapartida, sabe-se que as proteínas Cas1 e Cas2 estão presentes na maioria dos arranjos CRISPR conhecidos, formando um complexo encarregado pelo módulo de adaptação de CRISPR responsável pela aquisição e inserção de novos espaçadores na extremidade líder, além de participarem da síntese de uma nova sequência DR (BARRANGOU, 2013; BARRANGOU; HORVATH, 2017; ISHINO; KRUPOVIC; FORTERRE, 2018).

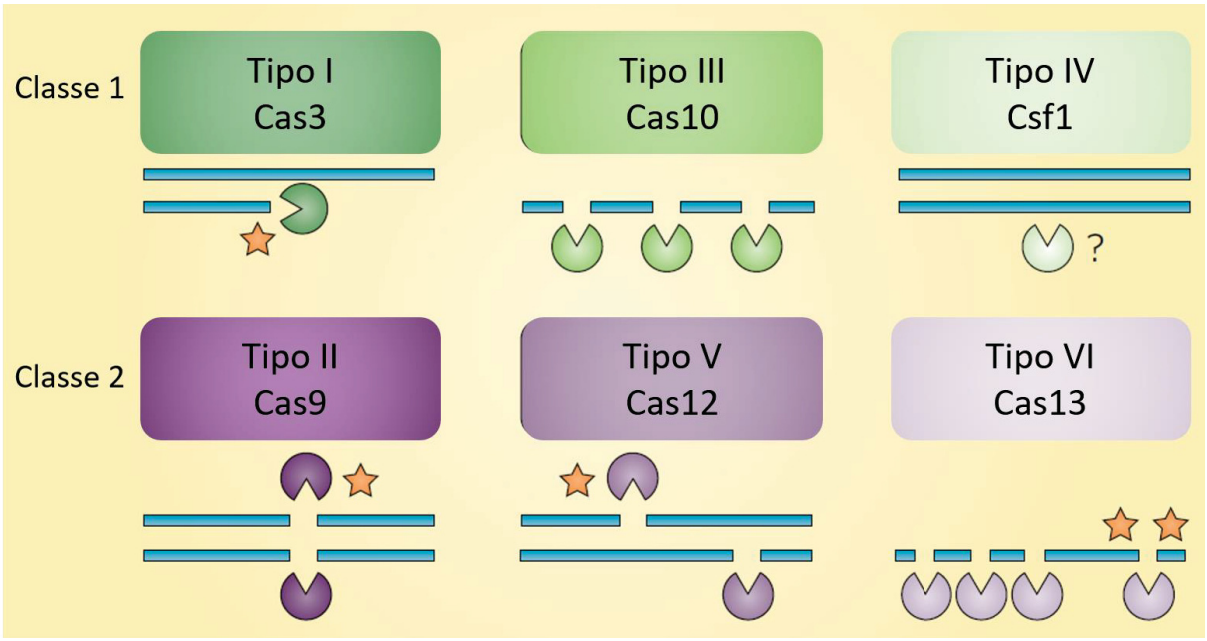
Os sistemas CRISPR/Cas podem ser divididos em duas classes principais de acordo com a nuclease que direciona ação de degradação do material genético invasor. A classe 1 responde a um complexo multiproteico chamado *Cascade* e a classe 2 atua a partir da assinatura de uma única proteína, como a endonuclease Cas9 para o tipo II, por exemplo (BARRANGOU; HORVATH, 2017).

Cascade (do inglês, *CRISPR-associated complex for antiviral defense*) é um complexo multiproteico composto por uma variação entre 4 a 7 proteínas Cas. Esse complexo é visto em ~90% de todos os *loci* CRISPR identificados até o momento;

apenas os ~10% restantes pertencem à classe 2 e estes são, quase que exclusivamente, do domínio bactéria (BURSTEIN et al., 2016). Entende-se que *Cascade* seja responsável pela maturação de pré-crRNAs em crRNAs, etapa necessária para a resposta antiviral, e que essa ligação crRNA-*Cascade* serviria como um guia para direcionar um novo complexo para degradação (BROUNS et al., 2008).

Sabe-se que a família de proteínas Cas é altamente polimórfica e possui diferentes funcionalidades envolvidas em diversas etapas da imunidade mediada por CRISPR, em especial à clivagem do DNA do contaminante por complementariedade de sequências (KARIMI et al., 2018). Uma ilustração da ação de algumas proteínas Cas é mostrada na FIGURA 3.

FIGURA 3: FORMA DE AÇÃO DE ALGUMAS PROTEÍNAS CAS DURANTE A IMUNIDADE MEDIADA POR CRISPR



FONTE: Adaptada de Barrangou & Horvath (2017).

NOTA: Entre os CRISPRs de classe 1, temos sistemas tipo I, III e IV, por exemplo. No tipo I, a exonuclease Cas3 retira a fita de DNA alvo e depois a "mastiga"; no tipo III, para a ação da nuclease Cas10 a sequência alvo deve ser transcrita para que Cas10 clive o mRNA; o mecanismo dos sistemas do tipo IV com Csf1 ainda não foi caracterizado. Nos sistemas CRISPR de classe 2 temos os sistemas tipo II, V e VI. No tipo II, a endonuclease Cas9 é responsável por fazer dois cortes paralelos para gerar uma quebra de dsDNA (do inglês, *double-stranded* DNA); no tipo V, Cas12 é usada para gerar dois pontos de clivagem não paralelos, diferente de Cas9; e os sistemas tipo VI usam Cas13 para clivar o mRNA alvo e, na sequência, processam esse mRNA de maneira não específica para gerar danos colaterais. As estrelas presentes em alguns sistemas indicam que a ação é dependente de PAM para o reconhecimento do alvo.

Além das informações dadas sobre as proteínas Cas na FIGURA 3, sabemos que Cas1 é um marcador universal de CRISPR, funcionando como uma endonuclease de DNA fita dupla, que está relacionada com o processo de imunização e é a única

proteína vista em todas as espécies que contêm o lócus CRISPR (MAKAROVA et al., 2006).

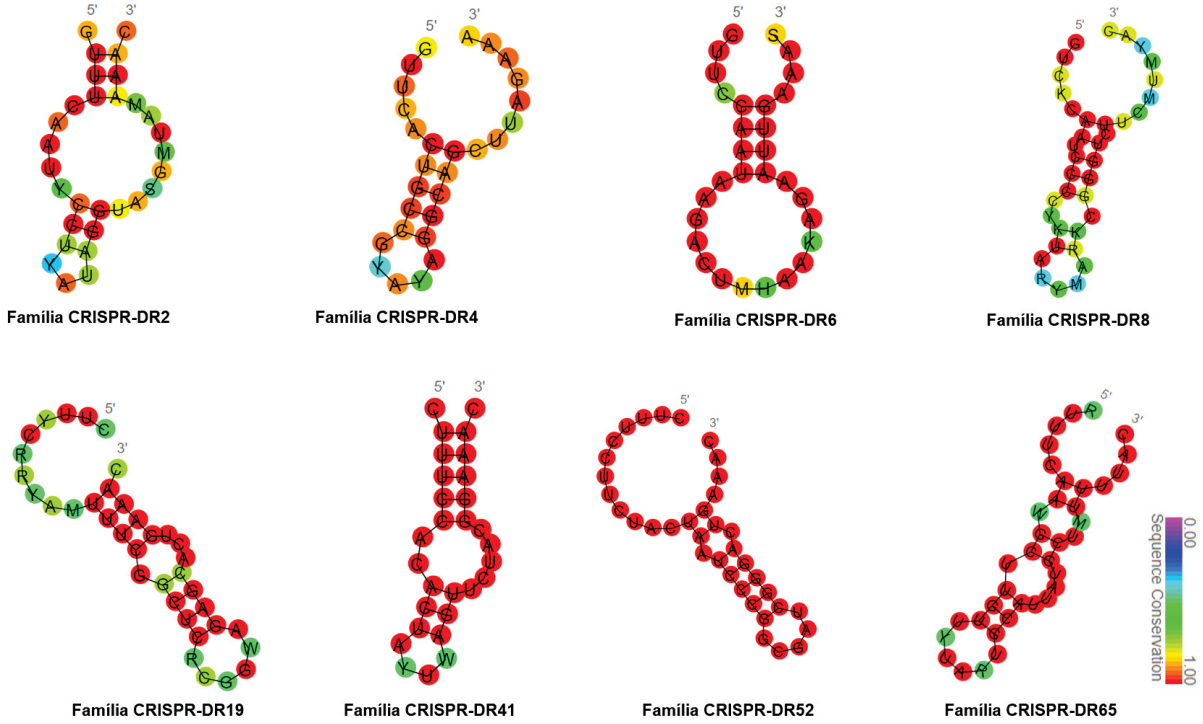
Cas2 atua como uma endoribonuclease que corta RNAs de fita simples e ricos em uracila. Cas3, por sua vez, contém 7 motivos característicos da família das helicases; já Cas4 se enquadra na família das exonucleases relacionadas a RecB, e tem uma sequência motivo rica em cisteína, o que sugere um alvo de ligação do DNA. Assim, sugere-se que Cas3 e Cas4 estão envolvidos no metabolismo do DNA, incluindo reparo, recombinação, regulação transcricional e segregação cromossômica (HAFT et al., 2005; HORVATH; BARRANGOU, 2010; JANSEN et al., 2002).

Cas6 também é uma endoribonuclease que age clivando a transcrição do pré-crRNA em crRNA. Esses crRNAs são responsáveis por orientar o complexo CRISPR/Cas até moléculas de DNA estranhas que invadem seu material, fazendo com que essas moléculas sejam degradadas (HORVATH; BARRANGOU, 2010).

2.3 FAMÍLIA CRISPR RNA

Não há na literatura dados que se refiram a uma categoria de RNA própria de CRISPR. Entretanto o Rfam, uma vasta base de dados de famílias de RNA, traz 64 famílias nomeadas como *CRISPR RNA direct repeat element*, identificadas como CRISPR-DR*n*, onde *n* abrange de 2 a 65. O número de membros, alinhamentos e espécies é variado entre as famílias, bem como o padrão da estrutura secundária. A FIGURA 4 traz alguns exemplos de estruturas secundárias encontradas no Rfam (KALVARI et al., 2018).

FIGURA 4: EXEMPLOS DE ESTRUTURA SECUNDÁRIA ENCONTRADAS NO BANCO DE DADOS RFAM



FONTE: A autora (2018); Adaptado de Rfam (KALVARI et al., 2018).
NOTA: Estrutura secundária de algumas famílias de RNA CRISPR encontradas no Rfam, no estilo "*sequence conservation*" (seqcons) que demonstra a conservação da composição nucleotídica da sequência em cada posição. A conservação de um nucleotídeo em dada posição é determinada pelas cores que variam de pouco conservadas (em lilás, mais próximas de 0,00), levemente conservadas (em verde) e altamente conservadas (em vermelho, mais próximas de 1,00). Assim é possível notar que a composição nucleotídicas das famílias de RNA CRISPR é bastante conservada, visto que a maioria dos seus nucleotídeos são indicados em vermelho ou cores próximas.

2.4 CRISPR E BIOINFORMÁTICA

O crescente interesse pelo sistema CRISPR, em especial como ferramenta de edição gênica, fez com que um grande número de *softwares* de bioinformática fossem desenvolvidos para sua análise (SOREK; KUNIN; HUGENHOLTZ, 2008). Entre os tipos de ferramentas desenvolvidos é possível citar *softwares* para detecção de CRISPR, como PILER-CR (EDGAR, 2007), CRISPRDetect (BISWAS et al., 2016), CRISPRdisco (CRAWLEY; HENRIKSEN; BARRANGOU, 2018), e CRISPRdigger (GE et al., 2016); *softwares* para detecção tanto de arranjos CRISPR e genes *Cas,* como CRISPRCasFinder (COUVIN et al., 2018) e HMMCAS (CHAI et al., 2017); *softwares* para comparar CRISPRs entre cepas de uma determinada espécie ou entre espécies relacionadas como o CRISPRcompar (GRISSA; VERGNAUD; POURCEL, 2008); repositórios *online* de CRISPRs conhecidos como o CRISPRdb (GRISSA;

VERGNAUD; POURCEL, 2007); visualização de CRISPRs como o CRISPRCasViewer (COUVIN et al., 2018), entre outros.

Entre esses programas, destacam-se especialmente o CRISPRdb e o CRISPRCasFinder. O CRISPRdb (disponível em http://crispr.i2bc.paris-saclay.fr/) é tido como padrão ouro entre os bancos de dados de matrizes CRISPR em genomas publicados desde sua implantação, além de fornecer ferramentas internas de análise (GRISSA; VERGNAUD; POURCEL, 2007). Entretanto, até a conclusão deste trabalho, sua última atualização datava de 09 de maio de 2017.

O CRISPRCasFinder está inserido na ferramenta CRISPR/Cas++ e foi desenvolvido pela *Université Paris-Sud*, mesma universidade desenvolvedora do CRISPRdb. Ele tem o diferencial entre as outras ferramentas de busca de CRISPR na procura combinada entre genes *Cas* e arranjos CRISPR, algo fortemente almejado para os pesquisadores da área (COUVIN et al., 2018; SOREK; KUNIN; HUGENHOLTZ, 2008). No momento da escrita deste trabalho, está em implementação no *software* um novo banco de dados chamado CRISPRCasdb, possivelmente vindo para substituir o CRISPRdb.

3 JUSTIFICATIVA

O sistema CRISPR é um tipo de imunidade adquirida por organismos procariotos (arqueas e bactérias) contra infecções causadas por elementos genéticos móveis como plasmídeos e bacteriófagos, que ocorre através da captura e inserção de trechos de sequências desses MGEs como um espaçador ao arranjo CRISPR. Esses espaçadores são separados entre si por sequências repetitivas palindrômicas, chamadas sequências DR, e são os responsáveis pelo reconhecimento da sequência caso o mesmo elemento genético móvel tente novamente infectar aquele organismo (BULT et al., 1996; GRISSA; VERGNAUD; POURCEL, 2007; HORVATH; BARRANGOU, 2010; ISHINO et al., 1987; JANSEN et al., 2002; SOREK; KUNIN; HUGENHOLTZ, 2008).

Além da separação dos espaçadores e da formação de grampos durante o reconhecimento do alvo, não se tem muitas outras informações a respeito da importância e função das sequências DR. Ainda que se tenha conceitos firmados sobre a principal função do sistema CRISPR e muito se fale dele como uma ferramenta de edição gênica dentro da Engenharia Genética, existem poucas

explorações acerca da origem, das características e dos padrões das sequências estruturais básicas (DR e espaçadores), mesmo havendo semelhanças com estruturas conhecidas como as Ilhas Genômicas, como apresentarem sequências de repetição, conferirem resistência e proteção ao organismo, entre outros (GRISSA; VERGNAUD; POURCEL, 2007; HORVATH; BARRANGOU, 2010; JANSEN et al., 2002; SOREK; KUNIN; HUGENHOLTZ, 2008).

Assim, a pesquisa exploratória sobre os padrões e características de sequências genômicas de CRISPR pode trazer correlações entre as sequências CRISPR e outras estruturas genômicas não descritas na literatura. E ainda, responder questões como se a estrutura é sintetizada pelo próprio organismo ou se é originada de transferência horizontal, e se existe relação com regiões regulatórias ou regiões de reconhecimento.

4 OBJETIVOS

4.1 OBJETIVO GERAL

Fazer uma análise exploratória *in silico* de sequências genômicas de regiões CRISPR (*Clustered Regularly Interspaced Palindromic Repeat*) em organismos procariotos.

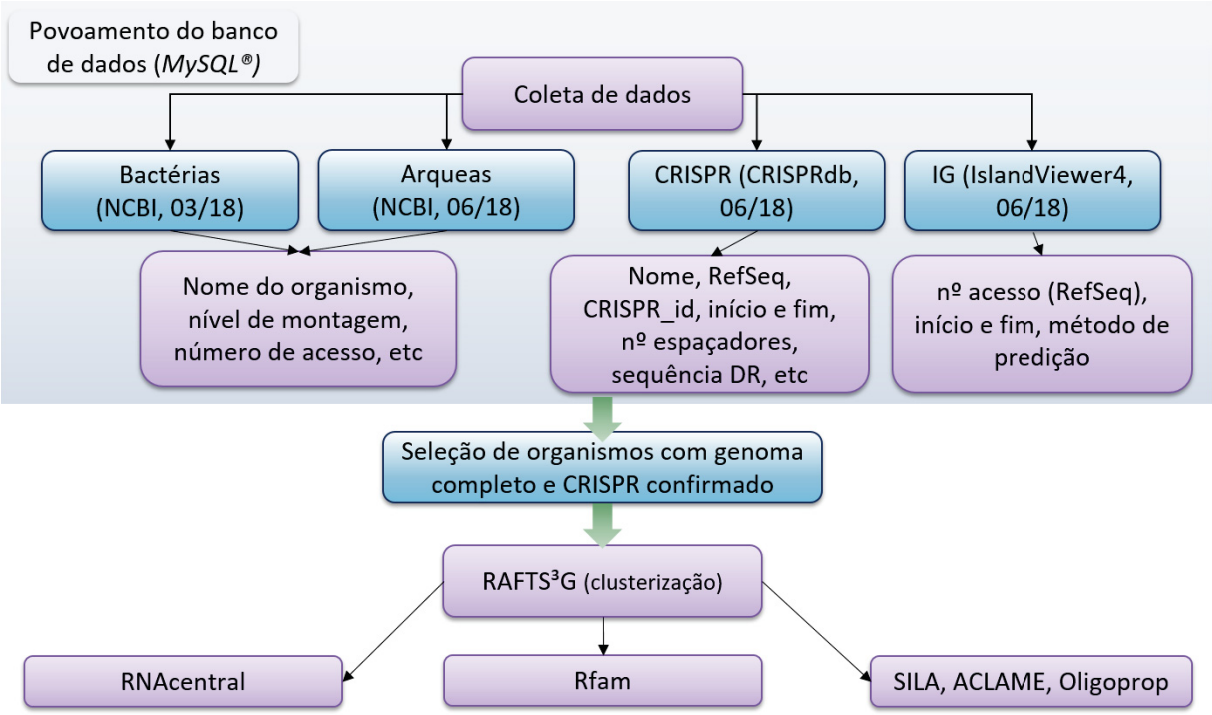
4.2 OBJETIVOS ESPECÍFICOS

- Selecionar organismos que apresentam CRISPR e genomas completos sequenciados;
- Verificar se há ocorrência simultânea de CRISPR e Ilhas Genômicas (IG);
- Separar as sequências repetitivas (DR) dos CRISPRs em organismos com genomas completos e agrupar conforme a identidade entre elas;
- Confrontar as sequências DR com bancos de dados específicos afim de inferir se pode haver uma função ainda não descrita para essas sequências;
- Analisar se existem organismos além de bactérias e arqueas que apresentam essa estrutura como artefato de resistência adquirida.

5 MATERIAL E MÉTODOS

Uma série de etapas foram percorridas durante a execução do trabalho e estão ilustrados na FIGURA 5. Nos tópicos a seguir estão descritos os passos trilhados.

FIGURA 5: FLUXOGRAMA DE ETAPAS SEGUIDAS NA METODOLOGIA



FONTE: a autora (2019).

NOTA: Fluxograma demonstrando as etapas seguidas na metodologia. Inicialmente, na etapa de coleta de dados, foram coletadas informações sobre bactérias e arqueas a partir do NCBI entre março e junho de 2018; esses dados trazem informações sobre nomes de organismos, número de acesso, nível de montagem, entre outros. Dados sobre CRISPR foram coletados do banco CRISPRdb contendo informações sobre o nome do organismo, CRISPR_id, número de espaçadores, DR consenso, etc. Dados referentes a número de acesso, posição de início e fim e método de predição de IGs foram obtidos através do IslandViewer4. Todas essas informações foram usadas para povoar um banco de dados local usando o sistema de gerenciamento de banco de dados MySQL. A partir disso, foi feita a seleção de organismos com genoma completo e CRISPR confirmado, e também de organismos com genoma completo e CRISPR e IGs simultâneos. Então, as sequências DR consenso e os espaçadores foram clusterizados usando a ferramenta RAFTS³G e os grupos com maior densidade de membros tiveram suas sequências DR distintas confrontadas com diferentes bases de dados como RNAcentral e Rfam, além de outras ferramentas como SILA, ACLAME e funções de extração de características de sequências como Oligoprop.

5.1 COLETA DE DADOS

Para este estudo, foram coletados metadados vindos do NCBI referentes a genomas completos de bactérias e arqueas, dados de CRISPR e dados de Ilhas Genômicas.

Os metadados se referem a informações sobre outros dados, ou seja, contêm elementos como o nome do organismo, nível de montagem (genoma completo, *contig*, *scaffold* ou cromossomo), o número de acesso à montagem, entre outros dados que se referem aos organismos de interesse (KITTS et al., 2015). Esses dados, obtidos entre março e junho de 2018, permitiram fazer uma seleção inicial dos organismos que apresentam genoma completo já depositado tanto para bactérias quanto para arqueas na base de dados do NCBI.

As informações quanto aos dados de CRISPR foram obtidas a partir do CRISPRdb (GRISSA; VERGNAUD; POURCEL, 2007) e do arquivo de instalação do *software* CRISPRCasFinder em junho de 2018 (COUVIN et al., 2018). Esse arquivo contém informações quanto ao nome da espécie, número de identificação no RefSeq, identificação do CRISPR (CRISPR_id), posição de início e fim, tamanho, número de espaçadores, a sequência DR e seu tamanho, e a indicação sobre ser um CRISPR confirmado (0) ou hipotético (1). Com esses dados foram organizadas planilhas referentes a CRISPRs confirmados de bactérias, CRISPRs confirmados de arqueas, CRISPRs hipotéticos de bactérias e CRISPRs hipotéticos de arqueas, sendo que os dois primeiros grupos são os que seguiram para as próximas etapas de análise.

Os dados referentes às IGs tiveram origem do banco de dados da ferramenta *IslandViewer4* em junho de 2018. Esses dados contêm informações sobre o número de acesso da ilha que é compatível com o identificador do RefSeq, a posição de início e fim e o método de predição (BERTELLI et al., 2017).

5.2 BANCO DE DADOS

Em posse das informações acima citadas, foi possível povoar um banco de dados através do sistema de gerenciamento de banco de dados MySQL (ORACLE, 2018). As consultas foram feitas através da linguagem SQL (*Structured Query Language*) e salvas em arquivos CSV (*Comma-Separated Values*).

5.3 RAPID ALIGNMENT FREE TOOL FOR SEQUENCES SIMILARITY SEARCH TO GROUPS (RAFTS³G)

Com base nas consultas feitas ao banco de dados, foi possível filtrar e selecionar as sequências DR que foram submetidas à ferramenta RAFTS³G no *software* MathWorks MATLAB® (NICHIO et al., 2018).

A ferramenta RAFTS³G foi desenvolvida com base em funções internas da *Bioinformatics Toolbox* da versão 2017a do *software* MATLAB, bem como no algoritmo RAFTS3 proposto por Vialle (2016), outra função desenvolvida pelo grupo de Inteligência Artificial Aplicada à Bioinformática do Programa de Pós-Graduação em Bioinformática da UFPR. O RAFTS³G busca por semelhanças entre sequências contidas em um arquivo de texto no formato multifasta através do seu método livre de alinhamento, usando uma função de filtro *Hash* para seleção de candidatos baseado em *k-mers* compartilhados e uma medida de comparação usando matriz de coocorrência (NICHIO et al., 2018).

A função recebe como entrada um arquivo multifasta com sequências de nucleotídeos ou aminoácidos a serem agrupados, e a identidade entre as sequências é determinada pelo valor de *self-score* (COIMBRA, 2015). Esse valor traz o mínimo de identidade avaliada, ou seja, um *self-score* de 0.6 retornará agrupamentos com identidade mínima de 60% entre uma sequência em relação a qualquer outra presente naquele *cluster* (COIMBRA, 2015; NICHIO et al., 2018).

A função é chamada através da linha de comando presente na FIGURA 6:

FIGURA 6: LINHA DE COMANDO DA FUNÇÃO RAFTS³G

```
rafts3groups(file, nself, varargin)
```

FONTE: Nichio (2018).
NOTA: Exemplo de linha de comando usada na ferramenta RAFTS³G, onde *file* se refere ao arquivo multifasta com as sequências, *nself* determina o mínimo de identidade desejado e *varargin* indica se são sequências em nucleotídeos (1) ou aminoácidos (2).

Os seguintes parâmetros de entrada são necessários para a utilização da função:

- *file*: arquivo de texto no formato multifasta contendo sequências de nucleotídeos ou aminoácidos;
- *nself*: valor de *self-score* que determina o mínimo de identidade no qual as sequências serão agrupadas;
- *varargin*: variação do argumento de entrada "*file*", definindo se o mesmo é composto por nucleotídeos (1) ou aminoácidos (2).

A função tem como retorno os seguintes itens:

- Um arquivo tabular com nome padrão de “Orthologs_Clustered_report” com o número de identificação do *cluster* de sequências e a quantidade de membros do grupo;
- Arquivos texto no formato multifasta nomeados como “allClusters.fasta”, com todos os *clusters* gerados, e “Cluster_n.fasta”, onde *n* representa o número de identificação de cada *cluster* gerado.

5.4 ACLAME (A CLASSIFICATION OF MOBILE GENETIC ELEMENTS)

ACLAME é uma base de dados referente a MGEs já sequenciados, como fagos, plasmídeos, transposons e sequências de inserção. Através dela é possível buscar por sequências em interface *web* com ajuda da ferramenta BLAST incorporada à sua busca (MCGINNIS; MADDEN, 2004). As pesquisas retornam informações a respeito de cada MGE, como a qual MGE corresponde àquela sequência (plasmídeo, fago, etc), a família de proteínas a que pertence, anotação funcional, o hospedeiro, *e-value, bit-score,* entre outros. Esses dados podem ser baixados e arquivados em arquivo de texto delimitado por tabulação (LEPLAE; LIMA-MENDEZ; TOUSSAINT, 2010). A ferramenta *online* está disponível através do link http://aclame.ulb.ac.be/.

5.5 RNACENTRAL

O RNAcentral é um consórcio entre bases de dados de RNA não-codificante (ncRNA), formado por 42 bases de dados especializadas das quais 28 já foram importadas para o RNAcentral. A importação de múltiplas bases de dados através do RNAcentral permite um acesso gratuito e integrado a um grande conjunto de informações através de busca por termos específicos e busca por sequências (RNACENTRAL CONSORTIUM, 2016).

O principal uso da ferramenta visa a busca por sequências executada através de alinhamento local pelo método *nhmmer* (WHEELER; EDDY, 2013). Essa opção de busca traz resultados como *e-value* (probabilidade de encontrar aquele alinhamento ao acaso), o percentual de identidade entre os nucleotídeos, a cobertura da sequência de consulta pelo alinhamento, a cobertura da sequência alvo pelo alinhamento e a porcentagem de *gaps* para todos os alinhamentos possíveis (RNACENTRAL CONSORTIUM, 2017).

Ao acessar um resultado, vemos informações quanto à descrição do organismo, a classificação e o *link* da família no Rfam, *link* para o ENA (*European Nucleotide Archive*), e a taxonomia dos organismos que compartilham daquela sequência em formato de árvore filogenética (RNACENTRAL CONSORTIUM, 2017).

5.6 CRISPRCASFINDER

CRISPRCasFinder é uma ferramenta que permite a identificação e previsão de orientação de arranjos CRISPR e de genes *Cas* com base em um acurado sistema de classificação, disponível para uso tanto em versão *online* quanto *standalone.* Para determinar a confiabilidade do CRISPR encontrado, a ferramenta traz consigo o nível de evidência que varia de 1 (baixa probabilidade de ser um CRISPR verdadeiro) a 4 (CRISPR verdadeiro), levando em consideração o comprimento da sequência DR e o número de espaçadores (COUVIN et al., 2018).

Como entrada, a ferramenta *online* recebe arquivos no formato multifasta de até 50MB e com até 100 sequências, enquanto que na versão local o único limite de processamento é a memória disponível na máquina usada. Além disso, é possível determinar se será feita a detecção de genes *Cas* e qual o modelo de agrupamento conforme o grau de rigor desejado. Os parâmetros padrão para detecção do arranjo CRISPR são definidos para detectar repetições com o mais alto grau de homologia, mas ainda assim é possível modificar alguns parâmetros que definem a repetição máxima e as propriedades CRISPR (COUVIN et al., 2018).

Finalmente, os arranjos CRISPR e os genes *Cas* são retornadas como arquivos formatados em .xls, GFF3, TSV, fasta e JSON, sendo que esse último pode ser visualizado no *software* CRISPRCasViewer desenvolvido pelos mesmos pesquisadores (COUVIN et al., 2018).

5.7 OUTRAS FERRAMENTAS USADAS NA ANÁLISE EXPLORATÓRIA

5.7.1 Sila

O *software* SILA é um sistema para anotação automática de genomas de procariotos que utiliza estratégias de *gene-finding* e comparação de sequências com o algoritmo RAFTS3, já comentado anteriormente (VIALLE, 2013; VIALLE et al.,

2016). Em sua interface *web*, o usuário submete apenas um arquivo fasta contendo a sequência e mais nenhum parâmetro é requisitado. O resultado da anotação fica disponível para visualização *web* em diferentes formatos, incluindo o formato *genbank* (VIALLE, 2013).

5.7.2 Oligoprop

Para determinar algumas propriedades das sequências DR dos arranjos CRISPR foi usada a função oligoprop.m também presente na *Bioinformatics Toolbox* da ferramenta MATLAB®. Em seu uso mais simples, tem como *input* uma sequência de nucleotídeos de qualquer comprimento e seu *output* traz dados como o conteúdo GC, o número de *hairpins,* a temperatura de *melting* (TM), entre outros (MATHWORKS, 2018). A função oligoprop.m foi otimizada em uma função chamada extract_oligoProps.m, visando facilitar a determinação do conteúdo GC e do número de *hairpins* em um grande número de sequências como as analisadas neste trabalho (KULIK, 2019).

5.7.3 Sweep

Sweep é uma ferramenta em desenvolvimento idealizada pelo grupo de Inteligência Artificial Aplicada à Bioinformática do Programa de Pós-Graduação em Bioinformática da UFPR que busca diminuir o custo computacional em análises filogenéticas a partir projeções em espaços vetoriais, visando a diminuição de dimensão e a geração de árvores filogenéticas (DE PIERRI, 2017). O *Sweep* é uma evolução da ferramenta *SVect* (UFPR, 2018) que foi explorada em estudos filogenéticos de proteomas mitocondriais (DE PIERRI, 2017), e proteomas de cloroplastos e outros plastídeos (CAMARGO, 2018), executando análises filogenéticas em larga escala e com redução de dimensionalidade. No *Sweep* é introduzido o conceito de projeção em um espaço vetorial, onde reduz o espaço de representação das sequências e preserva as características de distâncias entre os elementos vetorizados.

Neste desenvolvimento, foi testada uma nova abordagem da ferramenta originalmente idealizada para sequências de aminoácidos. Testamos pela primeira vez uma versão construída para análise de sequências nucleotídicas, buscando inferir

especialmente a relação entre *Cas1* e *Cas2* na geração de novas sequências DR, focando nas sequências distintas presentes no maior *cluster* entre as bactérias (*Cluster 1B*).

6 RESULTADOS E DISCUSSÃO

6.1 COLETA DE DADOS

As consultas feitas ao banco de dados construído a partir dos dados das bases CRISPRdb, NCBI e *IslandViewer*, retornaram quatro conjuntos de resultados contendo informações referentes às bactérias e arqueas em relação ao sistema CRISPR. As principais informações sobre os dados recuperados nas consultas estão contidas no QUADRO 1.

QUADRO 1: DADOS REFERENTES ÀS QUATRO TABELAS PRINCIPAIS ORIUNDAS DA CONSULTA AO BANCO DE DADOS

Nome	**all_data_arquea_crispr**	**all_data_bacteria_crispr**	**arq_gc_crispr_ig**	**bac_gc_crispr_ig**
nº resultados	751	7.081	172	2.629
Informação	Arqueas com genomas completos e com CRISPR confirmado	Bactérias com genomas completos e com CRISPR confirmado	Arqueas com genomas completos que apresentam CRISPR e IG simultaneamente	Bactérias com genomas completos que apresentam CRISPR e IG simultaneamente
Campos	Refseq, GCF, GCA, id do CRISPR, nome do organismo, posição de início e fim do CRISPR, tamanho do CRISPR, nº de espaçadores, sequência DR, taxonomia, nível de montagem, taxid	Refseq, GCF, GCA, id do CRISPR, nome do organismo, posição de início e fim do CRISPR, tamanho do CRISPR, nº de espaçadores, sequência DR, taxonomia, nível de montagem, taxid	Refseq, GCF, GCA, id do CRISPR, nome do organismo, posição de início e fim do CRISPR, tamanho do CRISPR, nº de espaçadores, sequência DR, taxonomia, nível de montagem, taxid, nº de acesso da IG, posição de início e fim da ilha, método de predição	Refseq, GCF, GCA, id do CRISPR, nome do organismo, posição de início e fim do CRISPR, tamanho do CRISPR, nº de espaçadores, sequência DR, taxonomia, nível de montagem, taxid, nº de acesso da IG, posição de início e fim da ilha, método de predição
Filo mais representativo	*Euryarchaeota* (526 membros → 70,03%)	*Proteobacteria* (3.535 membros → 49,92%) *Firmicutes* (1.709 membros → 24,14%)	*Euryarchaeota* (125 membros → 72,67%)	*Proteobacteria* (1.321 membros → 50,25%)
Sufixo adotado para identificação dos *clusters*	A	B	Aig	Big

FONTE: A autora (2018).

NOTA: o nome de arquivo “all_data_arquea_crispr” se refere a todos os dados de arqueas com genoma completo e CRISPR confirmado; “all_data_bacteria_crispr” se refere a todos os dados de bactérias com genoma completo e CRISPR confirmado; “arq_gc_crispr_ig“ se refere a todos os dados de arqueas com genoma completo, CRISPR confirmado e Ilhas Genômicas; “bac_gc_crispr_ig” se refere a todos os dados de bactérias com genoma completo, CRISPR confirmado e Ilhas Genômicas. Para facilitar a compreensão, no decorrer do texto os sufixos mostrados na última linha do quadro acima serão adotados para identificar a origem das sequências analisadas.

Com base nesses dados, as sequências DR de arqueas e bactérias presentes nos arquivos referentes a organismos com genomas completos e com CRISPR confirmado foram coletadas e transferidas para arquivos no formato multifasta.

Os arquivos multifasta foram submetidos à ferramenta RAFTS³G com o objetivo de organizar e agrupar as sequências de acordo com a identidade entre elas e avaliar a sua diversidade na composição nucleotídica e características estruturais.

6.2 RAPID ALIGNMENT FREE TOOL FOR SEQUENCES SIMILARITY SEARCH TO GROUPS (RAFTS³G)

A ferramenta RAFTS³G foi executada através da linha de comando vista na FIGURA 7, onde o parâmetro “seq_bac.fasta” identifica o nome do arquivo com as sequências DR das bactérias, 0,5 se refere ao mínimo de 50% identidade desejado entre o alinhamento das sequências, e o valor 1 representa sequências em nucleotídeos.

As sequências DR extraídas da tabela “all_data_bacteria_crispr” (ver QUADRO 1) relacionadas às bactérias que apresentam CRISPR confirmado, foram transferidas para o arquivo multifasta denominado “seq_bac.fasta”, contendo 7.081 sequências DR, que foi então submetido à ferramenta RAFTS³G através do *software MathWorks MATLAB*®.

FIGURA 7: LINHA DE COMANDO USADA PARA O AGRUPAMENTO DAS SEQUÊNCIAS

```
rafts3groups('seq_bac.fasta', 0.5, 1)
```

FONTE: A autora (2018).

NOTA: Linha de comando executada na ferramenta RAFTS³G, onde o parâmetro “seq_bac.fasta” identifica o nome do arquivo fasta com as sequências DR de bactérias com genoma completo e CRISPR confirmado; 0,5 se refere ao mínimo de 50% identidade; e o valor 1 representa que as sequências analisadas são nucleotídicas.

O agrupamento das 7.081 sequências DR referentes a bactérias com genoma completo e CRISPR confirmado realizado pelo RAFTS³G resultou em 1.547 *clusters*, sendo 878 deles *clusters* únicos (formados por apenas uma única sequência DR). A maior densidade entre os grupos foi vista no *Cluster 1B*, com 1.001 sequências, cerca de 14% do total de sequências, seguido do *Cluster 260B* com 140 sequências. Isso reflete que as sequências presentes nesses *clusters* têm, no mínimo, 50% de identidade com relação a qualquer outra presente naquele *cluster*. Estes dois *clusters*

foram os escolhidos para passar para as próximas análises, e suas informações mais relevantes se encontram no QUADRO 2.

QUADRO 2: INFORMAÇÕES RELEVANTES SOBRE OS DOIS MAIORES CLUSTERS ENTRE AS BACTÉRIAS

	Cluster 1B	*Cluster 260B*
nº de membros	1.001	140
Conteúdo GC médio	65,79	70,39
nº de sequências DR distintas	30	32
Maior nº de repetições de DR	425 (42,46%)	48 (34,29%)
Filo predominante	*Proteobacteria* (947 membros → 94,61%)	*Proteobacteria* (116 membros → 82,86%)
Classe predominante	*Gammaproteobacteria* (913 membros → 91,21%)	*Gammaproteobacteria* (111 membros → 79,29%)
Ordem predominante	*Enterobacterales* (888 membros → 88,71%)	*Enterobacterales* (105 membros → 75%)
Família predominante	*Enterobacteriaceae* (799 membros → 79,82%)	*Enterobacteriaceae* (99 membros → 70,71%)
Gêneros predominantes	*Salmonella* (441 membros → 44,06%) *Escherichia* (318 membros → 31,77%)	*Escherichia* (68 membros → 48,57%) *Salmonella* (25 membros → 17,86%)

FONTE: A autora (2018).

NOTA: O quadro acima traz as principais informações a respeito das características dos dois maiores *clusters* encontrados entre as bactérias, incluindo informações quanto ao número de membros, sequências distintas, conteúdo GC e predominância taxonômica.

A partir do QUADRO 2 é possível notar que há pouca diversidade de sequências dentro dos *clusters*, visto que no *Cluster 1B*, por exemplo, existem apenas 30 sequências repetidas distintas entre os 1.001 membros. Além disso, essas sequências apresentam um alto valor de conteúdo GC médio, o que favorece a discussão sobre a relação entre o conteúdo GC e a ocorrência de palíndromos como visto por Ninh (2012). Também, há uma grande conservação taxonômica em especial entre os filos em ambos os *clusters*. Essa predominância do filo *Proteobacteria* também foi visto em estudo feito por Makarova e colaboradores (2011), onde 50,92% das bactérias estudadas eram deste filo.

O resultado o RAFTS³G para as sequências DR de arqueas não apresentou um *cluster* com grande densidade. Entre os 250 *clusters* formados, os com maior número de membros foram o *Cluster 36A* com 55 sequências DR, e o *Cluster 38A* com 33 sequências, e entre os demais 131 são *clusters* únicos (com uma única sequência). No que se refere à análise taxonômica, nossos resultados corroboram com o que foi visto por Makarova (2011), onde 66,23% das arqueas pesquisadas em seu estudo eram membros do filo *Euryachaeota.* Nossos dados trazem valores que

demonstram que 70,03% das arqueas com CRISPR confirmado são também deste filo.

Cabe observar que o número elevado de *clusters* indica uma grande diversidade entre as sequências DR e, consequentemente, um baixo índice de identidade (<50%) entre as sequências disponíveis. Com o sequenciamento de novos genomas bacterianos que apresentem CRISPR, poderão surgir novas sequências DR que potencialmente apresentem similaridade equidistante entre sequências presentes em diferentes *clusters* observados atualmente, permitindo que ocorra a fusão entre dois ou mais *clusters* vistos nesse trabalho. Por enquanto, entende-se que as sequências DR são mais conservadas no domínio bactéria do que o observado entre as arqueas.

As sequências do *Cluster 1B* e do *Cluster 260B*, que juntas representam 16,11% das sequências DR de bactérias, foram escolhidas para serem confrontadas com bancos de dados públicos que disponibilizam ferramentas de busca por similaridade de sequências com outros organismos ou regiões específicas. O resultado dessa busca permitiu identificar alguns padrões de sequência não descritos na literatura quanto a algumas funções e estruturas das sequências DR e, ao mesmo tempo, foi possível confirmar o índice de identidade entre as sequências como critério de agrupamento.

Os resultados mais relevantes foram vistos nas bases ACLAME (LEPLAE; LIMA-MENDEZ; TOUSSAINT, 2010) e RNAcentral (RNA CENTRAL CONSORTIUM, 2017) descritos a seguir, e outras ferramentas foram utilizadas para dar base ao visto em ambas.

6.3 ANÁLISE QUANTO À PRESENÇA DE ELEMENTOS GENÉTICOS MÓVEIS

Todas as sequências DR presentes nos *clusters* 1B e 260B foram submetidos à base de dados ACLAME à procura de elementos genéticos móveis. Os dados referentes a essa consulta puderam ser exportados na forma de tabela e trazem informações como: identificador da sequência de consulta; tamanho da sequência alvo; identificação, nome, tipo e hospedeiro do MGE; *bit-score*; *e-value*; percentual de identidade, entre outros.

Para o *Cluster 1B* foram 645 ocorrências de similaridade entre as sequências DR e genes referentes a MGEs em 181 das 1.001 sequências pertencentes a esse

cluster. Entre esses, 96 são caracterizados como prófagos e 549 como plasmídeos. Entre os plasmídeos é possível identificar quatro MGEs distintos, que codificam seis genes diferentes em seis famílias de proteínas das quais quatro são classificadas como proteínas hipotéticas, uma sendo provável proteína reguladora transcricional da família GntR (relacionada ao *Rhizobium etli CFN 42*), e uma transposase pertencente à família IS605 OrfB. Além disso, os resultados mais representativos ocorrem 14 vezes e correspondem ao plasmídeo pIPL (id: 804) que tem como hospedeiro principal a proteobactéria *Legionella pneumophila str. Lens*, trazendo valores de *bit-score* de 58, *e-value* de $5{,}00^{-3}$, e um percentual de identidade de 100%.

Das 96 sequências classificadas como prófagos, foram identificados quatro MGEs diferentes, com quatro genes para quatro diferentes proteínas, sendo duas classificadas também como proteínas hipotéticas e duas como supostas proteínas associadas a fagos. Os quatro MGEs identificados têm como hospedeiro principal diferentes espécies do gênero *Neisseria*. Os valores de *bit-score*, *e-value* e percentual de identidade são de 52, $2{,}00^{-1}$ e 100%, respectivamente, para todos os resultados. Todos os resultados referentes à análise pelo ACLAME das sequências DR do *Cluster 1B* podem ser vistos no APÊNDICE 1.

O ACLAME classificou 24 das 140 sequências pertencentes ao *Cluster 260B*, retornando 25 ocorrências de similaridade entre MGEs e as sequências DR. Apenas plasmídeos foram identificados, sendo os resultados mais expressivos referentes ao *bit-score* de 58, *e-value* de $5{,}00^{-3}$ e percentual de identidade de 100% para três ocorrências do plasmídeo SCP1 presente em *Streptomyces coelicolor*. Os resultados obtidos na análise do *Cluster 260B* podem ser vistos no APÊNDICE 2.

Esses resultados apontam a possibilidade de uma discussão referente à origem das sequências DR, sendo que elas também poderiam ser derivadas de MGEs assim como os espaçadores.

6.4 RNACENTRAL *E RFAM*

As diferentes sequências dos *clusters* 1B e 260B foram submetidas à ferramenta de busca do RNAcentral. Nessa ferramenta, as buscas podem ser realizadas por sequências de DNA ou RNA, ou por códigos de identificação do próprio site.

6.4.1 *Cluster 1B* – Maior grupo de sequências DR de bactérias

Para o *Cluster 1B*, os melhores resultados vêm de duas sequências que se diferenciam entre si em apenas um nucleotídeo (sublinhado) presentes no FIGURA 8. O resultado dessas e das demais sequências pode ser visto no APÊNDICE 3.

FIGURA 8: SEQUÊNCIAS QUE APRESENTARAM OS MELHORES RESULTADOS ENCONTRADOS NO *CLUSTER 1B*

>Cluster_1_NC_017160_3
GTTCACTGCCGCACAGGCAGCTTAGAAA
>Cluster_1_NC_017626_3
GTTCACTGCCGTACAGGCAGCTTAGAAA

FONTE: RAFTS³G (2018).
NOTA: Quadro representativo contendo as duas sequências DR do *Cluster 1B* que apresentaram os melhores resultados na pesquisa feita na ferramenta RNAcentral. É possível notar a grande similaridade entre elas, sendo que se diferenciam apenas por um nucleotídeo (sublinhado). O resultado completo dessa e das demais sequências à consulta ao RNAcentral pode ser visto no Apêndice 3.

A primeira sequência da FIGURA 8 corresponde à sequência DR vista no CRISPR de *Yersinia pestis D182038* (GCF_000022825.1, GCA_000022825.1) e em outros 43 CRISPRs. A pesquisa no RNAcentral resultou em uma sequência correspondente e classificada como *CRISPR RNA direct repeat element*, encontrada em 28 diferentes espécies, apresentando 100% de identidade, cobertura de consulta e cobertura de alvo.

A segunda sequência da FIGURA 8 pertence ao CRISPR de *Escherichia coli 042* (GCF_000027125.1, GCA_000027125.1) e se repete em outros 40 CRISPRs. Assim como no caso anterior, houve correspondência exata com uma sequência também classificada como *CRISPR RNA direct repeat element*. Nesse caso, a sequência está presente em 39 espécies distintas. Os valores resultantes também foram de 100% para os três parâmetros citados anteriormente.

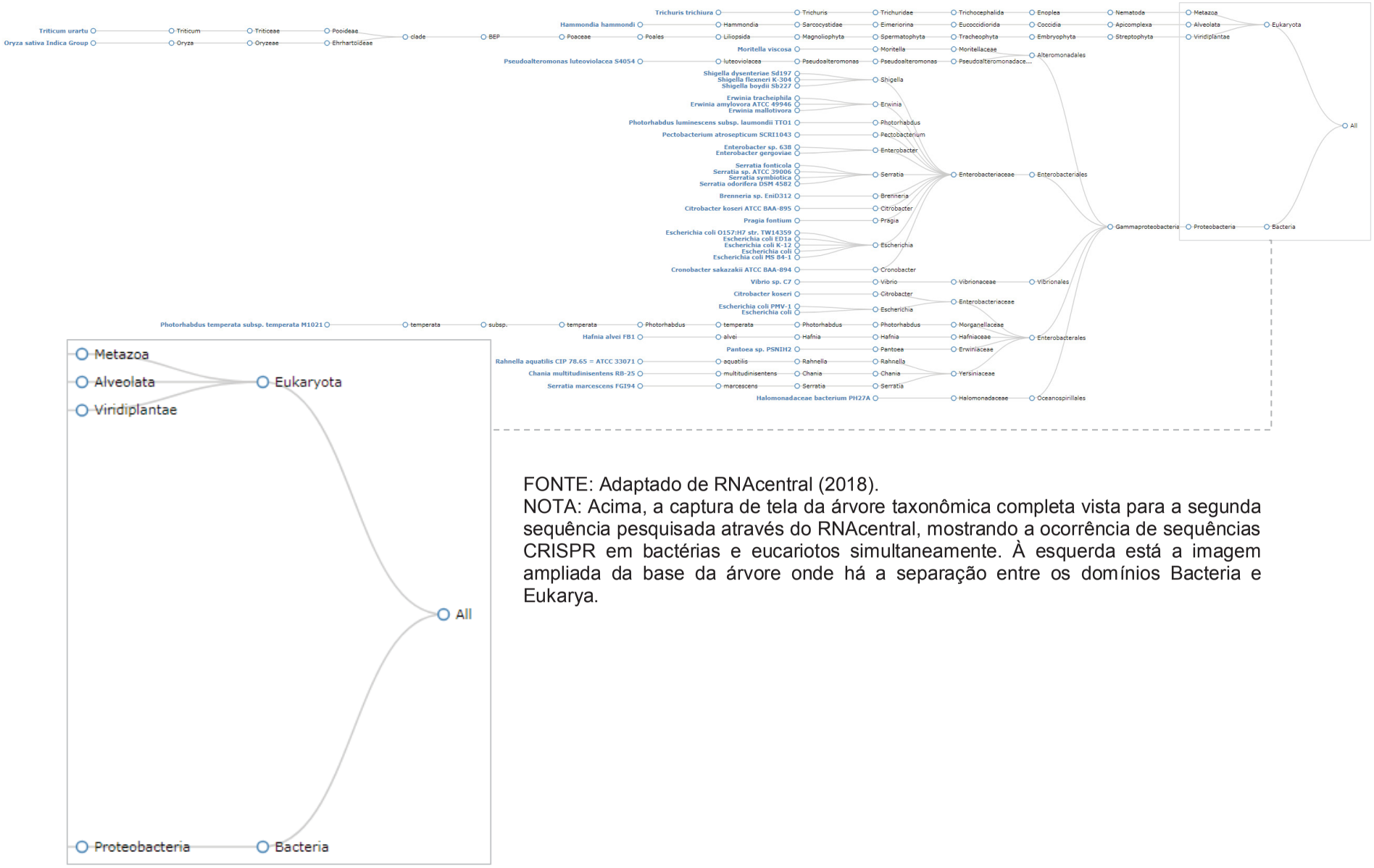
Além desses resultados numéricos, o RNAcentral traz uma árvore taxonômica dos organismos que apresentam a sequência, bem como a classificação no Rfam e no QuickGO (*Gene Ontology and GO Annotations*), ferramenta *web* que visa especialmente descrever funções moleculares, processos biológicos e localizações celulares relacionados a determinados genes (BINNS et al., 2009). Na consulta referente às duas sequências da FIGURA 8, o Rfam direciona para uma mesma

família identificada como RF01317 e para a mesma descrição no QuickGO, identificada como GO:0006952.

Segundo o QuickGO, esse RNA CRISPR é definido como um processo biológico de resposta de defesa a um corpo estranho que possa causar danos ao organismo atacado. No Rfam, o resultado corresponde à família CRISPR-DR4, que é fortemente ligada às sequências de busca visto seu baixo *e-value* ($2{,}2^{-12}$). Além disso, o Rfam fornece a visualização da estrutura secundária representativa da família, mostrando as regiões mais conservadas e a estrutura em grampo conhecida do sistema CRISPR, estrutura que pode ser revista na FIGURA 4.

Quanto à visualização da árvore taxonômica dada pelo RNAcentral, um resultado intrigante foi visto para a segunda sequência. Já na base da árvore ocorre a divisão entre bactérias e eucariotos como portadores da sequência, algo inesperado partindo do pressuposto de que apenas procariotos apresentam o sistema CRISPR (FIGURA 9).

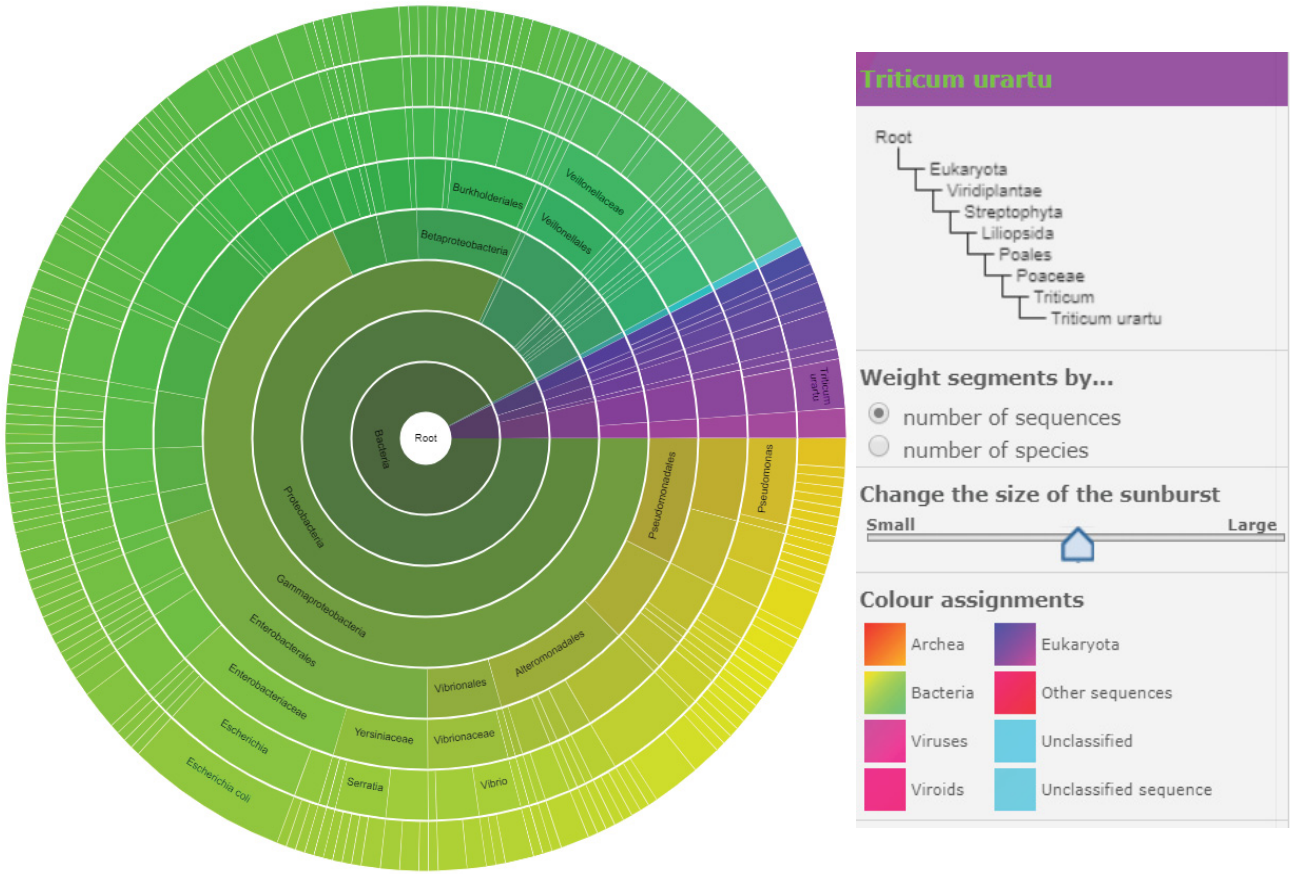
FIGURA 9: CAPTURA DE TELA REFERENTE À ARVORE TAXONÔMICA VISTA A PARTIR DA CONSULTA AO RNACENTRAL MOSTRANDO A OCORRÊNCIA DE SEQUÊNCIA CRISPR EM EUCARIOTO



FONTE: Adaptado de RNAcentral (2018).
NOTA: Acima, a captura de tela da árvore taxonômica completa vista para a segunda sequência pesquisada através do RNAcentral, mostrando a ocorrência de sequências CRISPR em bactérias e eucariotos simultaneamente. À esquerda está a imagem ampliada da base da árvore onde há a separação entre os domínios Bacteria e Eukarya.

Ao analisar o resultado do Rfam, a opção "*Species"* permite ver as espécies daquela família de RNAs na forma de *sunburst* (explosão solar). Essa representação mostra a distribuição taxonômica separando cada nó da árvore como um arco, e a coloração varia de acordo com o reino, sendo as arqueas representadas em tons laranjados, os procariotos identificados em tons verdes e eucariotos retratados em tons roxos. Ao analisar a árvore da família CRISPR-DR4 é possível notar novamente a presença de eucariotos junto às bactérias apresentando sequências DR de CRISPR. São 20 sequências distribuídas em 10 espécies diferentes de eucariotos, e 252 sequências em 185 espécies de bactérias. A FIGURA 10 mostra a distribuição em *sunburst*.

FIGURA 10: REPRESENTAÇÃO DA ÁRVORE TAXONÔMICA DA FAMÍLIA CRISPR-DR4 MOSTRANDO A OCORRÊNCIA EM EUCARIOTOS



FONTE: Adaptado de Rfam (2018).

NOTA: Captura de tela referente à distribuição de espécies para a família CRISPR-DR4 ilustrada na forma de *sunburst* (em português, explosão solar) trazida pela ferramenta Rfam. Nela os diferentes domínios são representados por cores e cada nó de taxonomia é representado por um arco. Eucariotos são representados em tons roxos, bactérias em tons verdes e arqueas em tons laranjas. Assim, é possível notar que o Rfam relaciona a presença de sequências de RNA CRISPR em eucariotos assim como foi visto no RNAcentral, sendo que 12 diferentes espécies de eucariotos apresentam sequências pertencentes à família CRISPR-DR4.

Uma das espécies com indicação da presença do sistema CRISPR é a *Triticum urartu* (em roxo na FIGURA 10), uma espécie de trigo selvagem (LING et al., 2013). No intuito de compreender o motivo de encontrar essa espécie entre os portadores de CRISPR, tomou-se o cuidado em se obter a sequência contida no ENA (KD046323.1) a partir do *link* vindo do próprio RNAcentral para garantir que a exata sequência onde aquele dado foi encontrado fosse analisada.

Inicialmente, a sequência foi analisada pelo *software* CRISPRCasFinder. Segundo essa análise, a sequência de *Triticum urartu* apresenta duas estruturas CRISPR com nível de evidência igual a quatro, o maior alcançado. Os dados referentes aos dois CRISPRs estão no QUADRO 3 e as imagens referentes às regiões estão nas FIGURAS 11 e 12.

QUADRO 3: DADOS REFERENTES AOS CRISPRS ENCONTRADO EM *Triticum urartu*

CRISPR	Início	Fim	Conservação DR	Sequência DR consenso	Número de espaçadores
KD046323.2	3757	4804	97,79%	TTTCTAAGCTGCCTGTACGGCAGTGAAC	17
KD046323.3	13353	14580	98,03%	TTTCTAAGCTGCCTGTACGGCAGTGAAC	20

FONTE: a autora (2018).

NOTA: Quadro contendo os dados referentes aos dois CRISPRs identificados pela ferramenta CRISPRCasFinder para *Triticum urartu*, como número de identificação, posição de início e fim, conservação da sequência DR, DR consenso e o número de espaçadores.

FIGURA 11: CAPTURA DE TELA REFERENTE À REGIÃO DO PRIMEIRO CRISPR ENCONTRADO EM *Triticum urartu*

FONTE: CRISPRCasFinder (2018).

NOTA: Captura de tela tirada a partir da ferramenta de busca CRISPRCasFinder ilustrando o primeiro CRISPR com alto grau de evidência (4) encontrado na sequência de *Triticum urartu* (KD046323.1). Na primeira e última linhas da sequência estão regiões flanqueadoras de 100 nucleotídeos, em amarelo estão as sequências DR do arranjo e em cores diversas os 17 espaçadores.

FIGURA 12: CAPTURA DE TELA REFERENTE À REGIÃO DO SEGUNDO CRISPR ENCONTRADO EM *Triticum urartu*

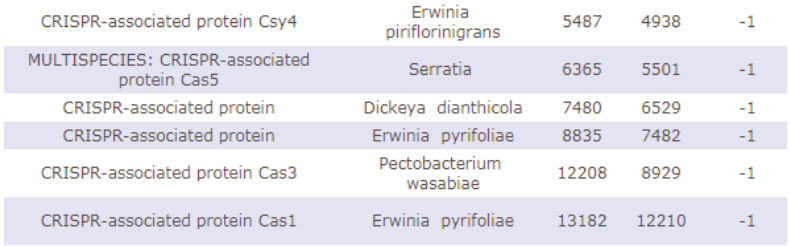
FONTE: CRISPRCasFinder (2018).

NOTA: Captura de tela tirada a partir da ferramenta de busca CRISPRCasFinder ilustrando o segundo CRISPR com alto grau de evidência (4) encontrado na sequência de *Triticum urartu* (KD046323.1). Na primeira e última linhas da sequência estão regiões flanqueadoras de 100 nucleotídeos, em amarelo estão as sequências DR do arranjo e em cores diversas os 20 espaçadores.

Buscando pelos espaçadores desses potenciais CRISPRs na base ACLAME, foi notada compatibilidade de pelo menos um deles com o plasmídeo pBBta01 encontrado em *Bradyrhizobium sp. BTAi 1*, uma bactéria fixadora de nitrogênio que em geral pode formar nódulos na haste e raiz da planta (GIRAUD et al., 2007), com valores de 5^{-3} para o *e-value* e 100% de identidade.

Além disso, essa sequência oriunda do *Triticum urartu* foi analisada pelo *software* SILA, responsável por fazer anotação automática de genes procarióticos (VIALLE, 2013). Essa anotação retornou um agrupamento de genes *Cas*, todos na mesma fita, como é possível ver na recuperação de tela mostrada na FIGURA 13.

FIGURA 13: CAPTURA DE TELA MOSTRANDO GENES *CAS* EM *Triticum urartu*



CRISPR-associated protein Csy4	Erwinia piriflorinigrans	5487	4938	-1
MULTISPECIES: CRISPR-associated protein Cas5	Serratia	6365	5501	-1
CRISPR-associated protein	Dickeya dianthicola	7480	6529	-1
CRISPR-associated protein	Erwinia pyrifoliae	8835	7482	-1
CRISPR-associated protein Cas3	Pectobacterium wasabiae	12208	8929	-1
CRISPR-associated protein Cas1	Erwinia pyrifoliae	13182	12210	-1

FONTE: SILA (Vialle, 2013).

NOTA: A imagem acima mostra a captura de tela referente ao *software* SILA, responsável pela anotação automática de genes procarióticos. A análise foi feita através da submissão da sequência de

Triticum urartu (KD046323.1) em formato fasta e seis diferentes genes *Cas* foram identificados. Na segunda coluna constam organismos dos quais houve maior similaridade de sequência entre os genes conhecidos e os genes encontrados, sendo que a grande maioria desses organismos são fitopatógenos, ou seja, são patógenos de plantas. Além disso, é possível notar que todos os genes encontrados estão na mesma fita (indicado pelo –1) e estão organizados em sequência como operon (colunas 3 e 4).

O que pode justificar essa ocorrência de genes *Cas* e provável presença do sistema CRISPRs nessa espécie de trigo é a capacidade do sistema CRISPR em conferir resistência a vírus em plantas, sendo que essas plantas têm o CRISPR inserido no seu genoma e isso faz com que elas não tenham mais sintomas de doenças causadas pelos vírus que às atingia (ALI et al., 2016). Outras pesquisas ainda são necessárias para elucidar melhor esse mecanismo.

6.4.2 *Cluster 260B* – Segundo maior grupo de sequências DR de bactérias

Os melhores resultados das análises realizadas nas sequências DR presentes no *Cluster 260B* foram obtidos de duas sequências que se diferenciam entre si pela adição de duas adeninas no final da primeira sequência (sublinhado), como mostra a FIGURA 14. O resultado dessas e das demais sequências podem ser vistos no APÊNDICE 4.

FIGURA 14: SEQUÊNCIAS QUE APRESENTARAM OS MELHORES RESULTADOS ENCONTRADOS NO *CLUSTER 260B*

>Cluster_260_NZ_CP010172_2
GTGTTCCCCGCGCCAGCGGGGATAAA
>Cluster_260_NZ_LT571437_2
GTGTTCCCCGCGCCAGCGGGGATA

FONTE: RAFTS³G (2018).

NOTA: Quadro representativo contendo as duas sequências DR do *Cluster 260B* que apresentaram os melhores resultados na pesquisa feita na ferramenta RNAcentral. É possível notar a grande similaridade entre elas, sendo que se diferenciam apenas por dois nucleotídeos no final da sequência (sublinhado). O resultado completo dessa e das demais sequências à consulta ao RNAcentral pode ser visto no Apêndice 4.

A primeira sequência corresponde à DR do CRISPR de *Escherichia coli H8* (GCF_001900835.1, GCA_001900835.1) e não se repete em outros CRISPRs nesse *cluster.* Essa pesquisa no RNAcentral teve correspondência com 100% de identidade

e cobertura da consulta com uma sequência identificada como *Escherichia coli 2-316-03_S4_C1 partial 16S ribosomal RNA*.

A segunda sequência é proveniente do CRISPR de *Salmonella enterica subsp. enterica serovar Java* (GCF_900086565.1, GCA_900086565.1), exclusivamente. É relevante registrar que há grande similaridade entre essas duas e outras 51 sequências, sendo que, em geral, apenas são acrescidas por poucos nucleotídeos no final da sequência. Essa sequência também obteve resultado de 100% de identidade e cobertura de consulta com a sequência identificada como *Escherichia coli 2-316-03_S4_C1 rRNA*. Apesar de identificadas em organismos diferentes, o resultado quase idêntico provavelmente se deve à proximidade taxonômica entre a *Escherichia coli* e a *Salmonella*, e há possibilidade de que tenha ocorrido transferência horizontal de genes entre esses gêneros próximos.

A ocorrência de similaridade com rRNA abre caminho para discussão sobre a conservação das sequências DR de forma semelhante à conservação de regiões de rRNA. Outras análises são necessárias para elucidar essa ocorrência.

6.5 ANÁLISE DE FILOGENIA PELO USO DA FERRAMENTA SWEEP

Buscando investigar se a composição nucleotídica dos genes *Cas1* e *Cas2* tem relação com o processo de inserção de novas sequências DR no arranjo CRISPR, obteve-se as sequências em nucleotídeos dos dois genes supracitados através da ferramenta CRISPRCasFinder para um representante de cada sequência DR distinta presente no *Cluster 1B* escolhido aleatoriamente.

Logo de início foi possível notar que, diferentemente do que é descrito na literatura (MAKAROVA et al., 2006), esses genes não são universais, uma vez que *Cas1* foi encontrado em 24 dos 30 representantes, enquanto *Cas2* foi visto em 15 dos mesmos organismos.

Ao comparar as árvores filogenéticas das sequências dos genes *Cas1* e *Cas2* com as árvores das sequências DR, observou-se baixa similaridade entre elas e consequente ausência de regiões motivos para identificar uma correlação direta entre os segmentos. Não foram identificadas regiões conservadas concomitantes nessas sequências, o que permite afirmar que a origem da composição da sequência DR não se encontra presente nas sequências desses genes ou nas regiões adjacentes.

Sendo assim, a origem e incorporação de novas sequências DR não têm relação direta com a composição nucleotídica das sequências presentes nos genes *Cas1* e *Cas2*. As árvores podem ser vistas nos APÊNDICES 5 (relativo a *Cas1*) e 6 (relativo a *Cas2*).

6.6 ANÁLISE SIMULTÂNEA ENTRE ARQUEAS E BACTÉRIAS

O agrupamento realizado com a ferramenta RAFTS³G usando todas as sequências DR de bactérias e de arqueas forneceu 1.333 *clusters*, dos quais dois apresentaram sequências referentes a arqueas e bactérias simultaneamente: o *Cluster 109AB*[1] (com 31 sequências de bactérias e uma sequência de arquea), e o *Cluster 1270AB* (com duas sequências de bactérias e uma de arquea). Análises realizadas em busca do ancestral entre esses pontos de intersecção não foram suficientes para inferir um ancestral comum entre os reinos. Dessa forma, é possível que dois eventos independentes de HGT tenham ocorrido entre esses organismos presentes nos dois *clusters.* Entretanto, faltam indícios que comprovem uma real relação de descendência entre os organismos desses dois *clusters.*

6.7 ANÁLISE DAS SEQUÊNCIAS DE DNA DOS ESPAÇADORES

Os dados referentes às sequências de DNA dos espaçadores foram obtidos pelo arquivo de instalação do *software* phageParser (BONSMA et al., 2016). Nesse arquivo os organismos estão identificados por um número compatível com o CRISPR_id proveniente do CRISPRdb, seguido de sua posição no arranjo CRISPR (ex. NC_019042_2_8). É importante dizer que, diferentemente das sequências DR, os espaçadores não estão separados entre arqueas e bactérias, uma vez que um mesmo espaçador pode estar presente em ambos os reinos.

O arquivo ‘spacerdatabase.txt’ apresenta 120.497 sequências de espaçadores, vários compartilhados por mais de um organismo ou arranjo CRISPR. Essas sequências, assim como as sequências DR, foram salvas em arquivo multifasta e agrupadas através da ferramenta RAFTS³G com critério equivalente ao alinhamento de no mínimo de 50% de identidade. Os resultados apresentaram grupos com baixa

[1] AB refere-se a agrupamentos entre arqueas e bactérias.

densidade. Foram formados 115.825 *clusters*, sendo o maior deles com 34 sequências; percentual baixo, consequência da pequena identidade entre as sequências, diferentemente do que é visto entre as sequências DR.

Quanto às características dos espaçadores, foi possível notar uma diferença em relação ao conteúdo GC quando comparado com o conteúdo GC observado nas sequências DR. Enquanto que nos *clusters* principais das sequências DR o conteúdo GC ultrapassa a faixa de 65%, entre os espaçadores o conteúdo GC médio é de 47,70%. Sabe-se que a maior hipótese sobre a origem dos espaçadores seja o genoma viral; tendo isso em vista, o fato do conteúdo GC dos espaçadores ser menor que o visto nas sequências DR é explicado pelo conceito afirmado na literatura de que o genoma viral tem, em geral, conteúdo GC 7% menor do que o conteúdo GC do genoma bacteriano (KUNIN; SOREK; HUGENHOLTZ, 2007; POURCEL; SALVIGNOL; VERGNAUD, 2006).

A análise quanto à presença de MGEs ficou comprometida devido ao número de sequências, tendo em vista que a ferramenta ACLAME permite apenas entrada de 1.000 sequências a cada consulta. Assim, um teste amostral com as 1.000 primeiras sequências retornou 258 MGEs para 150 diferentes espaçadores. Entre eles, são 196 MGEs diferentes que se dividem em 118 plasmídeos, 44 vírus e 34 profagos. Todos os resultados obtidos podem ser vistos no APÊNDICE 7.

Os resultados mais representativos (com menor valor de *e-value*) são todos para vírus que têm como hospedeiros organismos do gênero *Methanothermobacter*, e seu CRISPR de origem vem da arquea *Methanothermobacter thermautotrophicus str. Delta H*. É interessante dizer que os sete espaçadores com melhores resultados na pesquisa do ACLAME estão presentes no mesmo arranjo CRISPR do organismo citado acima.

Esperávamos que todos os espaçadores pesquisados através do ACLAME fossem compatíveis com algum MGE, partindo do conceito firmado de que os espaçadores são sequências curtas provenientes de material exógeno contaminante que invade o genoma da bactéria ou arqueas. Por esse motivo, o encontro de correspondências entre MGEs e espaçadores em apenas 150 das 1.000 sequências analisadas é surpreendente. Isso corrobora com o que foi visto por Shmakov e colaboradores (2017), onde para apenas 7% (em média) dos espaçadores pesquisados foi detectado um protoespaçador (sequência homóloga) compatível. Assim, os pesquisadores afirmam que os demais espaçadores são, em geral, "matéria

escura" e abrem a possibilidade de que eles sejam originados de mobilomas microbianos próprios de cada espécie, ideia que ainda deve ser desbravada (SHMAKOV et al., 2017).

7 CONCLUSÃO

A análise exploratória de sequências de DNA de regiões genômicas associadas ao sistema CRISPR foi realizada a partir de etapas que abrangeram desde a coleta de dados para o povoamento de um banco de dados local, a seleção de organismos com genomas completos e CRISPR confirmado e de suas sequências estruturais básicas (DR e espaçadores), até a análise dos resultados oriundos do agrupamento com base na identidade das sequências e confronto com bases de dados públicas. Com base nos dados encontrados e nas análises realizadas, conclui-se:

I. Os componentes principais do arranjo CRISPR – DR e espaçadores – estão relacionados com diferentes estruturas funcionais descritas

Isso pôde ser visto através da similaridade de sequências com alguns elementos genéticos móveis, bem como semelhanças com regiões de RNA CRISPR e compatibilidade de regiões com Ilhas Genômicas.

II. Há alta correlação entre sequências DR e RNAs conservados em um dos *clusters*

A grande similaridade entre algumas sequências DR do *Cluster 260B* de bactérias, o segundo de maior densidade, e regiões de RNA ribossomal 16S trouxe a possibilidade de estudar a conservação e disseminação do sistema CRISPR entre os organismos através da análise de 16S.

III. CRISPR está sendo transferido horizontalmente de bactérias para eucariotos

O encontro de arranjos CRISPR na espécie de trigo *Triticum urartu* trouxe à tona o fato de que bactérias têm a capacidade de transferir horizontalmente essa estrutura para eucariotos que, por sua vez, a usam como um artefato de resistência adquirida, se protegendo de lesões antes causadas por organismos fitopatogênicos.

IV. Existe a possibilidade de inferir o ancestral comum entre bactérias e arqueas

A ocorrência de dois *clusters* com organismos dos domínios arquea e bactéria simultaneamente trouxe a possibilidade de inferir um ancestral comum entre os dois

reinos; porém, as análises executadas neste desenvolvimento não foram suficientes para responder esta questão e novas sequências CRISPR provavelmente são a peça-chave para solucionar esse questionamento.

V. Falta de informações sobre a real origem dos espaçadores e de sequências DR

Existe na literatura o conceito firmado de que os espaçadores têm origem de elementos genéticos móveis como plasmídeos e bacteriófagos. Porém, os confrontos entre sequências de espaçadores e bancos de MGEs demonstraram que a grande maioria dos espaçadores não apresenta uma sequência homóloga com um MGE conhecido, trazendo o conceito de que os espaçadores são uma "matéria escura" que precisa ter sua origem melhor descrita. Já sobre as sequências DR, foi possível notar que a sua origem não tem relação com a composição nucleotídica dos genes *Cas1* e *Cas2*, sendo que as árvores filogenéticas não tiveram regiões de similaridade. A forma como *Cas1* e *Cas2* interferem na origem e inserção de novas sequências DR ainda precisa ser melhor estudada.

VI. Devido à baixa diversidade observada no maior *cluster* de DRs, era esperado uma relação de sintenia maior entre os genes *Cas*

Durante as análises foi possível notar que as sequências DR dentro dos *clusters* são bastante conservadas e pouco diversas dentro das espécies; com base nisso, esperava-se que o mesmo ocorresse entre os genes *Cas*, fazendo com que a relação de sintenia entre eles fosse ampla, porém percebeu-se que eles estão menos conservados quando comparados às DRs.

Para o futuro, espera-se estudar os demais *clusters* além dos dois de maior densidade, possibilitando assim descobrir novas relações com estruturas conhecidas. Ainda, conforme o número de genomas sequenciados aumente e novos CRISPRs sejam descritos, haverá a necessidade de repetição do estudo partindo de novo agrupamento, visto que poderá haver mudança na composição dos *clusters*. Por fim, faz-se extremamente necessária a criação de um banco de dados bem consolidado de elementos genéticos móveis, buscando aumentar o poder de resposta sobre qual é realmente a origem dos espaçadores e de quê os organismos se protegem ao adquiri-los.

BIBLIOGRAFIA

ALI, Z. et al. CRISPR/Cas9-Mediated Immunity to Geminiviruses: Differential Interference and Evasion. **Scientific Reports**, v. 6, mai, 2016.

BARRANGOU, R. et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. **Science**, v. 315, p. 1709–1712, 2007.

BARRANGOU, R. CRISPR-Cas systems and RNA-guided interference. **Wiley Interdisciplinary Reviews: RNA**, v. 4, n. 3, p. 267–278, mai, 2013.

BARRANGOU, R.; HORVATH, P. A decade of discovery: CRISPR functions and applications. **Nature Microbiology**, v. 2, jun, p. 1–9, 2017.

BERTELLI, C. et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. **Nucleic Acids Research**, v. 45, n. W1, p. W30–W35, 2017.

BINNS, D. et al. QuickGO: a web-based tool for Gene Ontology searching. **Bioinformatics**, v. 25, n. 22, p. 3045–3046, nov. 2009.

BISWAS, A. et al. CRISPRTarget: Bioinformatic prediction and analysis of crRNA targets. **RNA Biology**, v. 10, p. 817–827, 2013.

BISWAS, A. et al. CRISPRDetect: A flexible algorithm to define CRISPR arrays. **BMC Genomics**, v. 17, n. 1, p. 1–14, 2016.

BONSMA, M. et al. **phageParser.** Disponível em: <https://github.com/phageParser/phageParser> Acesso em: 16 out. 2018.

BROUNS, S. J. J. et al. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. **Science**, v. 321, n. 5891, p. 960–964, 15 ago. 2008.

BULT, C. J. et al. Complete genome sequence of the Methanogenic archaeon, Methanococcus jannaschii. **Science**, v. 273, n. 5278, p. 1058–1073, 1996.

BURSTEIN, D. et al. New CRISPR–Cas systems from uncultivated microbes. **Nature**, v. 542, n. 7640, p. 237–241, 22 dez. 2016.

CAMARGO, J. O. **Análise filogenética de proteoma de cloroplastos e outros plastídios:** uma abordagem livre de alinhamento utilizando o algoritmo Svect. 2018. 102f. Dissertação de Mestrado (Programa de Pós-Graduação em Bioinformática) - Universidade Federal do Paraná, 2018.

CHAI, G. et al. HMMCAS: a web tool for the identification and domain annotations of Cas proteins. **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, v. 5963, p. 1–1, 2017.

COIMBRA, N. A. D. R. **METODOLOGIA COMPUTACIONAL PARA ESTUDO DE GENES COM VIZINHAÇA CONECTADA:** análise do cluster nif. 2015. 80f. Dissertação de Mestrado (Programa de Pós-Graduação em Bioinformática) - Universidade Federal do Paraná, 2015.

CONSORTIUM, T. RNA. RNAcentral: a comprehensive database of non-coding RNA sequences. **Nucleic Acids Research**, v. 45, p. 128–134, 2016.

COUVIN, D. et al. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. **Nucleic Acids Research**, v. 46, n. W1, p. 246–251, 2018.

CRAWLEY, A. B.; HENRIKSEN, J. R.; BARRANGOU, R. CRISPRdisco: An Automated Pipeline for the Discovery and Analysis of CRISPR-Cas Systems. **The CRISPR Journal**, v. 1, n. 2, 2018.

DE PIERRI, C. R. **REPRESENTAÇÕES VETORIAIS DE PROTEOMAS:** um estudo de caso com sequências mitocondriais. 2017. 89f. Dissertação de Mestrado (Programa de Pós-Graduação em Bioinformática) - Universidade Federal do Paraná, 2017.

DEVEAU, H. et al. Phage Response to CRISPR-Encoded Resistance in Streptococcus thermophilus. **Journal of Bacteriology**, v. 190, n. 4, p. 1390–1400, 2008.

EDGAR, R. C. PILER-CR: Fast and accurate identification of CRISPR repeats. **BMC Bioinformatics**, 2007.

GE, R. et al. CRISPRdigger: detecting CRISPRs with better direct repeat annotations. **Scientific Reports**, v. 6, n. 1, 2016.

GIRAUD, E. et al. Legumes Symbioses: Absence of Nod Genes in Photosynthetic Bradyrhizobia. **Science**, v. 316, n. 5829, p. 1307–1312, jun. 2007.

GRISSA, I.; VERGNAUD, G.; POURCEL, C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. **BMC Bioinformatics**, v. 8, n. 172, 2007.

GRISSA, I.; VERGNAUD, G.; POURCEL, C. CRISPRcompar: a website to compare clustered regularly interspaced short palindromic repeats. **Nucleic Acids Research**, v. 36, p. 145–148, 2008.

HAFT, D. H. et al. A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. **PLoS Computational Biology**, v. 1, 2005.

HORVATH, P.; BARRANGOU, R. CRISPR/Cas, the Immune System of Bacteria and Archaea. **Science**, v. 327, n. 5962, p. 167–170, 2010.

ISHINO, Y. et al. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. **Journal of Bacteriology**, v. 169, n. 12, p. 5429–5433, 1987.

ISHINO, Y.; KRUPOVIC, M.; FORTERRE, P. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. **Journal of Bacteriology**, v. 200, n. 7, 22 jan. 2018.

JANSEN, R. et al. Identification of genes that are associated with DNA repeats in

prokaryotes. **Molecular Microbiology**, v. 43, n. 6, p. 1565–1575, 2002.

KALVARI, I. et al. Rfam 13.0: Shifting to a genome-centric resource for non-coding RNA families. **Nucleic Acids Research**, v. 46, n. D1, p. D335–D342, 2018.

KARIMI, Z. et al. Bacterial CRISPR Regions: General Features and their Potential for Epidemiological Molecular Typing Studies. **The Open Microbiology Journal**, v. 12, n. 18, p. 59–70, 2018.

KITTS, P. A. et al. Assembly: a resource for assembled genomes at NCBI. **Nucleic Acids Research**, v. 44, n. D1, p. 73–80, nov. 2015.

KULIK, M. G. **extract_oligoProps.m.** Disponível em: <https://github.com/mariane-g/bio-info-all/blob/master/extract_oligoProps.m>. Acesso em: 30 jan. 2019.

KUNIN, V.; SOREK, R.; HUGENHOLTZ, P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. **Genome Biology**, v. 8, 2007.

LEPLAE, R.; LIMA-MENDEZ, G.; TOUSSAINT, A. ACLAME: A CLAssification of Mobile genetic Elements, update 2010. **Nucleic Acids Research**, v. 38, n. SUPPL.1, p. 57–61, nov. 2010.

LETUNIC, I.; BORK, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. **Nucleic Acids Research**, v. 44, n. W1, p. W242–W245, 8 jul. 2016.

LILLESTOL, R. K. et al. A putative viral defence mechanism in archaeal cells. **Archaea**, v. 2, p. 59–72, 2006.

LING, H.-Q. et al. Draft genome of the wheat A-genome progenitor *Triticum urartu*. **Nature**, v. 496, n. 7443, p. 87–90, mar. 2013.

MAKAROVA, K. S. et al. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. **Biology Direct**, 2006.

MAKAROVA, K. S. et al. Evolution and classification of the CRISPR-Cas systems. **Nature Reviews Microbiology**, v. 9, n. 6, p. 467–477, 2011.

MARRAFFINI, L. A.; SONTHEIMER, E. J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. **Nature Reviews Genetics**, v. 11, n. 3, p. 181–190, 2 mar. 2010.

MATHWORKS. **Oligoprop**. Disponível em: <https://www.mathworks.com/help/bioinfo/ref/oligoprop.html>. Acesso em: 17 dez. 2018.

MCGINNIS, S.; MADDEN, T. L. BLAST: At the core of a powerful and diverse set of sequence analysis tools. **Nucleic Acids Research**, v. 32, p. 20–25, 2004.

MOJICA, F. J. M. et al. Intervening Sequences of Regularly Spaced Prokaryotic

Repeats Derive from Foreign Genetic Elements. **Journal of Molecular Evolution**, v. 60, p. 174–182, 2005.

NICHIO, B. T. DE L. et al. An efficient and versatile clustering software to analyses in large protein datasets. **bioRxiv**, 2018.

NINH, A. Correlation Between GC-content and Palindromes in Randomly Generated Sequences and Viral Genomes. 2012.

ORACLE. **MySQL.** Disponível em: <https://www.oracle.com/br/mysql/> Acesso em: 14 fev. 2018.

POURCEL, C.; SALVIGNOL, G.; VERGNAUD, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. **Microbiology**, v. 151, n. 3, p. 653–663, 2006.

RAN, F. A. et al. Genome engineering using the CRISPR-Cas9 system. **Nature Protocols**, v. 8, n. 11, p. 2281–2308, 2013.

SHMAKOV, S. A. et al. The CRISPR spacer space is dominated by sequences from species-specific mobilomes. **mBio**, v. 8, n. 5, p. 1–18, 2017.

SOREK, R.; KUNIN, V.; HUGENHOLTZ, P. CRISPR - a widespread system that provides acquired resistance against phages in bacteria and archaea. **Nature Reviews Microbiology**, v. 6, p. 181–186, 2008.

TANG, T.H. et al. Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. **PNAS**, v. 99, n. 11, p. 7536–7541, 2002.

TANG, T. H. et al. Identification of novel non-coding RNAs as potential antisense regulators in the archaeon Sulfolobus solfataricus. **Molecular Microbiology**, v. 55, n. 2, p. 469–481, 2005.

VIALLE, R. A. **SILA - Um sistema para anotação automática de genomas utilizando técnicas independentes de alinhamento**. 2013. 96f. Dissertação de Mestrado (Programa de Pós-Graduação em Bioinformática) - Universidade Federal do Paraná, 2013.

VIALLE, R. A. et al. RAFTS3: Rapid Alignment-Free Tool for Sequence Similarity Search. **bioRxiv**, p. 055269, 2016.

WHEELER, T. J.; EDDY, S. R. nhmmer: DNA homology search with profile HMMs. **Bioinformatics**, v. 29, n. 19, p. 2487–2489, jul. 2013.

UFPR. Wilczek A., de Oliveira, A. M. R., Nichio, B. T. L., de Pierri, C. R., Marchaukoski, J. N., Camargo, J. O., Santos, L. G. C., Kulik, M. G., Voyceik, R., Raittz, R. T. **SVect - Slide Vector.** BR n. 51 2018 000244-7. 23 nov 2016, 06 mar. 2018.

APÊNDICE 1 – RESULTADOS REFERENTES À ANÁLISE PELO ACLAME DAS SEQUÊNCIAS DR DO *CLUSTER 1B*

#HitId	QueryId	QuerySeqLength	HitSeqLength	MgeId	MgeName	MgeType	Hosts	BitScore	EvalScore	PercIdent	QueryRange	HitRange	Strand
gene:plasmid:119510	Cluster_1_NZ_CP009104_1	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	37-9	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009104_1	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	19-47	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010132_4	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	37-9	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010132_4	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	19-47	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010238_5	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	37-9	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010238_5	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	19-47	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010242_4	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	37-9	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010242_4	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	19-47	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010344_4	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	37-9	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010344_4	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	19-47	+/+
gene:plasmid:119510	Cluster_1_NZ_CP013029_3	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	37-9	+/-
gene:plasmid:119478	Cluster_1_NZ_CP013029_3	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	19-47	+/+
gene:plasmid:119510	Cluster_1_NZ_CP013662_5	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	37-9	+/-
gene:plasmid:119478	Cluster_1_NZ_CP013662_5	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	58	5,00E-03	100	1-29	19-47	+/+
gene:plasmid:119510	Cluster_1_NC_017626_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017626_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017626_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_017628_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017628_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017628_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_017628_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017628_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017628_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_017634_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017634_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017634_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_017634_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017634_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017634_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_017646_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017646_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017646_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_020064_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_020064_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_020064_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_020064_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_020064_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_020064_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_020064_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_020064_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_020064_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP005930_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP005930_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP005930_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP005930_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP005930_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP005930_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006636_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006636_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+

gene:plasmid:119478	Cluster_1_NZ_CP006636_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP007592_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP007592_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP007592_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP007744_10	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP007744_10	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP007744_10	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP007744_9	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP007744_9	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP007744_9	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP008788_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP008788_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP008788_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009104_1	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009859_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009859_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009859_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010132_4	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010160_7	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP010160_7	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP010160_7	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010200_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP010200_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP010200_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010200_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP010200_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP010200_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010238_5	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010242_4	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010344_4	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010377_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP010377_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP010377_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010377_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP010377_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP010377_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP011254_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP011254_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP011254_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP011254_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP011254_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP011254_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP011254_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP011254_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP011254_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP012633_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP012633_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP012633_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP012635_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP012635_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-

gene:plasmid:119478	Cluster_1_NZ_CP012635_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119512	Cluster_1_NZ_CP013029_3	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP013662_5	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119510	Cluster_1_NZ_CP013990_9	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP013990_9	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP013990_9	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014007_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP014007_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP014007_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014034_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP014034_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP014034_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014034_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP014034_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP014034_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014667_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP014667_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP014667_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP015379_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP015379_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP015379_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP015379_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP015379_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP015379_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP015750_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP015750_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP015750_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP015834_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP015834_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP015834_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP016007_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016007_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016007_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP016007_7	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016007_7	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016007_7	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP016358_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016358_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016358_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP017181_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP017181_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP017181_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP017184_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP017184_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP017184_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP018362_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP018362_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP018362_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP018957_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP018957_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+

gene:plasmid:119478	Cluster_1_NZ_CP018957_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP018995_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP018995_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP018995_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP019029_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP019029_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP019029_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP019777_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP019777_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP019777_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP019777_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP019777_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP019777_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP019944_7	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP019944_7	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP019944_7	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP020058_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP020058_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP020058_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_L>N554852_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_L>N554852_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_L>N554852_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	56	2,00E-02	100	1-28	19-46	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009786_1	29	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	96	1-29	70-98	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009786_1	29	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	96	1-29	79-107	+/+
gene:plasmid:119510	Cluster_1_NZ_CP015613_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	34-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP015613_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	13-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP015613_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	22-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP017180_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	34-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP017180_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	13-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP017180_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	22-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP017184_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	34-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP017184_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	13-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP017184_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	50	1,00E+00	100	4-28	22-46	+/+
gene:plasmid:119510	Cluster_1_NC_017160_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017160_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017160_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_017160_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017160_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017160_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_017168_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_017168_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_017168_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_018145_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_018145_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_018145_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NC_022900_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NC_022900_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NC_022900_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP003709_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP003709_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+

gene:plasmid:119478	Cluster_1_NZ_CP003709_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP003715_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP003715_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP003715_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006748_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006748_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006748_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006748_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006748_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006748_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006751_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006751_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006751_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006754_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006754_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006754_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006758_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006758_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006758_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006758_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006758_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006758_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006762_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006762_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006762_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP006792_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP006792_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP006792_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP007142_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP007142_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP007142_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP007142_7	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP007142_7	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP007142_7	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP007230_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP007230_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP007230_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP007230_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP007230_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP007230_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP008943_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP008943_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP008943_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP008943_8	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP008943_8	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP008943_8	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009364_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009364_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009364_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009364_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+

gene:plasmid:119512	Cluster_1_NZ_CP009364_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009364_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009704_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009704_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009704_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009704_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009704_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009704_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009712_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009712_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009712_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009712_9	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009712_9	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009712_9	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009715_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009715_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009715_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009715_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009715_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009715_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009715_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009715_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009715_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009723_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009723_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009723_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009723_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009723_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009723_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009723_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009723_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009723_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009757_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009757_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009757_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009757_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009757_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009757_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009759_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009759_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009759_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009759_7	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009759_7	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009759_7	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009780_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009780_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009780_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009780_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009780_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009780_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-

gene:plasmid:119510	Cluster_1_NZ_CP009785_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009785_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009785_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009785_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009785_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009785_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009785_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009785_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009785_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009786_1	29	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009786_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009786_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009786_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009786_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009786_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009786_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009792_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009792_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009792_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009792_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009792_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009792_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009840_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009840_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009840_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009840_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009840_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009840_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009840_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009840_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009840_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009844_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009844_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009844_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009844_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009844_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009844_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009844_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009844_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009844_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009973_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009973_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009973_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009973_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009973_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009973_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009973_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP009973_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP009973_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP009991_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-

gene:plasmid:119512	Cluster_1_NZ_CP009991_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009991_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009991_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009991_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009991_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009997_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009997_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009997_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP009997_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP009997_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP009997_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010023_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010023_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010023_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP010067_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP010067_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP010067_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010067_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010067_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010067_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP010247_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP010247_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP010247_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP010247_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010247_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010247_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP010247_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP010247_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP010247_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP011104_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP011104_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP011104_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014024_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP014024_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP014024_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014024_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP014024_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP014024_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014024_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP014024_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP014024_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP014544_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP014544_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP014544_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP015749_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP015749_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP015749_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP015749_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP015749_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP015749_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-

gene:plasmid:119510	Cluster_1_NZ_CP016273_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP016273_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP016273_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP016273_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016273_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016273_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP016868_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP016868_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP016868_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP016868_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP016868_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP016868_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP016870_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016870_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016870_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP016870_3	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016870_3	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016870_3	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP016872_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP016872_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP016872_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP016873_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016873_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016873_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_CP016874_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP016874_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP016874_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP016874_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP016874_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP016874_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP016876_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_CP016876_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_CP016876_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_L>N554846_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_L>N554846_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_L>N554846_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_L>N554847_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_L>N554847_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_L>N554847_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_LT596221_2	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119512	Cluster_1_NZ_LT596221_2	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119478	Cluster_1_NZ_LT596221_2	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	19-46	+/+
gene:plasmid:119510	Cluster_1_NZ_LT596221_4	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_LT596221_4	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_LT596221_4	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	48	4,00E+00	96	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_L>N908249_1	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	46	2,00E+01	96	1-27	10-36	+/+
gene:plasmid:119512	Cluster_1_NZ_L>N908249_1	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	46	2,00E+01	96	1-27	37-11	+/-
gene:plasmid:119478	Cluster_1_NZ_L>N908249_1	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	46	2,00E+01	96	1-27	46-20	+/-
gene:plasmid:91230	Cluster_1_NZ_CP014476_1	29	1407	mge:1976	pACRY04	plasmid	Acidiphilium cryptum JF-5	40	1,00E+03	92	1-28	1350-1323	+/-
gene:plasmid:119510	Cluster_1_NZ_CP013245_6	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	10-37	+/+

gene:plasmid:119512	Cluster_1_NZ_CP013245_6	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP013245_6	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP013479_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP013479_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP013479_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP015877_5	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP015877_5	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP015877_5	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	46-19	+/-
gene:plasmid:119510	Cluster_1_NZ_CP015877_7	28	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	10-37	+/+
gene:plasmid:119512	Cluster_1_NZ_CP015877_7	28	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	37-10	+/-
gene:plasmid:119478	Cluster_1_NZ_CP015877_7	28	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	40	1,00E+03	92	1-28	46-19	+/-
gene:plasmid:29356	Cluster_1_NC_013946_8	29	483	mge:656	SCP1	plasmid	Streptomyces coelicolor	40	1,00E+03	100	6-25	124-105	+/-
gene:plasmid:29356	Cluster_1_NZ_CP014476_1	29	483	mge:656	SCP1	plasmid	Streptomyces coelicolor	40	1,00E+03	100	6-25	124-105	+/-
gene:plasmid:91230	Cluster_1_NZ_CP012943_1	29	1407	mge:1976	pACRY04	plasmid	Acidiphilium cryptum JF-5	38	4,00E+03	92	1-27	1350-1324	+/-
gene:plasmid:119510	Cluster_1_NZ_CP015878_2	27	123	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	38	4,00E+03	92	1-27	10-36	+/+
gene:plasmid:119512	Cluster_1_NZ_CP015878_2	27	126	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	38	4,00E+03	92	1-27	37-11	+/-
gene:plasmid:119478	Cluster_1_NZ_CP015878_2	27	120	mge:804	pLPL	plasmid	Legionella pneumophila str. Lens	38	4,00E+03	92	1-27	46-20	+/-
gene:plasmid:140560	Cluster_1_NC_017160_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NC_017160_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NC_017168_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NC_018145_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NC_022900_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP003709_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP003715_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006748_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006748_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006751_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006754_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006758_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006758_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006762_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP006792_6	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP007142_6	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP007142_7	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP007230_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP007230_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP008943_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP008943_8	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009364_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009364_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009704_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009704_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009712_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009712_9	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009715_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009715_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009715_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009723_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009723_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009723_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-

gene:plasmid:140560	Cluster_1_NZ_CP009757_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009757_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009759_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009759_7	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009780_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009780_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009785_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009785_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009785_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009786_1	29	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009786_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009786_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009792_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009792_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009840_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009840_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009840_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009844_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009844_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009844_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009973_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009973_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009973_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP009991_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009991_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009997_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP009997_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP010023_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP010067_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP010067_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP010247_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP010247_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP010247_5	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP011104_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP014544_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP016273_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP016273_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP016868_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP016868_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP016870_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP016870_3	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP016872_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP016873_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_CP016874_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP016874_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_CP016876_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_L>N554846_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_L>N554847_1	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-
gene:plasmid:140560	Cluster_1_NZ_LT596221_2	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	6-24	524-542	+/+
gene:plasmid:140560	Cluster_1_NZ_LT596221_4	28	618	mge:1048	p42e	plasmid	Rhizobium etli CFN 42	38	4,00E+03	100	5-23	542-524	+/-

gene:proph:178606	Cluster_1_NZ_CP016665_9	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016665_9	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016668_7	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016668_7	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016669_5	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016669_5	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016669_6	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016669_6	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016669_7	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016669_7	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016673_6	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016673_6	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016674_4	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016674_4	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016674_5	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016674_5	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016674_6	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016674_6	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016674_7	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016674_7	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016674_8	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016674_8	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016675_10	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016675_10	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016675_11	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016675_11	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016675_9	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016675_9	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016677_6	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016677_6	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016678_5	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016678_5	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016678_6	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016678_6	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016678_7	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016678_7	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016681_7	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016681_7	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016681_8	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016681_8	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016681_9	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016681_9	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016684_5	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016684_5	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016684_6	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016684_6	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:178606	Cluster_1_NZ_CP016684_7	26	504	mge:2575	prophinder:44025	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	238-213	+/-
gene:proph:178597	Cluster_1_NZ_CP016684_7	26	369	mge:2574	prophinder:44024	prophage	Neisseria gonorrhoeae FA 1090	52	2,00E-01	100	1-26	103-78	+/-
gene:proph:168314	Cluster_1_NZ_CP016665_9	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016668_7	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+

gene:proph:168314	Cluster_1_NZ_CP016669_5	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016669_6	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016669_7	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016673_6	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016674_4	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016674_5	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016674_6	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016674_7	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016674_8	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016675_10	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016675_11	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016675_9	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016677_6	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016678_5	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016678_6	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016678_7	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016681_7	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016681_8	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016681_9	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016684_5	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016684_6	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:168314	Cluster_1_NZ_CP016684_7	26	210	mge:2287	prophinder:47307	prophage	Neisseria meningitidis 053442	52	2,00E-01	100	1-26	157-182	+/+
gene:proph:182405	Cluster_1_NZ_CP016665_9	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016668_7	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016669_5	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016669_6	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016669_7	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016673_6	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016674_4	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016674_5	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016674_6	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016674_7	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016674_8	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016675_10	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016675_11	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016675_9	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016677_6	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016678_5	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016678_6	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016678_7	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016681_7	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016681_8	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016681_9	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016684_5	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016684_6	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-
gene:proph:182405	Cluster_1_NZ_CP016684_7	26	528	mge:2680	prophinder:44629	prophage	Neisseria meningitidis MC58	52	2,00E-01	100	1-26	222-197	+/-

APÊNDICE 2 - RESULTADOS REFERENTES À ANÁLISE PELO ACLAME DAS SEQUÊNCIAS DR DO *CLUSTER 260B*

#HitId	QueryId	QuerySeqLength	HitSeqLength	MgeId	MgeName	MgeType	MgeSize	Hosts	BitScore	EvalScore	PercIdent	QueryRange	HitRange	Strand
gene:plasmid:29356	Cluster_260_NZ_CP012382_10	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	58	5,00E-03	100	1-29	101-129	+/+
gene:plasmid:29356	Cluster_260_NZ_CP013142_27	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	58	5,00E-03	100	1-29	101-129	+/+
gene:plasmid:29356	Cluster_260_NZ_CP019724_9	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	58	5,00E-03	100	1-29	101-129	+/+
gene:plasmid:29356	Cluster_260_NC_016582_1	26	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	52	2,00E-01	100	1-26	101-126	+/+
gene:plasmid:29356	Cluster_260_NZ_CP010407_20	25	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	50	8,00E-01	100	1-25	101-125	+/+
gene:plasmid:29356	Cluster_260_NZ_CP010407_21	25	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	50	8,00E-01	100	1-25	101-125	+/+
gene:plasmid:29356	Cluster_260_NZ_CP010407_25	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	50	1,00E+00	96	1-29	101-129	+/+
gene:plasmid:29356	Cluster_260_NZ_CP015622_2	28	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	48	4,00E+00	100	5-28	105-128	+/+
gene:plasmid:29356	Cluster_260_NZ_CP010407_17	28	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	44	6,00E+01	96	1-26	104-129	+/+
gene:plasmid:29356	Cluster_260_NZ_CP007574_10	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	44	7,00E+01	100	3-24	103-124	+/+
gene:plasmid:29356	Cluster_260_NZ_CP007574_11	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	44	7,00E+01	100	3-24	103-124	+/+
gene:plasmid:29356	Cluster_260_NZ_CP007574_9	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	44	7,00E+01	100	3-24	103-124	+/+
gene:plasmid:29356	Cluster_260_NC_013501_1	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	42	3,00E+02	96	5-29	105-129	+/+
gene:plasmid:29356	Cluster_260_NC_016785_1	28	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	40	1,00E+03	100	5-24	166-185	+/+
gene:plasmid:29356	Cluster_260_NC_016787_1	28	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	40	1,00E+03	100	5-24	166-185	+/+
gene:plasmid:29356	Cluster_260_NC_016788_1	28	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	40	1,00E+03	100	5-24	166-185	+/+
gene:plasmid:29356	Cluster_260_NZ_CP004350_2	28	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	40	1,00E+03	100	5-24	166-185	+/+
gene:plasmid:29356	Cluster_260_NZ_CP004350_3	28	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	40	1,00E+03	100	5-24	166-185	+/+
gene:plasmid:91230	Cluster_260_NZ_CP014360_1	29	1407	mge:1976	pACRY04	plasmid	37415	Acidiphilium cryptum JF-5	40	1,00E+03	92	2-29	1323-1350	+/+
gene:plasmid:29356	Cluster_260_NZ_CP014360_1	29	483	mge:656	SCP1	plasmid	356023	Streptomyces coelicolor	40	1,00E+03	100	5-24	105-124	+/+
gene:plasmid:91230	Cluster_260_NC_017574_2	29	1407	mge:1976	pACRY04	plasmid	37415	Acidiphilium cryptum JF-5	38	4,00E+03	92	3-29	1324-1350	+/+
gene:plasmid:91230	Cluster_260_NZ_CP010415_2	29	1407	mge:1976	pACRY04	plasmid	37415	Acidiphilium cryptum JF-5	38	4,00E+03	92	3-29	1324-1350	+/+
gene:plasmid:91230	Cluster_260_NZ_CP012939_2	29	1407	mge:1976	pACRY04	plasmid	37415	Acidiphilium cryptum JF-5	38	4,00E+03	92	3-29	1324-1350	+/+
gene:plasmid:91230	Cluster_260_NZ_LK391695_2	29	1407	mge:1976	pACRY04	plasmid	37415	Acidiphilium cryptum JF-5	38	4,00E+03	92	3-29	1324-1350	+/+
gene:plasmid:91230	Cluster_260_NZ_LK391695_3	29	1407	mge:1976	pACRY04	plasmid	37415	Acidiphilium cryptum JF-5	38	4,00E+03	92	3-29	1324-1350	+/+

APÊNDICE 3 – RESULTADOS REFERENTES À ANÁLISE PELO RNACENTRAL DAS SEQUÊNCIAS DR DO *CLUSTER 1B*

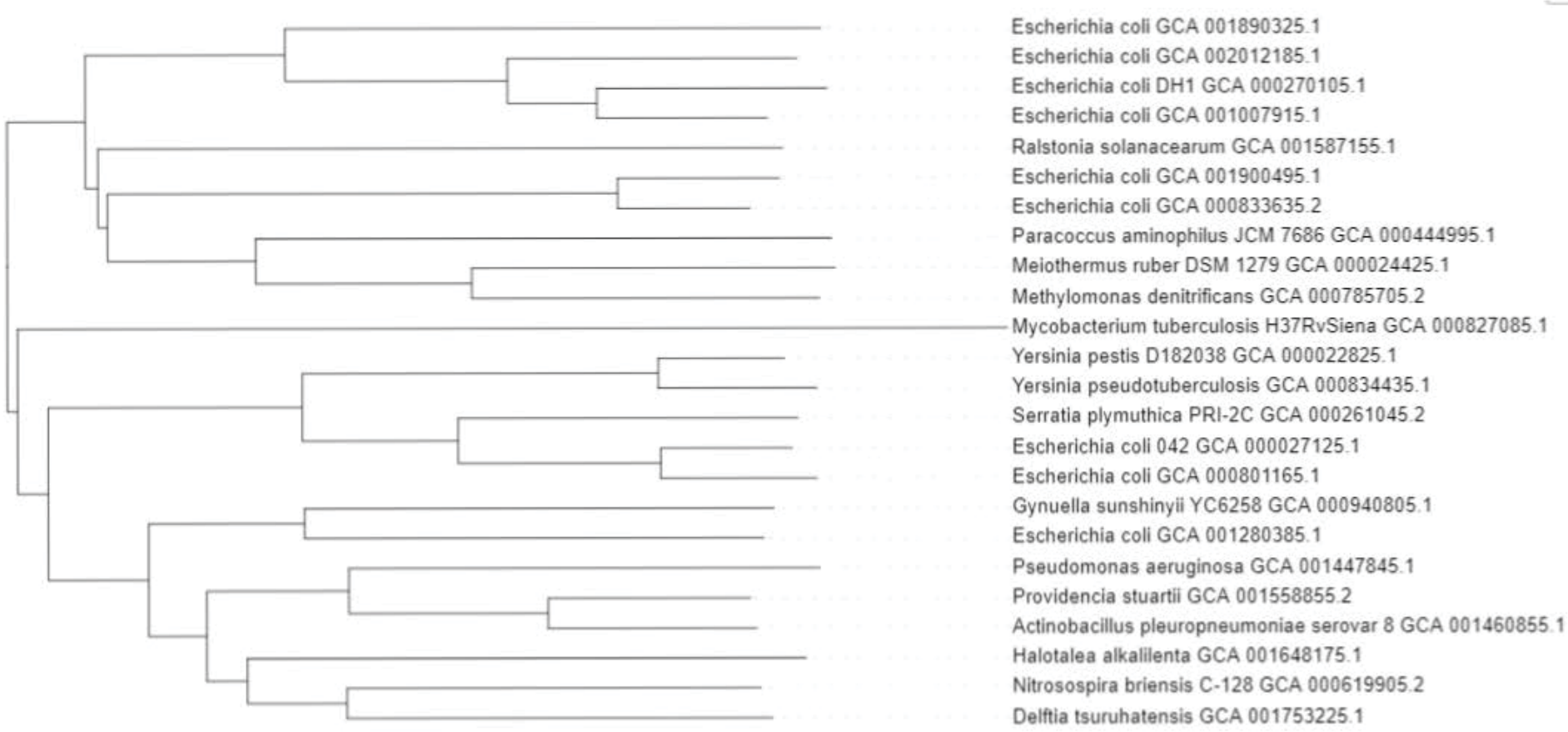
Cluster ID	DR	Melhor resultado	E-value	Identity	Query coverage	Target coverage	Link
Cluster_1_NC_017160_3	GTTCACTGCCGCACAGGCAGCTTAGAAA	CRISPR RNA direct repeat element other from 28 species	2.30e+0	100.0% (28/28)	100.0% (28/28)	100.0% (28/28)	https://rnacentral.org/rna/URS00006562F9
Cluster_1_NC_017626_3	GTTCACTGCCGTACAGGCAGCTTAGAAA	CRISPR RNA direct repeat element other from 39 species	1.50e+0	100.0% (28/28)	100.0% (28/28)	100.0% (28/28)	https://rnacentral.org/rna/URS00004200E4
Cluster_1_NZ_CP015613_1	GTGCACTGCCGTACAGGCAGCTTAGAAA	CRISPR RNA direct repeat element from 4 species	1.80e+0	100.0% (28/28)	100.0% (28/28)	100.0% (28/28)	https://rnacentral.org/rna/URS0000682811
Cluster_1_NZ_CP012346_3	GTGTTCCCCGCGCCAGCGGGGATAAAC	Escherichia coli 2-316-03_S4_C1 rRNA	1.60e+1	100.0% (27/27)	100.0% (27/27)	27.3% (27/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_1_NZ_CP009104_1	GTTCACTGCCGTACAGGCAGCTTAGAAAT	CRISPR RNA direct repeat element other from 39 species	1.70e+0	100.0% (28/28)	96.6% (28/29)	100.0% (28/28)	https://rnacentral.org/rna/URS00004200E4
Cluster_1_NZ_CP009786_1	GTTCACTGCCGCACAGGCAGCTTAGAAAA	CRISPR RNA direct repeat element other from 28 species	2.70e+0	100.0% (28/28)	96.6% (28/29)	100.0% (28/28)	https://rnacentral.org/rna/URS00006562F9
Cluster_1_NZ_CP016665_9	GATTCCCGCCTGCGCGGGAATGACGG	Neisseria meningitidis 80179 rRNA	3.90e+1	100.0% (25/25)	96.2% (25/26)	4.5% (25/556)	https://rnacentral.org/rna/URS00005392AE
Cluster_1_NZ_CP011124_2	GTGTTCCCCGCGCCAGCGGGGATAAACC	Escherichia coli 2-316-03_S4_C1 rRNA	8.70e+0	100.0% (27/27)	96.4% (27/28)	27.3% (27/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_1_NZ_CP020107_2	GAGTTCCCCGCGCCAGCGGGGATAAACCG	Escherichia coli other (RNA14_SPACER1_SPACER2)	7.10e+0	100.0% (28/28)	96.6% (28/29)	25.2% (28/111)	https://rnacentral.org/rna/URS00005BB5A6
Cluster_1_NZ_CP010117_5	GTGTTCCCCGCGCCAGCGGGGATAAACCG	Escherichia coli 2-316-03_S4_C1 rRNA	5.00e+0	100.0% (28/28)	96.6% (28/29)	28.3% (28/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_1_NZ_CP015243_1	TTTCTGAGCTGCCTATACGGCAGCGAAC	rRNA from 2 species	8.20e+3	81.5% (22/27)	96.4% (27/28)	7.2% (27/377)	https://rnacentral.org/rna/URS00003B300D
Cluster_1_NZ_CP020055_2	GGTTTATCCCCGCTGGCGCGGGGAACTCA	rRNA from 2 species	9.40e+3	83.3% (20/24)	82.8% (24/29)	2.8% (24/871)	https://rnacentral.org/rna/URS00000B4F3C
Cluster_1_NZ_CP015878_2	TTTCTTAGCTGCCTATACGGCAGTGAA	Brevipalpus sp. ARP-2015 rRNA	8.70e+3	84.0% (21/25)	92.6% (25/27)	0.6% (25/3989)	https://rnacentral.org/rna/URS00008DB402
Cluster_1_NZ_CP012943_1	CGGTTCATCCCCGCAGGCGCGGGGAACAC	Thermoproteus sp. CIS_19 rRNA	9.10e+3	84.0% (21/25)	86.2% (25/29)	0.7% (25/3377)	https://rnacentral.org/rna/URS00009EB5E6
Cluster_1_NC_017635_4	CGGTTTATCCCCGCTGGCGCGGGGAACTC	Thermoproteus sp. AZ2 rRNA	8.40e+3	84.0% (21/25)	86.2% (25/29)	0.8% (25/3017)	https://rnacentral.org/rna/URS000082B1FF
Cluster_1_NZ_CP018962_2	GGTTTATCCCCGCTGGCGCGGGGAACAC	Thermoproteus sp. AZ2 rRNA	9.40e+3	84.0% (21/25)	89.3% (25/28)	0.8% (25/3017)	https://rnacentral.org/rna/URS000082B1FF
Cluster_1_NZ_CP013245_6	TTTCTTAGCTGCCTATACGGCAGTGAAC	Brevipalpus californicus rRNA	9.20e+3	84.0% (21/25)	89.3% (25/28)	3.1% (25/798)	https://rnacentral.org/rna/URS0000B49365
Cluster_1_NC_012971_5	CGGTTTATCCCCGCTGGCGCGGGGAACAC	Cephaloticoccus capnophilus RNase P RNA	6.30e+3	85.2% (23/27)	93.1% (27/29)	5.4% (27/503)	https://rnacentral.org/rna/URS00009DA617
Cluster_1_NZ_CP017420_4	TTTCTGAGCTGCCTATGCGGCAGCGAAC	rRNA from 2 species	1.70e+3	85.2% (23/27)	96.4% (27/28)	7.2% (27/377)	https://rnacentral.org/rna/URS00003B300D
Cluster_1_NC_022041_2	CGGTTCATCCCCGCCGGTGCGGGGAACAG	Hevea brasiliensis misc RNA	3.00e+4	85.7% (18/21)	72.4% (21/29)	1.5% (21/1445)	https://rnacentral.org/rna/URS0000B4C556
Cluster_1_NZ_CP012371_3	TTTCTGAGCTGCCTATGCGGCAGTGAAC	Rhabditella axei rRNA	4.10e+3	86.4% (19/22)	78.6% (22/28)	0.7% (22/3372)	https://rnacentral.org/rna/URS00001D20CA
Cluster_1_NZ_CP012633_4	TTTCTAAGCTGCCTGTACGGCAGTGAAC	Fukomys damarensis lncRNA	1.00e+4	86.4% (19/22)	78.6% (22/28)	0.8% (22/2932)	https://rnacentral.org/rna/URS0000BC9123
Cluster_1_NZ_LN908249_1	TTTCTAAGCTGCCTATACGGCAGTGAAG	Homo sapiens lncRNA	5.80e+3	86.4% (19/22)	78.6% (22/28)	1.9% (22/1169)	https://rnacentral.org/rna/URS0000D5A3A4
Cluster_1_NZ_CP014024_3	TTTCTAAGCTGCCTATACGGCAGTGAAC	Homo sapiens lncRNA	4.70e+3	87.0% (20/23)	82.1% (23/28)	2.0% (23/1169)	https://rnacentral.org/rna/URS0000D5A3A4
Cluster_1_NZ_CP018239_3	GGTTTATCCCCACTGGCGCGGGGAACTC	uncultured bacterium rRNA	6.90e+3	87.0% (20/23)	82.1% (23/28)	2.2% (23/1035)	https://rnacentral.org/rna/URS0000C9BF32
Cluster_1_NZ_CP007534_2	CGGTTTATCCCCGCTGGCGCGGGGAAC	rRNA from 2 species	1.30e+4	87.0% (20/23)	85.2% (23/27)	2.6% (23/871)	https://rnacentral.org/rna/URS00000B4F3C
Cluster_1_NC_013946_8	CGGTTCATCCCCGCGGGTGCGGGGAATAC	Columba livia lncRNA	5.60e+3	87.0% (20/23)	79.3% (23/29)	4.3% (23/539)	https://rnacentral.org/rna/URS0000D1E8C6
Cluster_1_NZ_CP014476_1	CGGTTCATCCCCGCGGGTGCGGGGAACGC	Columba livia lncRNA	9.90e+3	87.0% (20/23)	79.3% (23/29)	4.3% (23/539)	https://rnacentral.org/rna/URS0000D1E8C6
Cluster_1_NZ_CP007142_6	TTTCTAAGCTGCCTGTGCGGCAGTGAAC	Euglena agilis rRNA	1.10e+4	90.9% (20/22)	78.6% (22/28)	25.9% (22/85)	https://rnacentral.org/rna/URS000021C5B7
Cluster_1_NZ_CP007027_4	GTTTCCGTCCCCTCTCGGGGTTTTGGGTCTGACGAC	Sem resultado					

APÊNDICE 4 – RESULTADOS REFERENTES À ANÁLISE PELO RNACENTRAL DAS SEQUÊNCIAS DR DO *CLUSTER 260B*

Cluster ID	DR	Melhor resultado	E-value	Identity	Query coverage	Target coverage	Link
Cluster_260_NZ_LT571437_2	GTGTTCCCCGCGCCAGCGGGGATA	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	3.80e+2	100.0% (24/24)	100.0% (24/24)	24.2% (24/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP013253_3	GAGTTCCCCGCGCCAGCGGGGATAA	Escherichia coli other RNA14_SPACER1_SPACER2	1.10e+2	100.0% (25/25)	100.0% (25/25)	22.5% (25/111)	https://rnacentral.org/rna/URS00005BB5A6
Cluster_260_NZ_CP010172_2	GTGTTCCCCGCGCCAGCGGGGATAAA	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	3.20e+1	100.0% (26/26)	100.0% (26/26)	26.3% (26/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP010833_21	GTGGACCCCGCGCCTGCGGGGATGGTCC	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	3.20e+1	100.0% (26/26)	100.0% (26/26)	26.3% (26/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NC_013850_2	ATGTTCCCCGCGCCAGCGGGGATAAACCG	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	1.20e+1	100.0% (27/27)	93.1% (27/29)	27.3% (27/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NC_017625_2	GTGTTCCCCGCGCCAGCGGGGATAAACC	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	8.70e+0	100.0% (27/27)	96.4% (27/28)	27.3% (27/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP014314_1	GAGTTCCCCGCGCCAGCGGGGATAAACC	Escherichia coli other RNA14_SPACER1_SPACER2	1.20e+1	100.0% (27/27)	96.4% (27/28)	24.3% (27/111)	https://rnacentral.org/rna/URS00005BB5A6
Cluster_260_NC_010468_1	GAGTTCCCCGCGCCAGCGGGGATAAACCG	Escherichia coli other RNA14_SPACER1_SPACER2	7.10e+0	100.0% (28/28)	96.6% (28/29)	25.2% (28/111)	https://rnacentral.org/rna/URS00005BB5A6
Cluster_260_NC_012947_1	GTGTTCCCCGCGCCAGCGGGGATAAACCG	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	5.00e+0	100.0% (28/28)	96.6% (28/29)	28.3% (28/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP013483_1	GTGTTCCCCGCGCCAGCGGGGATAAACCGA	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	5.00e+0	100.0% (29/29)	96.7% (29/30)	29.3% (29/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NC_013501_1	GGTGTCCCCGCACCCGCGGGGATAGTCCC	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	7.70e+4	79.2% (19/24)	82.8% (24/29)	24.2% (24/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NC_013131_12	GGGATCAGCCCCGCGCCTGCGGGGAGCAC	rRNA from 2 species large subunit ribosomal RNA	5.10e+3	81.5% (22/27)	93.1% (27/29)	4.2% (27/644)	https://rnacentral.org/rna/URS00004B2ACB
Cluster_260_NZ_CP013988_2	GTAGTCCCCGCACCAGCGGGGGTGATCC	rRNA from 2 species ribosomal RNA (16S)	1.40e+4	81.5% (22/27)	96.4% (27/28)	3.8% (27/716)	https://rnacentral.org/rna/URS000018900A
Cluster_260_NZ_CP019326_3	CGGTTGAGCCCCGCGCCTGCGGGGATCGG	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	8.00e+4	82.6% (19/23)	79.3% (23/29)	23.2% (23/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP012382_10	CTGCTCCCCGCACCCGCGGGGATGGTCCC	Mus caroli lncRNA	1.40e+4	83.3% (20/24)	82.8% (24/29)	2.1% (24/1124)	https://rnacentral.org/rna/URS0000BC8331
Cluster_260_NZ_CP015622_2	GTTTTCCCCGCACCCGCGGGGATGGTCC	Mus musculus lncRNA	2.40e+4	83.3% (20/24)	85.7% (24/28)	2.1% (24/1128)	https://rnacentral.org/rna/URS00008120B9
Cluster_260_NC_016785_1	GTTTTCCCCGCACCAGCGGGGATGAGCC	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	2.70e+3	84.0% (21/25)	89.3% (25/28)	25.3% (25/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP014222_1	GCGTTCCCCGCACCAGCGGGGCTGAACCG	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	1.90e+3	85.2% (23/27)	93.1% (27/29)	27.3% (27/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP014360_1	GCGTTCCCCGCACCCGCGGGGATGAACCG	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	2.10e+3	85.2% (23/27)	93.1% (27/29)	27.3% (27/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP010407_25	CTGCTCCCCGCACCTGCGGGGATGGTCCC	uncultured bacterium rRNA 16S ribosomal RNA	1.80e+4	85.7% (18/21)	72.4% (21/29)	9.3% (21/225)	https://rnacentral.org/rna/URS00001BF5D9
Cluster_260_NZ_CP010407_17	CTCCCTGCACCCGCGGGGATGGTCCCGA	Mus caroli lncRNA	3.20e+3	87.5% (21/24)	85.7% (24/28)	2.1% (24/1124)	https://rnacentral.org/rna/URS0000BC8331
Cluster_260_NC_013131_16	ATCAGCCCCGCGCCTGCCGGGAGCAC	Capra hircus lncRNA	2.10e+4	88.0% (22/25)	96.2% (25/26)	2.7% (25/942)	https://rnacentral.org/rna/URS0000D6397A
Cluster_260_NZ_CP004350_2	ATGTTCCCCGCACCAGCGGGGATGAGCC	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	1.90e+3	88.5% (23/26)	92.9% (26/28)	26.3% (26/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP009247_2	GTCTTCCCCGCGCCAGCGGGGATGAGCC	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	8.60e+2	88.5% (23/26)	92.9% (26/28)	26.3% (26/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP020100_1	GTGTTCCCCGCACCTGCGGGGATGAACCG	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	6.40e+2	89.3% (25/28)	96.6% (28/29)	28.3% (28/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP007574_10	CCGCTCCCCGCACCCGCGGGGATGAGCCC	rRNA from 2 species small subunit ribosomal RNA (16S)	2.10e+4	90.5% (19/21)	72.4% (21/29)	1.5% (21/1437)	https://rnacentral.org/rna/URS00000789CA
Cluster_260_NC_009952_3	GGCTCCCCCGCCCCCGCGGGGATAGACCC	uncultured Catonella sp. rRNA	1.50e+4	91.3% (21/23)	79.3% (23/29)	6.2% (23/371)	https://rnacentral.org/rna/URS0000AD7A64
Cluster_260_NC_016002_2	GTGATCCCCGCGCCAGCGGGGATGAACCG	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	1.10e+2	92.9% (26/28)	96.6% (28/29)	28.3% (28/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NC_017574_2	GTGTTCCCCGCGCCTGCGGGGATGAACCG	Escherichia coli 2-316-03_S4_C1 rRNA 16S ribosomal RNA	1.20e+2	92.9% (26/28)	96.6% (28/29)	28.3% (28/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NC_016582_1	CTGCTCCCCGCACCCGCGGGGATGGT	Rhinopithecus bieti lncRNA	8.70e+4	94.7% (18/19)	73.1% (19/26)	1.4% (19/1370)	https://rnacentral.org/rna/URS0000D5F56F
Cluster_260_NZ_LK391695_2	GTGTTCCCCGCGCCAGCGGGGATGAACCG	Escherichia coli 2-316-03_S4_C1 rRNA partial 16S ribosomal RNA	2.50e+1	96.4% (27/28)	96.6% (28/29)	28.3% (28/99)	https://rnacentral.org/rna/URS00007800F7
Cluster_260_NZ_CP010407_20	CTGCTCCCCGCACCCGCGGGGATGG	Sem resultado					

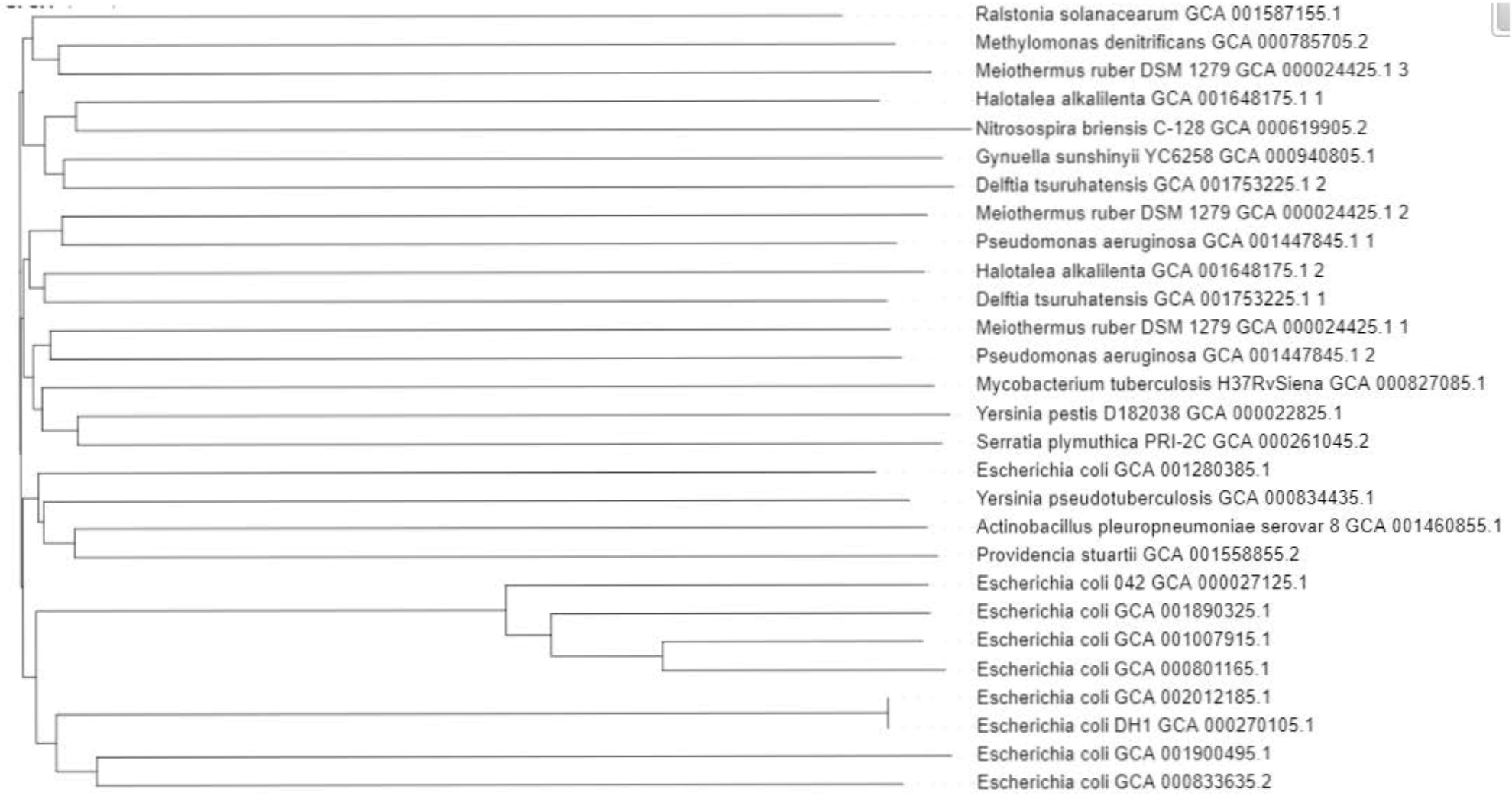
APÊNDICE 5 – COMPARAÇÃO ENTRE ÁRVORES FILOGENÉTICAS: SEQUÊNCIAS DR DO *CLUSTER 1B versus* SEQUÊNCIAS DO GENE *CAS1*

A figura[2] a seguir mostra a árvore referente às sequências DR dos organismos escolhidos que apresentam o gene *Cas1.*



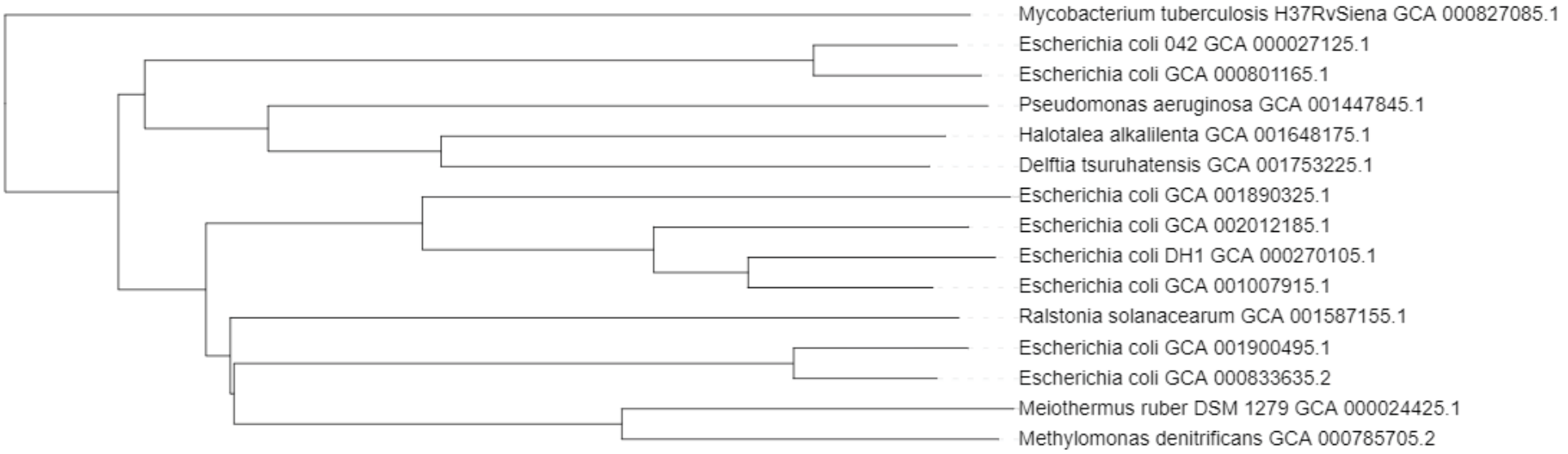
[2] Todas as imagens referentes às árvores foram retiradas do visualizador *online* iTOL (LETUNIC; BORK, 2016).

A figura˚ a seguir mostra a árvore referente às sequências do gene *Cas1* encontrados nas sequências distintas do *Cluster 1B.*

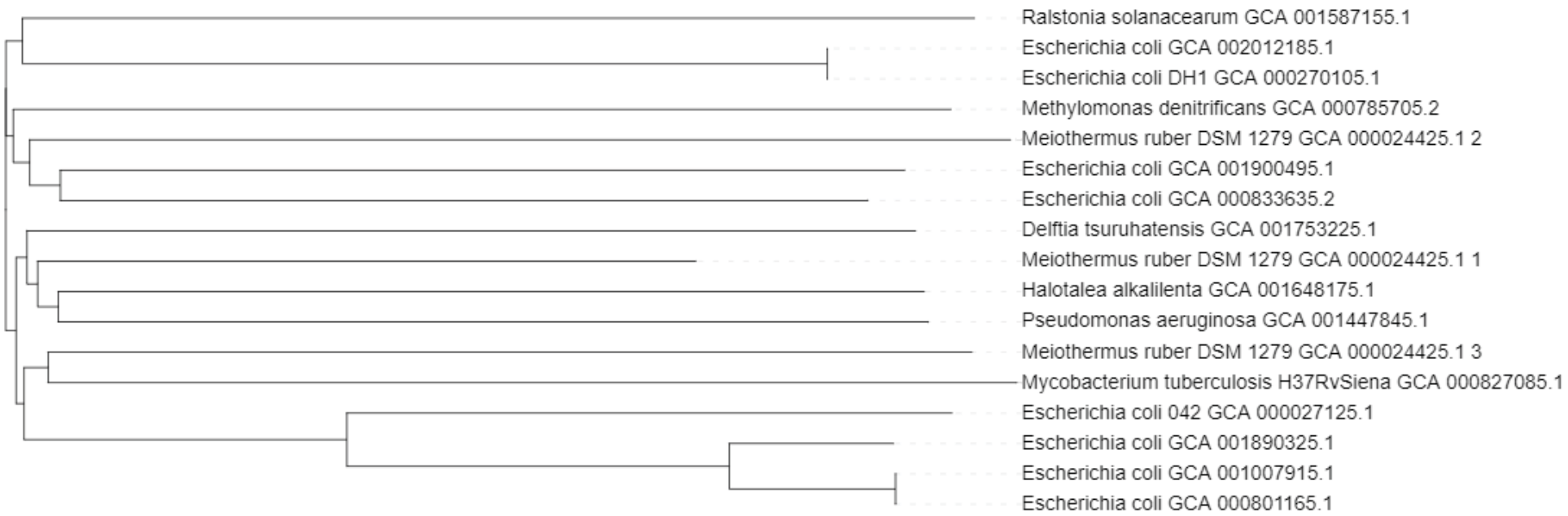


APÊNDICE 6 – COMPARAÇÃO ENTRE ÁRVORES FILOGENÉTICAS: SEQUÊNCIAS DR DO *CLUSTER 1B versus* SEQUÊNCIAS DO GENE *CAS2*

A figura[2] a seguir mostra a árvore referente às sequências DR dos organismos escolhidos que apresentam o gene *Cas2*.



A figura̋ a seguir mostra a árvore referente às sequências do gene *Cas2* encontrados nas sequências distintas do *Cluster 1B.*



APÊNDICE 7 - RESULTADOS REFERENTES À ANÁLISE PELO ACLAME DOS ESPAÇADORES

#HitId	QueryId	QuerySeqLength	HitSeqLength	MgeId	MgeName	MgeType	MgeSize	Hosts	BitScore	EvalScore	Percident	QueryRange	HitRange	Strand
gene:vir:4180	NC_000916_7_6	38	453	mge:88	psiM100	virus	28798	Methanothermobacter wolfeii	76	3,00E-08	100	1-38	273-310	+/+
gene:vir:4174	NC_000916_7_24	37	918	mge:87	psiM2	virus	26111	Methanothermobacter thermautotrophicus	74	1,00E-07	100	1-37	19-55	+/+
gene:vir:4152	NC_000916_7_22|NC_000916_7_28	36	354	mge:87	psiM2	virus	26111	Methanothermobacter thermautotrophicus	72	4,00E-07	100	1-36	137-102	+/-
gene:vir:4205	NC_000916_7_8	36	3474	mge:88	psiM100	virus	28798	Methanothermobacter wolfeii	72	4,00E-07	100	1-36	2866-2831	+/-
gene:vir:4180	NC_000916_7_7	38	453	mge:88	psiM100	virus	28798	Methanothermobacter wolfeii	70	2,00E-06	100	1-35	129-95	+/-
gene:vir:4163	NC_000916_7_110	37	420	mge:87	psiM2	virus	26111	Methanothermobacter thermautotrophicus	66	3,00E-05	97	1-37	151-187	+/+
gene:vir:4152	NC_000916_7_40	36	354	mge:87	psiM2	virus	26111	Methanothermobacter thermautotrophicus	62	4,00E-04	97	1-35	158-192	+/+
gene:plasmid:12465	NC_001880_1_1	36	402	mge:212	ece1	plasmid	39456	Aquifex aeolicus VF5	54	1,00E-01	100	1-27	27-1	+/-
gene:proph:182078	NC_002570_2_4	35	1290	mge:2670	prophinder:45952	prophage	42577	Staphylococcus aureus RF122	50	2,00E+00	96	5-33	407-379	+/-
gene:plasmid:120831	NC_000916_16_42	37	981	mge:920	p250	plasmid	6286	Avibacterium paragallinarum	42	4,00E+02	100	17-37	125-105	+/-
gene:vir:97165	NC_002570_2_4	35	1296	mge:1654	85	virus	44283	Staphylococcus aureus	42	4,00E+02	93	5-33	407-379	+/-
gene:proph:183865	NC_002570_2_4	35	1305	mge:2725	prophinder:43046	prophage	33203	Bacillus cereus subsp. cytotoxis NVH 391-98	42	4,00E+02	90	1-33	438-406	+/-
gene:proph:169763	NC_002570_2_4	35	1296	mge:2332	prophinder:45292	prophage	42628	Staphylococcus aureus subsp. aureus JH9	42	4,00E+02	93	5-33	407-379	+/-
gene:proph:173208	NC_002570_2_4	35	1296	mge:2430	prophinder:45258	prophage	42628	Staphylococcus aureus subsp. aureus JH1	42	4,00E+02	93	5-33	407-379	+/-
gene:vir:108667	NC_000909_16_2	43	1915	mge:1883	A511	virus	134494	Listeria monocytogenes	42	5,00E+02	100	21-41	1194-1214	+/+
gene:plasmid:153445	NC_000962_6_12|NC_008769_4_13|NC_(	41	198	mge:887	2	plasmid	338007	Polaromonas sp. JS666	42	5,00E+02	96	9-33	126-102	+/-
gene:plasmid:28926	NC_000909_17_5	47	552	mge:634	pSE-12228-02	plasmid	4679	Staphylococcus epidermidis ATCC 12228	42	6,00E+02	96	5-29	184-160	+/-
gene:plasmid:29896	NC_000917_1_13	35	681	mge:667	pHEN7	plasmid	7830	Sulfolobus islandicus	40	1,00E+03	95	1-24	632-655	+/+
gene:plasmid:134355	NC_000854_3_9	41	1554	mge:980	pSVH1	plasmid	12652	Streptomyces venezuelae	40	2,00E+03	95	2-25	1083-1060	+/-
gene:plasmid:13151	NC_000909_13_1	43	1539	mge:228	p1414	plasmid	7949	Bacillus subtilis	40	2,00E+03	95	7-30	428-405	+/-
gene:plasmid:13133	NC_000909_13_1	43	1449	mge:226	pTA1015	plasmid	5807	Bacillus subtilis	40	2,00E+03	95	7-30	338-315	+/-
gene:plasmid:13137	NC_000909_13_1	43	1449	mge:227	pTA1060	plasmid	8737	Bacillus subtilis	40	2,00E+03	95	7-30	338-315	+/-
gene:plasmid:111034	NC_000909_13_1	43	1449	mge:907	pBS608	plasmid	6611	Bacillus subtilis	40	2,00E+03	95	7-30	338-315	+/-
gene:plasmid:115796	NC_000909_13_1	43	1449	mge:874	pLS30	plasmid	6610	Bacillus subtilis	40	2,00E+03	95	7-30	338-315	+/-
gene:plasmid:139444	NC_000909_20_6	43	2628	mge:902	pMP118	plasmid	242436	Lactobacillus salivarius subsp. salivarius UCC118	40	2,00E+03	100	19-38	1777-1758	+/-
gene:proph:170345	NC_000916_16_14	36	669	mge:2348	prophinder:45092	prophage	38641	Dichelobacter nodosus VCS1703A	40	2,00E+03	100	14-33	620-601	+/-
gene:plasmid:90603	NC_000916_7_65	36	1383	mge:1972	pBVIE02	plasmid	265616	Burkholderia vietnamiensis G4	40	2,00E+03	100	7-26	34-15	+/-
gene:plasmid:120051	NC_000961_3_11	38	1476	mge:779	pRL12	plasmid	870021	Rhizobium leguminosarum bv. viciae 3841	40	2,00E+03	95	11-34	21-44	+/+
gene:plasmid:23348	NC_000853_1_22|NC_023151_4_19|NZ_(	37	891	mge:574	pNGR234a	plasmid	536165	Rhizobium sp. NGR234	38	6,00E+03	100	8-26	488-470	+/-
gene:vir:79170	NC_000853_5_1	36	2793	mge:1866	phiE202	virus	35741	Burkholderia thailandensis	38	6,00E+03	95	7-29	1264-1242	+/-
gene:vir:103976	NC_000853_5_1	36	2793	mge:1665	phi52237	virus	37639	Burkholderia pseudomallei	38	6,00E+03	95	7-29	1264-1242	+/-
gene:proph:165662	NC_000853_5_1	36	2793	mge:2216	prophinder:46259	prophage	38615	Burkholderia pseudomallei K96243	38	6,00E+03	95	7-29	1264-1242	+/-
gene:plasmid:82009	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	1050	mge:1901	pSMED02	plasmid	1245408	Sinorhizobium medicae WSM419	38	6,00E+03	100	18-36	73-91	+/+
gene:plasmid:135472	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	807	mge:1811	TC1	plasmid	328237	Arthrobacter aurescens TC1	38	6,00E+03	100	19-37	662-680	+/+
gene:plasmid:29083	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	963	mge:653	SAP1	plasmid	94287	Streptomyces avermitilis MA-4680	38	6,00E+03	100	14-32	477-459	+/-
gene:plasmid:21780	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	849	mge:570	pGMI1000MP	plasmid	2094509	Ralstonia solanacearum; Ralstonia solanacearum GMI1000	38	6,00E+03	100	11-29	273-291	+/+
gene:plasmid:21677	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	1614	mge:570	pGMI1000MP	plasmid	2094509	Ralstonia solanacearum; Ralstonia solanacearum GMI1000	38	6,00E+03	95	13-35	1076-1098	+/+
gene:plasmid:86825	NC_000909_11_11	36	357	mge:1939	pQBR103	plasmid	425094	Pseudomonas fluorescens SBW25	38	6,00E+03	100	17-35	116-134	+/+
gene:vir:80903	NC_000909_6_10	36	480	mge:1885	phiEF24C	virus	142072	Enterococcus faecalis	38	6,00E+03	95	3-25	45-67	+/+
gene:plasmid:121807	NC_000913_5_3|NC_007779_5_3|NC_01(	33	1536	mge:942	pSS_046	plasmid	214396	Shigella sonnei Ss046	38	6,00E+03	100	6-24	969-987	+/+
gene:vir:103124	NC_000916_16_15	37	2508	mge:1583	Syn9	virus	176847	Synechococcus sp. WH 8012; Prochlorococcus marinus	38	6,00E+03	100	6-24	1268-1250	+/-
gene:proph:160706	NC_000916_7_39	37	618	mge:2071	prophinder:46863	prophage	47746	Clostridium difficile 630	38	6,00E+03	100	4-22	154-136	+/-
gene:proph:160666	NC_000916_7_39	37	618	mge:2070	prophinder:46861	prophage	55514	Clostridium difficile 630	38	6,00E+03	100	4-22	154-136	+/-
gene:plasmid:12652	NC_000916_7_65	36	669	mge:216	pX02	plasmid	96231	Bacillus anthracis	38	6,00E+03	95	13-35	607-629	+/+
gene:plasmid:122459	NC_000916_7_65	36	675	mge:1096	pXO2	plasmid	94830	Bacillus anthracis str. 'Ames Ancestor'	38	6,00E+03	95	13-35	613-635	+/+
gene:plasmid:12939	NC_000916_7_65	36	675	mge:218	pXO2	plasmid	94829	Bacillus anthracis str. A2012	38	6,00E+03	95	13-35	613-635	+/+
gene:plasmid:141784	NC_000918_8_2	37	1872	mge:805	2	plasmid	223670	Azoarcus sp. EbN1	38	6,00E+03	100	3-21	922-904	+/-
gene:plasmid:154842	NC_000962_5_13|NC_002755_5_14|NC_(	37	1026	mge:1804	1	plasmid	653815	Paracoccus denitrificans PD1222	38	6,00E+03	100	10-28	218-200	+/-
gene:plasmid:92704	NC_000962_6_15|NC_008769_4_15|NC_(	37	1188	mge:1994	pNL2	plasmid	487268	Novosphingobium aromaticivorans DSM 12444	38	6,00E+03	95	8-30	1046-1024	+/-
gene:vir:96790	NC_000854_3_30	39	366	mge:1629	phiHSIC	virus	37966	Listonella pelagia	38	7,00E+03	100	2-20	270-288	+/+
gene:plasmid:122986	NC_000909_11_10|NC_000909_11_9	38	540	mge:1020	pE33L466	plasmid	466370	Bacillus cereus E33L	38	7,00E+03	95	8-30	44-22	+/-
gene:plasmid:122985	NC_000909_11_10|NC_000909_11_9	38	531	mge:1020	pE33L466	plasmid	466370	Bacillus cereus E33L	38	7,00E+03	95	8-30	510-488	+/-
gene:plasmid:129815	NC_000909_18_6	39	108	mge:957	pSciA	plasmid	7790	Spiroplasma citri	38	7,00E+03	95	8-30	12-34	+/+
gene:plasmid:14532	NC_000909_19_2	39	225	mge:293	pCP13	plasmid	54310	Clostridium perfringens str. 13	38	7,00E+03	100	14-32	225-207	+/-
gene:vir:104665	NC_000909_4_13	38	549	mge:1630	P-SSM2	virus	252401	Prochlorococcus marinus	38	7,00E+03	95	5-27	361-339	+/-
gene:vir:5607	NC_000909_5_20	38	531	mge:119	KVP40	virus	244834	Vibrio parahaemolyticus	38	7,00E+03	100	11-29	20-38	+/+
gene:plasmid:15583	NC_000909_8_11	38	1506	mge:355	pTEF3	plasmid	17963	Enterococcus faecalis V583	38	7,00E+03	100	14-32	692-674	+/-
gene:plasmid:127516	NC_000916_16_43	38	2469	mge:1827	pVir	plasmid	37473	Campylobacter jejuni subsp. jejuni 81-176	38	7,00E+03	100	5-23	84-66	+/-
gene:plasmid:128727	NC_000916_16_43	38	26241	mge:832	3	plasmid	194553	Lawsonia intracellularis PHE/MN1-00	38	7,00E+03	100	6-24	14324-14306	+/-
gene:plasmid:14194	NC_000916_16_43	38	2469	mge:285	pVir	plasmid	37468	Campylobacter jejuni; Campylobacter jejuni subsp. jejuni 81-176	38	7,00E+03	100	5-23	84-66	+/-
gene:proph:174413	NC_000916_16_43	38	1344	mge:2457	prophinder:44334	prophage	134822	Clostridium kluyveri DSM 555	38	7,00E+03	92	4-30	339-313	+/-
gene:proph:174405	NC_000916_16_43	38	1344	mge:2457	prophinder:44334	prophage	134822	Clostridium kluyveri DSM 555	38	7,00E+03	92	4-30	339-313	+/-

gene:proph:174278	NC_000916_16_43	38	1344	mge:2456	prophinder:44330	prophage	57911	Clostridium kluyveri DSM 555	38	7,00E+03	92	4-30	339-313	+/-
gene:plasmid:153473	NC_000916_7_66	38	1269	mge:887	2	plasmid	338007	Polaromonas sp. JS666	38	7,00E+03	100	6-24	544-526	+/-
gene:plasmid:112666	NC_000917_1_35	39	2373	mge:769	pNG600	plasmid	155300	Haloarcula marismortui ATCC 43049	38	7,00E+03	100	17-35	2104-2122	+/+
gene:vir:99683	NC_000917_3_10	39	2283	mge:1523	VP4	virus	39503	Vibrio	38	7,00E+03	100	13-31	1243-1225	+/-
gene:plasmid:28190	NC_000917_3_41|NZ_CP006577_6_89	38	1173	mge:608	pSymB	plasmid	1683333	Sinorhizobium meliloti; Sinorhizobium meliloti 1021	38	7,00E+03	100	4-22	815-797	+/-
gene:proph:165173	NC_000909_14_8	41	321	mge:2202	prophinder:44538	prophage	24183	Clostridium botulinum A str. ATCC 19397	38	8,00E+03	100	8-26	2-20	+/+
gene:vir:104561	NC_000909_20_6	43	975	mge:1548	P-SSM4	virus	178249	Prochlorococcus marinus	38	8,00E+03	100	18-36	455-437	+/-
gene:plasmid:120468	NC_000854_4_10	47	1311	mge:779	pRL12	plasmid	870021	Rhizobium leguminosarum bv. viciae 3841	38	1,00E+04	100	17-35	1257-1239	+/-
gene:plasmid:90126	NC_000909_15_1	50	891	mge:1969	pSMED01	plasmid	1570951	Sinorhizobium medicae WSM419	38	1,00E+04	100	23-41	413-395	+/-
gene:vir:104755	NC_000909_5_21	47	1458	mge:1630	P-SSM2	virus	252401	Prochlorococcus marinus	38	1,00E+04	100	1-19	1085-1067	+/-
gene:proph:183382	NC_000909_5_22	47	1047	mge:2708	prophinder:44919	prophage	41629	Haemophilus somnus 129PT	38	1,00E+04	100	8-26	566-548	+/-
gene:plasmid:135534	NC_000913_4_12|NC_007779_4_12|NC_(	32	1194	mge:1838	pSbal01	plasmid	116763	Shewanella baltica OS155	36	2,10E+04	100	1-18	944-961	+/+
gene:plasmid:15057	NC_000916_7_100	32	1107	mge:339	pQpH1	plasmid	37393	Coxiella burnetii RSA 493	36	2,10E+04	100	12-29	679-662	+/-
gene:plasmid:14988	NC_000916_7_100	32	1155	mge:337	QpH1	plasmid	37329	Coxiella burnetii	36	2,10E+04	100	12-29	727-710	+/-
gene:plasmid:22103	NC_000913_4_8|NC_007779_4_8|NC_01(	33	627	mge:570	pGMI1000MP	plasmid	2094509	Ralstonia solanacearum; Ralstonia solanacearum GMI1000	36	2,20E+04	100	4-21	563-580	+/+
gene:proph:165226	NC_000909_11_2	35	432	mge:2204	prophinder:44541	prophage	27608	Clostridium botulinum A str. ATCC 19397	36	2,30E+04	100	17-34	166-183	+/+
gene:plasmid:120815	NC_000916_16_44	34	969	mge:894	pSpnP1	plasmid	5413	Streptococcus pneumoniae	36	2,30E+04	100	13-30	863-846	+/-
gene:plasmid:27295	NC_000917_2_21	35	996	mge:608	pSymB	plasmid	1683333	Sinorhizobium meliloti; Sinorhizobium meliloti 1021	36	2,30E+04	100	17-34	918-935	+/+
gene:proph:159767	NC_000917_3_33	35	654	mge:2041	prophinder:42821	prophage	24628	Bacteroides vulgatus ATCC 8482	36	2,30E+04	100	17-34	413-396	+/-
gene:plasmid:11064	NC_002570_1_11	34	1053	mge:208	AT	plasmid	542869	Agrobacterium tumefaciens; Agrobacterium tumefaciens str. C58	36	2,30E+04	100	2-19	819-802	+/-
gene:plasmid:11809	NC_002570_1_11	34	1053	mge:210	AT	plasmid	542780	Agrobacterium tumefaciens str. C58	36	2,30E+04	100	2-19	819-802	+/-
gene:proph:169553	NC_002570_2_4	35	1296	mge:2328	prophinder:43279	prophage	37329	Streptococcus pyogenes MGAS9429	36	2,30E+04	92	5-30	425-400	+/-
gene:proph:159558	NC_002570_2_4	35	1290	mge:2033	prophinder:46657	prophage	40271	Streptococcus pyogenes M1 GAS	36	2,30E+04	92	5-30	419-394	+/-
gene:plasmid:89068	NC_002570_3_3	34	657	mge:1969	pSMED01	plasmid	1570951	Sinorhizobium medicae WSM419	36	2,30E+04	100	15-32	431-448	+/+
gene:plasmid:122987	NC_000853_1_15|NC_023151_4_26|NZ_(	36	357	mge:1020	pE33L466	plasmid	466370	Bacillus cereus E33L	36	2,40E+04	100	3-20	117-100	+/-
gene:plasmid:125230	NC_000853_4_14|NC_023151_1_11|NZ_(	36	2355	mge:849	unnamed plasmid	plasmid	130973	Silicibacter sp. TM1040	36	2,40E+04	100	5-22	1273-1290	+/+
gene:plasmid:149991	NC_000916_7_95	36	1401	mge:972	pWCFS103	plasmid	36069	Lactobacillus plantarum WCFS1	36	2,40E+04	100	12-29	110-93	+/-
gene:vir:95022	NC_000917_1_39	36	1284	mge:1547	KS7	virus	40794	Salmonella enterica subsp. enterica serovar Typhimurium	36	2,40E+04	100	16-33	250-233	+/-
gene:plasmid:11470	NC_000917_1_5	36	633	mge:208	AT	plasmid	542869	Agrobacterium tumefaciens; Agrobacterium tumefaciens str. C58	36	2,40E+04	100	19-36	504-521	+/+
gene:plasmid:12213	NC_000917_1_5	36	585	mge:210	AT	plasmid	542780	Agrobacterium tumefaciens str. C58	36	2,40E+04	100	19-36	456-473	+/+
gene:vir:96122	NC_000918_8_1	36	3006	mge:1602	37	virus	43681	Staphylococcus aureus	36	2,40E+04	100	8-25	555-538	+/-
gene:plasmid:16913	NC_000961_4_26	36	1077	mge:406	pHel4	plasmid	10970	Helicobacter pylori	36	2,40E+04	100	19-36	489-472	+/-
gene:plasmid:16894	NC_000961_4_26	36	915	mge:404	pHPM8	plasmid	7817	Helicobacter pylori	36	2,40E+04	100	19-36	327-310	+/-
gene:plasmid:129983	NC_000961_4_26	36	1077	mge:766	pAL202	plasmid	12120	Helicobacter pylori	36	2,40E+04	100	19-36	489-472	+/-
gene:plasmid:142545	NC_000961_6_13	36	4353	mge:1812	TC2	plasmid	300725	Arthrobacter aurescens TC1	36	2,40E+04	100	10-27	2636-2619	+/-
gene:plasmid:118383	NC_000962_6_10|NC_002755_6_11|NC_(	36	2847	mge:1042	1	plasmid	343931	Mesorhizobium sp. BNC1	36	2,40E+04	100	16-33	537-520	+/-
gene:plasmid:123401	NC_000853_1_19|NC_023151_4_22|NZ_(	37	1050	mge:909	1	plasmid	27048	Lawsonia intracellularis PHE/MN1-00	36	2,60E+04	96	7-32	1046-1022	+/-
gene:vir:78907	NC_000853_1_19|NC_023151_4_22|NZ_(	37	1557	mge:1859	A006	virus	38124	Listeria monocytogenes	36	2,60E+04	100	20-37	897-880	+/-
gene:proph:164727	NC_000853_1_19|NC_023151_4_22|NZ_(	37	1557	mge:2190	prophinder:46338	prophage	27776	Listeria innocua Clip11262	36	2,60E+04	100	20-37	897-880	+/-
gene:vir:7070	NC_000853_1_3|NC_023151_4_39|NZ_CI	37	516	mge:142	T4	virus	168903	Escherichia coli	36	2,60E+04	100	17-34	369-352	+/-
gene:plasmid:91603	NC_000853_4_18|NC_023151_1_7|NZ_CI	37	447	mge:1979	pSN254	plasmid	176473	Salmonella enterica subsp. enterica serovar Newport str. SL254	36	2,60E+04	100	10-27	353-336	+/-
gene:plasmid:91105	NC_000853_4_18|NC_023151_1_7|NZ_CI	37	447	mge:1975	pYR1	plasmid	158038	Yersinia ruckeri	36	2,60E+04	100	10-27	353-336	+/-
gene:plasmid:83827	NC_000853_4_18|NC_023151_1_7|NZ_CI	37	447	mge:1908	pIP1202	plasmid	182913	Yersinia pestis biovar Orientalis str. IP275	36	2,60E+04	100	10-27	353-336	+/-
gene:plasmid:109175	NC_000853_4_18|NC_023151_1_7|NZ_CI	37	447	mge:1803	pP91278	plasmid	131520	Photobacterium damselae subsp. piscicida	36	2,60E+04	100	10-27	353-336	+/-
gene:plasmid:136223	NC_000853_4_18|NC_023151_1_7|NZ_CI	37	447	mge:1802	pP99-018	plasmid	150157	Photobacterium damselae subsp. piscicida	36	2,60E+04	100	10-27	353-336	+/-
gene:plasmid:140697	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	930	mge:1048	p42e	plasmid	505334	Rhizobium etli CFN 42	36	2,60E+04	100	18-35	735-752	+/+
gene:plasmid:146235	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	2517	mge:731	1	plasmid	574127	Deinococcus geothermalis DSM 11300	36	2,60E+04	100	12-29	619-636	+/+
gene:vir:5136	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	6558	mge:114	PBC5	virus	57416	Sinorhizobium meliloti	36	2,60E+04	100	17-34	1432-1449	+/+
gene:proph:167924	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	1977	mge:2278	prophinder:42494	prophage	28915	Yersinia pseudotuberculosis IP 32953	36	2,60E+04	100	17-34	655-672	+/+
gene:proph:158960	NC_000853_8_3|NC_023151_5_11|NZ_CI	37	438	mge:2008	prophinder:46672	plasmid	38747	Rhodobacter sphaeroides ATCC 17025	36	2,60E+04	100	20-37	249-266	+/+
gene:plasmid:154105	NC_000868_1_19	37	3153	mge:1030	R478	plasmid	274762	Serratia marcescens	36	2,60E+04	100	18-35	1373-1356	+/-
gene:plasmid:13952	NC_000868_4_9	37	348	mge:269	cp32-8	plasmid	30885	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:plasmid:13988	NC_000868_4_9	37	348	mge:270	cp32-9	plasmid	30651	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:plasmid:13866	NC_000868_4_9	37	348	mge:267	cp32-6	plasmid	29838	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:plasmid:13908	NC_000868_4_9	37	357	mge:268	cp32-7	plasmid	30800	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:plasmid:13821	NC_000868_4_9	37	348	mge:266	cp32-4	plasmid	30299	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:plasmid:13735	NC_000868_4_9	37	348	mge:264	cp32-1	plasmid	30750	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:plasmid:13777	NC_000868_4_9	37	348	mge:265	cp32-3	plasmid	30223	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:plasmid:14045	NC_000868_4_9	37	348	mge:272	lp56	plasmid	52971	Borrelia burgdorferi B31	36	2,60E+04	95	16-37	93-72	+/-
gene:vir:100438	NC_000868_4_9	37	1068	mge:1488	25	virus	161475	Aeromonas salmonicida	36	2,60E+04	100	3-20	802-785	+/-
gene:plasmid:14558	NC_000909_1_6	37	1899	mge:293	pCP13	plasmid	54310	Clostridium perfringens str. 13	36	2,60E+04	100	19-36	1696-1713	+/+
gene:vir:100428	NC_000916_16_18	37	1523	mge:1488	25	virus	161475	Aeromonas salmonicida	36	2,60E+04	100	7-24	848-831	+/-
gene:proph:172759	NC_000916_16_21|NC_000916_16_26	37	1359	mge:2416	prophinder:43793	prophage	41339	Listeria monocytogenes EGD-e	36	2,60E+04	100	16-33	1219-1202	+/-

gene:plasmid:127527	NC_000916_7_41	37	432	mge:1827	pVir	plasmid	37473	Campylobacter jejuni subsp. jejuni 81-176	36	2,60E+04	100	15-32	209-226	+/+
gene:plasmid:14181	NC_000916_7_41	37	432	mge:285	pVir	plasmid	37468	Campylobacter jejuni; Campylobacter jejuni subsp. jejuni 81-176	36	2,60E+04	100	15-32	209-226	+/+
gene:vir:4635	NC_000916_7_41	37	999	mge:101	PVL	virus	41401	Staphylococcus aureus	36	2,60E+04	100	8-25	6-23	+/+
gene:proph:180833	NC_000916_7_41	37	1446	mge:2639	prophinder:45721	prophage	32988	Clostridium botulinum A str. Hall	36	2,60E+04	100	10-27	1394-1411	+/+
gene:proph:165140	NC_000916_7_41	37	1446	mge:2201	prophinder:44537	prophage	32988	Clostridium botulinum A str. ATCC 19397	36	2,60E+04	100	10-27	1394-1411	+/+
gene:plasmid:128771	NC_000916_7_47	37	2526	mge:832	3	plasmid	194553	Lawsonia intracellularis PHE/MN1-00	36	2,60E+04	100	8-25	367-384	+/+
gene:plasmid:12513	NC_000916_7_73	37	3963	mge:215	pX01	plasmid	181654	Bacillus anthracis	36	2,60E+04	100	1-18	2854-2871	+/+
gene:plasmid:156135	NC_000916_7_73	37	1041	mge:807	pLVPK	plasmid	219385	Klebsiella pneumoniae	36	2,60E+04	100	5-22	1022-1039	+/+
gene:plasmid:154355	NC_000916_7_73	37	3963	mge:729	pXO1	plasmid	181677	Bacillus anthracis str. 'Ames Ancestor'	36	2,60E+04	100	1-18	2854-2871	+/+
gene:plasmid:12745	NC_000916_7_73	37	3963	mge:217	pXO1	plasmid	181677	Bacillus anthracis str. A2012	36	2,60E+04	100	1-18	2854-2871	+/+
gene:plasmid:16829	NC_000917_1_14	37	330	mge:398	pNRC100	plasmid	191346	Halobacterium sp. NRC-1	36	2,60E+04	100	3-20	85-68	+/-
gene:plasmid:84048	NC_000917_3_15	37	1074	mge:1913	plasmid_153kb	plasmid	153140	Yersinia pseudotuberculosis IP 31758	36	2,60E+04	100	12-29	433-416	+/-
gene:vir:106995	NC_000917_3_15	37	888	mge:1459	S-PM2	virus	196280	Synechococcus sp.; Synechococcus sp. WH 7803	36	2,60E+04	100	15-32	241-224	+/-
gene:plasmid:133012	NC_000961_4_50	37	1158	mge:715	2	plasmid	115507	Arthrobacter sp. FB24	36	2,60E+04	95	2-23	772-751	+/-
gene:plasmid:88651	NC_000868_4_23	38	942	mge:1964	pCM1	plasmid	27357	Clavibacter michiganensis subsp. michiganensis NCPPB 382	36	2,70E+04	95	7-28	4-25	+/+
gene:plasmid:150024	NC_000909_1_11	38	618	mge:972	pWCFS103	plasmid	36069	Lactobacillus plantarum WCFS1	36	2,70E+04	95	11-32	185-164	+/-
gene:proph:174436	NC_000909_11_3	38	1521	mge:2457	prophinder:44334	prophage	134822	Clostridium kluyveri DSM 555	36	2,70E+04	100	17-34	1251-1268	+/+
gene:proph:174341	NC_000909_11_3	38	1521	mge:2457	prophinder:44334	prophage	134822	Clostridium kluyveri DSM 555	36	2,70E+04	100	17-34	1251-1268	+/+
gene:vir:108045	NC_000909_4_13	38	198	mge:2002	JS98	virus	170523	Escherichia coli	36	2,70E+04	95	1-22	92-71	+/-
gene:plasmid:114859	NC_000909_5_4	38	723	mge:719	A	plasmid	366354	Anabaena variabilis ATCC 29413	36	2,70E+04	100	16-33	392-409	+/+
gene:plasmid:150533	NC_000916_16_36	38	1269	mge:1800	pPRO1	plasmid	202397	Pelobacter propionicus DSM 2379	36	2,70E+04	95	12-33	697-676	+/-
gene:plasmid:85251	NC_000916_16_43	38	1686	mge:1934	pFN3	plasmid	11934	Fusobacterium nucleatum subsp. polymorphum ATCC 10953	36	2,70E+04	100	5-22	1591-1574	+/-
gene:plasmid:146695	NC_000916_16_43	38	1128	mge:905	lp44	plasmid	44010	Borrelia duttonii	36	2,70E+04	100	6-23	363-346	+/-
gene:plasmid:23414	NC_000916_7_66	38	1305	mge:574	pNGR234a	plasmid	536165	Rhizobium sp. NGR234	36	2,70E+04	100	4-21	582-599	+/+
gene:plasmid:120836	NC_000916_7_70	38	273	mge:920	p250	plasmid	6286	Avibacterium paragallinarum	36	2,70E+04	100	13-30	27-10	+/-
gene:plasmid:110784	NC_000917_1_31	38	129	mge:1790	lp60	plasmid	59958	Borrelia afzelii PKo	36	2,70E+04	100	2-19	68-51	+/-
gene:proph:174160	NC_000917_1_31	38	153	mge:2454	prophinder:44327	prophage	16913	Clostridium kluyveri DSM 555	36	2,70E+04	95	12-33	93-114	+/+
gene:vir:6600	NC_000917_1_8	38	795	mge:139	RB49	virus	164018	Escherichia coli	36	2,70E+04	100	10-27	1-18	+/+
gene:vir:80998	NC_000917_1_8	38	795	mge:1888	Phi1	virus	164270	Escherichia coli K12	36	2,70E+04	100	10-27	1-18	+/+
gene:plasmid:87906	NC_000917_2_23	38	2706	mge:1955	pBVIE01	plasmid	397868	Burkholderia vietnamiensis G4	36	2,70E+04	100	7-24	470-453	+/-
gene:plasmid:137714	NC_000917_2_23	38	2469	mge:1808	pNOCA01	plasmid	307814	Nocardioides sp. JS614	36	2,70E+04	100	4-21	278-261	+/-
gene:plasmid:127700	NC_000917_2_45	38	147	mge:935	pTT27	plasmid	232605	Thermus thermophilus HB27	36	2,70E+04	100	9-26	111-94	+/-
gene:plasmid:157510	NC_000962_6_16|NC_002945_5_13|NC_	38	5187	mge:1143	pRL9	plasmid	352782	Rhizobium leguminosarum bv. viciae 3841	36	2,70E+04	95	10-31	2049-2070	+/+
gene:plasmid:27611	NC_000854_3_30	39	339	mge:608	pSymB	plasmid	1683333	Sinorhizobium meliloti; Sinorhizobium meliloti 1021	36	2,80E+04	95	17-38	300-321	+/+
gene:plasmid:129609	NC_000854_3_30	39	960	mge:572	p42d	plasmid	371254	Rhizobium etli CFN 42	36	2,80E+04	95	17-38	299-320	+/+
gene:plasmid:18841	NC_000868_4_25	39	588	mge:497	pCC7120alpha	plasmid	408101	Nostoc sp. PCC 7120	36	2,80E+04	100	1-18	300-317	+/+
gene:vir:105317	NC_000909_12_3	39	1383	mge:1679	PH15	virus	44041	Staphylococcus epidermidis	36	2,80E+04	100	5-22	505-488	+/-
gene:vir:107104	NC_000909_12_3	39	1383	mge:1571	CNPH82	virus	43420	Staphylococcus epidermidis	36	2,80E+04	100	5-22	505-488	+/-
gene:vir:80530	NC_000909_19_2	39	504	mge:1883	A511	virus	134494	Listeria monocytogenes	36	2,80E+04	100	1-18	291-308	+/+
gene:plasmid:129818	NC_000909_6_2	39	1791	mge:957	pSciA	plasmid	7790	Spiroplasma citri	36	2,80E+04	100	18-35	851-868	+/+
gene:plasmid:154756	NC_000916_7_48	39	2793	mge:1804	1	plasmid	653815	Paracoccus denitrificans PD1222	36	2,80E+04	100	13-30	2396-2379	+/-
gene:plasmid:122231	NC_000909_4_10	40	1044	mge:1069	pBc10987	plasmid	208369	Bacillus cereus ATCC 10987	36	2,90E+04	100	16-33	934-917	+/-
gene:plasmid:118367	NC_000854_3_9	41	963	mge:1042	1	plasmid	343931	Mesorhizobium sp. BNC1	36	3,00E+04	100	14-31	699-682	+/-
gene:vir:6329	NC_000854_3_9	41	4014	mge:132	phiKMV	virus	42519	Pseudomonas aeruginosa	36	3,00E+04	100	7-24	838-821	+/-
gene:vir:99451	NC_000854_3_9	41	1128	mge:1595	BJ1	virus	42271	Halorubrum	36	3,00E+04	100	8-25	268-285	+/+
gene:proph:174395	NC_000909_6_9	41	444	mge:2457	prophinder:44334	prophage	134822	Clostridium kluyveri DSM 555	36	3,00E+04	100	10-27	139-122	+/-
gene:proph:183832	NC_000909_8_1	41	4491	mge:2725	prophinder:43046	prophage	33203	Bacillus cereus subsp. cytotoxis NVH 391-98	36	3,00E+04	100	12-29	357-340	+/-
gene:plasmid:112301	NC_000917_2_24	41	1866	mge:893	pMOL28	plasmid	171461	Ralstonia metallidurans CH34	36	3,00E+04	100	1-18	853-870	+/+
gene:plasmid:117630	NC_000917_2_24	41	1740	mge:811	2	plasmid	171459	Ralstonia metallidurans CH34	36	3,00E+04	100	1-18	727-744	+/+
gene:plasmid:146521	NC_000917_2_25	41	915	mge:731	1	plasmid	574127	Deinococcus geothermalis DSM 11300	36	3,00E+04	100	15-32	331-314	+/-
gene:vir:106043	NC_000962_6_12|NC_008769_4_13|NC_	41	6072	mge:1505	Cooper	virus	70654	Mycobacterium; Mycobacterium smegmatis	36	3,00E+04	95	13-34	4913-4892	+/-
gene:plasmid:118437	NC_000916_7_51	42	2982	mge:1042	1	plasmid	343931	Mesorhizobium sp. BNC1	36	3,20E+04	100	18-35	2441-2424	+/-
gene:vir:102772	NC_000909_14_1	43	1689	mge:1610	Y540	virus	152372	Thermus thermophilus	36	3,30E+04	100	24-41	583-600	+/+
gene:plasmid:124744	NC_000909_14_2	43	648	mge:941	pCD6	plasmid	6830	Clostridium difficile	36	3,30E+04	100	14-31	118-101	+/-
gene:plasmid:112759	NC_000909_20_3	43	615	mge:1155	pSSP2	plasmid	22870	Staphylococcus saprophyticus subsp. saprophyticus ATCC 15305	36	3,30E+04	100	19-36	542-559	+/+
gene:plasmid:86411	NC_000854_3_26	46	666	mge:1938	pACRY01	plasmid	203589	Acidiphilium cryptum JF-5	36	3,60E+04	100	8-25	619-602	+/-
gene:plasmid:147771	NC_000854_3_26	46	924	mge:1034	pRL10	plasmid	488135	Rhizobium leguminosarum bv. viciae 3841	36	3,60E+04	100	10-27	731-748	+/+
gene:plasmid:83559	NC_000854_4_10	47	1233	mge:1906	pRSPA02	plasmid	289489	Rhodobacter sphaeroides ATCC 17025	36	3,80E+04	100	12-29	1007-1024	+/+
gene:proph:173555	NC_000909_5_22	47	744	mge:2437	prophinder:45598	prophage	15546	Clostridium botulinum F str. Langeland	36	3,80E+04	95	3-24	649-628	+/-
gene:plasmid:16898	NC_002163_2_1|NC_018521_2_1|NZ_CP	30	519	mge:405	pHel5	plasmid	18291	Helicobacter pylori	34	7,20E+04	100	1-17	447-431	+/-
gene:plasmid:13264	NC_002163_2_2|NC_017279_2_1|NC_01	30	1290	mge:251	cp18-2	plasmid	21170	Borrelia burgdorferi	34	7,20E+04	100	14-30	720-704	+/-
gene:proph:176181	NC_000913_4_12|NC_007779_4_12|NC_	32	2220	mge:2505	prophinder:44575	prophage	26166	Xylella fastidiosa 9a5c	34	8,20E+04	100	16-32	317-301	+/-
gene:proph:176132	NC_000913_4_12|NC_007779_4_12|NC_	32	2220	mge:2503	prophinder:44572	prophage	25518	Xylella fastidiosa 9a5c	34	8,20E+04	100	16-32	317-301	+/-

gene:plasmid:135129	NC_000913_4_5|NC_007779_4_5|NC_01	32	5688	mge:1094	TM1040 mega plasmid	plasmid	821788	Silicibacter sp. TM1040	34	8,20E+04	95	11-31	3299-3319	+/+
gene:proph:168700	NC_000913_4_8|NC_007779_4_8|NC_01	33	1863	mge:2301	prophinder:47364	prophage	23904	Bordetella petrii DSM 12804	34	8,70E+04	95	2-22	1025-1005	+/-
gene:proph:176828	NC_000913_5_5|NC_007779_5_5|NC_01	33	1227	mge:2520	prophinder:43595	prophage	31258	Escherichia coli APEC O1	34	8,70E+04	89	1-29	1144-1172	+/+
gene:plasmid:23396	NC_000853_7_2|NC_023151_6_2|NZ_CPI	35	489	mge:574	pNGR234a	plasmid	536165	Rhizobium sp. NGR234	34	9,10E+04	95	9-29	352-372	+/+
gene:plasmid:97857	NC_000853_8_1|NC_023151_5_13|NZ_CI	35	297	mge:1160	pKEF9	plasmid	28930	Sulfolobus islandicus	34	9,10E+04	100	1-17	48-64	+/+
gene:plasmid:126869	NC_000853_8_1|NC_023151_5_13|NZ_CI	35	297	mge:1113	pHVE14	plasmid	35422	Sulfolobus islandicus	34	9,10E+04	100	1-17	48-64	+/+
gene:plasmid:150115	NC_000853_8_1|NC_023151_5_13|NZ_CI	35	297	mge:1053	pARN3	plasmid	26200	Sulfolobus islandicus	34	9,10E+04	100	1-17	48-64	+/+
gene:plasmid:112248	NC_000853_8_2|NC_023151_5_12|NZ_CI	35	1554	mge:893	pMOL28	plasmid	171461	Ralstonia metallidurans CH34	34	9,10E+04	100	2-18	879-863	+/-
gene:plasmid:117576	NC_000853_8_2|NC_023151_5_12|NZ_CI	35	1554	mge:811	2	plasmid	171459	Ralstonia metallidurans CH34	34	9,10E+04	100	2-18	879-863	+/-
gene:plasmid:84092	NC_000909_13_8	35	426	mge:1913	plasmid_153kb	plasmid	153140	Yersinia pseudotuberculosis IP 31758	34	9,10E+04	100	11-27	166-150	+/-
gene:plasmid:121276	NC_000909_13_8	35	1188	mge:1799	pALH1	plasmid	55939	Bacillus thuringiensis str. Al Hakam	34	9,10E+04	100	12-28	1098-1082	+/-
gene:plasmid:23243	NC_000916_16_16	35	879	mge:574	pNGR234a	plasmid	536165	Rhizobium sp. NGR234	34	9,10E+04	100	18-34	620-604	+/-
gene:plasmid:147800	NC_000916_16_19	35	1146	mge:1034	pRL10	plasmid	488135	Rhizobium leguminosarum bv. viciae 3841	34	9,10E+04	100	19-35	164-180	+/+
gene:plasmid:118785	NC_000916_16_19	35	456	mge:801	1	plasmid	257447	Rhodoferax ferrireducens T118	34	9,10E+04	100	15-31	291-275	+/-
gene:plasmid:114793	NC_000916_16_28	35	1647	mge:719	A	plasmid	366354	Anabaena variabilis ATCC 29413	34	9,10E+04	100	17-33	1487-1471	+/-
gene:plasmid:147410	NC_000916_7_108	34	636	mge:1796	1	plasmid	278942	Shewanella sp. ANA-3	34	9,10E+04	100	3-19	323-307	+/-
gene:plasmid:23605	NC_000917_2_21	35	1029	mge:574	pNGR234a	plasmid	536165	Rhizobium sp. NGR234	34	9,10E+04	95	7-27	169-189	+/+
gene:vir:78160	NC_000962_5_6|NC_002755_5_6|NC_00	35	1305	mge:1847	Min1	virus	46365	Microbacterium nematophilum	34	9,10E+04	100	12-28	129-145	+/+
gene:vir:78146	NC_000962_5_6|NC_002755_5_6|NC_00	35	2328	mge:1847	Min1	virus	46365	Microbacterium nematophilum	34	9,10E+04	100	12-28	1840-1824	+/-
gene:plasmid:20278	NC_002570_1_7	34	741	mge:556	Rts1	plasmid	217182	Proteus vulgaris	34	9,10E+04	100	8-24	556-540	+/-
gene:proph:184907	NC_002570_2_2	35	747	mge:2757	prophinder:45964	prophage	26663	Enterobacter sakazakii ATCC BAA-894	34	9,10E+04	95	2-22	692-712	+/+
gene:proph:180495	NC_002570_2_4	35	1296	mge:2632	prophinder:43253	prophage	45020	Streptococcus pyogenes MGAS8232	34	9,10E+04	89	2-30	428-400	+/-
gene:proph:184177	NC_000853_4_17|NC_023151_1_8|NZ_CI	36	738	mge:2735	prophinder:43895	prophage	50942	Gramella forsetii KT0803	34	9,60E+04	100	10-26	489-505	+/+
gene:plasmid:85650	NC_000853_5_1	36	1422	mge:1936	pRSPA01	plasmid	877879	Rhodobacter sphaeroides ATCC 17025	34	9,60E+04	100	14-30	75-59	+/-
gene:vir:104812	NC_000853_6_4|NC_023151_7_5|NZ_CPI	36	1413	mge:1630	P-SSM2	virus	252401	Prochlorococcus marinus	34	9,60E+04	100	1-17	74-90	+/+
gene:plasmid:119868	NC_000853_6_6|NC_010483_6_13|NC_0	36	3477	mge:779	pRL12	plasmid	870021	Rhizobium leguminosarum bv. viciae 3841	34	9,60E+04	95	16-36	2944-2964	+/+
gene:plasmid:127358	NC_000853_8_13|NC_023151_5_1|NZ_CI	36	2856	mge:763	pRL8	plasmid	147463	Rhizobium leguminosarum bv. viciae 3841	34	9,60E+04	100	20-36	2705-2721	+/+
gene:plasmid:131626	NC_000853_8_7|NC_023151_5_7|NZ_CPI	36	2214	mge:1936	pRSPA01	plasmid	877879	Rhodobacter sphaeroides ATCC 17025	34	9,60E+04	100	9-25	18-34	+/+
gene:plasmid:135998	NC_000853_8_8|NC_023151_5_6|NZ_CPI	36	768	mge:1120	pCRY	plasmid	21742	Yersinia pestis biovar Microtus str. 91001	34	9,60E+04	100	9-25	370-386	+/+
gene:plasmid:133876	NC_000868_4_11	36	1701	mge:762	pRHL2	plasmid	442536	Rhodococcus sp. RHA1	34	9,60E+04	100	10-26	138-122	+/-
gene:vir:199	NC_000909_11_11	36	189	mge:6	HK97	virus	39732	Escherichia coli	34	9,60E+04	95	4-24	44-64	+/+
gene:plasmid:13494	NC_000909_17_2	36	243	mge:258	lp28-4	plasmid	27323	Borrelia burgdorferi B31	34	9,60E+04	100	6-22	219-203	+/-
gene:plasmid:110020	NC_000909_19_3	36	1893	mge:908	pNOB8	plasmid	41229	Sulfolobus sp. NOB8H2	34	9,60E+04	100	20-36	800-816	+/+
gene:vir:100062	NC_000909_19_3	36	798	mge:1604	P-SSP7	virus	44970	Prochlorococcus marinus	34	9,60E+04	100	20-36	544-528	+/-
gene:plasmid:84113	NC_000909_4_4	36	1161	mge:1913	plasmid_153kb	plasmid	153140	Yersinia pseudotuberculosis IP 31758	34	9,60E+04	95	2-22	1008-988	+/-
gene:vir:5587	NC_000909_4_4	36	627	mge:119	KVP40	virus	244834	Vibrio parahaemolyticus	34	9,60E+04	100	19-35	394-378	+/-
gene:vir:102363	NC_000909_4_4	36	687	mge:1566	phi CD119	virus	53325	Clostridium difficile	34	9,60E+04	95	7-27	415-395	+/-
gene:plasmid:13388	NC_000909_5_2	36	1161	mge:256	lp28-2	plasmid	29766	Borrelia burgdorferi B31	34	9,60E+04	100	17-33	531-547	+/+
gene:proph:159125	NC_000909_8_13	36	648	mge:2018	prophinder:44556	prophage	19192	Carboxydothermus hydrogenoformans Z-2901	34	9,60E+04	100	1-17	370-386	+/+
gene:plasmid:14560	NC_000909_8_4	36	645	mge:293	pCP13	plasmid	54310	Clostridium perfringens str. 13	34	9,60E+04	100	18-34	308-324	+/+
gene:vir:9287	NC_000916_16_12	36	780	mge:165	phi 11	virus	43604	Staphylococcus aureus subsp. aureus NCTC 8325	34	9,60E+04	100	11-27	371-355	+/-
gene:vir:1212	NC_000916_16_12	36	780	mge:25	phi ETA	virus	43081	Staphylococcus aureus	34	9,60E+04	100	11-27	371-355	+/-
gene:vir:80017	NC_000916_16_12	36	780	mge:1875	tp310-3	virus	41966	Staphylococcus aureus	34	9,60E+04	100	11-27	371-355	+/-
gene:vir:97291	NC_000916_16_12	36	780	mge:1666	52A	virus	41690	Staphylococcus aureus	34	9,60E+04	100	11-27	371-355	+/-
gene:vir:94081	NC_000916_16_12	36	780	mge:1494	96	virus	43576	Staphylococcus aureus	34	9,60E+04	100	11-27	371-355	+/-
gene:proph:160281	NC_000916_16_12	36	780	mge:2053	prophinder:44149	prophage	34027	Staphylococcus aureus subsp. aureus NCTC 8325	34	9,60E+04	100	11-27	371-355	+/-
gene:vir:99497	NC_000916_16_14	36	657	mge:1559	Lj928	virus	38384	Lactobacillus johnsonii	34	9,60E+04	100	1-17	330-346	+/+
gene:vir:105458	NC_000916_16_14	36	657	mge:1502	KC5a	virus	38239	Lactobacillus gasseri	34	9,60E+04	100	1-17	330-346	+/+
gene:proph:167092	NC_000916_16_9	36	747	mge:2261	prophinder:47495	prophage	30918	Acholeplasma laidlawii PG-8A	34	9,60E+04	100	7-23	202-186	+/-
gene:plasmid:112662	NC_000916_7_102	36	1338	mge:769	pNG600	plasmid	155300	Haloarcula marismortui ATCC 43049	34	9,60E+04	100	4-20	302-318	+/+
gene:plasmid:18817	NC_000916_7_33	36	1956	mge:497	pCC7120alpha	plasmid	408101	Nostoc sp. PCC 7120	34	9,60E+04	100	20-36	1435-1419	+/-
gene:plasmid:148903	NC_000916_7_34	36	486	mge:778	pAYWB-I	plasmid	3972	Aster yellows witches'-broom phytoplasma AYWB	34	9,60E+04	100	4-20	298-314	+/+
gene:plasmid:86374	NC_000916_7_65	36	225	mge:1937	pCLI	plasmid	17531	Clostridium botulinum F str. Langeland	34	9,60E+04	100	10-26	73-89	+/+
gene:plasmid:137015	NC_000916_7_68	36	321	mge:1830	pYVe8081	plasmid	67721	Yersinia enterocolitica subsp. enterocolitica 8081	34	9,60E+04	100	15-31	28-12	+/-
gene:vir:7277	NC_000916_7_68	36	738	mge:142	T4	virus	168903	Escherichia coli	34	9,60E+04	100	15-31	224-240	+/+
gene:vir:7276	NC_000916_7_68	36	1878	mge:142	T4	virus	168903	Escherichia coli	34	9,60E+04	100	15-31	905-921	+/+
gene:vir:103281	NC_000916_7_74	36	285	mge:1605	JK06	virus	46072	Escherichia coli O157:H7	34	9,60E+04	100	6-22	164-180	+/+
gene:vir:103282	NC_000916_7_74	36	315	mge:1605	JK06	virus	46072	Escherichia coli O157:H7	34	9,60E+04	100	6-22	180-196	+/+
gene:plasmid:124105	NC_000916_7_84	36	663	mge:741	pRL11	plasmid	684202	Rhizobium leguminosarum bv. viciae 3841	34	9,60E+04	100	19-35	155-171	+/+
gene:plasmid:16914	NC_000917_2_19	36	1983	mge:406	pHel4	plasmid	10970	Helicobacter pylori	34	9,60E+04	100	16-32	1948-1932	+/-
gene:plasmid:129984	NC_000917_2_19	36	1983	mge:766	pAL202	plasmid	12120	Helicobacter pylori	34	9,60E+04	100	16-32	1948-1932	+/-
gene:plasmid:110008	NC_000917_3_5	36	333	mge:908	pNOB8	plasmid	41229	Sulfolobus sp. NOB8H2	34	9,60E+04	100	10-26	266-250	+/-
gene:vir:9081	NC_000918_4_1	36	1560	mge:163	phiKZ	virus	280334	Pseudomonas aeruginosa	34	9,60E+04	100	7-23	664-648	+/-

gene:plasmid:124543	NC_000961_4_28	36	2958	mge:741	pRL11	plasmid	684202	Rhizobium leguminosarum bv. viciae 3841	34	9,60E+04	100	3-19	345-361	+/+
gene:vir:93832	NC_000961_4_59	36	486	mge:1479	712	virus	30510	Lactococcus lactis	34	9,60E+04	95	4-24	398-418	+/+
gene:plasmid:122044	NC_000961_6_9	36	1113	mge:1069	pBc10987	plasmid	208369	Bacillus cereus ATCC 10987	34	9,60E+04	100	12-28	151-167	+/+
gene:vir:104499	NC_000961_6_9	36	657	mge:1548	P-SSM4	virus	178249	Prochlorococcus marinus	34	9,60E+04	95	9-29	59-79	+/+
gene:plasmid:145258	NC_000962_6_4|NC_002755_6_5|NC_00.	36	1476	mge:1006	megaplasmid	plasmid	634917	Ralstonia eutropha JMP134	34	9,60E+04	100	5-21	242-226	+/-
gene:plasmid:82947	NC_000962_6_5|NC_002945_5_7|NC_00	36	1308	mge:1903	5	plasmid	155098	Aeromonas salmonicida subsp. salmonicida A449	34	9,60E+04	100	11-27	938-954	+/+
gene:proph:174791	NC_002570_1_10	36	330	mge:2467	prophinder:45789	prophage	26282	Bacillus pumilus SAFR-032	34	9,60E+04	100	15-31	22-38	+/+