

FÁBIA GIOVANA DO VAL DE ASSIS

**GÊNOMICA COMPARATIVA DE *CLUSTERS* PUTATIVOS DE
BIOSSÍNTESE DE PEPTÍDEOS LAÇO EM PROCARIOTOS DO
RÚMEN**

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

VIÇOSA
MINAS GERAIS – BRASIL
2017

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

T

A848g
2017

Assis, Fábía Giovana do Val de, 1981-

Genômica comparativa de *clusters* putativos de biossíntese de peptídeos laço em procariotos do rúmen / Fábía Giovana do Val de Assis. – Viçosa, MG, 2017.

x, 99f. : il. (algumas color.) ; 29 cm.

Orientador: Hilário Cuquetto Mantovani.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Peptídeos. 2. Rúmen. 3. Microbiota. 4. Bacteriocinas.
I. Universidade Federal de Viçosa. Departamento de Microbiologia. Programa de Pós-graduação em Microbiologia Agrícola. II. Título.

CDD 22 ed. 572.65

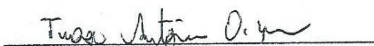
FÁBIA GIOVANA DO VAL DE ASSIS

**GÊNOMICA COMPARATIVA DE *CLUSTERS* PUTATIVOS DE
BIOSSÍNTESE DE PEPTÍDEOS LAÇO EM PROCARIOTOS DO
RÚMEN**

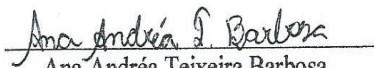
Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Microbiologia Agrícola, para obtenção do título de *Doctor Scientiae*.


APROVADA: 23 de fevereiro de 2017.


Mateus Ferreira Santana
(Coorientador)


Tiago Antônio de Oliveira Mendes


Maria Cristina Baracat Pereira


Ana Andréa Teixeira Barbosa


Hilário Cuquetto Mantovani
(Orientador)

À minha família e amigos por todo amor e apoio depositado,

OFEREÇO

A Deus, por ter me concedido esta conquista!

Aos meus familiares e amigos, pelo carinho e amizade!

Dedico

AGRADECIMENTOS

Agradeço a Deus pelo amor e misericórdia a mim concedidos para realização deste trabalho.

À Universidade Federal de Viçosa (UFV), em especial ao Programa de Pós-graduação em Microbiologia Agrícola, pela oportunidade da realização do curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão das bolsas de estudo.

Ao meu orientador, Professor Hilário Cuquetto Mantovani, pela orientação, competência, paciência e confiança depositadas em meu trabalho.

Aos meus coorientadores, Professora Denise Maria Soares Bazzoli e Professor Mateus Ferreira Santana, pelo exemplo de profissionalismo, ensinamentos e sugestões.

Aos Professores Maria Cristina Baracat Pereira e Tiago Antônio de Oliveira Mendes, pelos conhecimentos transmitidos e atenções a mim concedidas.

Aos demais professores do Departamento de Microbiologia, pela constante participação, apoio e ensinamentos fundamentais para minha formação acadêmica e profissional.

Aos membros da banca examinadora, pelas críticas e sugestões ao trabalho a fim de melhorá-lo.

Aos técnicos e funcionários do Departamento de Microbiologia, por toda ajuda concedida, amizade e carinho.

Aos amigos do Laboratório de Anaeróbios e Alimentos, em especial Analice, Bianca, Cláudia, Déborah, Elsa, Felipe, Juliana, Isabela, Marcílio, Sofia, pelos ensinamentos e exemplo de dedicação ao trabalho.

Aos meus queridos amigos de Viçosa Érica Lelis, Fernando Agostinho, Leandro Lopes pelos momentos de descontração e cumplicidade, e em especial Patrícia Leal, pelos ensinamentos e incentivo, e demais amigos, vocês fazem parte dessa conquista.

À minha pequena e doce sobrinha Larissa (*in memoriam*), pelo breve convívio, porém de grande aprendizado.

À minha mãe e irmão, que sempre foram o alicerce de cada etapa da minha vida, e com quem sempre irei compartilhar cada conquista.

Enfim, a todos que contribuíram de forma direta ou indiretamente para a realização deste trabalho.

BIOGRAFIA

Fábia Giovana do Val de Assis, filha de Tais das Graças Val de Assis, nasceu em 25 de abril de 1981, em Governador Valadares – Minas Gerais.

Iniciou o curso de Zootecnia na Universidade Federal de Viçosa em março de 2001, pela qual graduou-se em Bacharelado em Zootecnia em Maio de 2006.

Em 2011 ingressou no curso de Mestrado em Microbiologia Agrícola, na Universidade Federal de Lavras, defendendo a dissertação em fevereiro de 2013.

Em agosto de 2013 ingressou no curso de Doutorado em Microbiologia Agrícola, na Universidade Federal de Viçosa, defendendo a tese em fevereiro de 2017.

Sumário

Lista de tabelas.....	vi
Resumo.....	vii
Abstract.....	ix
Introdução geral.....	1
Capítulo 1 - Revisão De Literatura	
1.1 Características, estruturas e classificação dos peptídeos laço.....	3
1.2 Biossíntese e função de peptídeos laço.....	6
1.3 Prospecção de peptídeos laço por mineração genômica.....	7
1.4 Peptídeos antimicrobianos produzidos por bactérias ruminais.....	10
1.5 Referências Bibliográficas.....	14
Capítulo 2 - Mineração de <i>clusters</i> putativos de biossíntese de peptídeo laço em genomas da microbiota ruminal	
Resumo.....	26
Abstract.....	27
2.1 Introdução.....	27
2.2 Material e Métodos.....	30
2.3 Resultados.....	32
2.4 Discussão.....	66
2.5 Referências Bibliográficas.....	74

Lista de tabelas

Tabela 1. Características dos principais peptídeos laço reportados na literatura.....	3
Tabela 2: Genomas de bactérias com <i>clusters</i> putativos de biossíntese de peptídeos laço inferidos pelos softwares BAGEL3, antiSMASH, Banco de dados 1 e 2.....	34
Tabela 3: Domínios conservados das proteínas codificadas pelos genes putativos de biossíntese de peptídeos laço encontrados em <i>clusters</i> completos nos genomas de bactérias do rúmen.....	41
Tabela 4: Características dos peptídeos <i>core</i> inferidos em <i>cluster</i> putativos completos de biossíntese de peptídeos laço nos genomas de bactérias do rúmen.....	42
Tabela 5: Códigos atribuídos aos genomas de bactérias ruminais com <i>clusters</i> putativos completos de biossíntese de peptídeos laço utilizados na reconstrução da filogenia dos genes <i>lasABCD</i>	51
Tabela 6: Informações sobre os dados brutos do arquivo de metatranscriptoma SRR3169851 (2Gb).....	58
Tabela 7: Qualidade do arquivo SRR3169851 (2Gb) após eliminar <i>reads</i> de baixa qualidade, usando valor de <i>Phred quality</i> maior que 30%.....	58
Table S1: Rumen bacterial genome analyzed in this study (data collected from April to December 2016).....	80
Tabela S2: Banco de dados 1 (composto por sequências nucleotídicas de 4 genes)..	97
Tabela S3: Banco de dados 2 (composto por sequências de aminoácidos codificados por 20 genes relacionados a peptídeos laço).....	97

Resumo

ASSIS, Fábila Giovana do Val de Assis, D.Sc., Universidade Federal de Viçosa, fevereiro de 2017. **Gênômica comparativa de *clusters* putativos de biossíntese de peptídeos laço em procariotos do rúmen.** Orientador: Hilário Cuquetto Mantovani. Coorientadores: Denise Mara Soares Bazzolli e Mateus Ferreira Santana.

Peptídeos laço são uma classe de produtos naturais sintetizados ribossomicamente e modificados pós-tradução que apresentam um conjunto diversificado de atividades relevantes, incluindo inibição do crescimento bacteriano, antagonista de receptor, e inibição de enzimas. Estratégias de mineração genômica têm contribuído para a descoberta de novos peptídeos laço e também de novos microrganismos produtores de peptídeos desse grupo. No entanto, pouco se sabe sobre a produção de peptídeos laço pela microbiota do ecossistema ruminal. Neste trabalho, análises *in silico* foram realizadas para investigar a distribuição de *clusters* gênicos relacionados com a biossíntese de peptídeos laço (*lasA*, *lasB*, *lasC* e *lasD*) em genomas de 320 bactérias e 5 archaeas ruminais e de 25 bactérias isoladas de fezes de ruminantes. Para 42 genomas de bactérias ruminais foram inferidos *clusters* putativos incompletos e completos para peptídeos laço, utilizando as ferramentas BAGEL3, AntiSMASH, Banco de dados 1 e Banco de dados 2, sendo o Filo Firmicutes o de maior representatividade. De um total de 26 genomas de bactérias do rúmen que tiveram inferidos *clusters* gênicos completos de biossíntese de peptídeos laço (em sua maioria pertencentes à classe II), 53% foi representado por linhagens do gênero *Butyrivibrio*. A análise da organização dos 26 *clusters* putativos completos de biossíntese de peptídeos laço inferidos via mineração genômica revelou a presença de vários genes com funções moleculares e biológicas diversas, dentre esses, predominaram (38% dos *clusters*) os genes putativos para histidina cinase (HisKA). Os domínios conservados dos produtos preditos para os genes de biossíntese de peptídeos laço indicaram os genes *lasB* e *lasC* como os mais conservados e úteis na mineração genômica desses peptídeos. Árvores filogenéticas representativas dos genes *lasA*, *lasB*, *lasC* e *lasD* revelaram que a presença dos genes de biossíntese de peptídeos laço retratavam um processo evolutivo. A reconstrução filogenética com base no gene que codifica o rRNA 16S para 42 genomas de bactérias do rúmen que apresentaram *clusters* putativos

completos e incompletos para biossíntese de peptídeos laço, permitiu o agrupamento das linhagens em seus respectivos gêneros ou Família (Lachnospiraceae e Coriobacteriaceae). A partir das análises de metatranscriptoma, pôde-se inferir que os genes de biossíntese para peptídeos laço podem ser expressos com base no arquivo de metatranscriptoma SRR3169851 e reforçaram a evidência de que os genes putativos de biossíntese *lasB* e *lasC* são conservados entre linhagens da mesma espécie de bactérias do rúmen. Concluiu-se que a análise *in silico* forneceu evidências de novos *clusters* de genes biossintéticos em espécies bacterianas não previamente relacionadas com peptídeos laço, sugerindo que a microbiota ruminal representa uma fonte potencial de novos peptídeos laço.

Abstract

ASSIS, Fábيا Giovana do Val de Assis, D.Sc., Universidade Federal de Viçosa, February, 2017. **Comparative genome of putative clusters of biosynthesis of lasso peptide in rumen prokaryotes.** Advisor: Hilário Cuquetto Mantovani. Co-advisors: Denise Mara Soares Bazzolli and Mateus Ferreira Santana.

Lasso peptides are a class of naturally occurring ribosomally synthesized and post-translationally modified products that present a diverse set of relevant activities, including inhibition of bacterial growth, receptor antagonist, and inhibition of enzymes. Genomic mining strategies have contributed to the discovery of new lasso peptides and to new producer microorganisms in this group. However, little is known about the lasso peptide production by the ruminal ecosystem microorganisms. In this work, we performed genome mining of the complete and partial genome sequences of 320 ruminal bacteria and 5 ruminal archaea and 25 bacteria isolated from ruminant feces to determine the distribution and diversity of lasso peptide gene clusters. For 42 genomes of ruminal bacteria, incomplete and complete putative clusters for lasso peptides were obtained using BAGEL3, AntiSMASH tools, Database 1 and Database 2, with the Firmicutes being the most representative. Of a total of 26 genomes of ruminal bacteria that had inferred complete gene clusters of lasso peptide biosynthesis (mostly belonging to class II), 53% were represented by strains of the genus *Butyrivibrio*. The analysis of the organization of the 26 complete putative clusters of lasso peptide biosynthesis inferred via genomic mining revealed the presence of several genes with diverse molecular and biological functions, among which the putative genes for histidine kinase (HisKA) predominated (38% of the clusters). The conserved domains of the products predicted for lasso peptide biosynthesis genes indicated *lasB* and *lasC* genes as the most conserved and useful in the genomic mining of these peptides. Phylogenetic trees representative of the genes *lasA*, *lasB*, *lasC* and *lasD* revealed that the presence of lasso peptide biosynthesis genes portrayed an evolutionary process. Phylogenetic reconstruction based on the gene encoding the 16S rRNA for 42 genomes of ruminal bacteria that presented complete and incomplete putative clusters for lasso peptide biosynthesis allowed the grouping of the strains in their respective genera or Family (Lachnospiraceae and Coriobacteriaceae). From the metatranscriptome analyzes, it could be inferred

that the biosynthetic genes for lasso peptides can be expressed based on the metatranscriptome file SRR3169851 and reinforced the evidence that the putative biosynthetic genes *lasB* and *lasC* are conserved between lineages of the same rumen bacteria species. It was concluded that *in silico* analysis provided evidence of new clusters of biosynthetic genes in bacterial species not previously related to lasso peptides, suggesting that the ruminal microbiota represents a potential source of novel loop peptides.

INTRODUÇÃO GERAL

Peptídeos laço são moléculas protéicas antimicrobianas sintetizadas pelos ribossomos por uma ampla diversidade de bactérias, caracterizadas por sofrerem modificação pós-tradução, pela ciclização do peptídeo *core*. Esses peptídeos pertencem à classe I de bacteriocinas e apresentam uma topologia única. Tipicamente, a estrutura contém um anel macrolactâmico constituído por sete a nove resíduos de aminoácidos, formado por uma ligação amida entre o grupo α -amino N-terminal de resíduos de glicina/cisteína e o grupo carboxila de resíduos de aspartato/glutamato. A cauda C-terminal do peptídeo é presa ao anel, formando assim a estrutura de laço característica desses peptídeos.

A classificação dos peptídeos laço é determinada pelo número de ligações dissulfeto que conferem estabilidade estérica à topologia da molécula. A classe I é representada por peptídeos laço que apresentam duas ligações dissulfeto (RP 71955/aborycin; Siamycin I/MS-271/NP-06; Siamycin II; SSV-2083), enquanto os peptídeos da Classe II são desprovidos de ligações dissulfeto (Anantín; Capistruín; Lariatín A; Lariatín B; Microcin J25; Propeptín; RES-701-1; SRO15-2005). Com a descoberta do peptídeo laço BI-32169, contendo somente uma ligação dissulfeto, foi proposta a criação de uma nova classe, denominada Classe III.

A biossíntese de um peptídeo laço requer, em geral, três genes que codificam um peptídeo precursor A, uma cisteína protease B (homóloga à transglutaminase) e uma lactama sintetase ATP-dependente C (homóloga à asparagina sintetase). A exportação do peptídeo laço está relacionada ao produto do gene D que codifica um transportador ABC para bombear o peptídeo para fora da célula. Alguns *clusters* gênicos podem conter genes adicionais, mas esses genes geralmente não são necessários para o processo de maturação dos peptídeos laço.

O primeiro registro sobre peptídeos laço ocorreu em 1991, com a descoberta do anantín. Desde então, pelo menos, 40 peptídeos laço já foram identificados por meio de bioensaios e abordagens de mineração genômica, sendo a Microcin J25 (Mccj25), isolada de *Escherichia coli* AY25, considerada arquétipo nos estudos de biossíntese, maturação e exportação de peptídeos laço.

Além das atividades antimicrobianas contra bactérias, outras funções biológicas também têm sido atribuídas aos peptídeos laço, incluindo ação antiviral (incluindo

contra o vírus HIV), inibição de enzimas e antagonistas de receptores clinicamente relevantes.

Devido às diferentes atividades biológicas e o potencial industrial apresentado pelos peptídeos laço, vários trabalhos de identificação e caracterização desses peptídeos tem sido realizados. Entretanto, até agora nenhum trabalho explorou a distribuição ou diversidade genética dos peptídeos laço no ecossistema ruminal. O rúmen tem sido reconhecido por ser colonizado por uma comunidade microbiana altamente diversa e abundante em interações ecológicas que influenciam a colonização dos nichos ecológicos disponíveis. Estudos anteriores inferiram, por meio de análises *in silico*, que várias espécies de procariotos anaeróbios são produtores de peptídeos antimicrobianos, porém a análise de peptídeos laço em genomas de bactérias e archaeas do rúmen não foi investigada. Considerando que esses organismos são produtores de peptídeos antimicrobianos com potencial para serem aplicados na conservação de alimentos, manipulação da fermentação ruminal e na terapia clínica, há grande interesse em pesquisas acerca da diversidade e distribuição dos peptídeos antimicrobianos produzidos por bactérias ruminais. Com o aumento crescente de sequências genômicas de bactérias ruminais depositadas e disponíveis em bancos de dados públicos, nós hipotetizamos neste trabalho que a abordagem *in silico*, como a mineração de genomas, permitirá inferir a distribuição de *clusters* gênicos relacionados com a biossíntese de peptídeos laço em bactérias e archaeas do rúmen. Nós investigamos uma coleção de genomas completos e parciais de 320 bactérias e 5 archaeas ruminais e de 25 bactérias isoladas de fezes de ruminantes. para identificar *clusters* gênicos relacionados com a biossíntese de peptídeos laço.

CAPÍTULO 1

REVISÃO DE LITERATURA

1.1 Características, estruturas e classificação dos peptídeos laço

As bacteriocinas são peptídeos e proteínas antimicrobianas, sintetizados ribossomicamente por bactérias e algumas espécies do domínio Archaea (Nes et al., 2007; Riley 2011; Arnison et al., 2013). As bacteriocinas são, geralmente, ativas contra bactérias filogeneticamente relacionadas, entretanto, alguns peptídeos apresentam amplo espectro de ação (Cotter et al, 2005; Zacharof & Lovitt, 2012). Esses peptídeos antimicrobianos apresentam grande diversidade em termos de estrutura química e mecanismos de ação e têm sido classificados em quatro grupos principais de bacteriocinas (Riley, 2011): A Classe I se refere aos peptídeos antimicrobianos modificados após a tradução, tais como os lantibióticos, labirintopeptinas e sactibióticos que diferem entre si quanto à estrutura molecular, o mecanismo de ação e as enzimas envolvidas na modificação dos peptídeos precursores (Riley, 2011; Arnison et al., 2013); A Classe II é composta por peptídeos antimicrobianos não modificados pós traducionalmente e pode ser subdividida nas subclasses IIa (peptídeos semelhantes a pediocina), IIb (bacteriocinas de dois peptídeos), IIc (bacteriocinas circulares) e IId (bacteriocinas lineares não semelhantes a pediocina) (Rea et al., 2011); A Classe III é representada por proteínas termolábeis com mais de 10 kDa, também referidas como bacteriolisinas (Rea et al., 2011). Estas proteínas apresentam homologia com as endopeptidases e o domínio C-terminal parece estar envolvido na especificidade da proteína com as células-alvo (Nilsen et al., 2003); A classe IV foi atribuída para incluir as bacteriocinas cíclicas modificadas pós-traducionalmente que foram previamente agrupadas no subtipo IIc (Heng & Tagg, 2006).

Peptídeos laço são uma classe de moléculas antimicrobianas de síntese ribossomal que sofrem modificação pós-traducional, incluindo ciclização do peptídeo *core*, e apresentam diferentes tipos de atividade biológica (Maksimov et al., 2012a). Esses peptídeos pertencem à classe I de bacteriocinas, e apresentam, em geral, 15 a 24 resíduos de aminoácidos e massa molecular variando de 1500 a 2500 Da. (Arnison et al., 2013). Contudo, a característica peculiar aos peptídeos laço é sua topologia, com

presença de um anel macrolactâmico composto por sete a nove resíduos de aminoácidos, formado por uma ligação amida entre o grupo α -amino da região N-terminal de resíduos de glicina, cisteína, serina ou alanina e o grupo carboxila de resíduos de aspartato ou glutamato. A extremidade C-terminal gera o *loop* e a cauda do peptídeo laço, os quais são presos ao anel por interações covalentes, assistidas muitas vezes por ligações dissulfeto (McIntosh et al., 2009; Velasquez e Van der Donk, 2011; Alvarez-Sieiro et al., 2016) (Figura 1).

O resultado dessa topologia é uma estrutura compacta e estável, que confere aos peptídeos laço uma elevada resistência a proteases e a agentes desnaturantes, além de possuírem diferentes atividades biológicas (antimicrobiana, antagonista de receptor, inibição de enzimas) (Maksimov et al., 2012a; Hegemann et al., 2015). Outro aspecto interessante da estrutura dos peptídeos laço, segundo Maksimov et al. (2012a), é que, embora os peptídeos laço apresentem resíduos conservados, o tamanho da molécula, a composição dos resíduos de aminoácidos e a topologia dos peptídeos podem variar de forma considerável, e conseqüentemente, influenciar a bioatividade dos peptídeos laço (Tabela 1). Ducasse et al. (2012) realizaram mutações sítio-específicas no peptídeo laço Microcin J25 (MccJ25), a fim de identificar as sequências determinantes da topologia laço, e constataram que o tipo de resíduos abaixo do anel (Tyr²⁰, Gly²¹) e o tamanho do anel macrolactâmico foram críticos para a bioatividade da molécula, enquanto que o encurtamento ou aumento do número de resíduos de aminoácidos abaixo do anel não alterou a estrutura do laço, mas afetou diferencialmente a atividade antibacteriana.

A classificação dos peptídeos laço é determinada pelo número de ligações dissulfeto que conferem estabilidade estérica à topologia da molécula (Hegemann et al., 2015). A classe I é representada por peptídeos laço que apresentam duas ligações dissulfeto (RP 71955/aborycin; Siamycin I/MS-271/NP-06; Siamycin II; SSV-2083), enquanto os peptídeos da Classe II são desprovidos de ligações dissulfeto (Anantín; Capistruin; Lariatín A; Lariatín B; Microcin J25; Propeptin; RES-701-1; SRO15-2005). Com a descoberta do peptídeo laço BI-32169 (Potterat et al., 2004), contendo somente uma ligação dissulfeto, foi proposta a criação de uma nova classe, denominada Classe III (Knappe et al., 2010).

Peptídeos laço da classe I têm como principal característica a presença de duas ligações dissulfeto entre resíduos de cisteína, as quais contribuem para manutenção da topologia, uma vez que são responsáveis por manterem o *loop* e cauda presos ao anel. Os peptídeos da classe II, devido à falta de ligações dissulfeto, apresentam como

característica determinante para manutenção da topologia, a presença de *plugs* de aminoácidos, os quais são responsáveis por um bloqueio estérico, garantindo, portanto, que o *loop* e a cauda fiquem presos ao anel (Maksimov et al., 2012a). Para o peptídeo laço Mccj25, já foi verificado que os resíduos de fenilalanina acima do anel e o resíduo de tirosina abaixo do anel são responsáveis por preservar a topologia laço (Bayro et al., 2003; Rosengren et al., 2003; Wilson et al., 2003). Integrantes da classe III apresentam uma ligação dissulfeto que permite a ligação entre a calda e o anel, e ainda *plugs* de aminoácidos acima do anel com a finalidade de prender o *loop* (Figura 1).

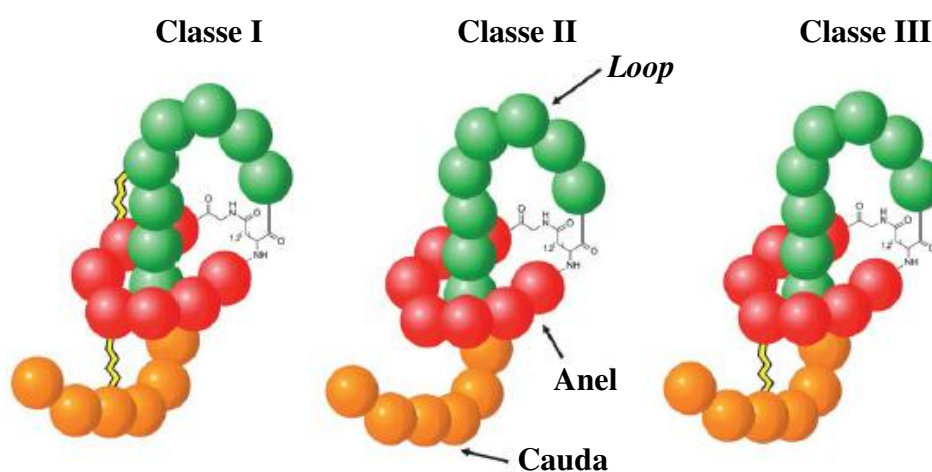


Figura 1: Esquema da estrutura e topologia das três diferentes classes de peptídeos laço. Ligação amida entre o grupo α -amino da região N-terminal de resíduos de glicina, cisteína, serina ou alanina e o grupo carboxila de resíduos de aspartato ou glutamato; a extremidade C-terminal origina o *loop* e a cauda do peptídeo. O anel, *loop*, cauda e ligações dissulfeto são representados pelas cores vermelho, verde, laranja e amarelo, respectivamente. Fonte: Maksimov et al. (2012a).

Tabela 1. Características dos principais peptídeos laço reportados na literatura

Peptídeos laço	Gene/Tamanho do precursor (aas)	Sequências de aa (peptídeo <i>core</i>) ¹	Nº de aa (anel)	P.I/Mw (Da) ²	Classes	Microrganismo produtor	Filos	Referência
Astexin I ³	<i>atxA1</i>	GLSQGVPEPDIGQTYFEESRINQD	23 (1-9)	3.83 / 25.82	II	<i>Asticcacaulis excentricus</i> CB 48	Proteobacteria	Zimmermann et al. (2013)
	<i>atxA2</i>	GLSQGVPEPDIGQTYFEESR	19 (1-9)	4.00 / 21,12				
BI-32169	N. A*	GLPWGCPDIPGWNTPWAC	19 (1-9)	3.80 / 2057	III	<i>Streptomyces</i> sp. DSM 14996	Actinobacteria	Potterat et al. (2004)
Capistruin	<i>capA/47</i>	GTPGFQTPDARVISRFGFN	19 (1-9)	9.60 / 2067	II	<i>Burkholderia thailandensis</i>	Proteobacteria	Knappe et al. (2008)
Caulonodin V ⁴	2 genes /40	SIGDSGLRESMSSQTYWP	18 (1-9)	4.37 / 2001	II	<i>Caulobacter</i> sp. K31	Proteobacteria	Zimmermann et al. (2014)
Caulosegnin 1	<i>csegA1/42</i>	GAFVQGPE AVNPLGREIQG	19 (1-8)	4,53/ 1939	II	<i>Caulobacter segnis</i>	Proteobacteria	Hegemann et al. (2013)
Lariatín B	<i>larA/44</i>	GSQLVYREWVGHNSVIKGGP	20 (1-8)	8.60/ 2223	II	<i>Rhodococcus jostii</i> K01-BD171	Proteobacteria	Inokoshi et al. (2012)
Lassomycin	<i>lasA/45</i>	GLRRLFADQLVGRRI	16 (1-8)	12.00/ 1885	II	<i>Lentzea kentuckyensis</i>	Actinobacteria	Gavriš et al. (2014)
Microcin J25	<i>mcjA/58</i>	GGAGHVPEYFVGIGTPISFYG	21 (1-8)	5.24/ 2125	II	<i>Escherichia coli</i>	Proteobacteria	Salomon e Farias (1992)
RES 701-1*	N. A*	GNWHGTXPDWFFNYXX	16 (1-9)	5.08 / 2027	II	<i>Streptomyces</i> sp. RE-701/RE-896	Actinobacteria	Katarita et al. (1995)
Siamycin*	N. A*	CLGVGSCNDFAGCGYAXVCFW	21 (1-9)	3.80/ 2183	I	<i>Streptomyces</i> sp.	Actinobacteria	Tsunakawa et al. (1995)
Streptomomicin	<i>StmA/42</i>	SLGSSPYNDILGYYPALIVIYP	22 (1-9)	3.80 / 2415	II	<i>Streptomonospora alba</i>	Actinobacteria	Metelev et al. (2015)
Sungsanpin*	N. A*	GFGSKPIDSFGLSWL	15 (1-8)	5.84 / 1610	II	<i>Streptomyces</i> sp.	Actinobacteria	Um et al. (2013)
Xanthomonin I	<i>xgaA1/46</i>	GGPLAGEEIGGFNVPG	16 (1-7)	3.79 / 1470	II	<i>Xanthomonas gardneri</i> ATCC 19865	Proteobacteria	Hegemann et al. (2013)
Xanthomonin II	<i>xgaA2/48</i>	GGPLAGEEMGGITT	14 (1-7)	3.79/ 1290	II	<i>Xanthomonas gardneri</i> ATCC 19865	Proteobacteria	Hegemann et al. (2013)

¹Sequência de resíduos de aminoácidos que compõem o peptídeo *core*, a qual é responsável pela formação do anel macrolactâmico, *loop* e cauda dos peptídeos laço; ² Ponto isoelétrico (p I) e massa de partículas atômicas (Da) do peptídeo *core*; ³*A. excentricus* contém dois diferentes *clusters* para peptídeo laço e com dois precursores cada; ⁴Caulonodin V apresenta dois genes que codificam dois peptídeos precursores porém não descritos na literatura. *N.A. – Não Aplicados

1.2 Biossíntese e função de peptídeos laço

O primeiro registro sobre peptídeos laço ocorreu em 1991, com a descoberta do anantín (Weber et al., 1991). Desde então, pelo menos, 40 peptídeos laço já foram identificados por meio de bioensaios e abordagens de mineração genômica, sendo a Microcín J25 (MccJ25), isolada de *Escherichia coli* AY25, considerada arquétipo nos estudos de biossíntese, maturação e exportação de peptídeos laço (Salomón & Farias 1992).

Segundo Maksimov et al. (2012a), quatro genes estão envolvidos nesses processos (denominados genericamente de *lasA*, *lasB*, *lasC*, *lasD*). O gene *mcjA* no cluster de MccJ25 codifica um precursor de 58 resíduos de aminoácidos (aas), que é modificado pós-tradução para a formação da estrutura laço (anel), a partir da ação de enzimas codificadas pelos genes *mcjB* e *mcjC*. A proteína precursora McjA contém uma sequência líder N-terminal de 37-aa que é removida durante o processamento do peptídeo laço. A atividade de peptidase requerida para esta transformação tem sido mapeada para o produto do gene *mcjB* (Pan et al., 2012; Yan et al., 2012). A enzima McjC apresenta homologia com enzimas formadoras de ligações amida dependentes de ATP (Wilson et al., 2003; Pan et al., 2012; Yan et al., 2012), e é responsável por catalisar a formação de ligações isopeptídicas para gerar a porção "anel" N-terminal da MccJ25. Estudos *in vitro* revelaram ainda que as enzimas McjB e McjC devem funcionar de forma conjunta para converter McjA em MccJ25, no processo de maturação do peptídeo (Duquesne et al., 2007; Yan et al., 2012).

O quarto gene do cluster, *mcjD*, codifica um transportador ABC que bombeia o MccJ25 para fora das células, e também serve como um fator de imunidade para as células que os produzem (Clarke & Campopiano, 2007; Duquesne et al., 2007). O transportador ABC é responsável pelo efluxo ativo do peptídeo para fora da célula garantindo que dentro da célula haja uma concentração mínima, a qual não afetaria a célula produtora (Delgado et al., 1999). No entanto, peptídeos laço de proteobactérias, incluindo astexin I, esfingopyxin I e caulosegnin I, são sintetizados a partir de *clusters* gênicos que apresentam uma isopeptidase específica de peptídeo laço ao invés de um transportador ABC (Hegemann et al., 2012; Hegemann et al., 2013; Zimmermann et al., 2013). Para o peptídeo paenonidín, foi observada uma organização gênica do tipo C A K B1 B2 D, em que as proteínas B1 e B2 partilham homologia com os domínios N- e C-terminais de uma proteína B "intacta", respectivamente, e o gene K codifica uma

cinase, possivelmente, relacionada com a clivagem dos peptídeos precursores e a fosforilação dos peptídeos laço (Zhu et al., 2016). Outros peptídeos laço como lariatin, svíceucin e streptomomicin também contêm os genes B1 e B2, e apresentam organização do tipo A C B1 B2 D (Iwatsuki et al., 2006; Inokoshi et al., 2012; Li et al., 2015; Metelev et al., 2015).

O aumento de estudos sobre peptídeos laço está relacionado, especialmente, às atividades antimicrobianas que estas moléculas apresentam. Tsunakawa et al. (1995) verificaram que Siamycin I e II foram igualmente ativos contra linhagens de *Bacillus subtilis*, *Staphylococcus aureus*, *Micrococcus luteus* e *Escherichia coli*. Kimura et al. (1997) demonstraram que propeptin possui atividade antimicrobiana contra *Pseudomonas aeruginosa*, *Mycobacterium phlei* e *Xanthomonas oryzae*, além de atividade inibitória da prolil endopeptidase, uma enzima do sistema renina angiotensina aldosterona. Knappe et al. (2008) constataram atividade antimicrobiana do peptídeo laço capistruin contra *Burkholderia caledonica*, *P. aeruginosa* e *E. coli*. Para esse mesmo peptídeo laço, foi verificada a ação como inibidor da RNA polimerase em *E. coli* (Kuznedelov et al., 2011). Gavrish et al. (2014) observaram ação antimicrobiana de lassomycin sobre *Mycobacterium tuberculosis*.

Para MccJ25, foi verificada atividade antimicrobiana contra várias espécies de bactérias Gram-negativas, incluindo *E. coli*, *Salmonella newport*, *Salmonella enteritidis* e *Shigella flexneri* (Salomón e Farías, 1992). Estes mesmos autores demonstraram que o uso de concentração mínima, na faixa de micromolar, de MccJ25 inibiu linhagens bacterianas sensíveis, e que estas apresentaram fenótipo filamentosos quando expostas ao peptídeo. Similar a outras bacteriocinas e vários bacteriófagos, a Mccj25 também interage com a proteína FhuA para atravessar a membrana externa das bactérias alvos. No periplasma, a MccJ25 utiliza as vias TonB e proteínas SbmA para atravessar a membrana interna (Salomón & Farías 1993, Salomón & Farías 1995). Esse peptídeo tem como principal alvo a RNA polimerase (RNAP), impedindo a atividade de transcrição do DNA (Delgado et al., 2001). Rintoul et al. (2001) revelaram que Mccj25 também atua alterando a permeabilidade da membrana citoplasmática de *Salmonella newport*, causando dissipação do gradiente iônico. Essas observações sugerem, portanto, que MccJ25 é capaz de comprometer o funcionamento da cadeia respiratória bacteriana, apresentando efeito bactericida.

Além da atividade antibacteriana, outras funções biológicas também têm sido relacionadas aos peptídeos laço, incluindo ação antiviral (incluindo contra o vírus HIV),

inibição de enzimas e antagonistas de receptores clinicamente relevantes. Helynck et al. (1993) demonstraram que o peptídeo laço RP 71955 inibiu a aspartil protease de HIV-1 e também a produção de transcriptase reversa do vírus HIV. Weber et al. (1991) verificaram que o peptídeo laço anantin, isolado de *Streptomyces coeruleus*, pode atuar como antagonista competitivo do fator natriurético atrial (ANF). Este último é um peptídeo circulante com propriedades natriuréticas, diuréticas e vasodilatadoras, descoberto em 1980 por A. J. de Bold, e considerado, na ocasião, como a molécula de ligação hormonal entre o coração e os rins. O peptídeo laço anantin foi o primeiro antagonista ANF de origem microbiana a ser descrito e caracterizado (Weber et al., 1991). Função de antagonista de receptor também foi relatada para o peptídeo laço RES-701-1, um inibidor competitivo eficaz da endotelina-1 que pode ser útil no tratamento de doenças cardiovasculares (Agapitov & Haynes, 2002).

O amplo espectro de atividade antimicrobiana que vem sendo reportado para os peptídeos laço indica que essas biomoléculas possuem aplicações biotecnológicas promissoras, justificando a intensificação nas investigações por novos microrganismos produtores e a pesquisa por novos peptídeos laço.

1.3 Prospecção de peptídeos laço por mineração genômica

Muitas informações sobre os tipos e características dos peptídeos laço têm sido obtidas por meio de estudos de mineração genômica, os quais têm possibilitado a descoberta de novas espécies microbianas produtoras, como representantes do Filo Firmicutes (Letzel et al., 2014; Zhu et al., 2016), e a identificação de novos peptídeos laço em genomas bacterianos, como o capistruin reportado em *Burkholderia thailandensis* (Knappe et al., 2008). Estima-se que mais da metade dos 40 peptídeos laço atualmente descritos foram descobertos a partir da mineração de genomas (Hegemann et al., 2015).

Estudos *in silico* levam em consideração a presença de genes conservados e que geralmente estão arranjados adjacentes nos *scaffolds* microbianos. Muito embora haja variação na ordem, complexidade e organização transcricional dos *clusters* de bacteriocinas, a presença de genes conservados têm sido demonstrada e supõe-se que os produtos da tradução destes genes apresentem a mesma função biológica (O'Sullivan et al., 2011; Wang et al., 2011). Dessa forma, alguns *softwares* específicos foram desenvolvidos baseados nos domínios conservados das proteínas envolvidas

principalmente na modificação e transporte de bacteriocinas. Algoritmos de acesso livre destinados especificamente à mineração de genes de bacteriocinas, como o BAGEL3 (van Heel et al., 2013) e antiSMASH, (Medema et al., 2011) já vêm sendo muito utilizados para analisar genomas de diferentes grupos microbianos. Além disso, o crescente aumento das sequências genômicas microbianas depositadas em bancos de dados de livre acesso vem permitindo a realização de estudos *in silico* para identificação de genes de bacteriocinas (Jong et al., 2010; O’Sullivan et al., 2011; van Heel et al., 2013; Sing e Saren, 2014; Azevedo et al., 2015).

No caso do peptídeo laço astexin-1, foi construído um algoritmo de mineração de genomas desenvolvido para identificar *clusters* gênicos de peptídeos laço em mais de três mil genomas sequenciados de procariotos (Maksimov et al., 2012b). Naquele estudo, pequenas ORFs, algumas não anotadas, contendo aminoácidos-padrão, consistentes com as sequências de peptídeos laço conhecidos, foram identificadas. Quando um precursor potencial de peptídeo laço era encontrado, regiões próximas do genoma eram pesquisadas para averiguar a homologia com as enzimas de maturação. Essas análises resultaram na primeira predição do peptídeo laço astexin-1 (detectado na bactéria *Asticcacaulis excentricus*, proveniente de água doce), além da identificação de 79 agrupamentos de genes putativos de diferentes peptídeos laço distribuídos nos filos bacterianos Proteobacteria, Actinobacteria, Chloroflexi, Firmicutes, Bacteroidetes, Espiroquetas, Cyanobacteria, Verrucomicrobia e Euryarchaeota (Archaea).

A combinação da mineração de genoma com a espectrometria de massas em tandem (MSⁿ) levou à descoberta de dois peptídeos laço (SSV-2083 e SRO15-200, ambos isolados de linhagens de *Streptomyces*) e seus agrupamentos gênicos (Kersten et al., 2011). A descoberta de SSV-2083 envolveu o tratamento do extrato contendo o peptídeo com NaBH₄/NiCl₂, a fim de aumentar a fragmentação da amostra e obter marcadores de sequências genéticas, que foram então submetidas ao MSⁿ. A partir das estruturas fornecidas pelo MSⁿ, foi possível identificar *clusters* gênicos contendo genes conservados e relacionados à biossíntese de peptídeos laço. O alinhamento com peptídeos laço de classe I conhecidos em combinação com dados gerados pelo MSⁿ permitiu a predição da estrutura de SSV-2083, representando o primeiro agrupamento de genes do peptídeo laço de classe I. Para detecção de SRO15-2005 e do *cluster* gênico correspondente, foi utilizada abordagem de mineração genômica guiada por MSⁿ, destacando a potencialidade desse método para descoberta de novos produtos peptídicos de origem microbiana.

A estratégia de mineração de genomas combinada com um sistema de expressão heteróloga otimizado também possibilitou identificar o conjunto de genes biossintéticos de peptídeos laço caulosegnins I, II e III, a partir de *Caulobacter segnis* (Hegemann et al., 2013). A abordagem de mineração de genomas também foi importante para o isolamento e caracterização de xanthomonins I–III, os primeiros peptídeos laço cujo anel macrolactâmico possui apenas sete aminoácidos (Hegemann et al., 2014), assim como para a descoberta e caracterização dos peptídeos laço caulonodin IV a VII (Zimmermann et al., 2014).

Azevedo et al. (2015) realizaram uma análise comparativa de sequências genômicas de 224 bactérias e 5 Archaeas do rúmen e identificaram 46 *clusters* biossintéticos de bacteriocinas em 33 genomas de bactérias ruminais. Vinte *clusters* foram relacionados à biossíntese de lantipeptídeos, enquanto 11 *clusters* foram associados com a produção sacpeptídeos e 7 pertenciam a bacteriocinas da classe II e 8 *clusters* gênicos foram classificados entre as bacteriocinas da classe III. Esses mesmos autores constataram ainda que a frequência de linhagens que codificaram precursores de peptídeos antimicrobianos putativos foi de 14,4% e creditaram às análises *in silico* a possibilidade de evidenciar novos grupos de genes biossintéticos de espécies de bactérias anteriormente não relacionados com a produção de bacteriocina.

1.4 Principais grupos de bactérias ruminais produtoras de peptídeos antimicrobianos

O ecossistema ruminal pode ser considerado um ambiente complexo e funcionalmente redundante, colonizado por diversos grupos de microrganismos que conferem aos animais ruminantes a capacidade de digerir material vegetal rico em fibras (Flint et al., 2008). Dentre os grupos microbianos do rúmen, as bactérias se destacam pela abundância e riqueza de espécies nesse ecossistema (Creevey et al., 2014).

Estudos de ecologia microbiana do rúmen sugerem que a capacidade de bactérias produzirem peptídeos antimicrobianos possui papel importante na competição entre espécies por nichos ecológicos do ecossistema ruminal (Kalmokoff et al., 1996; Teather & Forster, 1998; Mantovani et al., 2001). Dos peptídeos antimicrobianos produzidos pelas bactérias do rúmen, as bacteriocinas têm merecido destaque, por representarem uma alternativa aos antibióticos convencionais, os quais têm sido relacionados com a seleção de espécies bacterianas resistentes que podem ser prejudiciais à saúde pública

(Yang et al., 2014). Além disso, as bacteriocinas são reconhecidas por apresentarem baixa citotoxicidade e por apresentarem potencial de aplicação na preservação de alimentos e na produção animal (Russell & Mantovani, 2002; Paiva et al., 2013).

Em estudo anterior, Teather et al. (1999) indicaram que a frequência de microrganismos produtores de bacteriocinas, isolados de alimentos fermentados pode variar entre 1% e 10%. No entanto, entre isolados do rúmen a frequência de atividade inibitória encontrada entre algumas espécies bacterianas pode ser superior a 20%, sendo que para algumas espécies, estimou-se que cerca de 50% dos isolados apresentaram características bacteriocinogênicas (Mantovani et al., 2001; Whitford et al., 2001; Kalmokoff et al., 2003). Portanto, supõe-se que a produção de bacteriocinas seja uma característica vantajosa para as bactérias do rúmen.

Para alguns gêneros de bactérias ruminais, incluindo *Ruminococcus* (Odenyo et al., 1994; Chen et al., 2004), *Butyrivibrio* (Kalmokoff et al., 1999; Rychlik & Russell, 2002), *Enterococcus* (Laukova & Marekova, 2001; Nigutova et al., 2007) e *Streptococcus* (Mantovani et al., 2001; Whitford et al., 2001), a produção de bacteriocinas já foi demonstrada *in vitro* e os peptídeos foram caracterizados. Algumas espécies desse gênero foram isoladas de rúmen bovino e se destacam como produtoras de bacteriocinas incluindo as linhagens *Ruminococcus albus* 7 (Chen et al., 2004), *Ruminococcus albus* 8 (Chen et al., 2004) e *Ruminococcus flavefaciens* FD1 (Singh & Sareen, 2014). A produção de bacteriocinas por microrganismos ruminais pode desempenhar um papel importante na competição inter e intraespecífica, sendo que esses peptídeos podem ser úteis como agentes para a manipulação de populações microbianas do rúmen, favorecendo a colonização de algumas espécies (Teather et al., 1999). Segundo Odenyo et al. (1994), a produção de bacteriocinas por bactérias fibrolíticas ruminais representa um mecanismo importante na competição por nutrientes com outros microrganismos, desempenhando um papel relevante para a ecologia do ecossistema ruminal.

O gênero *Butyrivibrio* é constituído por bactérias Gram-positivas que apresentam morfologia de vibrio, são móveis, produzem ácido butírico como principal produto final da fermentação e podem ser encontradas no trato gastrointestinal de ruminantes (Forster et al., 1997; Balamurugan et al., 2009; Maia et al., 2010). Essas bactérias possuem diferentes funções na fermentação ruminal, atuando na degradação da fibra, na hidrólise de proteínas e na biohidrogenação de lipídios. Espécies de *Butyrivibrio* produtoras de compostos antimicrobianos também já foram identificadas e

caracterizadas (Wallace & Brammall, 1985; Kim et al, 2000; Fukuda et al, 2005; Pailard et al, 2007; Maia et al., 2010).

O gênero *Butyrivibrio* tem sido muito investigado em estudos de *screening* de atividade antimicrobiana no ecossistema ruminal (Kalmokoff & Theather, 1997; Rychlik & Russell, 2002; Cookson et al., 2004). No entanto, poucas bacteriocinas (butirivibriocina AR10 e butirivibriocina OR79A) produzidas por *Butyrivibrio* isolados do rúmen já foram caracterizadas e avaliadas quanto ao seu espectro de ação e potencial como agente de manipulação da fermentação ruminal ou controle de patógenos bacterianos (Kalmokoff & Teather, 1997 e Kalmokoff et al., 1999). As bacteriocinas butirivibriocina AR10 e butirivibriocina OR79A são sintetizadas, respectivamente, pelas linhagens *Butyrivibrio fibrisolvens* AR10 e *Butyrivibrio fibrisolvens* OR79A, ambas isoladas de rúmen de veado.

O gênero *Streptococcus* é representado por bactérias Gram-positivas que apresentam forma de diplococos ou estreptococos que produzem ácido láctico como principal produto de fermentação. O aumento da população de *Streptococcus* associado ao acúmulo de ácido láctico tem sido associado com acidose ruminal em bovinos alimentados com dietas ricas em carboidratos (Cotta & Whitehead, 1993; Ekinci et al., 1997). Em humanos, algumas espécies de *Streptococcus* também estão associadas com infecções do trato gastrointestinal (Gupta et al., 2010). Três bacteriocinas oriundas de *Streptococcus* isoladas do rúmen foram isoladas e caracterizadas: bovicina 255 (Whitford et al., 2001), bovicina HC5 (Mantovani et al., 2002) e bovicina HJ50 (Xiao et al., 2004).

A linhagem *Streptococcus equinus* HC5 produz uma bacteriocina denominada bovicina HC5 (Mantovani et al., 2002; Mantovani e Russell 2008). Esta bacteriocina apresenta espectro de ação contra diversas bactérias isoladas do leite, silagem, solo e rúmen, incluindo outras espécies de *Streptococcus*, *Clostridium*, *Listeria*, *Bacillus*, *Lactococcus*, *Alycyclobacillus* e *Staphylococcus* (Mantovani et al., 2002; de Carvalho et al., 2007, de Carvalho et al., 2008). A bovicina HC5 encontra-se principalmente aderida à célula produtora, e estudos anteriores indicaram que o peptídeo pode ser transferido diretamente para as bactérias-alvo (Xavier & Russell 2009).

Devido à capacidade de *Streptococcus equinus* HC5 inibir organismos comensais e patogênicos, conforme relatado por Mantovani et al. (2001) e Mantovani & Russell (2003), esta linhagem tem sido sugerida como veículo para entrega da bacteriocina no trato digestivo de animais de produção (Mantovani et al., 2002; Mantovani & Russell

2003; Xavier & Russell, 2009). Lima et al. (2009) demonstraram que a bovicina HC5 tem potencial para a otimização da fermentação ruminal, sendo capaz de inibir a atividade de desaminação de culturas mistas de bactérias ruminais em ensaios *in vitro*.

Alguns isolados do gênero *Enterococcus* também já foram identificados como produtores de bacteriocinas (Lauková & Czikková, 1998; Morovsky et al., 1998; Morovsky et al., 2001). A partir das linhagens *Enterococcus faecium* CCM isoladas de rúmen de bezerros e *Enterococcus faecium* BC25 isoladas de rúmen de bovinos, foram identificadas, respectivamente, as bacteriocinas Enterocina CCM 4231 (ação inibitória contra linhagens de *E. faecium*, *Streptococcus bovis*, *Escherichia coli* e *Listeria monocytogenes*) e Enterocina BC25 (bacteriocina do tipo IIa, homóloga à enterocin A, com ação inibitória contra as linhagens de *S. bovis*, *E. faecium*, *E. malodoratus* e *Listeria innocua*).

Apesar da produção de bacteriocinas pela microbiota ruminal ser bem documentada na literatura, os registros sobre a distribuição ou diversidade genética de peptídeos laço produzidos por bactérias ou archaeas do ecossistema ruminal são escassos. Os peptídeos laço pertencem à classe I de bacteriocinas, e são produzidos por bactérias dos filos Actinobacteria, Proteobacteria (Alpha e Beta-proteobacteria) e Firmicutes, sendo os gêneros *Streptomyces*, *Burkholderia*, *Rhodococcus*, *Escherichia*, *Microbispora*, *Asticcacaulis*, *Sphingopyxis*, *Xanthomonas*, *Bacillus*, *Butyrivibrio* e *Paenibacillus* os principais produtores de peptídeo laço (Maksimov et al., 2012b). Esses microrganismos têm sido encontrados em variados nichos ecológicos como solos e sedimentos marinhos, ambientes aquáticos naturais e ambientes contaminados (Li et al., 2014).

Recentemente, a presença de genes putativos de peptídeos laço em linhagens bacterianas do ecossistema ruminal foi identificada a partir de estudos baseados em mineração de genomas. Letzel et al. (2014) verificaram genes putativos de peptídeos laço em *Butyrivibrio proteoclasticum* B316 ao demonstrarem o potencial de bactérias anaeróbias produzirem peptídeos sintetizados ribosomicamente e modificados pós-tradução (RiPPs). Azevedo et al. (2015), ao avaliarem distribuição e diversidade genética de agrupamentos de genes de bacteriocinas em genomas microbianos do rúmen, reportaram a presença de *clusters* gênicos para peptídeos laço em três gêneros bacterianos (*Butyrivibrio*, *Lachnospiraceae* e *Ruminococcus*). Essas evidências, somadas à reconhecida capacidade da microbiota ruminal em produzir peptídeos antimicrobianos, reforçam a hipótese de que microrganismos do rúmen sejam potenciais

produtores de peptídeos laço. Assim, o objetivo do presente estudo foi investigar, por meio de abordagem *in silico* (mineração de genomas), uma coleção de 320 bactérias e 5 archaeas ruminais e de 25 bactérias isoladas de fezes de ruminantes, para identificar *clusters* gênicos relacionados com a biossíntese de peptídeos laço.

Referências Bibliográficas

Agapitov, A.A.; Haynes, W.G. Role of endothelin in cardiovascular disease. **Journal Renin-Angiotensin-Aldosterone System**, v.3, p.1–15, 2002.

Alvarez-Sieiro, P.; Montalbán-López, M.; Mu, D.; Oscar P. Bacteriocins of lactic acid bacteria: extending the family. **Applied Microbiology Biotechnology**, v.100, p. 2939–2951, 2016.

Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni T. S.; Bula, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, S.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, P. S.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, T. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. L.; Van Der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. **Natural Product Reports**, v.30, p. 108-160, 2013.

Azevedo, A. C.; Bento, C. B.; Ruiz, J. C.; Queiroz, M. V.; Mantovani, H. C. D. Distribution and genetic diversity of bacteriocin gene *clusters* in rumen microbial genomes. **Applied and Environmental Microbiology**, v. 81, p. 7290-7304, 2015.

Balamurugan, R.; Chittaranjan, S.P.; Chandragunasekaran, A. M.; Balakrishnan, S. R. Molecular detection of the ruminal bacterium, *Butyrivibrio fibrisolvens*, in feces from rural residents of southern India. **Microbial Ecology in Health and Disease**, v. 21, p. 38-43, 2009.

Bayro, M.J.; Mukhopadhyay, J.; Swapna, G.V.T.; Huang, J.Y.; Ma, L.C.; Sineva, E.; Dawson, P.E.; Montelione, G.T.; Ebricht, R.H. Structure of Antibacterial Peptide

Microcin J25: A 21-Residue Lariat Protoknot. **Journal of the American Chemical Society**, v. 125, p. 12382–12383, 2003.

Callaway, T. R.; Dowd, S. E.; Edrington, T. S. Evaluation of bacterial diversity in the rumen and feces of cattle fed different levels of dried distillers grains plus solubles using bacterial tag-encoded FLX amplicon pyrosequencing. **Journal Animal Science**, v. 88, p. 3977–3983, 2010.

Chen, J.; Stevenson, D.M.; Weimer, P.J. Albusin b, a bacteriocin from the ruminal bacterium *Ruminococcus albus* 7 that inhibits growth of *Ruminococcus flavefaciens*. **Applied and Environmental Microbiology**, v.70, p.3167-3170, 2004.

Clarke, D.J.; Campopiano, D.J. Maturation of McjA precursor peptide into active microcin MccJ25. **Organic & Biomolecular Chemistry**, v. 21, p.2564–1566, 2007.

Cookson, A. L.; Noel, S. J.; Kelly, W. J.; Attwood, G. T. The use of PCR for the identification and characterisation of bacteriocin genes from bacterial strains isolated from rumen or caecal contents of cattle and sheep. **FEMS Microbiology Ecology**, v. 48, p. 1199-1207, 2004.

Cotta, M.A.; Whitehead, T.R. Regulation and cloning of the gene encoding amylase activity of the ruminal bacterium *Streptococcus bovis*. **Applied and Environmental Microbiology**. v.59, p. 189–196, 1993.

Cotter, P. D.; Hill, C.; Ross, R. P. Bacteriocins: developing innate immunity for food. **Nature Review Microbiology**, v. 3, p. 777-788, 2005.

Creevey, C. J.; Kelly, W. J.; Henderson G.; Leahy, S. C. Determining the culturability of the rumen bacterial microbiome. **Microbial Biotechnology**, p. 1-13, 2014.

De Carvalho, A. A. T.; Mantovani, H.C.; Vanetti, M.C.D. Bactericidal effect of bovicin HC5 and nisin against *Clostridium tyrobutyricum* isolated from spoiled mango pulp. **Letters Applied Microbiology**, v. 45, p. 68–74, 2007.

De Carvalho, A. A.T.; Vanetti, M. C. D.; Mantovani, H.C. Bovicin HC5 reduces thermal resistance of *Alicyclobacillus acidoterrestris* in acidic mango pulp. **Journal Applied Microbiology**, v. 104, p.1685–169, 2008.

Delgado, M. A.; Rintoul, M. R.; Farias, R. N.; Salomon, R. A. *Escherichia coli* RNA polymerase is the target of the cyclopeptide antibiotic microcin J25. **Journal Bacteriology**, v.183, p.4543–4550, 2001.

Ducasse, R.; Yan, K.P.; Goulard, C.; Blond, A.; Li, Y.Y.; Lescop, E.; Guittet, E.; Rebuffat, S.; Zirah, S. Sequence determinants governing the topology and biological activity of a lasso peptide, microcin J25. **ChemBioChem**, v. 13, n. 3, p. 371–380, 2012.

Duquesne, S.; Destoumieux-Garzón, D.; Zirah, S.; Goulard, C.; Peduzzi, J.; Rebuffat, S. Two Enzymes Catalyze the Maturation of a Lasso Peptide in *Escherichia coli*. **Chemistry & Biology**, v.14, p.793–803, 2007.

Flint, H. J.; Bayer, E. A.; Rincon, M. T.; Lamed, R.; White, B. A. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. **Nature Reviews Microbiology**, v. 6, p. 121-31, 2008.

Forster, R. J.; Gong, J.; Teather, R. M. Group-specific 16S rRNA hybridization probes for determinative and community structure studies of *Butyrivibrio fibrisolvens* in the rumen. **Applied and Environmental Microbiology**, v. 7, p. 1256-1260, 1997.

Fukuda, S.; Furuya, H.; Suzuki, Y.; Asanuma, N.; Hino, T. A new strain of *Butyrivibrio fibrisolvens* that has high ability to isomerise linoleic acid to conjugated linoleic acid. **Journal Genetic and Applied Microbiology**, v. 51, p. 105-113, 2005.

Gavriš, E.; Sit, C. S.; Cao, S.; Kandrór, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K. Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. **Chemistry & Biology**, v. 21, p. 509–518, 2014.

Gupta, A.; Madani, R.; Mukhtar, H. *Streptococcus bovis* endocarditis, a silent sign for colonic tumour. **Colorectal Disease**, v.12, 164–171, 2010.

Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Lasso peptides: an intriguing class of bacterial natural products. **Accounts of Chemical Research**, v.48. p. 1909-1919, 2015.

Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Steuber, H.; Harms, K.; Xie, X.; Marahiel M. A.. Xanthomonins I–III: A new class of lasso peptides with a seven-residue macrolactam ring. *Angew. **Angewandte Chemie International Edition***, v.53, p.2230–2234, 2014.

Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Klug, D.; Marahiel, M. A. Lasso peptides from proteobacteria: Genome mining employing heterologous expression and mass spectrometry. ***Biopolymers***, v. 100, p. 527–542, 2013.

Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Caulosegnins I-III: a highly diverse group of lasso peptides derived from a single biosynthetic gene cluster. ***J. Journal of the American Chemical Society***, v.135, p.210–222, 2012.

Helynck, G.; Dubertret, C.; Mayaux, J.F.; Leboul, J. Isolation of RP 71955, a new anti-HIV-1 peptide secondary metabolite. ***Journal Antibiotics***, v. 46, p. 1756–1757, 1993.

Heng, N. C. K.; Tagg, J. R. What's in a name? Class distinction for bacteriocins. ***Nature Reviews Microbiology***, London, v. 4, 2006.

Inokoshi, J.; Matsuhama, M.; Miyake, M.; Ikeda, H.; Tomoda, H. Molecular cloning of the gene cluster for lariatins biosynthesis of *Rhodococcus jostii* K01-B0171. ***Applied Microbiology Biotechnology***, v. 95, p. 451–460, 2012.

Iwatsuki, M.; Tomoda, H.; Uchida, R.; Gouda, H.; Hirono, S.; Omura, S. Lariatins, antimycobacterial peptides produced by *Rhodococcus* sp. K01– B0171, have a lasso structure. ***Journal of the American Chemical Society***, v. 128, p. 7486–7491, 2006.

Jong, A.; Van Heel, A. J.; Kok, J.; Kuipers, O. P. BAGEL2: mining for bacteriocins in genomic data. ***Nucleic Acids Research***, v. 38, p.W647-W651, 2010.

Kalmokoff, M. L.; Cyr, T. D.; Hefford, M. A.; Whitford, M. F.; Teather, R. M. Butyrivibriocin AR10, a new cyclic bacteriocin produced by the ruminal anaerobe *Butyrivibrio fibrisolvens* AR10: characterization of the gene and peptide. ***Canadian Journal of Microbiology***, v. 49, p. 763-773, 2003.

Kalmokoff, M. L.; Whitford, M. F.; Teather, R. M. Evidence for production of a new lantibiotic (Butyrivibriocin OR79A) by the ruminal anaerobe *Butyrivibrio fibrisolvens*

OR79: Characterization of the structural gene encoding Butyriovibriocin OR79A. **Applied Environmental Microbiology**, v. 65, p. 52128-52135, 1999.

Kalmokoff, M. L.; Teather, R.M. Isolation and characterization of a bacteriocin (butyriovibriocin ar10) from the ruminal anaerobe *Butyriovibrio fibrisolvens* ar10: evidence in support of the widespread occurrence of bacteriocin-like activity among ruminal isolates of *B. fibrisolvens*. **Applied and Environmental Microbiology**, v. 63, p. 394-402, 1997.

Kalmokoff, M.L.; Bartlett, F.; Teather, R. M. Are ruminal bacteria armed with Bacteriocins? **Journal Dairy Science**, v. 79, p. 2297-2306, 1996.

Kersten, R. D.; Yang, Y. L.; Xu, Y.; Cimermancic, P.; Nam, S. J.; Fenical, W.; Fischbach, M. A.; Moore, B. S.; Dorrestein, P. C. A mass spectrometry-guided genome mining approach for natural product peptidogenomics. **Nature Chemical Biology**, v. 7, p. 794–802, 2011.

Kim, Y. J.; Liu, R. H.; Bond, D. R.; Russell, J. B. Effect of linoleic acid concentration on conjugated linoleic acid production by *Butyriovibrio fibrisolvens* A38. **Applied and Environmental Microbiology**, v. 66, p. 5226-5230, 2000.

Kimura, K.; Kanou, F.; Takahashi, H.; Esumi, Y.; Uramoto, M.; Yoshihama, M. Propeptin, a new inhibitor of prolyl endopeptidase produced by *Microbispora*. I. Fermentation, isolation and biological properties. **Journal Antibiotics**, v. 50, p. 373–378, 1997.

Knappe, T. A.; Linne, U.; Xie, X.; Marahiel, M. A. The glucagon receptor antagonist BI-32169 constitutes a new class of lasso. **FEBS Letters**, v. 584, p. 785–789, 2010.

Knappe, T. A.; Linne, U.; Zirah, S.; Rebuffat, S.; Xie, X.; Marahiel, M. A. Isolation and structural characterization of capistruin, a lasso peptide predicted from the genome sequence of *Burkholderia thailandensis* E264. **Journal of the American Chemical Society**, v. 130, p. 11446–11454, 2008.

Kuznedelov, K.; Semenova, E.; Knappe, T. A.; Mukhamedyarov, D.; Srivastava, A.; Chatterjee, S.; Ebright, R. H.; Marahiel, M. A.; Severinov, K. The antibacterial threaded-lasso peptide capistruin inhibits bacterial RNA polymerase. **Journal of Molecular Biology**, v. 412, p. 842–848, 2011.

Lauková, A.; Czikková, S. Inhibition effect of enterocin CCM 4231 in the rumen fluid environment. **Letters in Applied Microbiology**, v. 26, p. 215-218, 1998.

Letzel, A. C.; Pidot, S. J.; Hertweck, C. Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. **BMC Genomics**, v. 15, p. 983, 2014.

Li, Y.; Ducasse, R.; Zirah, S.; Blond, A.; Goulard, C.; Lescop, E.; Giraud, C.; Hartke, A.; Guittet, E.; Pernodet, J. L.; Rebuffat, S. Characterization of Sviceucin from *Streptomyces* Provides Insight into Enzyme Exchangeability and Disulfide Bond Formation in Lasso Peptides. **ACS Chemical biological**, v. 1, p. 2641–2649, 2015.

Li, Y.; Zirah, S.; Rebuffat, S. **Lasso Peptides**. Springer; New York: Biosynthesis, regulation and export of lasso peptides. p. 81-95, 2014.

Lima, J.R.; Ribon, A.O.; Russell, J.B.; Mantovani, H.C. Bovicin HC5 inhibits wasteful amino acid degradation by mixed ruminal bacteria in vitro. **FEMS Microbiology Letters**, v. 292, p. 78-84, 2009.

Maia, M. R. G.; Chaudhary, L. C.; Bestwick, C. S.; Richardson, A. J.; Mckain, N; Larson, T. R.; Graham, I. A.; Wallace, R. J. Toxicity of unsaturated fatty acids to the biohydrogenating ruminal bacterium, *Butyrivibrio fibrisolvens*. **BMC Microbiology**, v. 10, p. 52, 2010.

Maksimov, M. O.; Koos, J. D.; Zong, C.; Lisko, B.; Link, A. J. Elucidating the specificity determinants of the AtxE2 lasso peptide isopeptidase. **The Journal of Biological Chemistry**, v. 290, p. 30806–30812, 2015.

Maksimov, M. O.; Link, A. J. Discovery and characterization of an isopeptidase that linearizes lasso peptides. **Journal of the American Chemical Society**, v.135, p.12038–12047, 2013.

Maksimov, M. O.; Pan, S. J.; Link, J. A. Lasso peptides: Structure, function, biosynthesis, and engineering. **Natural Product Reports**, v. 29, p. 996–1006, (2012a).

Maksimov, M. O.; Pelczer, I.; Link, A. J. Precursor-centric genome-mining approach for lasso peptide Discovery. **PNAS, Proceedings of the National Academy of Sciences**, v. 109, p. 15223–15228, (2012b).

Mantovani, H. C.; Russell, J. B. Bovicin HC5, a Lantibiotic Produced by *Streptococcus bovis* HC5, Catalyzes the Efflux of Intracellular Potassium but Not ATP. **Antimicrobial Agents and Chemotherapy**, v. 52, p. 2247-2249, 2008.

Mantovani, H.C.; Russell J.B. Inhibition of *Listeria monocytogenes* by bovicin HC5, a bacteriocin produced by *Streptococcus bovis* HC5. **The International Journal of Food Microbiology**, v.89, p.77–83, 2003.

Mantovani, H.C.; Hu, H.; Worobo, R.W.; Russell, J.B. Bovicin HC5, a bacteriocin from *Streptococcus bovis* HC5. **Microbiology** v. 148, p. 3347-3352. 2002.

Mantovani H.C.; Kam, D.K.; Ha, J.K.; Russell J.B. The antibacterial activity and sensitivity of *Streptococcus bovis* strains isolated from the rumen of cattle. **FEMS Microbiology Ecology**, v. 37, p. 223-229, 2001.

Mcintosh, J. A.; Donia, M. S.; Schmidt, E. W. Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. **Natural Product Reports**, v. 26, p. 537–559, 2009.

Medema, M.H.; Blin, K.; Cimermancic, P.; Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic Acids Research**, v.39, p. W339-346, 2011.

Morovsky, M.; Pristas, P.; Javorský, P.; Nes, I.F.; Holo, H. Isolation and characterization of enterocin BC25 and occurrence of the *entA* gene among ruminal Gram-positive cocci. **Microbiological Research**, v. 156, p. 133-138, 2001.

Morovsky, M.; Pristas, P.; Czikkova, S.; Javorsky, P. A bacteriocin-mediated antagonism by *Enterococcus faecium* BC25 against ruminal *Streptococcus bovis*. **Microbiological Research**, v. 153, p. 277–281, 1998.

Nes, I. F.; Yoon, S.; Diep, D. B. Ribosomally synthesized antimicrobial peptides (bacteriocins) in lactic acid bacteria: a review. **Food Science and Biotechnology**, v. 16, n. 5, p. 675, 2007.

Nigutova, K.; Morovsky, M.; Pristas, P.; Teather, Rm.; Holo, H.; Javorsky, P. Production of enterolysin A by rumen *Enterococcus faecalis* strain and occurrence of

enlA homologues among ruminal Gram-positive cocci. **Journal Applied Microbiology**, v. 102, p. 563-569, 2007.

Nilsen, T.; Nes, I.F.; Holo, H. Enterolysin A, a cell wall-degrading bacteriocin from *Enterococcus faecalis* LMG 2333. **Applied Environment Microbiology**, v.69, p. 2975-2984, 2003.

O'sullivan, O.; Begley, M.; Ross, R. P.; Cotter, P. D.; Hill, C. F. Identification of novel lantibiotic operons using lanm-based genome mining. **Probiotics and Antimicrobial Proteins**, v. 3, p. 27-40, 2011.

Odenyo, A.A.; Mackie, R.I.; Stahl, D.A.; White, B.A. The use of 16S rRNA-targeted oligonucleotide probes to study competition between ruminal fibrolytic bacteria: development of probes for *Ruminococcus* species and evidence for bacteriocin production. **Applied Environmental Microbiology**, v. 60, p. 3688-3696, 1994.

Paiva, A.D.; Fernandes, K.M.; Dias, R.S.; Rocha, A.S.; Oliveira, L.L.; Neves, C.A.; Paula, S.O.; Mantovani, H.C. Safety evaluation of the antimicrobial peptide bovicin HC5 orally administered to a murine model. **BMC Microbiology**, v. 13, p. 69, 2013.

Pan, S. J; Rajniak, J.; Cheung, W. L.; Link A. J. Construction of a single polypeptide that matures and exports the lasso peptide microcin J25. **ChemBioChem**, v. 13, p.367–370, 2012.

Potterat, O.; Wagner, K.; Gemmecker, G.; Mack, J.; Puder, C.; Vettermann, R.; Streicher, R. BI-32169, a bicyclic 19-peptide with strong glucagon receptor antagonist activity from *Streptomyces* sp. **Journal of Natural Products**, v. 67, p. 1528–1531, 2004.

Rea, M.C.; Ross, R.P.; Cotter, P.D.; Hill, C. Classification of bacteriocins from Gram-positive bacteria. In Drider D, Rebuffat S (ed) Prokaryotic Antimicrobial Peptides: From Genes to Applications. **Springer Science**; New York, p.29-72, 2011.

Riley, M. A. Bacteriocin-Mediated Competitive Interactions of Bacterial Populations and Communities. In Prokaryotic Antimicrobial Peptides: From Genes to Applications. Edited by Drider D, Rebuffat S. New York: **Springer Science**; p. 13-26, 2011.

Rintoul, M. R.; Arcuri, B. F.; Salomon, R. A.; Farias, R. N.; Morero, R. D. The antibacterial action of microcin J25: evidence for disruption of cytoplasmic membrane energization in *Salmonella newport*. **FEMS Microbiology Letters**, v. 204, p. 265–270, 2001.

Rosengren, K.J.; Clark, R.J.; Daly, N.L.; Goransson, U.; Jones A.; Craik D.J. Microcin J25 Has a Threaded Sidechain-to-Backbone Ring Structure and Not a Head-to-Tail Cyclized Backbone. **Journal of the American Chemical Society**, v. 125, p. 12464–12474, 2003.

Russell, J. B.; Mantovani, H. C. The bacteriocins of ruminal bacteria and their potential as an alternative to antibiotics. **Journal of Molecular Microbiology and Biotechnology**, v.4, p. 347–355, 2002.

Rychlik, J. L.; Russell, J. B. Bacteriocin-like activity of *Butyrivibrio fibrisolvens* JL5 and its effect on other ruminal bacteria and ammonia production. **Applied and Environmental Microbiology**, v. 68, p. 1040-1046, 2002.

Salomon, R. A.; Farias, R. N. Microcin 25, a novel antimicrobial peptide produced by *Escherichia coli*. **Journal Bacteriology**, v.174, p.7428–7435, 1992.

Singh, M.; Sareen, D. Novel LanT Associated Lantibiotic Clusters Identified by Genome Database Mining. **PLoS ONE**, v. 9, p. 91352, 2014.

Teather, R. M.; Kalmokoff, M. L.; Whitford, M. F. The role of bacteriocins in rumen microbial ecology. **In: Proceedings of the 8 th International Symposium on Microbial Ecology**. Atlantic: Canada Society for Microbial Ecology, 1999.

Teather, R.M. And Forster, R.J. Manipulating the rumen microflora with bacteriocins to improve ruminant production. **Canadian Journal of Animal Science**, v.78, p. 57–69, 1998.

Tsunakawa, M.; Hu, S. L.; Hoshino, Y.; Detlefson, D. J.; Hill, S.E.; Furumai, T.; White, R. J.; Nishio, M.; Kawano, K.; Yamamoto, S.; Fukagawa, Y.; Oki, T. Siamycins I and II, new anti-HIV peptides: I. Fermentation, isolation, biological activity and initial characterization. **Journal Antibiotics**, v. 48, p. 433–434, 1995.

Um, S.; Kim, Y.J.; Kwon, H.; Wen, H.; Kim, S.H.; Kwon, H.C.; Park, S.; Shin, J. Sungsanpin, a lasso peptide from a deep-sea streptomycete. **Journal of Natural Products**, v. 76, p. 873–879, 2013.

Van Heel, A. J.; Jong, A.; Montalbán-López, M.; Kok, J.; Kuipers, O. P. Bagel3: automated identification of genes encoding bacteriocins and (non-) bactericidal posttranslationally modified peptides. **Nucleic Acids Research**, v. 41, p. 448-453, 2013.

Velasquez, J. E.; van der Donk. W. A. Genome mining for ribosomally synthesized natural products. **Current Opinion in Chemical Biology**, v.15, p. 11–21, 2011.

Wallace, R. J.; Brammall, M. L. The role of different species of rumen bacteria in the hydrolysis of protein in the rumen. **Jounal Genetic Microbiology**, v. 131, p. 821-832, 1985.

Wang, H.; Fewer, D.; Sivonen, K. Genome mining demonstrates the widespread occurrence of gene clusters encoding bacteriocins in cyanobacteria. **PLoS One**, v.6, p. 1-10, 2011.

Weber, W.; Fischli, W.; Hochuli, E.; Kupfer, E.; Weibel, E. K. Anantin—a peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity. **Journal Antibiotics**, v. 44, p. 164–171, 1991.

Whitford, M.F.; Mcpherson, M.A.; Forster, R.J.; Teather, R.M. Identification of bacteriocin-like inhibitors from rumen *Streptococcus spp.* and isolation and characterization of bovicin 255. **Applied Enviromental. Microbiology**, v. 67, p. 569-574, 2001.

Wilson, K. A.; Kalkum, M.; Ottesen, J.; Yuzenkova, J.; Chait, B. T.; Landick, R.; Muir, T.; Severinov, K.; Darst, S. A. Structure of Microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. **Journal of the American Chemical Society**, v. 125, p. 12475–12483, 2003.

Xavier, B.; Russel, J. B. The ability of non-bacteriocin producing *Streptococcus bovis* strains to bind and transfer bovicin HC5 to other sensitive bacteria. **Anaerobe**, v. 4, p. 168–172, 2009.

Xiao, H.; Chen, X.; Chen, M.; Tang, S.; Zhao, X.; Huan, L. Bovicin HJ50, a novel lantibiotic produced by *Streptococcus bovis* HJ50. **Microbiology**, v. 150, p. 103-108, 2004.

Yan, K. P.; Li, Y.; Zirah, S.; Goulard, C.; Knappe, T. A.; Marahiel, M. A.; Rebuffat, S. Dissecting the maturation steps of the lasso peptide microcin J25 in vitro. **ChemBioChem**, v. 13, p.1046– 1052, 2012.

Yang, S. C.; Lin, C. H.; Sung, C. T.; Fang, J. Y. Antibacterial activities of bacteriocins: application in foods and pharmaceuticals. **Frontiers in Microbiology**, Lausanne, v. 5, p. 1-10, 2014.

Zacharof, M. P; Lovitt, R. W. Bacteriocins produced by lactic acid bacteria-a review article. **APCBEE Procedia**, v. 2, p. 50-56, 2012.

Zhu, S.; Hegemann, J. D.; D. Fage, C. D.; Zimmermann, M.; Xiulan Xie, U. L.; Mohamed, A. Marahiel. Insights into the Unique Phosphorylation of the Lasso Peptide Paeninodin. **The Journal biological of chemistry**, v. 291, p. 13662-13678, 2016.

Zimmermann, M.; Hegemann, J. D.; Xie, X.; Marahiel, M. A. Characterization of caulonodin lasso peptides revealed unprecedented N-terminal residues and a precursor motif essential for peptide maturation. **Chemical Science**, v. 5, p. 4032–4043, 2014.

Zimmermann, M.; Hegemann, J. D.; Xie, X.; Marahiel, M. A. The astexin-1 lasso peptides: Biosynthesis, stability, and structural studies. **Chemistry & Biology**, v. 20, p. 558–569, 2013.

CAPÍTULO 2

Gênomica comparativa de *clusters* putativos de biossíntese de peptídeos laço em procariotos do rúmen

Resumo

Estratégias de mineração genômica têm contribuído para a descoberta de novos peptídeos laço e também de novos microrganismos produtores de peptídeos desse grupo. Estudos da microbiologia do rúmen sugerem que a capacidade de bactérias produzirem peptídeos antimicrobianos possui papel importante na competição entre espécies por nichos ecológicos do ecossistema ruminal. No entanto, pouco se sabe sobre a produção de peptídeos laço pela microbiota do ecossistema ruminal. Neste trabalho, análises *in silico* foram realizadas para investigar a distribuição de *clusters* gênicos relacionados com a biossíntese de peptídeos laço (*lasA*, *lasB*, *lasC* e *lasD*) em genomas de 320 bactérias e 5 archaeas ruminais e de 25 bactérias isoladas de fezes de ruminantes. Para 42 genomas de bactérias ruminais, foram inferidos *clusters* putativos incompletos e completos para peptídeos laço, utilizando as ferramentas BAGEL3, AntiSMASH, e dois Bancos locais construídos a partir de informações disponíveis no NCBI, sendo o filo Firmicutes o de maior representatividade. De um total de 26 genomas de bactérias do rúmen que tiveram inferidos *clusters* gênicos completos de biossíntese de peptídeos laço, 53% foi representado por linhagens do gênero *Butyrivibrio*. A análise da organização dos 26 *clusters* putativos completos de biossíntese de peptídeos laço inferidos via mineração genômica revelou a presença de vários genes com funções moleculares e biológicas diversas. Dentre esses, predominaram os genes putativos para histidina cinase (HisKA) em 38% dos *clusters* investigados. Os domínios conservados dos produtos preditos para os genes de biossíntese de peptídeos laço indicaram que os genes *lasB* e *lasC* foram os mais conservados entre os peptídeos analisados. Árvores filogenéticas representativas dos genes *lasA*, *lasB*, *lasC* e *lasD* revelaram que a presença dos genes de biossíntese de peptídeos laço retratavam um processo evolutivo. A reconstrução filogenética com base no gene que codifica o rRNA 16S para 42 genomas de bactérias do rúmen que apresentaram *clusters* putativos completos e incompletos para biossíntese de peptídeos laço, permitiu o agrupamento das linhagens em seus respectivos gêneros ou famílias (Lachnospiraceae e Coriobacteriaceae). A partir das

análises de metatranscriptoma, pôde-se inferir *reads* para genes de biossíntese de peptídeos laço. Os resultados demonstraram que a análise *in silico* revelou novos *clusters* de genes biossintéticos em espécies bacterianas não previamente relacionadas com peptídeos laço, sugerindo que a microbiota ruminal representa uma fonte potencial de novos peptídeos laço.

Palavras chaves: rúmen, mineração genômica, bacteriocinas, RiPPs.

Abstract

Genomic mining strategies have contributed to the discovery of new lasso peptides and new lasso peptide producing organisms. Studies of rumen microbial ecology suggest that the ability of bacteria to produce antimicrobial peptides plays an important role in species competing for ecological niches in the ruminal ecosystem. However, little is known about the production of lasso peptides by ruminal microorganisms. In this work, we performed genome mining in complete and partial genome sequences of 320 ruminal bacteria, 5 ruminal archaea and 25 bacteria isolated from ruminant feces to determine the distribution and diversity of lasso peptide gene clusters. Incomplete and complete putative clusters for lasso peptides were obtained for 42 genomes of ruminal bacteria using BAGEL3, AntiSMASH tools, Local Database 1 and Local Database 2, with the Firmicutes being the most representative phylum. Complete gene clusters of lasso peptides biosynthesis were inferred for 26 genomes of ruminal bacteria, which were mostly (53%) represented by strains of the genus *Butyrivibrio*. The analysis of the organization of the 26 complete putative clusters of lasso peptides biosynthesis inferred via genomic mining revealed the presence of several genes with diverse molecular and biological functions, among which predominated the putative genes for histidine kinase (HisKA) in 38% of the clusters. The conserved domains of the products predicted for lasso peptide biosynthetic genes indicated *lasB* and *lasC* genes as the most conserved among the peptides that were investigated. Phylogenetic trees representative of the genes *lasA*, *lasB*, *lasC* and *lasD* revealed that the presence of lasso peptide biosynthesis genes portrayed an evolutionary process. Phylogenetic reconstruction based on the gene encoding the 16S rRNA for 42 genomes of ruminal bacteria harboring complete and incomplete putative clusters for lasso peptide biosynthesis allowed the grouping of the strains in their respective genera or family (Lachnospiraceae and Coriobacteriaceae).

From the metatranscriptome analyzes, reads could be inferred that the biosynthetic genes for lasso peptides. These results demonstrated that *in silico* analysis revealed new biosynthetic gene clusters in bacterial species not previously related to lasso peptides, suggesting that the ruminal microbiota represents a potential source of novel lasso peptides.

Keywords: genome mining, bacteriocin, RiPPs, rumen.

2.0 Introdução

Os peptídeos laço compõem uma classe de moléculas sintetizadas ribossomicamente e modificadas pós-tradução (RiPPs) que apresentam diversas atividades biológicas, incluindo inibição de enzimas citoplasmáticas, inibição do crescimento bacteriano, ação antiviral, antagonismo de receptores de peptídeos de origens animal e humana (Weber et al., 1991; Helynck et al., 1993; Kuznedelov et al., 2011; Gavrish et al., 2014). A característica peculiar aos peptídeos laço é sua topologia estrutural, com presença de um anel macrolactâmico composto por sete a nove resíduos de aminoácidos, formado por uma ligação amida entre o grupo α -amino da região N-terminal de resíduos de glicina, cisteína, serina ou alanina e o grupo carboxila de resíduos de aspartato ou glutamato; a extremidade C-terminal origina o *loop* e a cauda do peptídeo laço, os quais são presos ao anel por interações estéricas, assistidas muitas vezes por ligações dissulfeto (McIntosh et al., 2009; Velasquez & Van der Donk, 2011; Alvarez-Sieiro et al., 2016). Muitas informações sobre os tipos e características dos peptídeos laço têm sido obtidas por meio de estudos de mineração genômica, os quais têm possibilitado a descoberta de novas espécies microbianas produtoras, como representantes do Filo Firmicutes (Letzel et al., 2014; Zhu et al., 2016), e a identificação de novos peptídeos laço em genomas bacterianos, como o capistruin reportado em *Burkholderia thailandensis* (Knappe et al., 2008). Estima-se que mais da metade dos 40 peptídeos laço atualmente descritos foram descobertos a partir da mineração de genomas (Hegemann et al., 2015).

A presença de genes conservados e, geralmente arranjados adjacentes nos *scaffolds* microbianos, tem permitido a mineração de genomas e descobertas de novos peptídeos laço, a exemplo de Capistruin (Knappe et al., 2008), SSV-2083 e SRO15-2005 (Kersten et al., 2011), Astexin-1 (Zimmermann et al., 2013) e Caulonodin IV a VII (Zimmermann et al., 2014). A mineração de genomas para descoberta de novos

peptídeos laço utiliza *softwares* específicos com base nos domínios conservados das proteínas envolvidas, principalmente, na modificação e transporte desses peptídeos. Estes *softwares* são então empregados em um *screening* molecular de *clusters* gênicos de interesse, a partir de centenas a milhares de sequências genômicas microbianas depositadas em bancos de dados de livre acesso (Hegemann et al., 2015). Estima-se que mais da metade dos 40 peptídeos laço já descritos foram descobertos a partir da mineração de genomas (Hegemann et al., 2015).

No caso do peptídeo laço astexin-1, foi utilizado um algoritmo de mineração de genoma desenvolvido para identificar *clusters* gênicos de peptídeos laço em mais de três mil genomas sequenciados de procaríotos (Maksimov et al., 2012b). Naquele estudo, pequenas ORFs, algumas não anotadas, contendo aminoácidos padrões consistentes com as sequências de peptídeos laço conhecidos, foram identificadas. Quando um precursor potencial de peptídeo laço era encontrado, regiões próximas do genoma eram pesquisadas para homologia com as enzimas de maturação. Essas análises resultaram na predição do peptídeo laço astexin-1, detectado na bactéria *Asticcacaulis Excentricus*, proveniente de água doce (Maksimov et al., 2012b).

Estudos com base em mineração genômica têm gerado também contribuições importantes para descoberta de novas espécies microbianas produtoras de peptídeos laço que, até recentemente, era restrita aos filos Actinobacteria e Proteobacteria (Alpha e Beta-proteobacteria) (Zhu et al., 2016). Atualmente, a síntese de peptídeos laço foi identificada em representantes de outros filos microbianos, como Chloroflexi, Firmicutes, Bacteroidetes, Espiroquetas, Cyanobacteria, Verrucomicrobia e Euryarchaeota (Archaea) (Maksimov et al., 2012b).

Microrganismos produtores de peptídeos laço têm sido relacionados a diversos habitats, tais como solos e sedimentos marinhos, ambientes hídricos naturais, áreas contaminadas (Li et al., 2014) e rúmen (Letzel et al., 2014). O primeiro registro na literatura a reportar o ecossistema ruminal como habitat de origem de um microrganismo produtor de peptídeos laço inferiu, via mineração genômica, genes putativos para peptídeos laço para *Butyrivibrio proteoclasticum* B316, uma linhagem bacteriana oriunda do rúmen e pertencente ao Filo Firmicutes (Letzel et al., 2014). Mais tarde, esse estudo foi corroborado por Azevedo et al. (2015) que também verificaram a distribuição de *clusters* gênicos para peptídeos laço em genomas microbianos do rúmen referentes aos gêneros bacterianos *Butyrivibrio*, *Lachnospiraceae* e *Ruminococcus*.

Ainda que os estudos de Letzel et al. (2014) e Azevedo et al. (2015) não tenham direcionado suas investigações, exclusivamente, para mineração de *clusters* putativos de biossíntese de peptídeo laço em genomas da microbiota ruminal, os resultados obtidos permitem sugerir que o rúmen contém linhagens microbianas produtoras desses peptídeos. Essa hipótese pode ser sustentada pela reconhecida capacidade da microbiota ruminal produzir bacteriocinas, que são peptídeos antimicrobianos que pode favorecer a competição entre espécies por nichos ecológicos do ecossistema ruminal (Kalmokoff et al., 1996; Teather & Forster, 1998; Mantovani et al., 2001). Desta forma, neste trabalho investigou-se, por meio de análises *in silico*, a distribuição de *clusters* gênicos relacionados com a biossíntese de peptídeos laço em genomas de 320 bactérias e 5 archaeas ruminais e de 25 bactérias isoladas de fezes de ruminantes. Os resultados confirmaram que o ecossistema ruminal representa uma importante fonte potencial de microrganismos produtores de peptídeos laço.

2.1 Material e Métodos

2.1.1 Genomas de bactérias e Archaeas

Os genomas completos e parciais de 320 bactérias e 5 archaeas ruminais e 25 bactérias isoladas de fezes de ruminantes foram obtidos em formato fasta (fa) a partir dos bancos de dados Joint Genome Institute (JGI) (<http://genome.jgi.doe.gov/>) e National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/genome/>). Para a obtenção dos genomas de bactérias e archaeas foram utilizadas as palavras chaves: "Hungate 1000", para o banco de dados JGI e, "rumen" para o banco de dados do NCBI.

2.1.2 Mineração de peptídeos laço e análises dos produtos dos genes

Os genomas no formato fasta foram carregados, individualmente, nos *softwares* BAGEL3 (Van Heel et al., 2013) e AntiSMASH (Medema et al., 2011). Os genomas foram também curados manualmente usando Blastn e Blastx contra os bancos de dados locais que foram organizados com informações disponíveis sobre peptídeos laços no NCBI. O Banco de dados local 1 foi montado com as 4 sequências nucleotídicas referentes aos genes *mcjABCD* previamente identificados, e o Banco de dados local 2

foi elaborado com 20 sequências de aminoácidos as quais estavam disponíveis no NCBI, codificados por genes envolvidos na biossíntese de outros peptídeos laço (Tabelas S2 e S3). Para essas análises comparativas dos genomas com os Bancos de dados 1 e 2 foi utilizado como critério de corte *E-value* <0.00001 e identidade de aminoácidos >30%. Após essas análises, foi calculada a porcentagem de *clusters* gênicos inferidos pelos *softwares* e bancos de dados.

Os genomas completos ou parciais com *clusters* putativos incompletos inferidos via mineração genômica (BAGEL3 e AntiSMASH) para peptídeos laço, foram submetidos à análises no BLAST via Putty (<http://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>) e WinSCP (<https://winscp.net/eng/index.php>) com o objetivo de buscar nesses genomas os genes de biossíntese que não haviam sido inferidos via *softwares*, uma vez que poderiam estar em outros *scaffolds* em função de montagens incompletas e incorretas dos genomas. Para essas análises comparativas foi utilizado como critério de corte um *E-value* <10⁻¹⁰. A partir dessas análises, foi calculada a distribuição percentual dos genes de biossíntese nos genomas de cada gênero bacteriano contendo *clusters* completos e incompletos de peptídeos laço.

A partir dos *clusters* gênicos putativos completos para peptídeos laço, foi realizada análise dos produtos codificados pelos genes, genericamente denominados, *lasA*, *lasB*, *lasC* e *lasD* identificados nos genomas bacterianos. Os domínios conservados de cada proteína foram identificados através do banco de dados “Conserved domains” (disponível no NCBI). Alinhamentos múltiplos foram realizados (Clustal Omega) dos produtos dos genes *lasB* e *lasC* para as linhagens da mesma espécie e, posteriormente, com o auxílio da ferramenta *Color Align Conservation* (Stothard, 2000) foi realizada a verificação das regiões consensos para as sequências. Para essas análises comparativas foi utilizado como critério de corte um *E-value* <10⁻¹⁰ e uma identidade de aminoácidos >30%.

2. 1. 3 Características dos peptídeos *core*

A partir dos peptídeos precursores obtidos via mineração pelas ferramentas Bagel3 e antiSMASH foram realizadas análises manuais para inferência dos peptídeos *core*, com base em características relatadas na literatura como: primeiros resíduos de aminoácidos das sequências de peptídeo *core* podem ser glicina, cisteína, serina ou

alanina; o anel macrolactâmico é composto por sete a nove resíduos de aminoácidos, formado por uma ligação entre resíduos de glicina, cisteína, serina ou alanina e resíduos de aspartato ou glutamato; o tamanho das sequências de peptídeo *core* variam de 14 a 21 resíduos de aminoácidos. Posteriormente essa metodologia foi validada com peptídeos laço já conhecidos. Também foi avaliada a presença de resíduos de cisteína nas sequências dos peptídeos *core* para inferência de possíveis ligações dissulfeto entre esses resíduos,

2. 1. 4 Análise filogenética/Evolutiva

Com o objetivo de verificar se as sequências dos genes de biossíntese de peptídeos laço revelavam evidências de evolução de sequências ou transferência horizontal de genes entre linhagens da mesma espécie, foi realizada a reconstrução filogenética dos genes de biossíntese de peptídeos laço (*lasA*, *lasB*, *lasC*, *lasD*) das linhagens que apresentaram *clusters* completos e nas sequências parciais do gene rRNA 16S de 168 genomas cujas sequências estavam disponíveis no GenBank. Os alinhamentos foram realizados utilizando a plataforma de Análise Genética de Evolução Molecular Mega 7 (Kumar et al., 2016), utilizando a opção “MUSCLE” para "Multiple Alignment" (Edgar, 2004). A matriz de distância das árvores filogenéticas foi estimada pelo método "Maximum likelihood" e o algoritmo BioNJ (Gascuel, 1997) foi utilizado na construção de filogramas, os quais foram processados também pelo programa Mega 7, aplicando um "bootstrap" de 1.000 vezes para inferência estatística.

2. 1. 5 Análise de Metatranscriptoma

Para avaliar se os genes de biossíntese de peptídeos laço são expressos pela microbiota ruminal, foram realizadas análises comparativas das sequências nucleotídicas referentes aos genes putativos *lasA*, *lasB*, *lasC* e *lasD* com dados de metatranscriptoma. Primeiramente, foi realizado o *download* dos arquivos de metatranscriptoma SRX1585089: *Rúmen bovino microbiota metatranscriptome* (mRNA-enriched RNA), cujo número de acesso é SRR3169851 (bovino dieta mista), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 10.5M spots, 3.9G bases, tamanho 2Gb; SRX195368: *Rumen metatranscriptome* , cujo

número de acesso é SRR594193(bovino 2155 com dieta controle), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 14.3M spots, 2.9G bases, 1.6Gb *downloads*; SRX196213, *Rumen metatranscriptome*, cujo número de acesso é SRR594215 (bovino 2155 com dieta rica em óleo), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 14.2M spots, 2.8G bases, 1.6Gb *downloads*; SRX196409: *Rumen metatranscriptome*, cujo número de acesso é SRR594642 (bovino 6021 com dieta controle), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 14.3M spots, 2.9G bases, 1.6Gb *downloads*; SRX196410: *Rumen metatranscriptome*, cujo número de acesso é SRR594654 (bovino 6021 com dieta rica em óleo), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 14.3M spots, 2.9G bases, 1.6Gb *downloads*; SRX286794: *metatranscriptomic sample*, cujo número de acesso é SRR873462 (bovino com baixa produção de metano), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 13.6M spots, 4.1G bases, 2.5Gb *downloads*; SRX445849: *metatranscriptomic sample*, cujo número de acesso é SRR1138697 (bovino com alta produção de metano), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 22.4M spots, 6.7G bases, 3.9Gb *downloads*; SRX1647089: *Bovine rumen microbiome L167 metatranscriptome*, cujo número de acesso é SRR3257011 (bovino oriundo do Canadá), e que apresentava as seguintes características: codificador ILLUMINA (Illumina MiSeq): 29.5M spots, 5.9G bases, 4Gb *downloads*. Em todos os arquivos de metatranscriptomas tiveram os adaptadores previamente eliminados

Os arquivos foram depositados entre os anos de 2013 a 2017 no banco de dados do Sequence Read Archive (SRA) do *National Center for Biotechnology Information* (NCBI), disponível no link:<http://www.ncbi.nlm.nih.gov/sra>. De posse dos arquivos de metatranscriptoma, foram feitas análises de matriz de distância, seguida da eliminação das sequências de *reads* compartilhadas ou que tinham distância menor de 0,5, com objetivo de selecionar apenas as *reads* únicas para os genes de interesse.

Com o auxílio do *software Bowtie* (Langmead et al., 2009) foi realizado o alinhamento das sequências nucleotídicas referentes aos genes *lasA*, *lasB* e *lasC* e *lasD* de cada espécie e linhagem bacteriana com os arquivos de metatranscriptoma. A expressão gênica foi quantificada calculando-se o número de *reads* por kilobase por milhões de *reads* mapeadas (Reads Per Kilobase Million - RPKM), para cada genoma

em função do gene analisado. O cálculo do RPKM considera o número de reads, o tamanho dos genes onde as sequências foram mapeadas e o número total de reads.

Onde:

$$\text{RPKM} = C/LN$$

- C: Número de reads mapeados para uma sequência gênica (ex. transcrito, exon, etc).
- L: Tamanho da sequência do gene (em kb).
- N: Número total de reads mapeados na amostra (em milhões).

2.3 Resultados

Genomas de 320 bactérias e 5 archaeas ruminais e de 25 bactérias isoladas de fezes de ruminantes foram obtidos dos bancos de dados do Joint Genome Institute (JGI) e do GenBank. Os genomas bacterianos foram classificados em 74 gêneros distintos, sendo que *clusters* putativos incompletos e completos para peptídeos laço foram encontrados em genomas de 11 gêneros de bactérias pertencentes aos filos Firmicutes (39 linhagens), Bacteroidetes (1 linhagem), Proteobacteria (1 linhagem) e Actinobacteria (1 linhagem), totalizando 42 linhagens bacterianas ou 12% dos genomas avaliados (Tabela 2). Nenhum *cluster* putativo gênico para biossíntese de peptídeos laço foi identificado em genomas de archaeas do rúmen e de bactérias isoladas de fezes de ruminantes.

O número de *clusters* gênicos para peptídeos laço identificados individualmente e simultaneamente nos genomas de bactérias do rúmen variou de acordo com a ferramenta utilizada para a mineração dos dados. O *software* BAGEL3 foi a ferramenta mais eficiente para identificar *clusters* gênicos de peptídeos laço nos genomas bacterianos (n = 26) quando comparado às demais ferramentas utilizadas (Tabela 2). Os percentuais de *clusters* putativos para peptídeos laço das 42 linhagens, inferidos via mineração genômica utilizando as ferramentas BAGEL3, AntiSMASH, Banco de dados local 1 e Banco de dados local 2, foram, respectivamente, 63, 43, 58 e 53%. Os genomas de *Bacillus* sp. MB2021, *Butyrivibrio* sp. IN11a14, *Butyrivibrio* sp. MC2013, *Butyrivibrio* sp. MC2021, *Butyrivibrio* sp. WCE2006, *Clostridium beijerinckii* HUN142, *Lachnospira multipara* LB2003, *Lachnospira multipara* MC2003, *Lachnospiraceae*

bacterium YSD2013 tiveram *clusters* gênicos de peptídeos laço preditos por todas as ferramentas de mineração genômica utilizadas neste estudo (Tabela 2). Contudo, *clusters* gênicos de peptídeos laços para três linhagens (*Butyrivibrio* sp. FC2001; *Butyrivibrio* sp. XPD 2006; *Lachnospiraceae bacterium* MA2020) foram detectados, exclusivamente, pela ferramenta AntiSMASH, enquanto que para a linhagem *Butyrivibrio* sp. Su6, o *cluster* gênico putativo de biossíntese de peptídeo laço foi inferido somente a partir da comparação de sequências com os genes de biossíntese de microcina J25 (genes *mcjABCD*) organizados no Banco de dados local 1.

Tabela 2: Genomas de bactérias com *clusters* putativos de biossíntese de peptídeos laço inferidos pelos softwares BAGEL3, AntiSMASH e por comparação de sequências com os Bancos de dados 1 (genes *mcjABCD*) e Banco de dados 2 (sequências de aminoácidos de peptídeos laço)

Linhasgens ^a (n° = 42)	Filos	Bagel	AntiSMASH	Banco de dados local1	Banco de dados local 2
<i>Bacillus cereus</i> KPR-7 ^a	Firmicutes	1	1	0	0
<i>Bacillus</i> sp. MB2021	Firmicutes	1	1	1	1
<i>Butyrivibrio</i> sp. FC2001	Firmicutes	0	1	0	0
<i>Butyrivibrio fibrisolvens</i> MD2001	Firmicutes	1	0	0	0
<i>Butyrivibrio fibrisolvens</i> WTE3004	Firmicutes	1	0	0	0
<i>Butyrivibrio proteoclasticus</i> B316	Firmicutes	1	0	1	0
<i>Butyrivibrio proteoclasticus</i> FD2007	Firmicutes	1	0	0	0
<i>Butyrivibrio</i> sp INlla14	Firmicutes	1	1	1	1
<i>Butyrivibrio</i> sp. LC3010	Firmicutes	1	0	1	1
<i>Butyrivibrio</i> sp. MC2013	Firmicutes	1	1	1	1
<i>Butyrivibrio</i> sp. MC2021	Firmicutes	1	1	1	1
<i>Butyrivibrio</i> sp. NC3005	Firmicutes	1	1	0	0
<i>Butyrivibrio</i> sp. Su6	Firmicutes	0	0	1	0
<i>Butyrivibrio</i> sp. WCE2006	Firmicutes	1	1	1	1
<i>Butyrivibrio</i> sp. XBB 1001	Firmicutes	1	1	0	0
<i>Butyrivibrio</i> sp. XPD 2006	Firmicutes	0	1	0	0
<i>Clostridium algidicarnis</i> B3	Firmicutes	0	0	1	1
<i>Clostridium beijerinckii</i> HUN142	Firmicutes	1	1	1	1
<i>Coriobacteriaceae bacterium</i> KH1P3	Bacteroidetes	1	0	0	0
<i>Enterococcus mundtii</i> C2	Firmicutes	0	0	1	1
<i>Eubacterium celulosolvans</i> LD2006	Proteobacteria	1	1	0	0
<i>Lachnospira multipara</i> D15d	Firmicutes	1	0	0	0
<i>Lachnospira multipara</i> LB2003	Firmicutes	1	1	1	1
<i>Lachnospira multipara</i> MC2003	Firmicutes	1	1	1	1
<i>Lachnospira mutipara</i> ATCC 19207	Firmicutes	1	1	0	1
<i>Lachnospiraceae bacterium</i> MA2020	Firmicutes	0	1	0	0
<i>Lachnospiraceae bacterium</i> C7	Firmicutes	0	0	1	1
<i>Lachnospiraceae bacterium</i> G11	Firmicutes	1	0	0	0
<i>Lachnospiraceae bacterium</i> ND2006	Firmicutes	1	0	0	0

<i>Lachnospiraceae bacterium</i> NK4A136	Firmicutes	1	0	0	0
<i>Lachnospiraceae bacterium</i> NKA144	Firmicutes	1	0	0	0
<i>Lachnospiraceae bacterium</i> YSD2013	Firmicutes	1	1	1	1
<i>Propionibacterium</i> sp. MB3007	Actinobacteria	0	0	1	1
<i>Pseudobutyrvibrio</i> sp. AR14	Firmicutes	0	0	1	0
<i>Pseudobutyrvibrio</i> sp. NOR37	Firmicutes	0	0	1	1
<i>Ruminococcus albus</i> 8	Firmicutes	1	1	0	0
<i>Ruminococcus flavefaciens</i> 17	Firmicutes	0	0	1	1
<i>Ruminococcus flavefaciens</i> FD1	Firmicutes	0	0	1	1
<i>Staphylococcus epidermidis</i> AG42	Firmicutes	0	0	1	1
<i>Streptococcus gallolyticus</i> LMG 15572	Firmicutes	0	0	1	1
<i>Streptococcus gallolyticus</i> VTM1R27	Firmicutes	0	0	1	1
<i>Streptococcus gallolyticus</i> VTM2R47	Firmicutes	0	0	1	1
Total		42	26	18	24

^aOs genomas foram obtidos a partir do portal de Genomas do JGI (Projeto Hungate1000) e do GenBank, durante o período de abril a dezembro de 2016 . As linhagens destacadas em negrito tiveram *clusters* gênicos putativos inferidos para peptídeo laço por todas as ferramentas de predição de sequências utilizadas neste estudo.

Para avaliar se os genes ausentes nos *clusters* putativos incompletos de biossíntese de peptídeos laço estavam distribuídos em *scaffolds* diferentes (decorrente da montagem incompleta ou incorreta dos genomas), efetuou-se uma análise manual dos mesmos no BLAST. Alguns genes de biossíntese que não puderam ser inferidos nas linhagens de *Bacillus cereus* KPR-7A, *Butyrivibrio* sp. LC3010, *Butyrivibrio* sp. MC2013, *Butyrivibrio* sp. NC3005, *Butyrivibrio* sp. FC2001, *Butyrivibrio* sp. WCE2006, *Lachnospiraceae bacterium* NKA144 e *Lachnospiraceae bacterium* YSD2013 com as ferramentas de bioinformática utilizadas para a mineração inicial dos dados, foram localizados em *scaffolds* diferentes. O número total de genomas contendo *clusters* putativos completos após a análise manual no BLAST foi de 26 genomas, cada qual com 1 *clusters* putativo inferido cada.

As linhagens do gênero *Butyrivibrio* e da família Lachnospiraceae (representados por 50 e 35 linhagens, respectivamente) foram as que apresentaram maior número de genomas contendo *clusters* gênicos putativos para peptídeos laço. Foram identificados *clusters* completos de peptídeos laço em genomas de 14 linhagens de *Butyrivibrio* e para família de Lachnospiraceae foram observados *clusters* putativos completos em 3 linhagens e *clusters* putativos incompletos em 4 linhagens (Figura 1). *Clusters* completos de biossíntese de peptídeos laço também foram identificados nos genomas bacterianos dos gêneros *Lachnospira*, *Eubacterium*, *Ruminococcus* e *Bacillus*. Enquanto os genomas das linhagens de *Streptococcus*, *Staphylococcus*, *Pseudobutyrovibrio*, *Propionibacterium*, *Enterococcus* e a linhagem pertencente à família Coriobacteriaceae apresentaram apenas *clusters* gênicos incompletos para biossíntese de peptídeos laço (Figura 1).

A distribuição percentual dos genes *lasA*, *lasB*, *lasC* e *lasD* entre os genomas de bactérias do rúmen indicou que 11 gêneros e 2 famílias (Lachnospiraceae e Coriobacteriaceae) apresentaram linhagens com *cluster* completo ou incompleto para peptídeo laço. Os gêneros *Bacillus*, *Butyrivibrio*, *Lachnospira* e *Eubacterium* apresentaram *clusters* completos e proporções homogêneas dos genes *lasA*, *lasB*, *lasC* e *lasD*. Os gêneros *Propionibacterium*, *Pseudobutyrovibrio*, *Staphylococcus*, *Streptococcus* apresentaram apenas *clusters* incompletos, contendo os genes *lasA* e *lasB* e a linhagem da família Coriobacteriaceae apresentou apenas *clusters* incompletos, contendo os genes *lasA* e *lasD*.

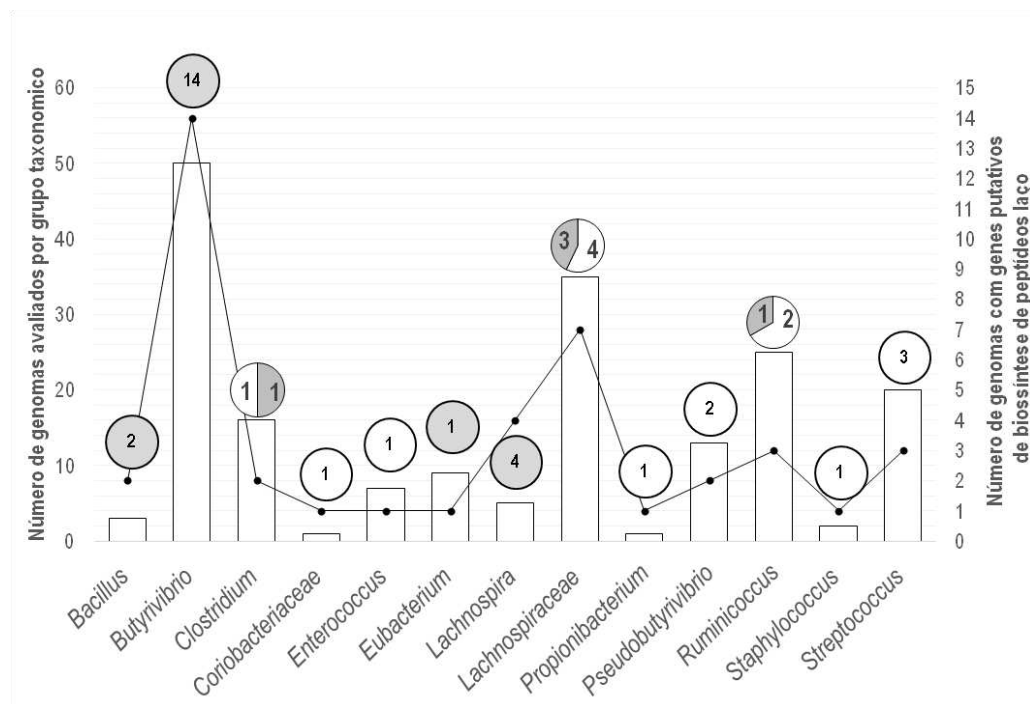
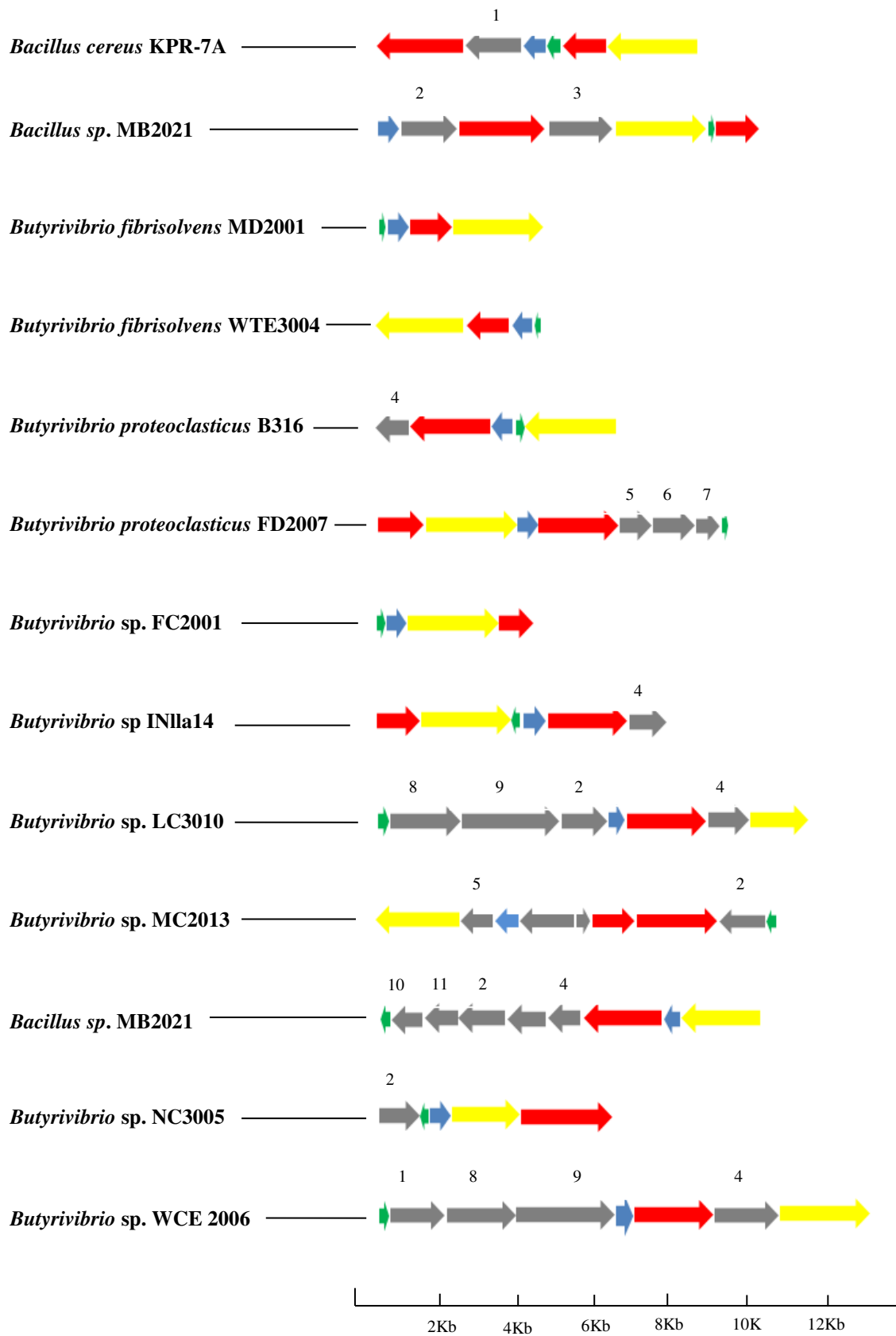


Figura 1: Distribuição de *clusters* gênicos putativos para biossíntese de peptídeos laço em genomas de bactérias ruminais. Barras representam o número total de genomas analisados e linhas representam o número total de genomas onde foram identificados pelo menos um dos genes *lasABCD*. Números dentro de círculos/parcelas cinzas e círculos/parcelas brancas indicam o total de genomas com *cluster* completo e incompleto, respectivamente.

As sequências de aminoácidos preditas a partir dos genes de biossíntese das linhagens com *clusters* putativos completos para peptídeos laço apresentaram tamanhos variados entre os genomas das linhagens bacterianas avaliadas (Tabela 3). Os peptídeos precursores codificados pelos genes *lasA* variaram de 30 aa (*Butyrivibrio* sp. MC2013) a 97 aa (*Bacillus cereus* KPR-7A), enquanto os produtos de *lasB* apresentaram de 117 aa (*Butyrivibrio* sp. WCE 2006) a 161 aa (*Clostridium beijerinckii* HUN142). As proteínas putativas codificadas pelo gene *lasC* variaram de 430 aa (*Butyrivibrio* sp. LC3010) a 669 aa (*Lachnospira mutipara* D15d) e os produtos do gene *lasD* variaram de 219 aa (*Lachnospiraceae bacterium* NKA144) a 615 aa (*Bacillus cereus* KPR-7A).

A análise da organização dos 26 *clusters* putativos completos de biossíntese de peptídeos laço inferidos via mineração genômica revelou vários genes com funções moleculares e biológicas diversas nos *clusters* de biossíntese (Figura 2). Desses, predominaram genes putativos para histidina cinase (HisKA) (38% dos *clusters*); glicosiltransferase (Glycos_transf) (30%); coenzima pirroloquinolina quinona (PqqD) (19%); proteína reguladora da transcrição do carbono C terminal (Trans_reg_C) (15,4%); proteínas de biossíntese de polissacarídeos (Polysacc_synt) (15,4%) e nucleotidil transferase (NTP_transf) (11,54%).



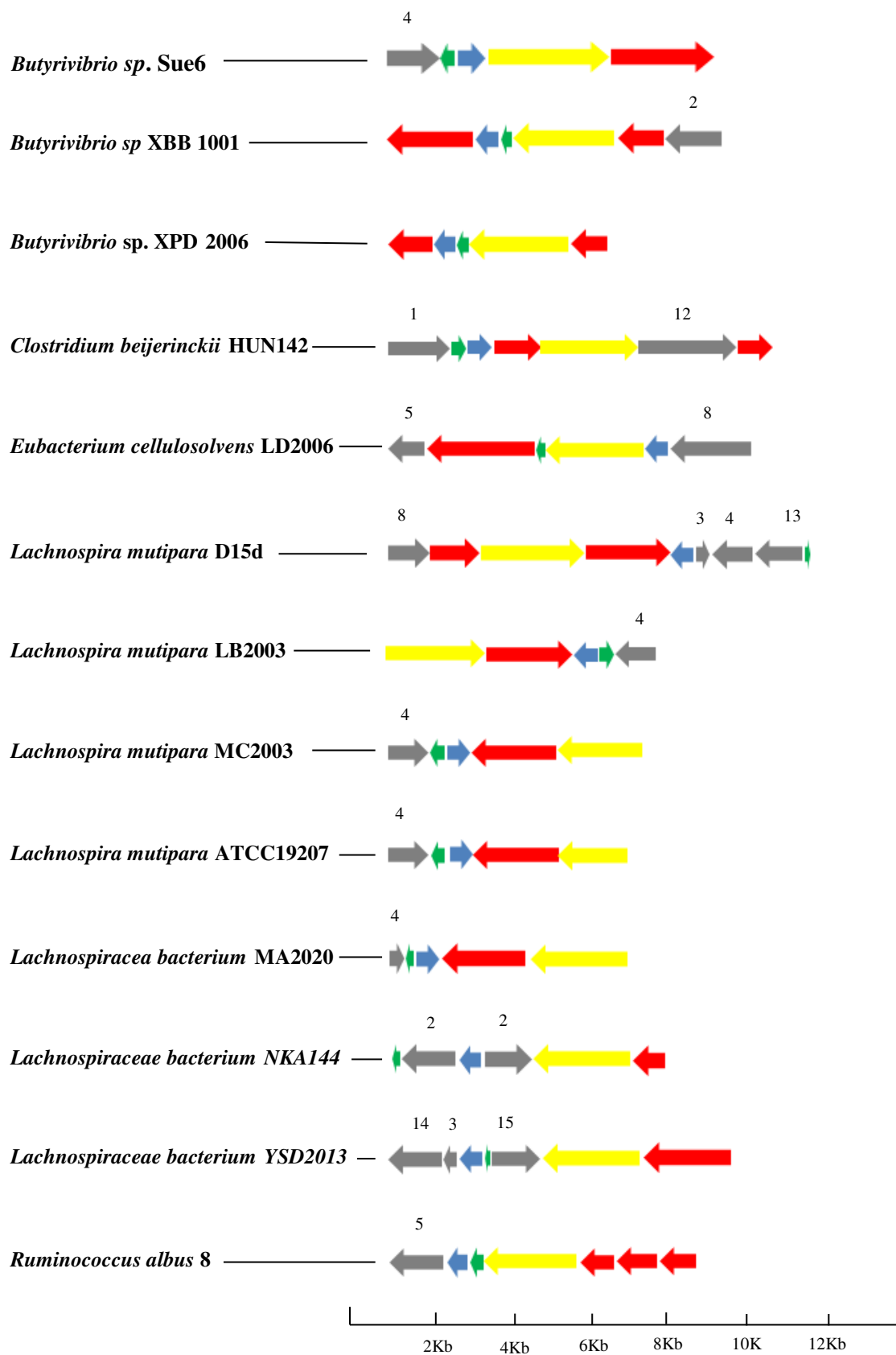


Figura 2: Esquema representativo dos *clusters* putativos completos de peptídeos laço identificados em 26 linhagens de bactérias do rúmen. Os genes envolvidos na biossíntese *lasA*, *lasB*, *lasC* e *lasD* foram representados pelas cores verde, azul, amarelo e vermelho, respectivamente. Os genes presentes nos *clusters*, porém não relacionados à biossíntese de peptídeos laço, foram representados pela cor cinza. Os

produtos dos genes foram representados por números nos *clusters*, genes que não tiveram produtos identificados não foram enumerados. 1 - NTP_transf; 2 - Glyco_transf; 3 - Pqqd; 4 - HisKA; 5- Trans_reg_c; 6 - Spoc; 7 - Sec-C; 8 - Polyysacc_synt_C; 9 - Ps_pyruv_trans; 10 - YebO; 11 - VTC; 12 - Peptidade_S8, 13 - Wave; 14 - Sulfotransfer; 15 - Ras.

Os domínios conservados dos produtos preditos para os genes *lasA*, *lasB*, *lasC* e *lasD* também foram analisadas para os genomas com *clusters* putativos completos (Tabela 3). O produto do gene *lasA* não apresentou domínios conservados para a maioria das linhagens, exceto para *Bacillus cereus* KPR-7A, *Butyrivibrio* sp. XPD 2006, *Clostridium beijerinckii* HUN142, *Lachnospira multipara* LB2003, *Lachnospira multipara* MC2003 *Lachnospira mutipara* ATCC 19207 e *Ruminococcus albus* 8, os quais apresentaram domínios conservados para a proteína PqqD e *Lachnospiraceae bacterium* NKA144 que apresentou domínios conservados para a proteína 4Fe-4S. Para os produtos dos genes *lasB* e *lasC*, foram verificados domínios conservados comuns à todas as linhagens, sendo que o domínio Transglut_core3, com função transglutaminase, foi conservado em *lasB* e o domínio Asn_synthase, com função de biossíntese de asparagina, foi conservado em *lasC*. Para o produto do gene *lasD*, foram identificados domínios conservados entre os genomas analisados para a superfamília dos transportadores ABC e HPr serina cinase.

Tabela 3: Domínios conservados das proteínas codificadas pelos genes putativos de biossíntese de peptídeos laço encontrados em *clusters* completos nos genomas de bactérias do rúmen.

Proteínas	Tamanho (aas)	Domínios conservados	Frequência (%)**
<i>lasA</i>	33-97	PF05402, Coenzyme PQQ synthesis protein D (PqqD)	27
	44	PF00037, 4Fe-4S binding domain	3,84
	21-88	ND*	69,16
<i>lasB</i>	117-161	PF13471, Transglutaminase-like superfamily	100
<i>lasC</i>	430-669	PF00733, Asparagine synthase	100
<i>lasD</i>	219-615	PF00005, ABC transporter	23
	293-576	PF00664, ABC transporter transmembrane region	65
	301-705	PF07475, HPr Serine kinase C-terminal domain	12

*Domínio conservado não verificado **Valor de Frequência (%) se refere ao número de vezes em que um determinado domínio ocorreu entre as 26 linhagens com *clusters* putativos completos para peptídeo laço.

Os peptídeos *core* das linhagens para as quais foram inferidos *clusters* putativos completos para peptídeos laço variaram de 14 a 21 resíduos de aminoácidos e apresentaram massa molecular variando de 1500 a 2491 Da (Tabela 4). Na maioria das linhagens avaliadas, o primeiro resíduo de aminoácido do peptídeo *core* foi a glicina, com exceção de *Bacillus cereus* KPR-7A, *Bacillus* sp. MB2021, *Butyrivibrio*

fibrisolvens WTE3004, *Butyrivibrio proteoclasticus* B316, *Butyrivibrio* sp. LC3010 e *Butyrivibrio* sp. MC2013, que apresentaram serina como primeiro resíduo de aminoácido do peptídeo precursor. Para a linhagem *Clostridium beijerinckii* HUN142, o primeiro resíduo de aminoácido foi à alanina. De acordo com a sequência inferida dos peptídeos *core*, todos os peptídeos laço podem pertencer à classe II, uma vez que a maioria das sequências não apresentam resíduos de cisteína, porém há necessidades de mais estudos sobre a classificação da topologia dessas sequências. As únicas exceções foram verificadas para *Bacillus cereus* KPR-7A, *Butyrivibrio fibrisolvens* MD2001, *Lachnospiraceae bacterium* NKA144, *Lachnospiraceae bacterium* YSD2013, cujas sequências de peptídeos *core* inferidas apresentaram resíduos de cisteína (Tabela 4).

Tabela 4: Características dos peptídeos *core* inferidos em *cluster* putativos completos de biossíntese de peptídeos laço nos genomas de bactérias do rúmen

Genomas	Peptídeo <i>core</i>	AAs (anel)	Nº de cisteína	p I/MM
<i>Bacillus cereus</i> KPR-7A	SKYNIERSDCEEQVLSFLDCL	21(1-9)	2	4,18/24,9 1
<i>Bacillus</i> sp. MB2021	STFDNNSLDNTYPAGTAWGDL	21 (1-9)	0	3,42/22,5 9
<i>Butyrivibrio fibrisolvens</i> MD2001	GFGGIITSDRVLEILKKPFST	21 (1-9)	0	8,59/22,7 8
<i>Butyrivibrio fibrisolvens</i> WTE3004	STCVTDNYFELRIKNLND	18 (1-6)	1	4,56/21,4 5
<i>Butyrivibrio proteoclasticus</i> B316	SGNNVDDTATPGEVIIGNP	19 (1-7)	0	3,49/18,6 9
<i>Butyrivibrio proteoclasticus</i> FD2007	GEAMRREISFPEFISMLSLA	20 (1-7)	0	4,79/22,8 4
<i>Butyrivibrio</i> sp. FC2001	GGIYDTEELDSEKWFDEEHK	20(1-8)	0	4,12/24,2 7
<i>Butyrivibrio</i> sp. IN11a14	GRQVPSEQYAYVSGEQHA	18 (1-6)	0	5,40/20,0 6
<i>Butyrivibrio</i> sp. LC3010	SYRWFLNMDMDKFNEYPM	18 (1-9)	0	4,56/23,8 7
<i>Butyrivibrio</i> sp. MC2013	SSRIISDPSMLYSQSKYQELS	21 (1-7)	0	5,79/24,1 9
<i>Butyrivibrio</i> sp. MC2021	GEGDPNEDEWNVHPQH	18 (1-8)	0	4,17/18,5 9
<i>Butyrivibrio</i> sp. NC3005	GMAPDENFDGDWVEINGKWYR	21 (1-9)	0	4,11/24,9 9
<i>Butyrivibrio</i> sp. Su6	GSGNDIWDLIQEDITVEEIIK	21 (1-8)	0	3,71/23,8 7
<i>Butyrivibrio</i> sp. XBB 1001	GITKRCITDIELDFWRSPS	21(1-9)	0	6,11/23,5 0
<i>Butyrivibrio</i> sp. XPD 2006	GTAVDVWDNIQSEITIAELLE	21 (1-8)	0	3,39/23,1 6
<i>Butyrivibrio</i> sp. WCE2006	GEEMQSKSESDISIITNQTD	21 (1-9)	0	3,71/23,2 8
<i>Clostridium beijerinckii</i> HUN142	ATCERNVLDYLGVLNEELIT	21 (1-9)	0	3,91/23,9 4
<i>Eubacterium cellulosolvens</i>	GQASGTTGETGTPAV	15 (1-9)	0	4,00/13,3

LD2006				4
<i>Lachnospira mutipara</i> LB2003	GAANDIWDIIQEDITVGDIIIN	21 (1-8)	0	3,28/22,8 6
<i>Lachnospira mutipara</i> MC2003	GAANDIWDIIQEDITVGDIIIN	21 (1-8)	0	3,28/22,8 5
<i>Lachnospira mutipara</i> ATCC 19207	GAANEIWDIIQEDTKVEDIVL	21 (1-8)	0	3,71/23,7 1
<i>Lachnospira mutipara D15D</i>	GFVDLDDSGSYVKVE	15 (1-7)	0	3,84/16,2 9
<i>Lachnospiraceae bacterium</i> MA2020	GEWITPDEGKSWWAGTVIVSG	21 (1-7)	0	4,14/22,7 5
<i>Lachnospiraceae bacterium</i> NKA144	GFVYPVIDDDNCIGCGMCIKV	21 (1-8)	2	3,93/22,6 1
<i>Lachnospiraceae bacterium</i> YSD2013	GILGICTIECMVCE	14 (1-9)	3	3,79/15,0 0
<i>Ruminococcus albus</i> 8	GKNTDEIIDAVAEQFNISKNQ	21 (1-6)	0	4,32/23,3 4

Somente as linhagens pertencentes à mesma espécie ou gênero apresentaram alinhamentos entre os genes putativos *lasB* e *lasC* com identidade acima de 30% (Figura 3). O alinhamento das sequências de aminoácidos codificadas pelos genes *lasB* putativos revelou que a identidade entre as sequências variou de 35% para *Lachnospiraceae bacterium* MA2020, *Lachnospiraceae bacterium* NKA144 e *Lachnospiraceae bacterium* YSD2013; a 94% para *Butyrivibrio fibrisolvens* WTE3004 e *Butyrivibrio fibrisolvens* MD2001, (Figura 3a). A identidade dos genes *lasB* de *Bacillus* sp. MB2021 e *Bacillus cereus* KPR-7A foi de 53% e de 58% para as sequências de *Butyrivibrio proteoclasticus* FD2007 e *Butyrivibrio proteoclasticus* B316, respectivamente. Entre *Lachnospira mutipara* D15d, *Lachnospira mutipara* MC2003, *Lachnospira mutipara* MB2003 e *Lachnospira mutipara* ATCC 19207 a identidade foi de 72%.

Quando as sequências de aminoácidos codificadas pelos genes *lasC* putativos foram alinhadas, os seguintes valores de identidade entre linhagens da mesma espécie foram observados (Figura 3b): 99% entre *Bacillus* sp. MB2021 e *Bacillus cereus* KPR-7A; 68% entre *Butyrivibrio fibrisolvens* WTE3004 e *Butyrivibrio fibrisolvens* MD2001; 94% entre *Butyrivibrio proteoclasticus* FD2007 e *Butyrivibrio proteoclasticus* B316; 57% entre *Lachnospira mutipara* D15d, *Lachnospira mutipara* MC2003, *Lachnospira mutipara* MB2003 e *Lachnospira mutipara* ATCC 19207; 30% entre *Lachnospiraceae bacterium* MA2020 *Lachnospiraceae bacterium* NKA144 e *Lachnospiraceae bacterium* YSD2013. Observou-se ainda que o alinhamento dos genes *lasC* putativos das linhagens pertencentes ao gênero *Lachnospira* indicou uma região consenso na porção C-terminal das sequências de aminoácidos (Figura 3b).

A

Bac. sp. MB2021 -----LARGRYLKLLEFKKVPATLGEES 23
 Bac. cereus KPR-7A MNIVKRLRVFLSLNMDTKLLLLLEAFIELGWARVLRKSLTEFSKVAPSLGEY 50

Bac. sp. MB2021 IETPFVILISQDEKVIKDVSEAVQTMSTRYTLWESACLVCATAAATRLNRRN 73
 Bac. cereus KPR-7A NETLVAQIEPHGETLKKVSEAIISTMSRYTFWESQCLVRAIAGMKMLEKRH 100

Bac. sp. MB2021 IETTLYLGIARDTKGIMHAHAWVRSGSLYVTGADAKQNFVVKNFAGGRE 123
 Bac. cereus KPR-7A IESTLYLGTAKDSHGELIAHAWLRSGSEYITGSEGMKFTVVGVSFAKRLS 150

Bac. sp. MB2021 MRGQSFNKTSRH 135

Buty. prot. FD2007 MSTIARGYRFIRYNEHKELTLKAWILSARFRFQMLYEDTAKLNKKWGIEG 50
 Buty. prot. B316 MCTIARGYRFIRYNEHKELTLKAWVLSARFRFQMLYEDTAKLNKKWGIEG 50

Buty. prot. FD2007 EETSSEATLDEYRFCRRVAYAVNQVCNKTFWESKCLVRLTAQRIMAEEG 100
 Buty. prot. B316 EETSEATLDEYRFCRRVAYAVNQVCNKTFWESKCLVRLTAQRIMAEKG 100

Buty. prot. FD2007 IESTMYLGCKELDGKMAHAWIRVGRMYVTGGNGAADGYGVVAKFKSRMK 150
 Buty. prot. B316 IESTMYLGCKELDGKMAHAWIRVGRMYVTGGNGTAEYGVVAKFKSRMK 150

Lach. bact. YSD2013 VVMRNPFTRVYRFFRYNNDKFLSIQCWYYSAYYRFKILHMNTEKLDKNWG 50
 Lach. bact. MA2020 -----MNIRSFLENNKKEITLKAYIYSFYREIVKHTSTKQLQYRMG 43
 Lach. bact. NKA144 -----MIDFYSLFTKNGEKGLTIKAYIYSFYREIVMKVPMKQLQKKEG 44

Lach. bact. YSD2013 ERGVESPATDTEENYTYAWKVSYNVSQVCGKTSWESKCLVRAWTAQTLIA 100
 Lach. bact. MA2020 RRDEESTFDEPDDKIAISKIYAEHVNRIITERLPWECKCFVRALITLRKILM 93
 Lach. bact. NKA144 IRGEESTYDFPEGMGRYVRLCAEHVNRIITEHLPWDAKCFVRALITLRKILN 94

Lach. bact. YSD2013 KKGTHSTLYLGCYDENKFMVAHSWLRCCGEKYITGGNGEGYAMVDKFSK- 149
 Lach. bact. MA2020 EKKISCTIYMGVQK-EDGFMKAHAWLRSGKYYLSGGNGAGFVTVAKFATL 142
 Lach. bact. NKA144 EKNISSTIYLGVME-EEGKLAHAWLRCCSYLITGGNGDGFVVAKER-- 141

Lac. mult. LB2003 MFGRLYQMGKIKKTIKLVFNFFKNHEKFMTIQAVIVSAYYFHCIKKSPM 50
 Lac. mult. MC2003 -----MGKIKKTIKLVFNFFKNHEKFMTIQAVIVSAYYFHCIKKSPM 43
 Lac. mult. ATCC19207 -----MRKIKRLELFCFNFIKFNHEKLMTIQAVIVSAYYFHCYKIPA 43
 Lac. mult. D15d -----MRKIKRLELFCFNFIKFNHEKLMTIQAVIVSAYYFHCYKIPA 43

Lac. mult. LB2003 KKLEIKFGERGKESADTESVENLKIAMHVSDRVARTTNKAPWEMKCLCRA 100
 Lac. mult. MC2003 KKLEIKFGERGKESADTESVENLKIAMHVSDRVARTTNKAPWEMKCLCRA 93
 Lac. mult. ATCC19207 KKLEERFGTRGVESPEKEDLENLKLAMHVGDRVARVTTKAPWETKCLCRA 93
 Lac. mult. D15d KKLEERFGTRGVESPEKEDLENLKLAMHVGDRVARVTTKAPWETKCLCRA 93

Lac. mult. LB2003 WTAQYLLKSRNIHTMYLGVGLDENKMKIAHAWLRCGQLYVTGGDGHEYT 150
 Lac. mult. MC2003 WTAQYLLKSRNIHTMYLGVGLDENKMKIAHAWLRCGQLYVTGGDGHEYT 143
 Lac. mult. ATCC19207 WTAQYLLKRRHVHSTMYLGIKDEETGKMAHAWLRCGQVFTGGDGS DYT 143
 Lac. mult. D15d WTAQYLLKRRHVHSTMYLGIKDEETGKMAHAWLRCGQVFTGGDGS DYT 143

Lac. mult. LB2003 TVACFYK 157
 Lac. mult. MC2003 TVACFYK 150
 Lac. mult. ATCC19207 TVARFYV 150
 Lac. mult. D15d TVARFYV 150

B

Bac. sp. MB2021	MSAIAAGIYNKKNEHVSDEHIHLMMSGFQQFPADDIQVFTMDHIFMGCHAQ	50
Bac. cereus KPR-7A	MSAIAAGIYNKKNEHVSDEHIHLMMSGFQQFPADDIQVFTMDHIFMGCHAQ	50
Bac. sp. MB2021	WITPESIGEKLPYYDSKRQIVITADAIIDNRTELISKLNIDAYDRKEITD	100
Bac. cereus KPR-7A	WITPESIGEKLPYYDSKRQIVITADAIIDNRTELISKLNIDAYDRKEITD	100
Bac. sp. MB2021	SQLILLAYEKWQEETPKYLIGDFAFVIWDKTKQRLFAARDFSGSRTLYYN	150
Bac. cereus KPR-7A	SQLILLAYEKWQEETPKYLIGDFAFVIWDKTKQRLFAARDFSGSRTLYYN	150
Bac. sp. MB2021	DQDGRFSFSTTMEPLLNLNPNVRKELNEEWLVEYLAISSAADVAVVAITPY	200
Bac. cereus KPR-7A	DQDGRFSFSTTMEPLLNLNPNVRKELNEEWLVEYLAISSAADVAVVAITPY	200
Bac. sp. MB2021	KALAQLPSSHLYLMVEGGYKQFRYSQLSQEIQPIRYKNSQDYVDAFMDEVY	250
Bac. cereus KPR-7A	KALAQLPSSHLYLMVEGGYKQFRYSQLSQEIQPIRYKNSQDYV-----Y	244
Bac. sp. MB2021	QTAIEARLRITFKKVGSQLSGGLDSGSVVSIAANKLKEKSNGLTITTSYIIP	300
Bac. cereus KPR-7A	QTAIEARLRITFKKVGSQLSGGLDSGSVVSIAANKLKEKSNGLTITTSYIIP	294
Bac. sp. MB2021	PKDFQPFVPKSLLSNETPFIKETVDFIGGLDNHYLDFAGKDSYSDIDTFL	350
Bac. cereus KPR-7A	PKDFQPFVPKSLLSNETPFIKETVDFIGGLDNHYLDFAGKDSYSDIDTFL	344
Bac. sp. MB2021	DTLEMPYKFFENTFWINGIYEEAKKQDIGILLNGNRGNLSISWGSAQDYF	400
Bac. cereus KPR-7A	DTLEMPYKFFENTFWINGIYEEAKKQDIGILLNGNRGNLSISWGSAQDYF	394
Bac. sp. MB2021	PRILKQLNWLKLYQEITSYSRLIKRPRRRVANMVIKELPPSLLSKNKAI	450
Bac. cereus KPR-7A	---LKQLNWLKLYQEITSYSRLIKRPRRRVANMVIKELPPSLLSKNKAI	441
Bac. sp. MB2021	TKGVSPLVHPAFAEKMSVYDRMRLFGISETGWFLSNNAYEHRTRHFDDLE	500
Bac. cereus KPR-7A	TKGVSPLVHPAFAEKMSVYDRMRLFGISETGWFLSNNAYEHRTRHFDDLE	491
Bac. sp. MB2021	HWNATNTFSGKLSLKYGIISRDPTNDLRVIRYCLSIPEQYVQNGMDRGL	550
Bac. cereus KPR-7A	HWNATNTFSGKLSLKYGIISRDPTNDLRVIRYCLSIPEQYVQNGMDRGL	541
Bac. sp. MB2021	IRRATKGYLPDKVRLNQQVRGVQGTDWVHRMIPQWSSFLQEAQEMLKDKKE	600
Bac. cereus KPR-7A	IRRATKGYLPDKVRLNQQVRGVQGTDWVHRMIPQWSSFLQEAQEMLKDKKE	591
Bac. sp. MB2021	IWSYMNKEAVIKGWNKAKSGVRTDEFAVDADIRILMRSIILFRFMKKMS	648
Bac. cereus KPR-7A	IWSYMNKEAVIKGWNKAKSGVRTDEFAVDADIRILMRSIILFRFMKKMS	639

Buty. fibr. WTE3004 MSAIWGCIDLSCADIDQDIPDRMSKCTYKYRIDHRETFFDNNVYMACGLQ 50
 Buty. fibr. MD2001 MSAIWGCVDLSCETVDEPDIPDRMSECTKYRIDQSQSVLFEENVYMACGLQ 50

Buty. fibr. WTE3004 YINKESRCEKMEFVKEDIYFTADVILDNRQQLISETSATESATDGELLEK 100
 Buty. fibr. MD2001 YINKESRYERLEIIVQDDIYFTADVLLDNRREELIKETLSQESSSDGELLEN 100

Buty. fibr. WTE3004 AWKKWGDDFGDHVGLGFSFAVYDKRKNEFRLYTDHLSNRCIHYYVIGSKV 150
 Buty. fibr. MD2001 AWKRWGEKFGDHVGLGLYSFAIYDKRKNEFRLYTDHLSNRCIHYYVESGSKV 150

Buty. fibr. WTE3004 FFSTLTSSITDIPVPGIEICEKWSASLITDMAYCFLFDGLTPFEGMILP 200
 Buty. fibr. MD2001 YFSTLTSSITDAVSDIDICDKWIRTSLSNMAYCFLFDGLTPFEGIKILP 200

Buty. fibr. WTE3004 YGTGVVFKNLSGNVSVNNEFRYYDPLSHVEQMKEDDEKCRELFRTFEIKC 250
 Buty. fibr. MD2001 YGTGIVVKKNERGIEVEKVRYYDPLSKTEENKKYDDEKCKEEFRTFEIKC 250

Buty. fibr. WTE3004 VRDVVRTDGCIGILLSGGLDSTVGSIAAEYLSKDNKNLYSYTSVPIKLE 300
 Buty. fibr. MD2001 VRDAIRTDGDCIGILLSGGLDSTVGSVAARKYLEKENKNLYSYTSVPIKRY 300

Buty. fibr. WTE3004 VDEFGSTNGYVYVDETKDVMDFCKYYPNIKESFIACEGENLFTHVDEACD 350
 Buty. fibr. MD2001 VDENGRTNGYKLEDESKDVMDFCNHFPNIKEEFLSCDGENCLTHVDELCD 350

Buty. fibr. WTE3004 YLELPYKALINEVMMHNSYAKAHEQGVKILLSGASGNDTISYGMMENTLA 400
 Buty. fibr. MD2001 YLELPFKSLINVWLRDSFIMARKQGVKIVILSGAYGNDTISFGYMENTLA 400

Buty. fibr. WTE3004 RLIKGFHFVFCYKQFKTYTDKERIICKRYKRFLEKLFVVRERENIYCDN 450
 Buty. fibr. MD2001 RLIKGFHFIEAYKQEMTFIKKEGIIKKRYSKRFLEKFISSFRKQEVIFDE 450

Buty. fibr. WTE3004 RGYKKEIIDRFNSKSEWEKVLQGTAGFIRTNKGYGROLFATGILQLIGII 500
 Buty. fibr. MD2001 RGYKKEYLNKYDIESEWDKLFRTGNTNIRTRDAYQNSVFATGILQLIGIV 500

Buty. fibr. WTE3004 NTRKGLYFTVMMRDPTMDKRTLELCLTLPYKFAWNGVERRLIREYLDY 550
 Buty. fibr. MD2001 NTRKGLYHGVMRDPSMDKRMLELCLSLPYKFAWNGVERRLVREYLDY 550

Buty. fibr. WTE3004 VPDEIRLNTRFRGRQSADAVLRLNREFGISDEKKLSDAVSDKLGKYYDLER 600
 Buty. fibr. MD2001 VPDEIRLNTRFRGRQSADAVFRLLLYGVKKVKKLSELVSDSLGKYYDLNL 600

Buty. fibr. WTE3004 VKEMDEPNKDENLDWKAKIVSCSIFLDKYAK----- 632
 Buty. fibr. MD2001 VKEMDEPNDEDNLDWKAKIVSFGVFLSKYSAKKEGD 637

Buty. prot. FD2007 MSAIWGYIKLKKEDGESSSDQIKEKCLMKKPYEECVIDRFEDTFDDGFF 50
 Buty. prot. B316 MSAIWGYINLKNNTSEDY---TKLSELMKKPYDECVIERIGIEQFENGFF 47

Buty. prot. FD2007 SCGIQYFNTRARNEKLPVYDESEITVFTADIVLNGRIILIKELKACGQTG 100
 Buty. prot. B316 ACGLQCFNKRSHNERLPIYDSESCIVFTADVVLNDRKKLIEELTWAGVGG 97

Buty. prot. FD2007 IDENTPDGELAFRAWRIWKODEVDHIHGLFAIAIYEKRTKEFYLFTHMG 150
 Buty. prot. B316 LNEETPDGELSEFYAWKRWGHEFTDHIIGLFAIAVYDKKDNSFYLFTHMG 147

Buty. prot. FD2007 TRCVDYSVCGDALYFSMLTKSILSVMPDEYKFIIDEKEMTGCEASSSPYMY 200
 Buty. prot. B316 SRCVDYSICGNELFFSTLSKPIINVMPEYKFIIDEQETVGCESLTPYMY 197

Buty. prot. FD2007 VFPTGTPFENIRHTFRGNYVKAQDKKGNIEQTEYFNPGKKNDRPKTD 250
 Buty. prot. B316 VFPDRTPFRNVYHSVRGNYVKAEPGKDKWYRSVVYMPGITR--KLDLV 245

Buty. prot. FD2007 WPKDEPDRVYREKFRFTFFECVEDAMDTDGEIAAMISSGLDSSSVATVAA 300
 Buty. prot. B316 WPKNPKDSAYREAFRDVFFKCVEDAMDTDGEVAATISSGLDSSSVATVAA 295

Buty. prot. FD2007 RMLAKDNRNLYGFTSTPMKDFEFTGGFEKTDSESEGVRIIAGYPNIHE 350
 Buty. prot. B316 RSLQKGRKRLYGFTSVPIKGFDRSSTKWEITDESSGVRNIIAGYPNIEQE 345

Buty. prot. FD2007 FCSCEGLSPFTEMNRLIHMYEVPKACINQVWFYIVRSAAQKKCKVLLN 400
 Buty. prot. B316 FCDQYGMSPFTEMDELVHMFEIPKAFVNQVWMKYIIKTASQKCKVLLT 395

Buty. prot. FD2007 GQFGNFTISKGSMELEFFQKLLSGQILEARNQLARFGKRGVPRKVLFDG 450
 Buty. prot. B316 GEFGNFTISRGEIREYFFQKFLSGHVLEAKQLGIYCGKCGIPRKALFQG 445

Buty. prot. FD2007 VISELGECSFHFGANKRFSAVIDKKLIKRDIVKKYGIKKLIKRRRFYFG 500
 Buty. prot. B316 MIDEIVSEVTFKLGINNAYKKIIDEDELIKPAISKYRINKVLIKRRRFYFG 495

Buty. prot. FD2007 FTSVLPVRKMFERNLGDVSTITQQLYDTKESLYYGVLEFRDPTDRKRMVEF 550
 Buty. prot. B316 FTSVLSPKMYRNLDINILQQLYDTKESLYYGVLEFRDPTDKRRLLEL 545

Buty. prot. FD2007 ISLSEPECFVYDGLERRLVREYLDLDPYDRLOIRYKGRQSDQVMRCQ 600
 Buty. prot. B316 CASFSPDQYVFNGLERRLVREYLDLDPDERLQYKYGRQSDQVMRRLS 595

Buty. prot. FD2007 KFGDDHKQIALDNRIFDYMVEDRVREVMQDIAADNVADIVRIDALNTFL 650
 Buty. prot. B316 SFGEDKRQSKINEAVYGFEMNRDQVEALMKEDINEKNALNVVRIEALGRFF 645

Buty. prot. FD2007 FEYA----- 654
 Buty. prot. B316 DEESEVINI 654

Lac. mult. ATCC19207 ----- 0
Lac. mult. D15d MAAISGVIKTTGGLEDLEAIKSAMTKSLQGYKIDKINAQTNPFFVVSGE 50
Lac. mult. LB2003 MGAISGIINTNFDLENRDDYCKKITSTLAPYKIDSILTNETNNAFFVYAN 50
Lac. mult. MC2003 ----- 0

Lac. mult. ATCC19207 ----- 0
Lac. mult. D15d QYFTKEDHYSISPYDEKRQIVFATDCMIDNREELIPLVKKNLVTDMSEF 100
Lac. mult. LB2003 QVFTKEDNYNIGAYYHKETNNFMIADCMIDNRDSLFLKIDPSYL----- 94
Lac. mult. MC2003 -----MIADCMIDNRDSLFLKIDPSYL----- 22

Lac. mult. ATCC19207 ----- 0
Lac. mult. D15d FKEDVADSMSDSFKKACEKISDQELIYAAYLSFGPDFPKKVLGIFSFVAVY 150
Lac. mult. LB2003 -----HNKPKAQVSDQELLYAAYLSFGTKFTDYVCGIFSFVIY 132
Lac. mult. MC2003 -----HNKPKAQVSDQELLYAAYLSFGTKFTDYVCGIFSFVIY 60

Lac. mult. ATCC19207 ----- 0
Lac. mult. D15d HLDSKEFYLYTDHTSTRTIHYALDGDYIYFGLTLLNSVKASTKETTCKSYK 200
Lac. mult. LB2003 NDKKKECLIFTDHTSTRTIHYMIDNGIYFATLFETIKK----CVNKPLE 178
Lac. mult. MC2003 NDKKKECLIFTDHTSTRTIHYMIDNGIYFATLFETIKK----CVNKPLE 106

Lac. mult. ATCC19207 -----MITYKVNLSPEYEGIYTV EAGHYIFIKNGKIKSHSY 34
Lac. mult. D15d LNEEWISLASALANPDMITYKVNLSPEYEGIYTV EAGHYIFIKNGKIKSHSY 250
Lac. mult. LB2003 LNEEWLSFADALPNPDLIYRVEQSEYIIGIYTV EAGHYVSIKKNNEIKAHCY 228
Lac. mult. MC2003 LNEEWLSFADALPNPDLIYRVEQSEYIIGIYTV EAGHYVSIKKNNEIKAHCY 156

Lac. mult. ATCC19207 YNPKKAKENKLYKYSDEDWREAFVSTLNTCVKSVLRSEKNTGCTLSSGL 84
Lac. mult. D15d YNPKKAKENKLYKYSDEDWREAFVSTLNTCVKSVLRSEKNTGCTLSSGL 300
Lac. mult. LB2003 YNPSKVKL--HLKTKTDEEYKNEFINLLDTCTTSLLRSNKNTGCTLSSGL 276
Lac. mult. MC2003 YNPSKVKL--HLKTKTDEEYKNEFINLLDTCTTSLLRSNKNTGCTLSSGL 204

Lac. mult. ATCC19207 DSSSVACLALKEIKKEDKKLYSYTSVPLKDYLASA--DSDRYAIDDESFG 132
Lac. mult. D15d DSSSVACLALKEIKKEDKKLYSYTSVPLKDYLASA--DSDRYAIDDESFG 348
Lac. mult. LB2003 DSSSVACLAAARALEKRGETLHSHYTSVPLKDFNAATLNTKGVYEIEDESFG 326
Lac. mult. MC2003 DSSSVACLAAARALEKRGETLHSHYTSVPLKDFNAATLNTKGVYEIEDESFG 254

Lac. mult. ATCC19207 EKIIRKYYDNIEFNFVDCKGKTAFSDLENTIKITELPHKSLVNYTWMSEI 182
Lac. mult. D15d EKIIRKYYDNIEFNFVDCKGKTAFSDLENTIKITELPHKSLVNYTWMSEI 398
Lac. mult. LB2003 EIAIRKYYDNIEFNFVDCKGKTAFSGMRNTIKINELPHKSLVNYTWMSEI 376
Lac. mult. MC2003 EIAIRKYYDNIEFNFVDCKGKTAFSGMRNTIKINELPHKSLVNYTWMSEI 304

Lac. mult. ATCC19207 YEELASNDCKVMLKGQYGNSTISNGKILSRVYQDLRKCKFATALKEAKAF 232
Lac. mult. D15d YEELASNDCKVMLKGQYGNSTISNGKILSRVYQDLRKCKFATALKEAKAF 448
Lac. mult. LB2003 YDLASNDCKVMLKGQYGNSTISNGKILSRVYQELLSFKFISAIKLSKEF 426
Lac. mult. MC2003 YDLASNDCKVMLKGQYGNSTISNGKILSRVYQELLSFKFISAIKLSKEF 354

Lac. mult. ATCC19207 CNYRKVSKKYLKTFIKTFMKNKFE LDISAIDEHEL RADLIEKYKIKQLIK 282
Lac. mult. D15d CNYRKVSKKYLKTFIKTFMKNKFE LDISAIDEHEL RADLIEKYKIKQLIK 498
Lac. mult. LB2003 IRVNKIPKKLFTKNFINSYKAKFKL DLSELDFYLRKDLIDKYHIRKELK 476
Lac. mult. MC2003 IRVNKIPKKLFTKNFINSYKAKFKL DLSELDFYLRKDLIDKYHIRKELK 404

Lac. mult. ATCC19207 ENFTREGGSYNDSRKQFERYVYDTRILQHLGLFDTRL SLIYGLVIRDPAK 332
Lac. mult. D15d ENFTREGGSYNDSRKQFERYVYDTRILQHLGLFDTRL SLIYGLVIRDPAK 548
Lac. mult. LB2003 KNFTKEGGTYIDSKAQFERVYVYDTRILQHLGLFDTRL SLIYGVVIRDPAK 526
Lac. mult. MC2003 KNFTKEGGTYIDSKAQFERVYVYDTRILQHLGLFDTRL SLIYGVVIRDPAK 454

Lac. mult. ATCC19207 DKRITDFCFR ADESCFVKNGVERRMVRDYMGRGIVPDEILDIVNRRGLQSS 382
Lac. mult. D15d DKRITDFCFR ADESCFVKNGVERRMVRDYMGRGIVPDEILDIVNRRGLQSS 598
Lac. mult. LB2003 DKRMIDFVFNAPNSCMVHNGVERRMIREYMKGIIPDEILEVTNRRGRQSS 576
Lac. mult. MC2003 DKRMIDFVFNAPNSCMVHNGVERRMIREYMKGIIPDEILEVTNRRGRQSS 504

Lac. mult. ATCC19207 DYAYRVNQVWKHELKSKVLSLGNPHELLEYFDKKEKIENBKKELEAREQIN 432
Lac. mult. D15d DYAYRVNQVWKHELKSKVLSLGNPHELLEYFDKKEKIENBKKELEAREQIN 648
Lac. mult. LB2003 DYAYRVNNAFNLGIKKEILTALENKPKLYDYLDASKIE LLKKRIENSNKLD 626
Lac. mult. MC2003 DYAYRVNNAFNLGIKKEILTALENKPKLYDYLDASKIE LLKKRIENSNKLE 554

Lac. mult. ATCC19207 SDQIQETLCLVSLSEFGNLDVN--- 453
Lac. mult. D15d SDQIQETLCLVSLSEFGNLDVN--- 669
Lac. mult. LB2003 EDEIQNILCICSLSVFLQIHYGNC 650
Lac. mult. MC2003 EDEIQNILCICSLSVFLQIHYGNC 578

```

Lach. bact. MA2020 MSAIWGCIIDFSGSIIIDSGIIPDIMECTKLYKIDRTEHLIENNVYMACGLQ 50
Lach. bact. NKA144 MSAIWGCIIDFSGSIIIDSGIIPDIMECTKLYKIDRTEHLIENNVYMACGLQ 50
Lach. bact. YSD2013 MSAIWGCIIDFSGKEITEKIKSVLKNAFDECAIDKTSIENKKNVYMGCGLQ 50

Lach. bact. MA2020 YITKESRYEKLPCKIKDGLYFSADCIIDNRQELIDSLGIDTDA--ADGEVL 98
Lach. bact. NKA144 YITKESRYEKLPCKIKDGLYFSADCIIDNRQELIDSLGIDTDA--ADGEVL 98
Lach. bact. YSD2013 EFTFESKHEELPTKKGDIYYTADAVLDNRDEMCARLGIINTPELIGDAEII 100

Lach. bact. MA2020 FSAWQKWGEIDFGNYVIGLFSFAVYDRKKSQFYLYTDHTSSRCIEYFVKGS 148
Lach. bact. NKA144 FSAWQKWGEIDFGNYVIGLFSFAVYDRKKSQFYLYTDHTSSRCIEYFVKGS 148
Lach. bact. YSD2013 RQCFEREGKIDCLNDLLGAYVFRFDDAAKGTVDLIVADAVGNRYIICYFCEDD 150

Lach. bact. MA2020 KVVYFATITKCIITDAAGGIDISEKWLITACMSTNFPSMLLYEGLSPFENVFI 198
Lach. bact. NKA144 KVVYFATITKCIITDAAGGIDISEKWLITACMSTNFPSMLLYEGLSPFENVFI 198
Lach. bact. YSD2013 ILYFASIRAPLESIKKEVRLNKKWIAYYYGVNTLDTFNDSESTIIDGIYR 200

Lach. bact. MA2020 LPYGCQVKIEYIDNSIVVEKKRYWDPINVTVTQIKKFDNMCNCRKMFREVHN 248
Lach. bact. NKA144 LPYGCQVKIEYIDNSIVVEKKRYWDPINVTVTQIKKFDNMCNCRKMFREVHN 248
Lach. bact. YSD2013 IAP---ASHVVVIONGRVAEKNEYWDTNINIKRCRRHKNDSDYRKEFVDLYR 247

Lach. bact. MA2020 KCVQDAIRTDGNIIGILLSSGLDSSAIGCTAARMNLKRLKLYSYTSTPLK 298
Lach. bact. NKA144 KCVQDAIRTDGNIIGILLSSGLDSSAIGCTAARMNLKRLKLYSYTSTPLK 298
Lach. bact. YSD2013 KCVSDTLRTEEEVAIEILSGGYDSTSVACTAAPILKERGKLYSFTSVPE 297

Lach. bact. MA2020 EEDGNSDNKQKFWIIDDESEGKVICNHYKNIKPKFLECRGKSVLSYVEEF 348
Lach. bact. NKA144 EEDGNSDNKQKFWIIDDESEGKVICNHYKNIKPKFLECRGKSVLSYVEEF 348
Lach. bact. YSD2013 DYVFDGPEETK---ERDEYESVQRNMEFLGNVECDTILSKGVNLWDIRHEY 344

Lach. bact. MA2020 VKYLEEPGKALVNHVVMKDAYEMAREIDGRIVMYGSYGNMSISYGSIFEN 398
Lach. bact. NKA144 VKYLEEPGKALVNHVVMKDAYEMAREIDGRIVMYGSYGNMSISYGSIFEN 398
Lach. bact. YSD2013 NKVVPLEYKSPENMIWLYEGYKAREKGRILILSGAFNGTVSYNNILQY 394

Lach. bact. MA2020 TWLMIKCGDIIIGAVKQIFRLABKEHLSKKNLLKMFVKKLFNNVDYTDSSH 448
Lach. bact. NKA144 TWLMIKCGDIIIGAVKQIFRLABKEHLSKKNLLKMFVKKLFNNVDYTDSSH 448
Lach. bact. YSD2013 LVWLVKHFKFKKFIIEVNGFNRRLGSSSRKQIIKATLDGVFKPSSL---PST 441

Lach. bact. MA2020 IDDELFSKSLVVKYINIRKEWENRLKREGIE-----KMTLSQR 485
Lach. bact. NKA144 IDDELFSKSLVVKYINIRKEWENRLKREGIE-----KMTLSQR 485
Lach. bact. YSD2013 IK-----FDAYKDCYAYESFLKENGAI DAVKENDMITRKSYPHNAKH 483

Lach. bact. MA2020 RNALVSGILFQLVGLIYETKDGLYNGILIRDPKDKRMIELCRLRPFKCF 535
Lach. bact. NKA144 RNALVSGILFQLVGLIYETKDGLYNGILIRDPKDKRMIELCRLRPFKCF 535
Lach. bact. YSD2013 QDMFLPAINFRHYGEFPLKNSLYIGVFFRDPTRDKRMIEYVKSLEPDQFT 533

Lach. bact. MA2020 WNGVERRLVRDYLDDVVPDEIRL INNRGRQSGDAVIRYDMYGLPGEKT 585
Lach. bact. NKA144 WNGVERRLVRDYLDDVVPDEIRL INNRGRQSGDAVIRYDMYGLPGEKT 585
Lach. bact. YSD2013 HNGYTRRLIREYMEDIMPPKFF-IRHPFGVQSADMKYREFEYKGDVAVL-EE 581

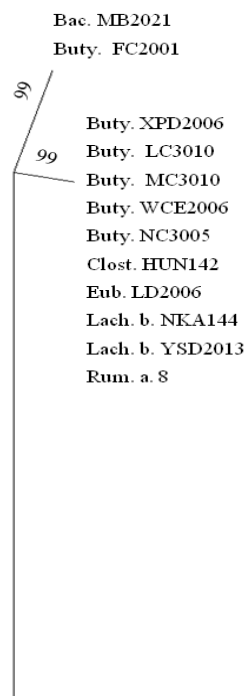
Lach. bact. MA2020 WDVLIIGRMKSYFNIDKSIELLKSKTDAANLINKAKIMS--CNYFLKEFGK 633
Lach. bact. NKA144 WDVLIIGRMKSYFNIDKSIELLKSKTDAANLINKAKIMS--CNYFLKEFGK 633
Lach. bact. YSD2013 WRRIVSKTD-----GKGI VDTARLLKDIERKNVVKDFGY 614

Lach. bact. MA2020 DDLV----- 637
Lach. bact. NKA144 DDLV----- 637
Lach. bact. YSD2013 GDIIIRLFYTTINALEYISNYEL 635

```

Figura 3: Alinhamentos múltiplos e regiões de similaridade gerados pelas ferramentas de bioinformática Clustal Omega e Color Align Conservation das sequências de aminoácidos codificadas pelos genes *lasB* e *lasC* identificados em linhagens de bactérias do rúmen da mesma espécie com *cluster* putativo completo para biossíntese de peptídeo laço. (A) Alinhamento de sequências de aminoácidos codificadas pelo gene *lasB* indicando as regiões consenso. (B) Alinhamento das sequências de aminoácidos codificadas pelo gene *lasC*, indicando regiões consenso. Fundo preto indica regiões do consenso; Fundo cinza claro, similaridade de sequência de aminoácidos superior a 50%; Fundo cinza escuro, similaridade de sequência de aminoácidos superior a 75%; Pontos e lacunas foram introduzidos para melhorar o alinhamento das sequências.

Para avaliar se a presença dos genes de biossíntese de peptídeos laço retratavam evidência de evolução ou transferência horizontal de genes entre linhagens do mesmo gênero ou espécie, foram reconstruídas árvores filogenéticas baseadas em: (1) genes de biossíntese para peptídeos laço (*lasA*, *lasB*, *lasC* e *lasD*) das linhagens que apresentaram *clusters* completos (Figuras 6 a 9); (2) sequências parciais do gene que codifica o rRNA16S, totalizando 168 genomas de linhagens bacterianas (Figura 8). Somente valores de *bootstrap* acima de 50% foram reportados nas árvores filogenéticas. Valores de *bootstrap* acima de 90% foram considerados altos, enquanto *bootstrap* entre 80% e 90% foram definidos como medianos e baixos quando inferiores a 80% (Efron & Tibshirani, 1993). A Tabela 5 indica os códigos de identificação referentes a cada linhagem apresentada nas árvores filogenéticas.



1

Figura 4: Relações filogenéticas dos genes *lasA* de genomas de bactérias do rúmen que apresentaram *clusters* putativos completos para peptídeos laço. A árvore filogenética, baseada na análise *Maximum Likelihood*, foi construída com 1000 réplicas de *bootstrap* usando o programa MEGA 7.0. A reconstrução filogenética foi baseada em sequências de aminoácidos codificados pelos genes *lasA* inferidos. Os valores mostrados na árvore representam a percentagem de *bootstrap*. Somente os ramos com valores de *bootstrap* > 50% são destacados. A barra de escala representa 100 alterações de nucleotídeos por 100 nucleotídeos analisados.

Tabela 5: Códigos atribuídos aos genomas de bactérias ruminais com *clusters* putativos completos de biossíntese de peptídeos laço utilizados na reconstrução da filogenia dos genes *lasABCD*.

Linhagens	Códigos
<i>Bacillus cereus</i> KPR-7A	Bac. KPR-7A
<i>Bacillus</i> sp. MB2021	Bac. MB2021
<i>Butyrivibrio fibrisolvens</i> MD2001	Buty. f. MD2001
<i>Butyrivibrio fibrisolvens</i> WTE3004	Buty. f. WTE3004
<i>Butyrivibrio proteoclasticus</i> B316	Buty. p. B316
<i>Butyrivibrio proteoclasticus</i> FD2007	Buty. p. FD2007
<i>Butyrivibrio</i> sp. IN11a14	Buty. IN11a14
<i>Butyrivibrio</i> sp. FC2001	Buty. FC2001
<i>Butyrivibrio</i> sp. LC3010	Buty. LC3010
<i>Butyrivibrio</i> sp. MC2013	Buty. MC2013
<i>Butyrivibrio</i> sp. MC2021	Buty. MC2021
<i>Butyrivibrio</i> sp. NC3005	Buty. NC3005
<i>Butyrivibrio</i> sp. Su6	Buty. Sue6
<i>Butyrivibrio</i> sp. WCE 2006	Buty. WCE2006
<i>Butyrivibrio</i> sp. XBB 1001	Buty. XBB1001
<i>Butyrivibrio</i> sp. XPD 2006	Buty. XPD2006
<i>Clostridium beijerinckii</i> HUN142	Clost. HUN142
<i>Eubacterium cellulosolvens</i> LD2006	Eub. LD2006
<i>Lachnospira multipara</i> LB2003	Lac. m. LB2003
<i>Lachnospira multipara</i> MC2003	Lac. m. MC2003
<i>Lachnospira mutipara</i> ATCC 19207	Lac. m. ATCC19207
<i>Lachnospira mutipara</i> D15d	Lac. m. D15d
<i>Lachnospiraceae bacterium</i> MA2020	Lach. b. MA2020
<i>Lachnospiraceae bacterium</i> NKA144	Lach. b. NKA144
<i>Lachnospiraceae bacterium</i> YSD2013	Lach. b. YSD2013
<i>Ruminococcus albus</i> 8	Rum. a. 8

A árvore filogenética representativa do gene *lasA* agrupou as linhagens bacterianas em dois clados com valores de *bootstrap* igual a 99% (Figura 4). Um dos clados foi formado pelas linhagens *Bacillus* sp. MB2021 e *Butyrivibrio* sp. FC2001 e outro clado agrupou dez linhagens dos gêneros *Butyrivibrio*, *Clostridium*, *Eubacterium* e *Ruminococcus* e duas linhagens da família Lachnospiraceae.

As árvores filogenéticas dos genes *lasB* (Figura 5) e *lasC* (Figura 6) apresentaram clados, com valores de *bootstrap* acima de 50%, que agruparam linhagens da mesma espécie, apontando portanto para uma hipótese evolutiva de distribuição desses genes. A análise filogenética de *lasB* agrupou no mesmo clado *Butyrivibrio fibrisolvens* MD2001 (Buty. f. MD2001) e *Butyrivibrio fibrisolvens* WTE3004 (Buty. f. WTE3004), espécies relacionadas com a degradação de celulose, hemicelulose, pectina, amido e açúcares no rúmen. As linhagens *Butyrivibrio proteoclasticus* B316 (Buty. f. B316) e *Butyrivibrio proteoclasticus* FD2007 (Buty. f. FD2007) também foram agrupadas no mesmo clado e no rúmen são responsáveis, especialmente, pela biohidrogenação de ácidos graxos insaturados.

A análise filogenética do gene *lasC* agrupou *Bacillus cereus* KPR-7A (Bac. KPR-7A) e *Bacillus* sp. MB2021 (Bac. MB2021), duas linhagens anaeróbias facultativas e degradadoras de amido no rúmen. De modo similar, bactérias com atividade proteolítica também foram agrupadas em clados comuns como *Lachnospira multipara* LB2003 (Lac. m. LB2003) e *Lachnospira multipara* MC2003 (Lac. m. MC2003); *Lachnospira mutipara* ATCC 19207 (Lac. m. ATCC 19207) e *Lachnospira mutipara* D15d (Lac. m. D15d) (Figura 5).

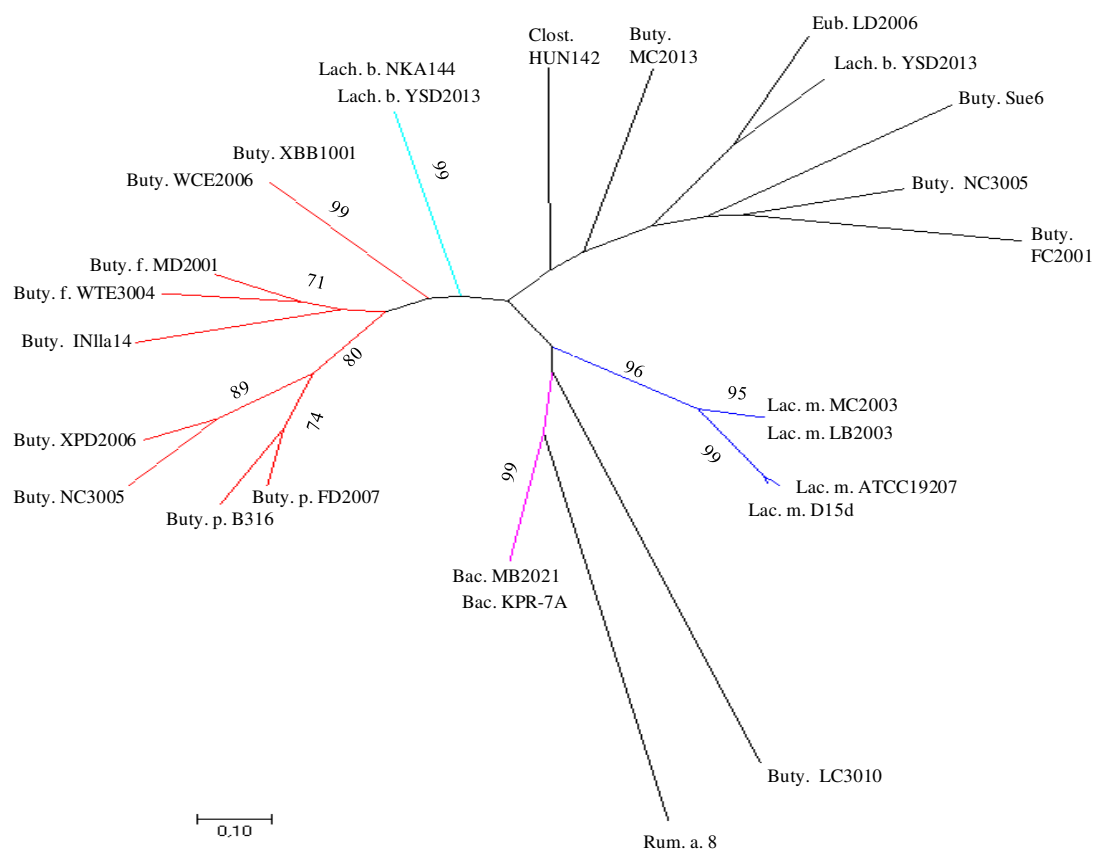


Figura 6: Relações filogenéticas dos genes *lasC* de genomas de bactérias do rúmen que apresentaram clusters putativos completos para peptídeos laço. A árvore filogenética, baseada na análise *Maximum Likelihood*, foi construída com 1000 réplicas de *bootstrap* usando o programa MEGA 7.0. A reconstrução da árvore foi baseada em sequências de aminoácidos codificados pelos genes *lasC*. Os valores mostrados na árvore são dados em percentagem de *bootstrap*. Somente os ramos com valores de *bootstrap* > 50% das linhagens pertencentes à mesma espécie ou gênero são destacados. A barra de escala representa 10 alterações de nucleotídeos por 100 nucleotídeos analisados. Clados coloridos indicam agrupamentos de linhagens bacterianas da mesma espécie ou gênero com valores de *bootstrap* > 70%.

As sequências de aminoácidos previstas para o produto do gene *lasD* foram menos conservadas em relação aos outros genes de biossíntese de peptídeos laço (*lasB* e *lasC*), resultando em árvores filogenéticas com agrupamentos suportados por valores de *bootstrap* baixos (Figura 7). Ainda assim, foi possível observar clados com valores de *bootstrap* acima de 90% reunindo linhagens da mesma espécie (*Butyrivibrio fibrisolvens*) cujas funções no rúmen envolvem a degradação de celulose, hemicelulose, pectina, amido e açúcares. Outros agrupamentos, porém com valores de *bootstrap* inferiores a 90%, ocorreram entre linhagens das mesmas espécies com funções no rúmen relacionadas à biohidrogenação de ácidos graxos insaturados (*Butyrivibrio proteoclasticus*) e proteolítica (*Lachnospira multipara*) (Figura 7).

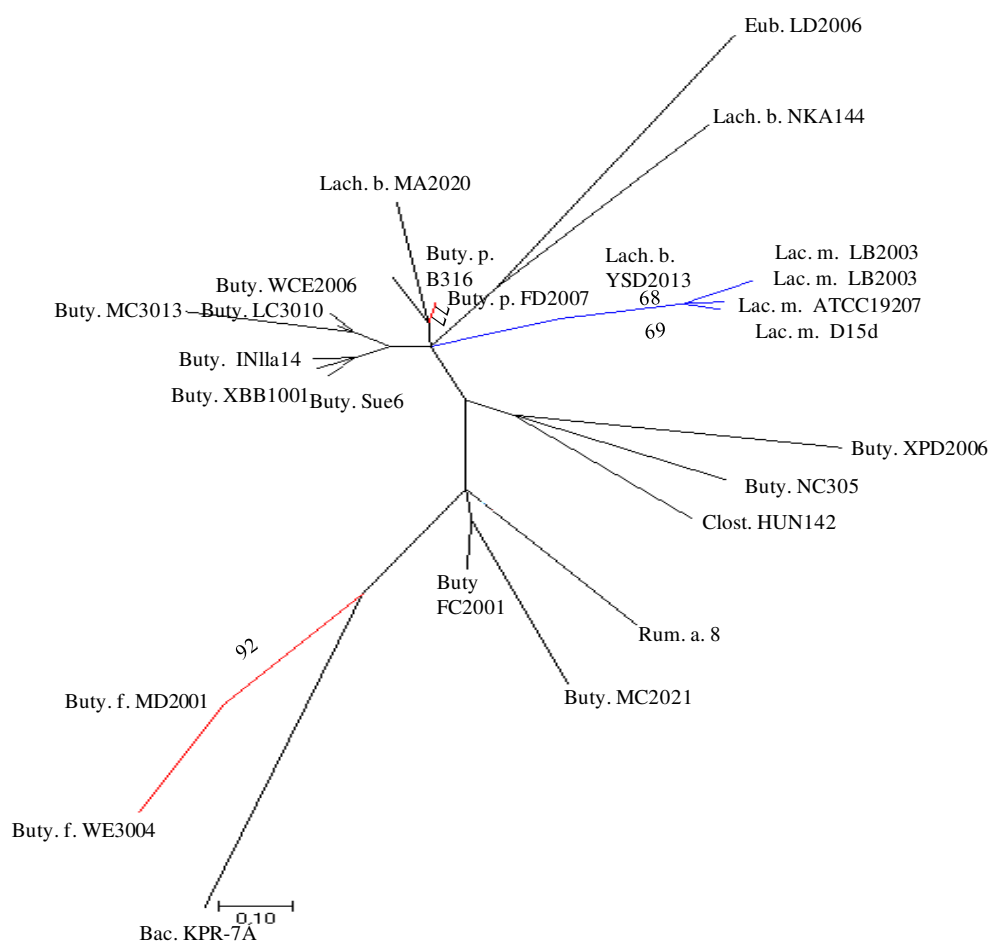
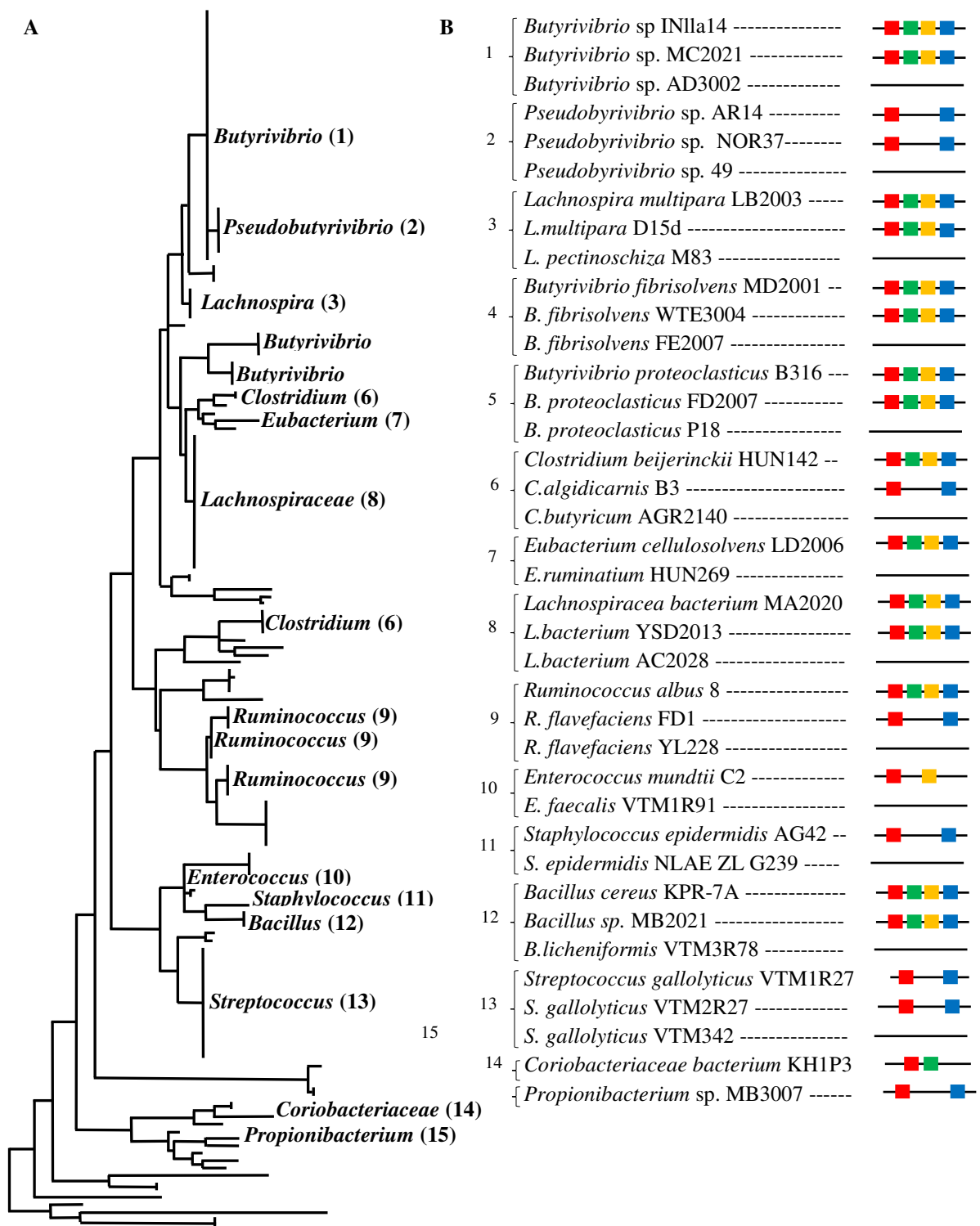


Figura 7: Relações filogenéticas dos genes *lasD* dos genomas de bactérias do rúmen que apresentaram *clusters* putativos completos para peptídeos laço. A árvore filogenética, baseada na análise *Maximum Likelihood*, foi construída com 1000 réplicas de *bootstrap* usando o programa MEGA 7.0. A reconstrução filogenética foi baseada em sequências de aminoácidos codificados pelos genes *lasD* inferidos. Os valores mostrados na árvore representam a percentagem de *bootstrap*. Somente os ramos com valores de *bootstrap* > 50% das linhagens pertencentes à mesma espécie ou gênero são destacados. A barra de escala representa 10 alterações de nucleotídeos por 100 nucleotídeos analisados. Clados coloridos indicam agrupamento de linhagens bacterianas da mesma espécie ou gênero, com valores de *bootstrap* > 60%.

Dos 350 genomas de linhagens de bactérias utilizados nas análises deste trabalho, apenas 168 apresentavam as sequências do gene que codifica o rRNA 16S disponíveis no *GenBank*. Desse total, 42 genomas de bactérias do rúmen apresentaram *clusters* putativos completos e incompletos para biossíntese de peptídeos laço, sendo então selecionados para a reconstrução filogenética com base no gene que codifica o rRNA 16S (Figura 8). A árvore consenso segregou as 42 linhagens em seus respectivos gêneros ou famílias (Lachnospiraceae e Coriobacteriaceae). Observou-se em clados comuns, o agrupamento de membros da mesma família: *Butyrivibrio* e *Pseudobutyrvibrio* (Lachnospiraceae); *Ruminococcus*, *Streptococcus* e *Bacillus*

(Ruminococcaceae, Streptococcaceae e Bacillaceae, respectivamente). No nível de gênero, foi observado o agrupamento em clados comuns, de linhagens de *Butyrivibrio* e *Ruminococcus* (Figura 8a).

A distribuição dos *clusters* putativos de biossíntese de peptídeos laço (completos e incompletos) em genomas de bactérias do rúmen foi indicada paralelamente à representação da reconstrução filogenética do rRNA 16S dessas linhagens (Figura 7b). Para o gênero *Butyrivibrio*, foram inferidos 14 *clusters* putativos completos para biossíntese de peptídeo laço em 50 genomas analisados. Para os demais genomas de linhagens do gênero *Butyrivibrio* (n = 36) não foi possível inferir nenhum gene relacionado à biossíntese de peptídeo laço, sendo esse conjunto de genomas representado, a fim de exemplificação, pela linhagem *Butyrivibrio* sp. AD3002 (Figura 8b). Para alguns casos, a exemplo do gênero *Pseudobutyrvibrio*, somente *clusters* incompletos contendo genes de biossíntese de peptídeos laço foram inferidos. Assim, nenhuma linhagem com *cluster* completo de biossíntese de peptídeos laço foi exemplificada na Figura 8b.



0,2

Figura 8: Relações filogenéticas dos genes que codificam o rRNA 16S de genomas de bactérias do rúmen que apresentaram *clusters* putativos completos e incompletos para peptídeos laço (A) e exemplificação de genomas com *clusters* completos e incompletos cujos gêneros estiveram destacados na árvore, onde quadros vermelhos, verdes, amarelos e azuis, indicam respectivamente os genes *lasA*, *lasB*, *lasC* e *lasD* (B). A árvore filogenética, baseada na análise *Maximum Likelihood*, foi construída com 1000 réplicas de *bootstrap* usando o programa MEGA 7.0. A reconstrução filogenética foi baseada em sequências parciais e completas de rRNA 16S disponíveis no NCBI.

A fim de avaliar a expressão dos genes de biossíntese de peptídeos laço, foram realizadas análises de metatranscriptomas utilizando os arquivos SRR3169851 (bovino com dieta mista), SRR594193 (bovino 2155 com dieta controle), SRR594215 (bovino 2155 com dieta rica em óleo), SRR594642 (bovino 6021 com dieta controle), SRR594654 (bovino 6021 com dieta rica em óleo), SRR873462 (bovino com baixa produção de metano), SRR1138697 (bovino com alta produção de metano), SRR3257011 (bovino oriundo do Canadá) para verificação do número de *reads* por kilobase por milhões de *reads* mapeadas (RPKM) para cada genoma em função do gene analisado.

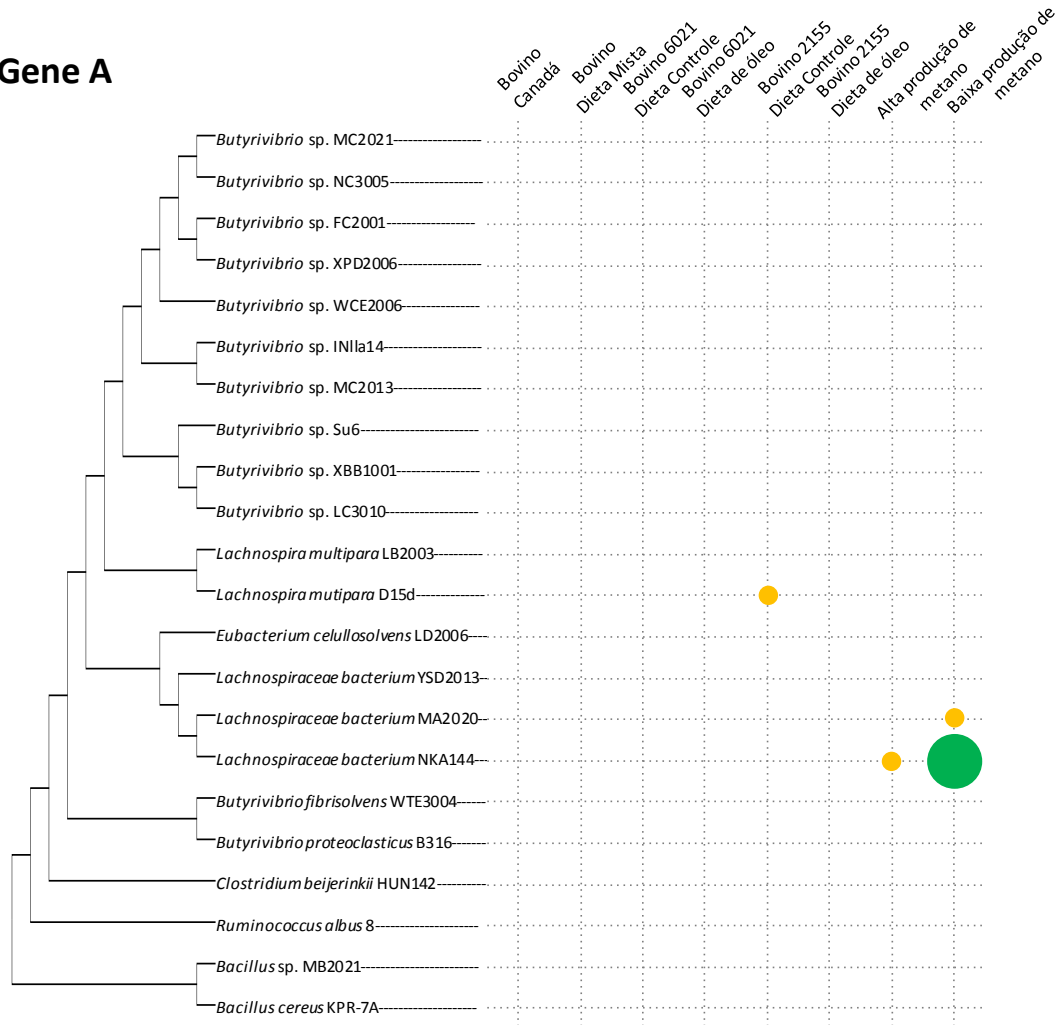
Os círculos nas figuras abaixo representam o número de *reads* por kilobase por milhões de *reads* mapeadas (RPKM). É uma forma de normalização dos números de *reads* transcritas, uma vez que, essas possuem sequência de tamanhos diferentes. O valor real de RPKM dos resultados é elevado a 10^{-7} . Para o gene putativo *lasA* foram inferidas *reads* somente para os genomas de *Lachnospira mutipara* D15d, *Lachnospiraceae bacterium* NKA144, *Lachnospiraceae bacterium* MA2020. Observa-se que o maior valor de RPKM para o genoma de *Lachnospiraceae bacterium* NKA144 a partir da análise do arquivo SRR1138697 (bovino com alta produção de metano). Para os gene putativo *lasB* foram inferidas *reads* somente para os genomas de *Lachnospira mutipara* LB2003 a partir do arquivo SRR1138697 (bovino com alta produção de metano), e *Ruminococcus albus* 8 a partir do arquivo SRR594215 (bovino 2155 com dieta rica em óleo), o número de RPKM para ambos genomas variou de 1 a 10.

Para os genes *lasC* e *lasD* foram inferidas *reads* a partir de um número maior de genomas. Para o gene *lasC* foram inferidas *reads* a partir dos genomas de *Butyrivibrio* sp. NC3005 (arquivo SRR873462 - bovino com baixa produção de metano), *Butyrivibrio* sp. FC2001 (arquivos SRR594193- bovino 2155 com dieta controle e SRR873462 - bovino com baixa produção de metano), *Butyrivibrio* sp. XPD2006 (arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano), *Butyrivibrio* sp. Su6 (arquivos SRR594642 - bovino 6021 com dieta controle e SRR873462 - bovino com baixa produção de metano), *Butyrivibrio* sp. LC3010 (arquivos SRR594654 - bovino 6021 com dieta rica em óleo, SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano), *Eubacterium cellulosolvens* LD2006 (arquivos SRR873462 - bovino com baixa produção de metano e SRR873462 - bovino com alta produção de metano), *Butyrivibrio proteoclasticus* FD2007 (arquivos SRR594193 - bovino 2155

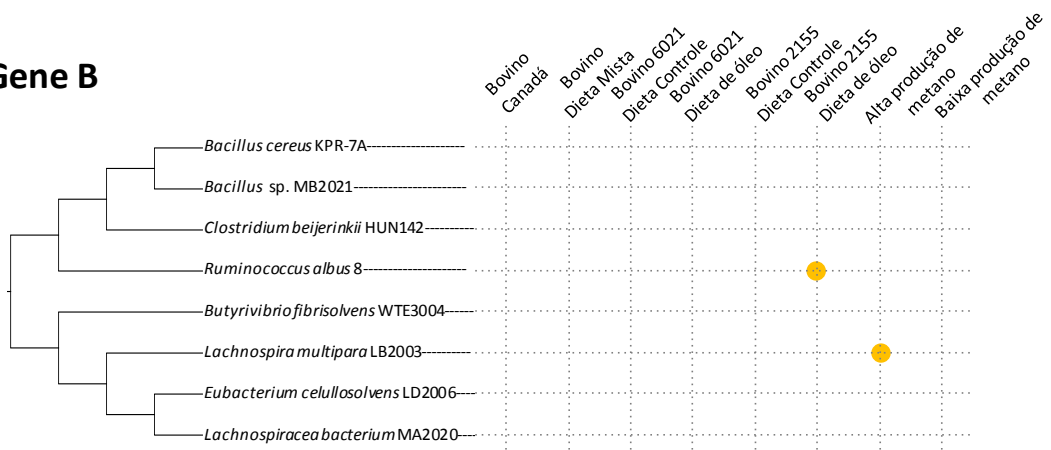
com dieta controle, SRR594215 - bovino 2155 com dieta rica em óleo, SRR594642 - bovino 6021 com dieta controle, SRR594654 - bovino 6021 com dieta rica em óleo), *Ruminococcus albus* 8 (arquivo SRR1138697 - bovino com alta produção de metano) e *Bacillus* sp. MB2021(arquivo SRR1138697 - bovino com alta produção de metano). As *reads* inferidas para o gene *lasC* que apresentaram valores de RPKM de 0,1 a ≥ 20 .

Para o gene *lasD* foram inferidas *reads* a partir dos genomas de *Lachnospiraceae bacterium* NKA144 (arquivos SRR3169851 - bovino com dieta mista e arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano), *Lachnospiraceae bacterium* YSD2013 (arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano), *Lachnospira mutipara* LB2003 (arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano), *Eubacterium cellulosolvens* LD2006 (arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano, e SRR3169851 - bovino com dieta mista), *Butyrivibrio proteoclasticus* FD2007 (arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano), *Butyrivibrio fibrisolvens* WTE3004 (arquivo SRR3169851 - bovino com dieta mista), *Clostridium beijerinckii* HUN142 (arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano), *Bacillus cereus* KPR-7A (arquivos SRR873462 - bovino com baixa produção de metano e SRR1138697 - bovino com alta produção de metano, e SRR3169851 - bovino com dieta mista). As *reads* inferidas para o gene *lasD* apresentaram valores de RPKM de 0,1 a ≥ 20 . Observou-se que o maior número de *reads* para o genes *lasC* e *lasD* foram inferidas a partir dos arquivos de metatranscriptoma de bovinos de alta e baixa produção de metano.

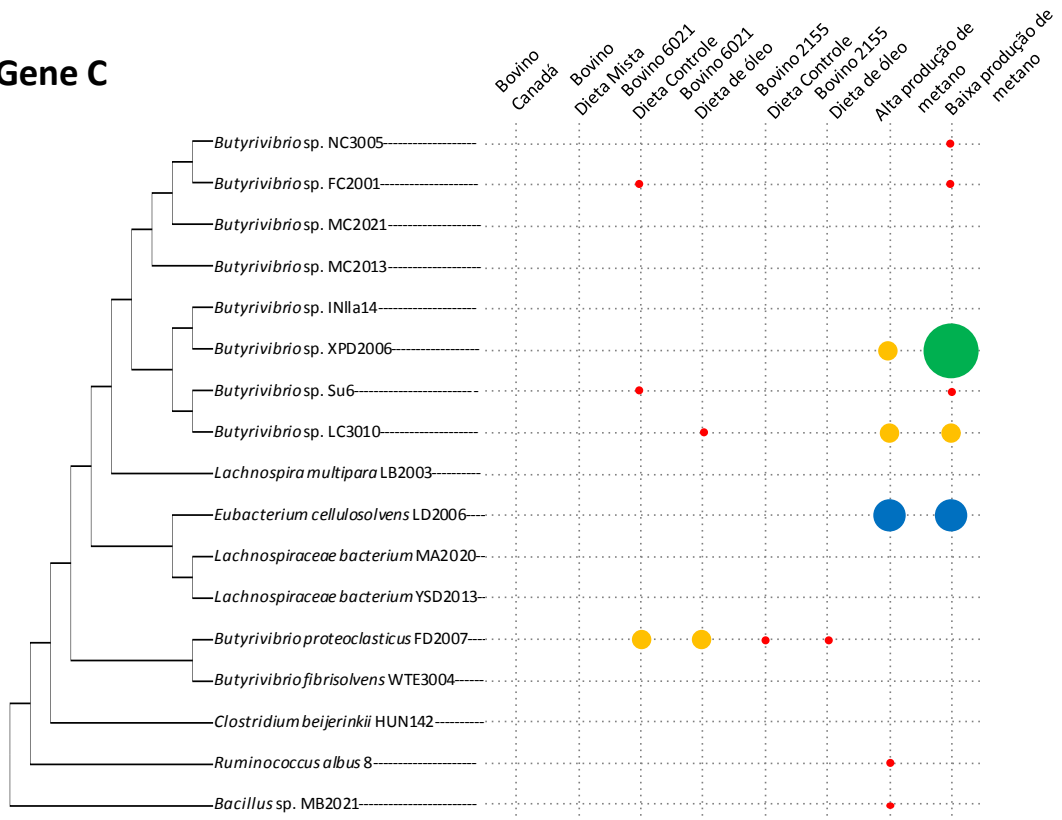
Gene A



Gene B



Gene C



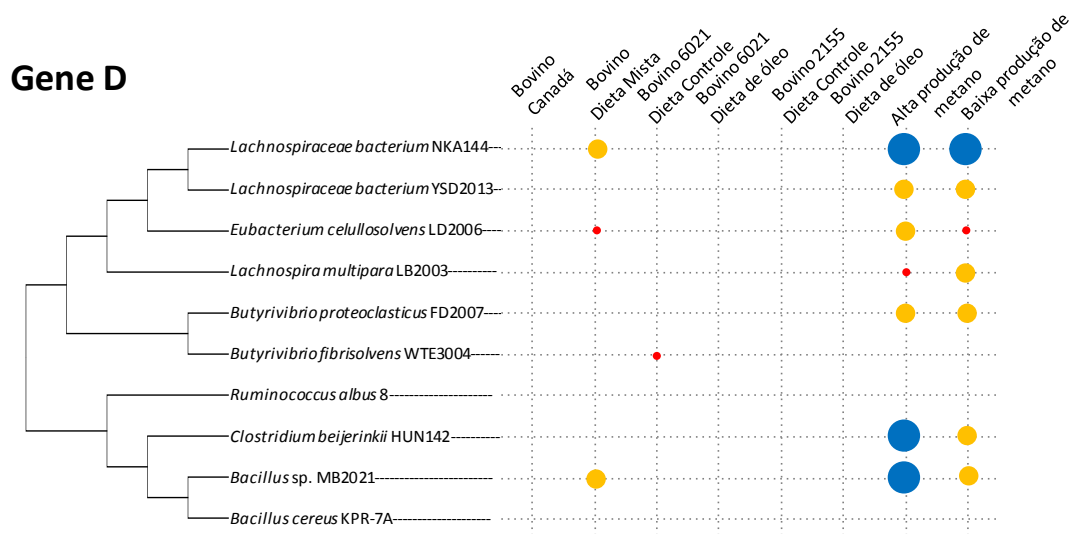


Figura 10: Números de reads por kilobase por milhões de *reads* mapeadas (RPKM) obtidos para cada genoma de bactérias do rúmen com *cluster* putativo completo para peptídeos laço. As análises foram realizadas com auxílio do software *Bowtie* a partir dos diferentes arquivos de metatranscriptomas analisados. (A) Número de *reads* por kilobase por milhões de *reads* mapeadas inferidos a partir do gene putativo *lasA*. (B) Número de *reads* por kilobase por milhões de *reads* mapeadas inferidos a partir do gene putativo *lasB*. (C) Número de *reads* por kilobase por milhões de *reads* mapeadas inferidos a partir do gene putativo *lasC*. (D) Número de *reads* por kilobase por milhões de *reads* mapeadas inferidos a partir do gene putativo *lasD*.

2.4 Discussão

Várias aplicações biotecnológicas têm sido atribuídas aos peptídeos laço mediante a diversidade de funções biológicas apresentadas, incluindo ação antiviral (ex. RP 71955 contra o vírus HIV), inibição de enzimas (ex. propeptin com atividade inibitória da prolil endopeptidase) e antagonistas de receptor (ex. anantin como antagonista competitivo do fator natriurético atrial - ANF) (Weber et al., 1991; Helynck et al., 1993; Kimura et al., 1997). No entanto, a atividade antimicrobiana parece ser a característica mais comum entre os peptídeos laço, como lariantin A (Inokoshi et al., 2012), lassomycin (Gavrish et al., 2014), microcina J25 (Salomón e Farías, 1992), propeptin (Kimura et al., 1997), siamycin (Yamasaki et al., 1996) e streptomomicin (Metelev et al., 2015). Alguns desses peptídeos apresentam amplo espectro de ação. A MccJ25, peptídeo laço mais bem estudado, apresentou atividade antimicrobiana contra várias espécies de bactérias Gram-negativas, incluindo *E. coli*, *Salmonella newport*, *Salmonella enteritidis* e *Shigella flexneri* (Salomón & Farías, 1992). Tsunakawa et al. (1995) demonstraram que siamycin I e II foram igualmente ativos contra linhagens de *Bacillus subtilis*, *Staphylococcus aureus*, *Micrococcus luteus* e *Escherichia coli*, enquanto propeptin inibiu linhagens de *Pseudomonas aeruginosa*, *Mycobacterium phlei* e *Xanthomonas oryzae* (Kimura et al., 1997). Posteriormente, a atividade antimicrobiana de capistrin foi demonstrada contra linhagens de *Burkholderia caledonica*, *P. aeruginosa* e *E. coli* (Knappe et al., 2008; Kuznedelov et al., 2011).

O interesse pelos peptídeos laço é crescente e desde a descoberta do primeiro peptídeo desse grupo (anantin) em 1991, pelo menos outros 39 peptídeos foram descobertos, sendo estes, em sua maioria, pertencentes à Classe I (Hegemann et al., 2015). A descoberta de novos peptídeos laço tem sido ampliada por análises *in silico* que utilizam ferramentas de bioinformática para mineração de sequências genômicas, que por sua vez, tem sido cada vez mais abundantes nos bancos de dados de acesso livre (Kersten et al., 2011; Hegemann et al., 2013; Hegemann et al., 2014; Maksimov & Link, 2014). A mineração genômica tem contribuído também para a identificação de microrganismos produtores de peptídeos laço, ao permitir a detecção de *clusters* gênicos potencialmente envolvidos na biossíntese de peptídeos laço em genomas de microrganismos isolados de diferentes habitats (solo, sedimentos, ambientes aquáticos) (Duquesne et al., 2007; Knappe et al., 2008; Letzel et al., 2014). Contudo, poucos estudos tem investigado a ocorrência de peptídeos laço em microrganismos do trato

gastrointestinal animal, incluindo o estômago de ruminantes. Este estudo baseou-se na premissa de que a produção de agentes antimicrobianos é vantajosa para a sobrevivência de bactérias no ecossistema ruminal (Kalmokoff et al., 1996; Teather e Forster, 1998; Mantovani et al., 2001), sendo avaliada a hipótese de que os microrganismos do rúmen são potenciais produtores de peptídeos laço. Identificou-se *clusters* putativos (completos e incompletos) para peptídeos laço em 12% dos genomas de bactérias do rúmen analisados no presente estudo (n = 345 genomas). A ocorrência de *clusters* putativos de biossíntese de peptídeos laço em genomas de bactérias ruminais foi maior do que o reportado em outros estudos, indicando que o ecossistema ruminal pode representar uma fonte potencial desses peptídeos. Em estudo anterior, Maksimov et al. (2012b) reportaram a ocorrência de *clusters* putativos completos para peptídeos laço em 2,64% dos 3 mil genomas microbianos estudados, aplicando um algoritmo desenvolvido para detectar precursores dos peptídeos laço, seguido pela busca por homólogos de *mcjB* e *mcjC*. De acordo ainda com esses autores, a ocorrência de *clusters* predominou entre genomas de bactérias dos filos Proteobacteria e Actinobacteria, quando comparada aos filos Firmicutes, Cianobacteria, Espiroqueta, Verrucomicrobia, Bacteroidete, Fusobactéria, Tenericutes e um filo de Eurychaeta pertencente domínio de Archaea.

Quando Letzel et al. (2014) realizaram a mineração genômica de *clusters* gênicos que codificavam peptídeos antimicrobianos sintetizados ribossomicamente e que sofrem modificações pós tradução (RiPPs), utilizando as ferramentas de bioinformática AntiSMASH, BAGEL e Bactibase, verificou-se que apenas 3,8% dos 211 genomas de bactérias anaeróbias analisadas apresentavam *clusters* gênicos putativos completos para biossíntese de peptídeos laço. Tal ocorrência de *clusters* predominou em genomas de bactérias classificadas no filo Proteobacteria, uma vez que a presença de *cluster* de peptídeo laço somente foi identificada no genoma de uma linhagem do filo Firmicutes (*Butyrivibrio proteoclasticus* B216). Nenhum *cluster* para peptídeo laço foi inferido em genomas referentes aos outros 14 filos que também foram alvos de investigação por aqueles autores (Actinobacteria, Chloroflexi, Bacteroidetes, Chrysiogenetes, Deferribacteres, Deinococcus-Thermus, Elusibacteria, Fibrobacteres, Fusobacteria, Spirochaetes, Synergistetes, Thermodesulfobacteria, Thermotogae, Verrucomicrobia).

Bactérias dos filos Proteobacteria e Actinobacteria, oriundas de diversos habitats, têm sido apontadas como as principais produtoras de peptídeos laço em estudos de mineração genômica. Em estudos anteriores, *Streptomyces roseosporus*

NRRL15998, produtora de SRO15-2005, e *Streptomyces sviveus* ATCC 29083, produtora de sviveucin, foram isoladas de amostras de solo (Kersten et al., 2011; Ducasse et al., 2012). *Asticcacaulis excentricus*, produtora de astexin I, II e III foi isolada em água doce (Maksimov et al., 2012b; Zimmerman et al., 2013), enquanto *Caulobacter* sp. K31, produtora de caulonodin I, II e III, foi proveniente de águas subterrâneas contaminadas com clorofenicol (Hegemann et al., 2013). *Sphingopyxis alaskensis* RB2256 DSM 13593, produtora de sphingopyxin, foi isolada a partir de amostras coletadas da água do mar (Hegemann et al., 2013).

Os resultados obtidos no presente estudo indicaram *clusters* putativos de biossíntese de peptídeos laço (completos ou incompletos) em genomas de bactérias do rúmen distribuídos em 11 gêneros (*Bacillus*, *Butyrivibrio*, *Clostridium*, *Enterococcus*, *Eubacterium*, *Lachnospira*, *Propionibacterium*, *Pseudobutyribrio*, *Ruminococcus*, *Staphylococcus* e *Streptococcus*) e duas famílias (Lachnospiraceae e Coriobacteriaceae). Esses resultados sugerem que as bactérias do filo Firmicutes são os principais produtores de peptídeos laço no rúmen, uma característica que ainda não havia sido demonstrada em outros ecossistemas onde a biossíntese de peptídeos laço foi avaliada por análises *in vitro* ou *in silico*.

Dentre os genomas de bactérias do rúmen em que foram detectados *clusters* gênicos putativos completos e incompletos para peptídeos laço, aqueles pertencentes ao gênero *Butyrivibrio* (filo Firmicutes) foram os que apresentaram o maior número de *clusters* (n = 14). No rúmen, bactérias do gênero *Butyrivibrio* possuem papel importante na digestão dos ruminantes, e conseqüentemente, na nutrição desses animais, uma vez que essas bactérias estão envolvidas na degradação de carboidratos estruturais das plantas, principalmente a hemicelulose, além de biohidrogenação de lipídeos. A produção de bacteriocinas também foi inferida em genomas de linhagens do gênero *Butyrivibrio*, conforme relatado por Azevedo et al. (2015). Estes autores reportaram, a partir da mineração de genomas de bactérias e archaea do rúmen (n = 229), a presença de *clusters* putativos para biossíntese de bacteriocinas (Classe III) em genomas de 16 linhagens de *Butyrivibrio*, além da presença de *clusters* putativos de peptídeos laço em genomas de 9 linhagens desse gênero.

Vinte e seis genomas de bactérias do rúmen apresentaram *clusters* putativos completos de biossíntese de peptídeos laço. Além dos genes putativos *lasA*, *lasB*, *lasC*, *lasD*, esses *clusters* apresentaram genes não relacionados com a biossíntese de peptídeos laço, incluindo genes para uma histidina cinase (HisKA) e a coenzima

pirroloquinolina quinona (PqqD). Segundo Varughese (2002), HisKA é uma proteína sensora que atua juntamente com a proteína reguladora de resposta (RR) na cascata de fosforilação de sistemas de dois componentes. A presença de genes codificadores de HisKA no *cluster* de biossíntese do peptídeo laço sviveucin (produzido por *Streptomyces sviveus*) também foi descrita por Li et al. (2015), que os consideraram como genes reguladores da expressão de peptídeos laço. Azevedo et al. (2015) reportaram, pela primeira vez, a existência de genes codificadores de histidina cinase em *clusters* de biossíntese de sactipeptídeos em genomas de bactérias do rúmen. A coenzima PqqD está envolvida na via de biossíntese da pirroloquinolina quinona (PQQ), tendo sido previamente associada ao gene *albA32* da via de biossíntese da subtilosina (sacpeptídeo) e a uma proteína de função desconhecida na biossíntese do peptídeo laço lariatin (Li et al., 2015).

Domínios conservados de PqqD também foram identificados nas sequências das proteínas codificadas pelos genes *lasA* presentes nos *clusters* putativos completos encontrados nos genomas de bactérias do rúmen. Esse domínio foi identificado em sete genomas de bactérias do rúmen, sendo sugerido que o mesmo atue como um elemento de reconhecimento do peptídeo precursor dos peptídeos antimicrobianos sintetizados ribossomicamente e que sofrem modificações pós-tradução (Burkhart, et al., 2015).

Além de HisKA e PqqD, outros produtos de genes não relacionados à biossíntese de peptídeos laço foram inferidas a partir dos *clusters* completos, como as proteínas nucleotidiltransferases (NTP_transf), glicosiltransferase (Glyco_transf_), proteína de regulação de transcrição (Trans_reg_C), proteína envolvida na transcrição (Spoc), dentre outras. Para alguns desses produtos gênicos a função na célula microbiana é conhecida, a exemplo das nucleotidiltransferases que são proteínas envolvidas na transferência de monofosfato de adenosina de uma molécula de trifosfato de adenosina para uma proteína, ácidos nucleicos ou pequenas moléculas (Kuchta et al., 2009). A presença de genes que codificam nucleotidiltransferases já foi verificada anteriormente, via predição, para o peptídeo laço paeninodin isolado de *Paenibacillus polymyxa* CR1. No entanto, sua função na biossíntese do peptídeo não foi esclarecida (Zhu et al., 2016).

A enzima glicosiltransferase atua na transferência de um monossacarídeo de um açúcar doador para um sacarídeo, proteínas, lipídeos, DNA ou pequenas moléculas aceptoras (Radomska-Pandya et al., 1999; Williams et al., 2009). De acordo com Vilches et al. (1992), as glicosiltransferases em microrganismos produtores de

antibióticos protegem a célula produtora a partir da inativação do antibiótico por glicosilação no espaço intracelular. Depois de secretado, o antibiótico é reativado por glicosidases hidrolíticas, as quais removem as moléculas de hexoses.

Os domínios conservados dos produtos dos genes *lasB* (PF13471 e transglutaminase) e *lasC* (PF00733 e asparagina sintetase) foram identificados em todas as 26 linhagens de bactérias do rúmen para as quais foram inferidos *clusters* putativos completos de peptídeos laço. De acordo com Hegemann et al. (2015), esses domínios têm sido amplamente utilizados em estudos de mineração gênica para predição de peptídeos laço com base na busca por proteínas homólogas aos produtos desses genes nos genomas analisados e na identificação de genes precursores de peptídeos laço flanqueando homólogos putativos de *lasB* e *lasC* (Maksimov et al., 2012a).

As proteínas codificadas pelo gene *lasD* estão relacionados, de forma geral, com a secreção dos peptídeos laço e podem conferir “imunidade” para as bactérias que as produzem (Clarke & Campopiano, 2007; Duquesne et al., 2007). No presente trabalho, o produto putativo do gene *lasD*, identificado em 17 genomas de bactérias do rúmen, apresentava o domínio conservado ABC transportador transmembranar (ABC transporter transmembrane region), que tem sido relacionado à exportação dos peptídeos laço para o meio extracelular, conforme reportado para a microcina de *E. coli* (Gu et al., 2015). Além desse domínio, verificou-se, para três genomas de bactérias do rúmen, o domínio conservado para HPr cinase. Segundo Zhu et al. (2016), que também verificaram esse domínio em *clusters* gênicos do peptídeo laço paeninodin isolado de *Paenibacillus polymyxa* CR1, a HPr cinase tem como função fosforilar a serina (aminoácido conservado presente no peptídeo *core* do paeninodin) no C-Terminal do peptídeo precursor.

A característica peculiar aos peptídeos laço é sua topologia estrutural, que apresenta um anel macrolactâmico composto por aminoácidos, o *loop* e a cauda, sendo estes dois últimos originados pela extremidade C-terminal. A formação do anel se dá pela ligação amida entre o grupo α -amino da região N-terminal de resíduos de glicina, cisteína, serina e alanina e o grupo carboxila de resíduos de aspartato/glutamato, nas posições 8 ou 9 (McIntosh et al., 2009; Velasquez & Van der Donk, 2011; Alvarez-Sieiro et al., 2016). Esses resíduos conservados de aminoácidos foram verificados nas sequências inferidas para os peptídeos *core* de 26 genomas de bactérias do rúmen que apresentaram *clusters* completos para peptídeos laço. As análises *in silico* indicaram que o tamanho dos anéis variaram de sete a nove resíduos de aminoácidos. Segundo

Hegemann et al. (2014), os anéis dos peptídeo laço com tamanho menor que sete ou maior que nove resíduos de aminoácidos comprometem a topologia do peptídeo, com conseqüente perda de atividade biológica. Desta forma, pode-se sugerir que os peptídeos laço inferidos neste estudo para 26 genomas de bactérias do rúmen podem desempenhar algum papel importante na ecologia ruminal, uma vez que o tamanho dos anéis dos peptídeos putativos não comprometeria as funções biológicas dos mesmos.

Conforme observado a maioria dos peptídeos laço inferidos para 26 genomas de bactérias do rúmen não apresentaram em suas sequências de peptídeo *core* resíduos de cistéina, com exceção de *Lachnospiraceae bacterium* YSD2013, cuja sequência apresentou dois resíduos de cistéina. A classificação dos peptídeos laço é determinada pelo número de ligações dissulfeto que conferem estabilidade estérica à topologia da molécula (Hegemann et al., 2015). Atualmente, a classe I dos peptídeos laço são representadas, por quatro peptídeos (RP 71955/aborycin, Siamycin I, Siamycin II e SSV-2083), enquanto a classe III é representada por um peptídeo (BI-32169) (Hegemann et al., 2015). Peptídeos laço de classe II têm sido os mais frequentemente descobertos, totalizando 35 dos 40 peptídeos já conhecidos, sendo frequentemente encontrados em bactérias do Filo Proteobacteria (Maksimov & Link, 2014). No entanto, no presente trabalho, tem-se evidências que os peptídeos laço inferidos podem pertencer a Classe II, sinalizando, portanto, para um novo Filo bacteriano (Filo Firmicutes) produtor de peptídeos laço da Classe II.

A reconstrução filogenética dos produtos putativos codificados pelos genes *lasB*, *lasC* e *lasD* revelou ramos monofiléticos compostos por linhagens da mesma espécie. Esses ramos monofiléticos agruparam, por exemplo, *Butyrivibrio fibrisolvens* MD2001 e *Butyrivibrio fibrisolvens* WTE3004 (árvore filogenética do gene *lasB*), *Butyrivibrio proteoclasticus* B316 e *Butyrivibrio proteoclasticus* FD2007 (árvore filogenética do gene *lasC*), e *Lachnospira multipara* LB2003 e *Lachnospira multipara* MC2003 (árvore filogenética do gene *lasD*). Os ramos monofiléticos compostos por linhagens da mesma espécie evidenciam um processo evolutivo e não a transferência horizontal de genes relacionados a biossíntese de peptídeos laço. A transferência vertical de genes ocorre de um ancestral para a geração seguinte. Isso também foi corroborado pelos resultados obtidos a partir dos alinhamentos múltiplos que evidenciaram as regiões consenso das sequências dos aminoácidos codificados pelos genes *lasB* para linhagens da mesma espécie, sendo o mesmo observado com os produtos putativos do gene *lasC*.

A reconstrução da filogenia dos genes que codificam o rRNA 16S de bactérias do rúmen com *clusters* putativos completos e incompletos para peptídeos laço segregou as 42 linhagens em seus respectivos gêneros ou famílias (Lachnospiraceae e Coriobacteriaceae). A distribuição dos *clusters* putativos de biossíntese de peptídeos laço (completos e incompletos) em genomas de bactérias do rúmen demonstrou que, pelo menos uma linhagem de cada gênero bacteriano não apresentou *cluster* de biossíntese de peptídeos laço inferidos pela análise do genoma. Neste trabalho para o gênero *Butyrivibrio* foram analisados 50 genomas quanto à presença de *cluster* gênicos de peptídeos laço, porém em apenas 14 genomas foi possível constatada a presença de *clusters* completos. Por outro lado, em 36 genomas de linhagens do gênero *Butyrivibrio* não foram inferidos os genes *lasA*, *lasB*, *lasC* ou *lasD*. A ausência de *clusters* completos ou incompletos pode ser justificada pela provável perda de genes da biossíntese de peptídeos laço ocorrida durante a evolução desses microrganismos, uma vez que não são essenciais a viabilidade celular. Todavia, deve-se reconhecer que a impossibilidade de constatar genes da biossíntese de peptídeos laço para algumas das linhagens do rúmen pode ter sido decorrente de montagens incorretas e/ou incompletas dos genomas.

Além de inferir os genes de biossíntese de peptídeos laço em genomas de bactérias do rúmen, buscou-se neste trabalho avaliar a expressão desses genes em metatranscriptomas do rúmen. Com base nas análises de metatranscriptoma, pôde-se inferir a expressão de genes de biossíntese de peptídeos laços em genomas de bactérias do rúmen, uma vez que *reads* foram inferidas, porém não foi possível inferir *reads* para os quatro genes de biossíntese de peptídeo laço, simultaneamente, a partir de nenhum genoma. Além disso, observou-se que a inferência sobre expressão de genes de biossíntese de peptídeos laço não foi proeminente para um determinado gênero ou linhagem de bactérias do rúmen, o que permite sugerir que a produção de peptídeos laço é uma importante vantagem adaptativa que contribui, de forma geral, para a sobrevivência e atividade de bactérias no ecossistema ruminal. Esses resultados representam os primeiros relatos na literatura quanto à expressão de genes de biossíntese de peptídeos laço, via análises de metatranscriptoma, porém testes de validação experimental se fazem necessários. Estudos de metatranscriptomas são interessantes para que se possa entender a correlação entre a presença de um gene e as características de um determinado ambiente.

As análises *in silico* realizadas forneceram evidências da presença de novos *clusters* de genes biossintéticos em espécies bacterianas não previamente relacionadas

com peptídeos laço, sugerindo que a microbiota ruminal representa uma fonte potencial de novos peptídeos laço. Os resultados apresentados neste estudo indicaram que a genes de biossíntese de peptídeos laço são amplamente distribuídos em genomas de bactérias ruminais. Estudos posteriores destes peptídeos antimicrobianos podem ser úteis para desenvolver novas estratégias para manipular a fermentação ruminal e melhorar a produtividade dos ruminantes.

2.5 Conclusões

Os resultados apresentados neste estudo indicaram que a genes de biossíntese de peptídeos laço são amplamente distribuídos em genomas de bactérias ruminais. A partir de 350 genomas analisados foram inferidos 26 clusters putativos completos para biossíntese de peptídeo laço. O maior número de *cluster* putativos completos inferidos foi relatado para o gênero *Butyrivibrio*. Para os produtos dos genes *lasB* e *lasC*, foram verificados domínios conservados comuns à todas as linhagens, sendo que o domínio Transglut_core3, com função transglutaminase, foi conservado em *lasB* e o domínio Asn_synthase, com função de biossíntese de asparagina, foi conservado em *lasC*. O produtos dos genes *lasB* e *lasC* apresenta similaridade entre espécie linhagens da mesma espécie. Maior número de *reads* foi observado para os genes *lasC* e *lasD*, inferidas a partir dos arquivos de metatranscriptoma.

2.6 Referências Bibliográficas

- Alvarez-Sieiro, P.; Montalbán-López, M.; Mu, D.; Oscar P. Bacteriocins of lactic acid bacteria: extending the family. **Applied Microbiology Biotechnology**, v.100, p. 2939–2951, 2016.
- Azevedo, A. C.; Bento, C. B.; Ruiz, J. C.; Queiroz, M. V.; Mantovani, H. C. D. Distribution and genetic diversity of bacteriocin gene *clusters* in rumen microbial genomes. **Applied and Environmental Microbiology**, v. 81, p. 7290-7304, 2015.
- Burkhart, B. J.; Hudson, G.A.; Dunbar, K. L.; Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. **Nature Chemical Biology**, v. 11, p. 564–70, 2015.
- Clarke, D.J.; Campopiano, D.J.) Maturation of MccJ25 precursor peptide into active microcin MccJ25. **Organic & Biomolecular Chemistry**, v. 21, p.2564–1566, 2007.
- De Cristobal, R. E.; Solbiati, J. O.; Zenoff, A. M.; Vincent, P. A.; Salomón, R. A.; Yuzenkova, J.; Severinov, K.; Farías, R. N. Microcin J25 uptake: His5 of the MccJ25 lariat ring is involved in interaction with the inner membrane MccJ25 transporter protein SbmA. **Journal Bacteriology**, v.188, p. 3324-3328, 2006.
- Destoumieux-Garzón, D.; Duquesne, S.; Peduzzi, J.; Goulard, C.; Desmadril, M.; Letellier, L.; Rebuffat, S.; Boulanger, P. The iron-siderophore transporter FhuA is the receptor for microcin J25. Role of the microcin Val¹¹-Pro¹⁶ β -hairpin region in the recognition mechanism. **Biochemical Journal.**, v. 389, p. 869–876, 2005.
- Ducasse, R.; Yan, K.P.; Goulard, C.; Blond, A.; Li, Y.Y.; Lescop, E.; Guittet, E.; Rebuffat, S.; Zirah, S. Sequence determinants governing the topology and biological activity of a lasso peptide, microcin J25. **ChemBioChem**, v. 13, p. 371–380, 2012.
- Duquesne, S.; Destoumieux-Garzón, D.; Zirah, S.; Goulard, C.; Peduzzi, J.; Rebuffat, S. Two Enzymes Catalyze the Maturation of a Lasso Peptide in *Escherichia coli*. **Chemistry & Biology**, v.14, p.793–803, 2007.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Res.**, v.32, p. 1792-1797.

Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*. Chapman & Hall, New York, 1993.

FastQC Disponível em < <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>>
Acesso em janeiro e fevereiro de 2017.

Gascuel O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol.*, v.14, p.685-95, 1997.

Gavriš, E.; Sit, C. S.; Cao, S.; Kandrór, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K. Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. **Chemistry & Biology**, v. 21, p. 509–518, 2014,

Gu, R. X.; Corradi, V.; Singh, G.; Choudhury, H. G.; Beis, K.; Tieleman, D. P. Conformational Changes of the Antibacterial Peptide ATP Binding Cassette Transporter McjD Revealed by Molecular Dynamics Simulations. **Biochemistry**, v.29, p. 5989-5998, 2015.

Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Lasso peptides: an intriguing class of bacterial natural products. **Accounts of Chemical Research**, v. 48, p. 1909-1919, 2015.

Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Steuber, H.; Harms, K.; Xie, X.; Marahiel, M. A. Xanthomonins I–III: A new class of lasso peptides with a seven-residue macrolactam ring. **Angewandte Chemie International**, v.53, p.2230-2234, 2014.

Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Klug, D.; Marahiel, M. A. Lasso peptides from proteobacteria: Genome mining employing heterologous expression and mass spectrometry. **Biopolymers**, v. 100, p. 527–542, 2013.

Helynck, G.; Dubertret, C.; Mayaux, J.F.; Leboul, J. Isolation of RP 71955, a new anti-HIV-1 peptide secondary metabolite. **Journal Antibiotics**, v. 46, p. 1756–1757, 1993.

Inokoshi, J.; Matsuhama, M.; Miyake, M.; Ikeda, H.; Tomoda, H. Molecular cloning of the gene cluster for lariat biosynthesis of *Rhodococcus jostii* K01-B0171. **Applied Microbiology Biotechnology**, v. 95, p. 451–460, 2012.

Kalmokoff, M.L.; Bartlett, F.; Teather, R. M. Are ruminal bacteria armed with Bacteriocins? **Journal Dairy Science**, v.79, p. 2297-2306, 1996.

Kersten, R. D.; Yang, Y. L.; Xu, Y.; Cimermancic, P.; Nam, S. J.; Fenical, W.; Fischbach, M. A.; Moore, B. S.; Dorrestein, P. C. A mass spectrometry-guided genome mining approach for natural product peptidogenomics. **Nature Chemical Biology**, v. 7, p. 794–802, 2011.

Kimura, K.; Kanou, F.; Takahashi, H.; Esumi, Y.; Uramoto, M.; Yoshihama, M. Propeptin, a new inhibitor of prolyl endopeptidase produced by *Microbispora*. I. Fermentation, isolation and biological properties. **Journal Antibiotics**, v. 50, p. 373–378, 1997.

Knappe, T. A.; Linne, U.; Zirah, S.; Rebuffat, S.; Xie, X.; Marahiel, M. A. Isolation and structural characterization of capistruin, a lasso peptide predicted from the genome sequence of *Burkholderia thailandensis* E264. **Journal of the American Chemical Society**, v. 130, p. 11446–11454, 2008.

Kuchta, K., Knizewski, L., Wyrwicz, L.S., Rychlewski, L., and Ginalski, K. Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. **Nucleic Acids Research**, v. 37, p. 7701–7714, 2009.

Kuznedelov, K.; Semenova, E.; Knappe, T. A.; Mukhamedyarov, D.; Srivastava, A.; Chatterjee, S.; Ebricht, R. H.; Marahiel, M. A.; Severinov, K. The antibacterial threaded-lasso peptide capistruin inhibits bacterial RNA polymerase. **Journal of Molecular Biology**, v. 412, p. 842–848, 2011.

Langmead, B.; Salzberg, S. L. Fast gapped-read alignment with bowtie 2. **Nature Methods**, v.9, p. 357-359, 2012.

Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. **Genome Biology**, v. 10, artigo R25, 2009.

Letzel, A. C.; Pidot, S. J.; Hertweck, C. Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. **BMC Genomics**, v. 15, p. 983, 2014.

Li, Y.; Ducasse, R.; Zirah, S.; Blond, A.; Goulard, C.; Lescop, E.; Giraud, C.; Hartke, A.; Guittet, E.; Pernodet, J. L.; Rebuffat, S. Characterization of Sviceucin from *Streptomyces* Provides Insight into Enzyme Exchangeability and Disulfide Bond Formation in Lasso Peptides. **ACS Chemical biological**, v. 1, p. 2641–2649, 2015

Li, Y.; Zirah, S.; Rebuffat, S. Lasso Peptides. Springer; New York: Biosynthesis, regulation and export of lasso peptides. p. 81-95, 2014.

Maksimov, M. O.; Link, A. J. Prospecting genomes for lasso peptides. **Journal of Industrial Microbiology and Biotechnology**, v.41, p. 333–344, 2014.

Maksimov, M. O.; Pan, S. J.; Link, J. A. Lasso peptides: Structure, function, biosynthesis, and engineering. **Natural Product Reports**, v. 29, p. 996–1006, (2012a).

Maksimov, M. O.; Pelczer, I.; Link, A. J. Precursor-centric genome-mining approach for lasso peptide Discovery. PNAS, **Proceedings of the National Academy of Sciences of United States of America**, v. 109, p. 15223–15228, (2012b).

Mantovani H.C.; Kam, D.K.; Ha, J.K.; Russell J.B. The antibacterial activity and sensitivity of *Streptococcus bovis* strains isolated from the rumen of cattle. **FEMS Microbiology Ecology**, v. 37, p. 223-229, 2001.

Mcintosh, J. A.; Donia, M. S.; Schmidt, E. W. Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. **Natural Product Reports**, v. 26, p. 537–559, 2009.

Marre, S. C.; Parisot, N.; Lepercq, P.; Chaucheyras-Durand F.; Mosoni, P.; Peyretailade, E.; Bayat, A. R.; Shingfield, K. J.; Peyre, P.; Forano, E. Metatranscriptomics Reveals the Active Bacterial and Eukaryotic Fibrolytic Communities in the Rumen of Dairy Cow Fed a Mixed Diet. **Frontiers in Microbiology**, v. 8, p. 1-13, 2017.

Medema, M.H.; Blin, K.; Cimermancic, P.; Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic Acids Research**, v.39, p. 339-346, 2011.

Metelev, M.; Tietz, J. I.; Melby, J. O.; Blair, P. M.; Zhu, L.; Livnat, I.; Severinov, K.; Mitchell, D. A. Structure, Bioactivity, and Resistance Mechanism of Streptomomicin, an Unusual Lasso Peptide from an Understudied Halophilic Actinomycete. **Chemistry & Biology**, v. 22, p. 241–250, 2015.

Radomska-Pandya, A.; Czernik, P.J.; Little, J.M.; Structural and functional studies of UDP-glucuronosyltransferases. **Drug Metabolism & Disposition**, v. 31, p. 817–899, 1999.

Salomón, R. A.; Farias, R. N. Microcin 25, a novel antimicrobial peptide produced by *Escherichia coli*. **Journal Bacteriology**, v. 174, p. 7428–7435, 1992.

Solbiati, J. O.; Ciaccio, M.; Farias, R. N.; Gonzalez-Pastor, J. E.; Moreno, F.; Salomon, R. A. Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25. **Journal Bacteriology**, v.181, p. 2659–2662, 1999.

Stothard, P. The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. **Biotechniques**, v. 28, p.1102-1104, 2000.

Sudhir, K.; Glen, S.; Koichiro, T. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular Biology and Evolution**, v. 33, p. 1870-1874, 2016.

Teather, R.M.; Forster, R.J. Manipulating the rumen microflora with bacteriocins to improve ruminant production. **Canadian Journal of Animal Science**, v.78, p. 57–69, 1998.

Tsunakawa, M.; Hu, S. L.; Hoshino, Y.; Detlefson, D. J.; Hill, S.E.; Furumai, T.; White, R. J.; Nishio, M.; Kawano, K.; Yamamoto, S.;Fukagawa, Y.; Oki, T. Siamycins I and II, new anti-HIV peptides: I.Fermentation, isolation, biological activity and initial characterization. **Journal Antibiotics**, v. 48, p. 433–434, 1995.

Van Heel, A. J.; Jong, A.; Montalbán-López, M.; Kok, J.; Kuipers, O. P. Bagel3: automated identification of genes encoding bacteriocins and (non-) bactericidal posttranslationally modified peptides. **Nucleic Acids Research**, v. 41, p. 448-453, 2013.

Varughese, K. I. Molecular recognition of bacterial phosphorelay proteins. **Current Opinion in Microbiology**, v. 5, p. 142–148, 2002.

Velasquez, J. E.; Van Der Donk W. A. Genome mining for ribosomally synthesized natural products. **Current Opinion in Microbiology**, v.15, p. 11–21, 2011.

Vilches, C.; Hernandez, C.; Mendez, C.; Salas, J. A. Role of glycosylation and deglycosylation in biosynthesis of and resistance to oleandomycin in the producer organism, *Streptomyces antibioticus*. **Journal of Bacteriology**, v. 174, p.161-165, 1992.

Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, p.57–63, 2009.

Weber, W.; Fischli, W.; Hochuli, E.; Kupfer, E.; Weibel, E. K. Anantin—a peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity. **Journal Antibiotics**, v. 44, p. 164–171, 1991.

Williams, G. J.; Thorson, J. S. "Natural product glycosyltransferases: properties and applications. **Advances in Enzymology and Related Areas of Molecular Biology**, v.76, p. 55–119, 2009.

Zhu, S.; Hegemann, J. D.; D. Fage, C. D.; Zimmermann, M.; Xiulan Xie, U. L.; Mohamed, A. Marahiel. Insights into the Unique Phosphorylation of the Lasso Peptide Paeninodin. **The Journal biological of chemistry**, v. 291, p. 13662-13678, 2016.

Zimmermann, M.; Hegemann, J. D.; Xie, X.; Marahiel, M. A. Characterization of caulonodin lasso peptides revealed unprecedented N-terminal residues and a precursor motif essential for peptide maturation. **Chemical Science**, v. 5, p. 4032–4043, 2014.

Zimmermann, M.; Hegemann, J. D.; Xie, X.; Marahiel, M. A. The astexin-1 lasso peptides: Biosynthesis, stability, and structural studies. **Chemistry & Biology**, v. 20, p. 558–569, 2013.

Tabela S1 - Genomas de bactérias do rúmen analisados (dados coletados de Abril a dezembro de 2016)

Linhagens	Source	Accession Number	Genome size (Mb)	Contigs/scaffolds	Database
<i>Acetivomaculum ruminis</i> DSM 5522	Rúmen bovino	FOJY00000000.1	3.09	54	Genbank
<i>Acidaminococcus fermentans</i> PGA-4	Rúmen bovino	FOWJ00000000.1	2.35	57	Genbank
<i>Acidaminococcus fermentans</i> WCC6	Rúmen bovino	PRJNA323320	2.29	40	BioProject/ Hungate1000*
<i>Acinetobacter</i> sp. DSM 11652	Rúmen de ovelha	FOBK00000000.1	3.05	39	Genbank
<i>Actinobacillus succinogenes</i> 130Z	Rúmen bovino	NC_009655.1	2.32	NA	Genbank
<i>Actinomyces denticolens</i> PA	Rúmen bovino	NZ_FQYL00000000.1	2.75	33	Genbank
<i>Actinomyces nasicola</i> KPR-1	Rúmen bovino	FNQV00000000.1	2.54	29	Genbank
<i>Actinomyces ruminicola</i> KPR-7B	Rúmen bovino	FNHU00000000.1	3.30	38	Genbank
<i>Allisonella histaminiformans</i> DSM15230	Rúmen bovino	PRJNA254896	1.70	41	BioProject/ Hungate1000*
<i>Anaerovibrio lipolyticus</i> LB2005	Rúmen bovino	JHYA00000000.1	2.67	31	Genbank
<i>Anaerovibrio</i> sp. RM50	Rumen	JHWV00000000.1	2.78	40	Genbank
<i>Bacillus cereus</i> KPR-7A	Rúmen bovino	NZ_FMZR01000004.1	4.62	54	Genbank
<i>Bacillus licheniformis</i> VTM3R78	Rúmen de alce	NZ_FOF01000004.1	4.02	23	Genbank
<i>Bacillus</i> sp. MB2021	Rúmen	JNJJ00000000.1	5.03	166	Genbank
<i>Bacteroidales bacterium</i> KHT7	Rúmen bovino	FNBQ00000000.1	3.8	61	Genbank
<i>Bacteroidales bacterium</i> WCE2004	Rúmen bovino	Gp0112426	2.33	5	BioProject/ Hungate1000*
<i>Bacteroidales bacterium</i> WCE2008	Rúmen bovino	Gp0112427	2.38	8	BioProject/ Hungate1000*

<i>Bacteroides ovatus</i> NLAE-ZL-C500	Fezes de bovino	NZ_FMYE00000000.1	5.86	125	Genbank
<i>Bacteroides</i> sp. AR20	Rúmen de ovelhaRúmen de ovelha	FOBY00000000.1	6.07	28	Genbank
<i>Bacteroides</i> sp. AR29	Rúmen	PRJNA254898	4.44	61	BioProject/ Hungate1000*
<i>Bacteroides</i> sp. GA6A1	Rúmen	JNLA01000001.1	3.84	79	Genbank
<i>Bacteroides</i> sp. GA6A2	Rúmen	JNKX00000000.1	3.85	82	Genbank
<i>Bacteroides thetaiotaomicron</i> KPPR-3	Rúmen bovino	NZ_FOAL01000008.1	3.23	38	Genbank
<i>Bacteroides thetaiotaomicron</i> NLAE ZL C579	Fezes de bovino	FNVL00000000.1	6.33	78	Genbank
<i>Bacteroides xylanisolvens</i> NLAE-ZL- C202	Fezes de bovino	FOUM00000000.1	5.87	93	Genbank
<i>Bacteroides xylanisolvens</i> NLAE-ZL- G339	Fezes de cabrito	NZ_FNRP00000000.1	6.14	64	Genbank
<i>Basfia succiniciproducens</i> DSM 22022	Rúmen bovino	FMUQ00000000.1	2.23	29	Genbank
<i>Basfia succiniciproducens</i> KPR-2	Rúmen bovino	FOFH00000000.1	2.33	57	Genbank
<i>Bifidobacterium adolescentis</i> DSM 20087	Rúmen bovino	JNKM00000000.1	2.05	10	Genbank
<i>Bifidobacterium bifidum</i> CALF96	Fezes de bezerro	FOLU00000000.1	2.16	15	Genbank
<i>Bifidobacterium boum</i> DSM 20432	Rúmen bovino	JHWO00000000.1	2.16	26	Genbank
<i>Bifidobacterium breve</i> RP2	Fezes de bezerro	NZ_FNFW00000000.1	2.27	25	Genbank
<i>Bifidobacterium longum</i> AGR2137	Fezes de bezerro	ATWX00000000.1	2.27	49	Genbank
<i>Bifidobacterium merycicum</i> DSM 6492	Rúmen bovino	JDTL00000000.1	2.27	15	Genbank
<i>Bifidobacterium pseudologum</i> AGR2145	Fezes de bezerro	ATWW00000000.1	1.99	30	Genbank
<i>Bifidobacterium pseudologum globosum</i> DSM 20092	Rúmen bovino	JHWN00000000.1	1.91	40	Genbank

<i>Bifidobacterium ruminantium</i> DSM 6489	Rúmen bovino	JHWQ00000000.1	2.22	43	Genbank
<i>Bifidobacterium thermophilum</i> DSM 20212	Rúmen bovino	JHWM00000000.1	2.25	50	Genbank
<i>Bifidobacterium</i> sp. AGR2158	Fezes de bezerro	AUJM00000000.1	2.19	19	Genbank
<i>Blautia schinkii</i> DSM10518	Rúmen de ovelha	JNKJ00000000.1	6.68	65	Genbank
<i>Blautia</i> sp. SF50	Rúmen	SAMN02910433	3.69	78	BioSample/ Hungate1000*
<i>Blautia wexlerae</i> AGR2146	Fezes de bezerro	AUJF00000000.1	3.58	190	Genbank
<i>Butyrivibrio fibrisolvens</i> AB2020	Rúmen bovino	ATVZ00000000.1	4.74	170	Genbank
<i>Butyrivibrio fibrisolvens</i> AR40	Rúmen de ovelha	NZ_FOGJ00000000.1	5.02	77	Genbank
<i>Butyrivibrio fibrisolvens</i> FE2007	Rúmen bovino	AUJW00000000.1	4.31	67	Genbank
<i>Butyrivibrio fibrisolvens</i> MD2001	Rúmen bovino	AUKD00000000.1	4.73	67	Genbank
<i>Butyrivibrio fibrisolvens</i> WTE3004	Rúmen bovino	AUJV00000000.1	4.66	90	Genbank
<i>Butyrivibrio fibrisolvens</i> YRB2005	Rúmen bovino	AUJA00000000.1	4.65	82	Genbank
<i>Butyrivibrio hungatei</i> NK4A153	Rúmen de ovelha	AUJY00000000.1	3.37	38	Genbank
<i>Butyrivibrio hungatei</i> XBD2006	Rúmen	PRJNA254909	3.40	36	BioProject/ Hungate1000*
<i>Butyrivibrio hungatei</i> DSM 14810	Rúmen	NZ_FRDH00000000.1	3.4	22	Genbank
<i>Butyrivibrio proteoclasticus</i> B316	Rúmen	CP001812.1	3.62	NA	Genbank
<i>Butyrivibrio proteoclasticus</i> FD2007	Rúmen	JHYH00000000.1	3.88	43	Genbank
<i>Butyrivibrio proteoclasticus</i> P18	Rúmen de ovelha	FOXO00000000.1	4.17	76	Genbank
<i>Butyrivibrio</i> sp. AC2005	Rúmen bovino	AUJI00000000.1	5.12	92	Genbank

<i>Butyrivibrio</i> sp. AD3002	Rúmen bovino	ATVV00000000.1	4.32	73	Genbank
<i>Butyrivibrio</i> sp. AE2015	Rúmen bovino	ATVR00000000.1	3.71	73	Genbank
<i>Butyrivibrio</i> sp. AE2032	Rúmen bovino	JNLF00000000.1	3.69	6	Genbank
<i>Butyrivibrio</i> sp. AE 3003	Rúmen bovino	JMLX00000000.1	3.46	1	Genbank
<i>Butyrivibrio</i> sp. AE3009	Rúmen bovino	ATVS00000000.1	4.20	103	Genbank
<i>Butyrivibrio</i> sp. AE3004	Rúmen bovino	JNLQ00000000.1	4.48	6	Genbank
<i>Butyrivibrio</i> sp. AE3006	Rúmen bovino	AUJJ00000000.1	4.14	45	Genbank
<i>Butyrivibrio</i> sp. FC2001	Rúmen bovino	NZ_AUJH00000000.1	4.57	45	Genbank
<i>Butyrivibrio</i> sp. FCS006	Rúmen bovino	AUKB00000000.1	3.79	16	Genbank
<i>Butyrivibrio</i> sp. INIA14	Rúmen de ovelha	PRJNA254902	4.19	79	BioSample/ Hungate1000*
<i>Butyrivibrio</i> sp. INIA16	Rúmen	SAMN02910263	4.79	133	BioSample/ Hungate1000*
<i>Butyrivibrio</i> sp. INIA18	Rúmen	PRJNA243945	3.25	34	BioSample/ Hungate1000*
<i>Butyrivibrio</i> sp. INIA21	Rúmen de ovelha	PRJNA254905	3.35	73	BioSample/ Hungate1000*
<i>Butyrivibrio</i> sp. LB2008	Rúmen bovino	JHXZ00000000.1	3.72	30	Genbank
<i>Butyrivibrio</i> sp. LC3010	Rúmen bovino	AUJU00000000.1	4.58	69	Genbank
<i>Butyrivibrio</i> sp. MB2005	Rúmen bovino	AUJP00000000.1	4.17	47	Genbank
<i>Butyrivibrio</i> sp. MC2013	Rúmen bovino	AUKE00000000.1	3.81	50	Genbank
<i>Butyrivibrio</i> sp. MC2021	Rúmen bovino	JHXX00000000.1	5.03	32	Genbank
<i>Butyrivibrio</i> sp. NC2002	Rúmen	JNKZ00000000.1	3.40	8	Genbank
<i>Butyrivibrio</i> sp. NC2007	Rúmen bovino	ATWY00000000.1	4.00	54	Genbank
<i>Butyrivibrio</i> sp. NC3005	Rúmen bovino	AUKC00000000.1	3.87	97	Genbank
<i>Butyrivibrio</i> sp. OB235	Rúmen bovino	FOBE00000000.1	4.73	83	Genbank

<i>Butyrivibrio</i> sp. SU6	Rúmen	SAMN02910276	3.69	23	BioSample/ Hungate1000*
<i>Butyrivibrio</i> sp. TB	Rúmen	SAMN02910382	4.40	56	BioSample/ Hungate1000*
<i>Butyrivibrio</i> sp. VCB2001	Rúmen bovino	AUIZ00000000.1	4.01	44	Genbank
<i>Butyrivibrio</i> sp. VCB2006	Rúmen bovino	ATVX00000000.1	3.95	40	Genbank
<i>Butyrivibrio</i> sp. WCD2001	Rúmen bovino	AUJZ00000000.1	4.27	60	Genbank
<i>Butyrivibrio</i> sp. WCD3002	Rúmen bovino	AUJO00000000.1	4.25	54	Genbank
<i>Butyrivibrio</i> sp. WCE2006	Rúmen bovino	JNKH00000000.1	4.52	94	Genbank
<i>Butyrivibrio</i> sp. XBB1001	Rúmen bovino	AUKA00000000.1	3.91	41	Genbank
<i>Butyrivibrio</i> sp. XPD2002	Rúmen bovino	AUIY00000000.1	4.44	52	Genbank
<i>Butyrivibrio</i> sp. XPD2006	Rúmen bovino	ATVT00000000.1	3.88	31	Genbank
<i>Butyrivibrio</i> sp. YAB3001	Rúmen	SAMN02910398	4.58	56	BioSample/ Hungate1000*
<i>Butyrivibrio</i> sp. VCD2006	Rúmen bovino	AUIX00000000.1	4.70	81	Genbank
<i>Cellulomonas</i> sp. KH9	Rúmen bovino	FOSE00000000.1	4.17	11	Genbank
<i>Citrobacter</i> sp. NLAE-ZL-C269	Fezes de bovino	FOJV00000000.1	4.96	24	Genbank
<i>Clostridiales bacterium</i> NK3B98	Rúmen de ovelha	ATWA00000000.1	3.04	211	Genbank
<i>Clostridiales bacterium</i> WTE2008	Rúmen bovino	Gp0104616	3.41	16	Hungate1000*
<i>Clostridium aerotolerans</i> DSM 5434	Rúmen de ovelha	JHWJ00000000.1	4.73	50	Genbank
<i>Clostridium algidicarnis</i> B3	Rúmen	JNLN00000000.1	3.06	1	Genbank
<i>Clostridium aminophilum</i> DSM 10710	Rúmen bovino	PRJNA234874	3.12	37	BioSample/ Hungate1000*
<i>Clostridium aminophilum</i> F	Rúmen bovino	NZ_FOZC01000002.1	3.26	31	Genbank

<i>Clostridium aminophilum</i> KH1P1	Rúmen	NZ_FOIL00000000.1	3.2	108	Genbank
<i>Clostridium beijerinckii</i> HUN142	Rúmen	JH XK00000000.1	6.10	53	Genbank
<i>Clostridium butyricum</i> AGR2140	Fezes de bezerro	AUJN00000000.1	4.55	41	Genbank
<i>Clostridium cadaveris</i> AGR2141	Fezes de bezerro	AUJL00000000.1	3.54	45	Genbank
<i>Clostridium clostridioforme</i> AGR2157	Fezes de bezerro	AUJR00000000.1	4.94	138	Genbank
<i>Clostridium clostridioforme</i> ATCC 25537	Rúmen	Gp0088045	5.46	148	Hungate1000*
<i>Clostridium glycolicum</i> KPPR-9	Rúmen	SAMN02910355	3.98	18	BioSample/ Hungate1000*
<i>Clostridium innocuum</i> NLAE-ZL-C381	Fezes de bovino	FOTN00000000.1	4.24	56	Genbank
<i>Clostridium lundense</i> DSM 17049	Rúmen de ovelha	JHVC00000000.1	4.75	52	Genbank
<i>Clostridium manganotii</i> LM2	Fezes de cordeiro	JIAA00000000.1	3.02	1	Genbank
<i>Clostridium paraputrificum</i> AGR2156	Fezes de bezerro	AUJC00000000.1	3.56	32	Genbank
<i>Coriobacteriaceae bacterium</i> KH1P3	Rúmen bovino	Gp0104622	2.03	15	Hungate1000*
<i>Corynebacterium vitaeruminis</i> GA6A13	Rúmen	JNKV00000000.1	2.90	69	Genbank
<i>Denitrobacterium detoxificans</i> DSM 21843	Rúmen bovino	NZ_FOEC00000000.1	2.41	38	Genbank
<i>Desulfotomaculum ruminis</i> DSM 2154	Rúmen de ovelha	CP002780.1	3.97	NA	Genbank
<i>Desulfovibrio desulfuricans desulfuricans</i> ATCC 27774	Rúmen de ovelha	CP001358.1	2.87	NA	Genbank
<i>Desulfovibrio desulfuricans</i> DSM 7057	Rúmen bovino	NZ_FPIW00000000.1	3.23	42	Genbank
<i>Desulfovibrio legallii</i> KHC7	Rúmen bovino	FNBX00000000.1	2.71	36	Genbank
<i>Dorea longicatena</i> AGR2136	Fezes de bezerro	AUJS00000000.1	3.59	144	Genbank
<i>Dorea</i> sp. AGR2135	Fezes de bezerro	ATVU00000000	3.50	73	Genbank
<i>Enterobacter</i> sp. KPR 6	Rúmen bovino	FOYH00000000.1	4.49	18	Genbank

<i>Enterococcus casseliflavus</i> NLAE-ZL-C414	Fezes de bovino	NZ_FNFS00000000.1	3.57	44	Genbank
<i>Enterococcus casseliflavus</i> NLAE-ZL-G268	Fezes de cabrito	FOMP00000000.1	3.69	28	Genbank
<i>Enterococcus faecalis</i> 68a	Rúmen de ovelha	JOOC00000000.1	3.04	9	Genbank
<i>Enterococcus faecalis</i> VTM1R91	Rúmen de alce	NZ_FNIP01000007.1	3.06	39	Genbank
<i>Enterococcus gallinarum</i> SKF1	Rúmen de ovelha	JNLR00000000.1	3.47	6	Genbank
<i>Enterococcus mundtii</i> C2	Rúmen bovino	NZ_FOUC01000012.1	3.25	49	Genbank
<i>Enterococcus</i> sp. KPPR-6	Rúmen bovino	FOHR00000000.1	4.45	68	Genbank
<i>Erysipelotrichaceae bacterium</i> NK3D112	Rúmen de ovelha	JNJQ00000000.1	3.14	97	Genbank
<i>Escherichia coli</i> PA-3	Rúmen bovino	NZ_FOEU01000001.1	3.58	78	Genbank
<i>Eubacterium celulosolvens</i> 6	Rúmen de ovelha	AEOA00000000.1	3.38	7	Genbank
<i>Eubacterium celulosolvens</i> LD2006	Rúmen	JHXY00000000.1	3.33	77	Genbank
<i>Eubacterium eligenes</i> ATCC 27750	Rúmen	CP001104.1	2.15	NA	Genbank
<i>Eubacterium oxidoreducens</i> DSM 3217	Rúmen bovino	FMXR00000000.1	2.92	33	Genbank
<i>Eubacterium pyruvativorans</i> KHPC4	Rúmen bovino	FOWF00000000.1	2.27	75	Genbank
<i>Eubacterium rectale</i> ATCC 33656	Rúmen	NC_012781.1	3.45	NA	Genbank
<i>Eubacterium ruminantium</i> FB3002	Rúmen bovino	FNUU00000000.1	3.51	40	Genbank
<i>Eubacterium ruminatum</i> HUN269	Rúmen bovino	PRJNA254842	2.88	36	BioSample/ Hungate1000*
<i>Eubacterium</i> sp. AB3007	Rúmen bovino	NZ_JIAD00000000.1	2.30	1	Genbank
<i>Fibrobacter succinogenes</i> HM2	Rúmen	1034376 (JGI)	3.51	9	Hungate1000

<i>Fibrobacter succinogenes</i> S85	Rúmen	CP002158.1	3.84	NA	Genbank
<i>Fusobacterium necrophorum</i> HUN048	Rúmen	JHWT00000000.1	2.03	17	Genbank
<i>Kandleria vitulina</i> DSM 20405	Rúmen bovino	JNKN00000000.1	2.12	75	Genbank
<i>Kandleria vitulina</i> MC3001	Rúmen	JHWI00000000.1	2.04	54	Genbank
<i>Kandleria vitulina</i> S3B	Rúmen de ovelha	NZ_FNNF00000000.1	2.13	61	Genbank
<i>Kandleria vitulina</i> WCE2011	Rúmen	JHXA00000000.1	2.03	41	Genbank
<i>Lachnobacterium bovis</i> AE2004	Rúmen bovino	NZ_JHWB00000000.1	2.85	22	Genbank
<i>Lachnobacterium bovis</i> C6A12	Rúmen	JHWP00000000.1	2.60	45	Genbank
<i>Lachnobacterium bovis</i> DSM 14045	Rúmen	SAMN02910414	2.71	58	BioSample/ Hungate1000*
<i>Lachnobacterium bovis</i> NK4B19	Rúmen de ovelha	ATWB00000000.1	2.51	130	Genbank
<i>Lachnobacterium bovis</i> S1B	Rúmen de ovelha	NZ_FOGW00000000.1	2.54	66	Genbank
<i>Lachnoclostridium aminophilum</i> KH1P1	Rúmen bovino	NZ_FOIL00000000.1	3.12	108	Genbank
<i>Lachnoclostridium clostridioforme</i> NLAE-ZL-C196	Fezes de bovino	NZ_FOIO00000000.1	5.23	164	Genbank
<i>Lachnoclostridium lavalense</i> NLAE-ZL-G277	Fezes de cabrito	FOIM00000000.1	6.4	97	Genbank
<i>Lachnoclostridium oroticum</i> NLAE-ZL-C242	Fezes de bovino	PRJNA322955	5.72	33	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium</i> AC2028	Rúmen bovino	PRJNA223007	2.78	14	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium</i> MA2020	Rúmen bovino	JQKK00000000.1	3.72	50	Genbank
<i>Lachnospiraceae bacterium</i> XPB1003	Rúmen bovino	FMVK00000000.1	3.15	42	Genbank

<i>Lachnospiraceae bacterium A10</i>	Rúmen bovino	FNYN00000000.1	2.52	13	Genbank
<i>Lachnospiraceae bacterium AB2028</i>	Rúmen bovino	JHYJ00000000.1	3.47	25	Genbank
<i>Lachnospiraceae bacterium AC2012</i>	Rúmen bovino	JNIN00000000.1	2.55	94	Genbank
<i>Lachnospiraceae bacterium AC 2014</i>	Rúmen bovino	JNJE00000000.1	2.75	14	Genbank
<i>Lachnospiraceae bacterium AC2029</i>	Rúmen bovino	NZ_JNLK00000000.1	3.86	1	Genbank
<i>Lachnospiraceae bacterium AC2031</i>	Rúmen bovino	JIAB00000000.1	3.19	1	Genbank
<i>Lachnospiraceae bacterium AC3007</i>	Rúmen	JHYK00000000.1	2.94	51	Genbank
<i>Lachnospiraceae bacterium AD3010</i>	Rúmen bovino	JHWU00000000.1	3.54	46	Genbank
<i>Lachnospiraceae bacterium C10</i>	Rúmen bovino	PRJNA254845	2.95	70	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium C6A11</i>	Rúmen	JNKY00000000.1	2.89	17	Genbank
<i>Lachnospiraceae bacterium C7</i>	Rúmen bovino	FOPE00000000.1	2.85	91	Genbank
<i>Lachnospiraceae bacterium FD2005</i>	Rúmen	JNKT00000000.1	2.44	9	Genbank
<i>Lachnospiraceae bacterium FE2018</i>	Rúmen	JNKK00000000.1	3.21	119	Genbank
<i>Lachnospiraceae bacterium G11</i>	Rúmen bovino	FMWZ00000000.1	3.05	31	Genbank
<i>Lachnospiraceae bacterium G41</i>	Rúmen bovino	FNEU00000000.1	2.6	22	Genbank
<i>Lachnospiraceae bacterium MC2017</i>	Rúmen bovino	JNJK00000000.1	4.25	56	Genbank
<i>Lachnospiraceae bacterium MD2004</i>	Rúmen bovino	JHWK00000000.1	2.53	10	Genbank
<i>Lachnospiraceae bacterium NC2008</i>	Rúmen bovino	JHWS00000000.1	2.71	26	Genbank
<i>Lachnospiraceae bacterium ND2006</i>	Rúmen	SAMN02910369	3.94	103	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium NE2001</i>	Rúmen	SAMN02910369	3.72	62	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium NK4A136</i>	Rúmen de ovelha	ATVW00000000.1	2.73	78	Genbank
<i>Lachnospiraceae bacterium NK4A179</i>	Rúmen de	ATWC00000000	3.22	73	Genbank

	ovelha				
<i>Lachnospiraceae bacterium NKA144</i>	Rúmen de ovelha	AUJT00000000.1	4.09	72	Genbank
<i>Lachnospiraceae bacterium P6A3</i>	Rúmen	NZ_JNKP00000000.1	2.98	9	Genbank
<i>Lachnospiraceae bacterium P6B14</i>	Rúmen	JHWZ00000000.1	2.68	36	Genbank
<i>Lachnospiraceae bacterium RM5</i>	Rúmen	AB730673.1	2.35	23	Genbank
<i>Lachnospiraceae bacterium V9D3004</i>	Rúmen	JHWW00000000.1	3.41	32	Genbank
<i>Lachnospiraceae bacterium XBB2008</i>	Rúmen	SAMN02910292	3.52	24	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium XBD2001</i>	Rúmen	SAMN02910301	2.48	17	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium YSB2008</i>	Rúmen	JHWR00000000.1	2.68	29	Genbank
<i>Lachnospiraceae bacterium YSD2013</i>	Rúmen	SAMN02910339	3.24	54	BioSample/ Hungate1000*
<i>Lachnospiraceae bacterium NC2004</i>	Rúmen	NZ_JHXB00000000.1	3.38	56	Genbank
<i>Lachnospira multipara LB2003</i>	Rúmen bovino	JNKW00000000.1	2.57	21	Genbank
<i>Lachnospira multipara MC2003</i>	Rúmen bovino	JHWY00000000.1	2.60	26	Genbank
<i>Lachnospira mutipara ATCC 19207</i>	Rúmen	AUJG00000000.1	2.87	32	Genbank
<i>Lachnospira multipara D15d</i>	Rúmen bovino	FNUL00000000.1	2.85	35	Genbank
<i>Lachnospira pectinoschiza M83</i>	Estômago de camelo	NZ_FNHZ00000000.1	2.71	12	Genbank
<i>Lactobacillus mucosae WCC8</i>	Rúmen bovino	NZ_FNIH00000000.1	1.89	39	Genbank
<i>Lactobacillus mucosae AGR63</i>	Rúmen bovino	SAMN02744693	1.94	9	BioSample/ Hungate1000*
<i>Lactobacillus plantarum AG30</i>	Rúmen de ovelha	JHWA00000000.1	3.44	48	Genbank

<i>Lactobacillus ruminis</i> ATCC 27782	Rúmen bovino	CP003032.1	2.07	NA	Genbank
<i>Lactobacillus ruminis</i> DSM 20403	Rúmen bovino	PRJNA254854	2.04	58	BioSample/ Hungate1000*
<i>Lactobacillus ruminis</i> ATCC 25644	Rúmen bovino	AFYE00000000	2.07	50	Genbank
<i>Lactobacillus brevis</i> AG48	Rúmen de ovelha	JAGR00000000.1	2.59	9	Genbank
<i>Lactococcus garvieae</i> M79	Estômago de camelo	FOTJ00000000.1	2.17	36	Genbank
<i>Lactococcus lactis</i> 511	Rúmen bovino	JNLP00000000.1	2.48	NA	Genbank
<i>Mannheimia succiniproducens</i> MBEL55E	Rúmen de cabrito	NC_006300.1	2.31	NA	Genbank
<i>Megamonas</i> sp. CALF98-2	Fezes de bezerros	FOCY00000000.1	2.22	68	Genbank
<i>Megasphaera elsdenii</i> DSM 20460	Rúmen de ovelha	NC_015873.1	2.47	NA	Genbank
<i>Megasphaera elsdenii</i> YE34	Rúmen bovino	FOQP00000000.1	2.42	34	Genbank
<i>Megasphaera elsdenii</i> j1	Rúmen de ovelha	NZ_FRAK01000016.1	2.46	33	Genbank
<i>Megasphaera elsdenii</i> t81	Rúmen	JHXC00000000.1	2.50	36	Genbank
<i>Methanobrevibacter millerae</i> SMA	Rúmen	NZ_CP011266.1	2.55	NA	Genbank
<i>Methanobrevibacter millerae</i> DSM 16643	Rúmen	SAMN02910315	2.72	48	BioSample/ Hungate1000*
<i>Methanobrevibacter olleyae</i> DSM 16632	Rúmen de ovelha	FOTL00000000.1	2.13	49	Genbank
<i>Methanobrevibacter wolinii</i> SH	Rúmen	JHWX00000000.1	2.04	26	Genbank

<i>Methanobrevibacter ruminantium</i> M1	Rúmen	NC_013790.1	2.94	NA	Genbank
<i>Methanomicrobium mobile</i> DSM 1539	Rúmen bovino	JOMF00000000.1	1.71	12	Genbank
<i>Methanosarcina</i> sp. DSM 11855	Rúmen de ovelha	PRJNA254857	3.10	23	BioSample/ Hungate1000*
<i>Methanospirillum hungatei</i> JF-1	Rúmen	CP000254.1	3.55	NA	Genbank
<i>Mitsuokella jalaludinii</i> DSM 13811	Rúmen bovino	JNKR00000000.1	2.41	66	Genbank
<i>Morganella morgani</i> NLAE-ZL-C84	Rúmen	NZ_FRBZ00000000.1	3.92	18	Genbank
<i>Olsenella</i> sp. KH2P3	Rúmen bovino	NZ_FPKD00000000.1	2.42	29	Genbank
<i>Olsenella umbonata</i> DSM 22619	Rúmen de ovelha	FMZL00000000.1	2.22	49	Genbank
<i>Olsenella umbonata</i> KHGC19	Rúmen bovino	FOGP00000000.1	2.06	13	Genbank
<i>Oribacterium</i> sp. FC2011	Rúmen bovino	JNJM00000000.1	4.16	67	Genbank
<i>Oribacterium</i> sp. NK2B42	Rúmen de ovelha	AUJX00000000	3.79	102	Genbank
<i>Oribacterium</i> sp. P6A1	Rúmen	NZ_JNKO00000000.1	3.69	95	Genbank
<i>Oscillibacter ruminantium</i> GH1	Rúmen bovino	BAGW01000001	3.11	29	Genbank
<i>Pediococcus acidilactici</i> AGR201	Rúmen de ovelha	JAGU00000000.1	1.92	5	Genbank
<i>Pepstreptococcus anaerobius</i> C	Rúmen	PRJNA254856	2.11	60	BioSample/ Hungate1000*
<i>Peptostreptococcaceae bacterium</i> pGA-8	Rúmen bovino	FONK00000000.1	1.79	24	Genbank
<i>Peptostreptococcaceae bacterium</i> VA2	Fezes de cervo Fezes de cervo	JMMB00000000.1	3.58	11	Genbank
<i>Peptostreptococcus</i> sp. D1	Faeces of deer	FONP00000000.1	2.28	37	Genbank
<i>Preovetella</i> sp. RM4	Rumen	JNKG00000000.1	3.35	15	Genbank

<i>Prevotella albensis</i> DSM 11370	Rúmen de ovelha	PRJNA195800	2.67	45	BioSample/ Hungate1000*
<i>Prevotella brevis</i> P6B11	Rúmen	JHXG00000000.1	3.21	22	Genbank
<i>Prevotella bryantii</i> B14	Rúmen bovino	ADWO00000000.1	3.59	98	Genbank
<i>Prevotella bryantii</i> C21A	Rúmen bovino	AUKF00000000.1	3.24	27	Genbank
<i>Prevotellaceae bacterium</i> HUN156	Rúmen bovino	NZ_FPIT00000000.1	2.8	19	Genbank
<i>Prevotellaceae bacterium</i> MN60	Rúmen de ovelha	Gp0110665	2.23	24	Hungate
<i>Prevotella ruminicola</i> 23	Rúmen	CP002006.1	3.62	NA	Genbank
<i>Prevotella ruminicola</i> GA6B6	Rúmen	JHXD00000000.1	3.47	52	Genbank
<i>Prevotella</i> sp. AGR2160	Fezes de bezerro	AUJK00000000.1	2.90	50	Genbank
<i>Prevotella</i> sp. FD3004	Rúmen bovino	JQKR00000000.1	2.95	14	Genbank
<i>Prevotella</i> sp. HUN102	Rúmen bovino	JIAF00000000.1	3.31	4	Genbank
<i>Prevotella</i> sp. KH1P2	Rúmen bovino	FOHY00000000.1	3.29	40	Genbank
<i>Prevotella</i> sp. KHP7	Rúmen bovino	FOIV00000000.1	3.09	8	Genbank
<i>Prevotella</i> sp. LC2012	Rúmen bovino	FNUF00000000.1	2.89	9	Genbank
<i>Prevotella</i> sp. MA2016	Rúmen bovino	JHUW00000000.1	3.38	12	Genbank
<i>Prevotella</i> sp. NE3005	Rúmen bovino	FOCK00000000.1	3.57	21	Genbank
<i>Prevotella</i> sp. P6B1	Rúmen	JNKF00000000.1	3.61	18	Genbank
<i>Prevotella</i> sp. P6B4	Rúmen	JHVE00000000.1	3.55	49	Genbank
<i>Propionibacteriaceae bacterium</i> P6A17	Rúmen	JHVD00000000.1	2.73	48	Genbank
<i>Propionibacterium</i> sp. MB3007	Rúmen	JMLB00000000.1	2.53	17	Genbank
<i>Proteinielasticum ruminis</i> DSM 24773	Rúmen	JNKC00000000.1	3.11	21	Genbank
<i>Pseudobutyrvibrio ruminis</i> ACV9	Rúmen	SAMN02910377	2.76	26	BioSample/ Hungate1000*
<i>Pseudobutyrvibrio ruminis</i> AD2017	Rúmen bovino	AUJQ00000000	2.91	22	Genbank

<i>Pseudobutyrvibrio ruminis</i> DSM 9787	Rúmen	SAMN02910411	3.02	18	BioSample/ Hungate1000*
<i>Pseudobutyrvibrio ruminis</i> HUN009	Rúmen de ovelha	JNLH00000000.1	2.81	2	Genbank
<i>Pseudobutyrvibrio</i> sp. 49	Rúmen bovino	FNDP00000000.1	3.40	44	Genbank
<i>Pseudobutyrvibrio</i> sp. ACV-2	Rúmen bovino	FNQZ00000000.1	2.67	45	Genbank
<i>Pseudobutyrvibrio</i> sp. AR14	Rúmen bovino	PRJNA254864	3.05	46	BioSample/ Hungate1000*
<i>Pseudobutyrvibrio</i> sp. C4	Rúmen bovino	SAMN029104	2.91	13	BioSample/ Hungate1000*
<i>Pseudobutyrvibrio</i> sp. LB2021	Rúmen bovino	NZ_JNJS00000000.1	2.81	9	Genbank
<i>Pseudobutyrvibrio</i> sp. MD2005	Rúmen bovino	JHXE00000000.1	3.08	26	Genbank
<i>Pseudobutyrvibrio</i> sp. NOR37	Rúmen bovino	FOZB00000000.1	2.95	8	Genbank
<i>Pseudobutyrvibrio</i> sp. UC1225	Rúmen bovino	FOVQ00000000.1	3.38	32	Genbank
<i>Pseudobutyrvibrio</i> sp. YE44	Rúmen	SAMN02910298	3.24	28	BioSample/ Hungate1000*
<i>Ruaniaceae bacterium</i> KH17	Rúmen bovino	Gp0110682	2.73	72	Hungate
<i>Ruminococcus albus</i> 7	Rúmen	CP002403.1	4.48	NA	Genbank
<i>Ruminobacter amylophilus</i> DSM 1361	Rúmen	SAMN02910344	2.81	117	BioSample/ Hungate1000*
<i>Ruminobacter</i> sp. RM87	Rúmen bovino	JNKD00000000.1	2.86	123	Genbank
<i>Ruminococcaceae bacterium</i> YAD3003	Rúmen	SAMN02910264	2.67	30	BioSample/ Hungate1000*
<i>Ruminococcaceae bacterium</i> AB4001	Rúmen bovino	JHXJ00000000.1	3.24	54	Genbank
<i>Ruminococcaceae bacterium</i> AE2021	Rúmen bovino	JIAE00000000.1	3.08	NA	Genbank
<i>Ruminococcaceae bacterium</i> D5	Rúmen de cervo	FORK00000000.1	2.64	49	Genbank

<i>Ruminococcaceae bacterium</i> FB2012	Rúmen	SAMN02910317	3.75	52	BioSample/ Hungate1000*
<i>Ruminococcaceae bacterium</i> KHP2	Rúmen bovino	Gp0104832	2.98	20	Hungate1000
<i>Ruminococcaceae bacterium</i> P7	Rúmen bovino	FMUA00000000.1	3.06	87	Genbank
<i>Ruminococcaceae bacterium</i> YRB3002	Rúmen	SAMN02910456	3.01	51	BioSample/ Hungate1000*
<i>Ruminococcus albus</i> 8	Rúmen	ADKM00000000	4.05	136	Genbank
<i>Ruminococcus albus</i> AD2013	Rúmen bovino	JAGS00000000.1	4.23	6	Genbank
<i>Ruminococcus albus</i> AR67	Rúmen	SAMN02910406	4.27	103	BioSample/ Hungate1000*
<i>Ruminococcus albus</i> DSM 20455	Rúmen	PRJNA221059	4.34	120	BioSample/ Hungate1000*
<i>Ruminococcus albus</i> KH2T6	Rúmen bovino	NZ_FOAT00000000.1	3.85	47	Genbank
<i>Ruminococcus albus</i> SY3	Rúmen	JE0B00000000.1	4.28	81	Genbank
<i>Ruminococcus bromii</i> YE282	Rúmen bovino	FMUV00000000.1	2.54	81	Genbank
<i>Ruminococcus flavefaciens</i> 007C	Rúmen bovino	ATAX00000000	3.65	39	Genbank
<i>Ruminococcus flavefaciens</i> 17	Rumen bovino	AFNE00000000	3.45	489	Genbank
<i>Ruminococcus flavefaciens</i> AE3010	Rúmen bovino	JAGT00000000.1	3.70	3	Genbank
<i>Ruminococcus flavefaciens</i> ATCC 19208	Rúmen	KI912489.1	3.58	25	Genbank
<i>Ruminococcus flavefaciens</i> FD1	Rúmen bovino	ACOK00000000	4.57	119	Genbank
<i>Ruminococcus flavefaciens</i> MA2007	Rúmen bovino	JHXF00000000.1	3.39	29	Genbank
<i>Ruminococcus flavefaciens</i> ND2009	Rúmen bovino	JHXI00000000.1	3.63	33	Genbank
<i>Ruminococcus flavefaciens</i> SAB67	Rúmen	PRJNA254877	3.84	28	BioSample/ Hungate1000
<i>Ruminococcus flavefaciens</i> XPD3002	Rúmen bovino	NZ_FPJT00000000.1	3.65	41	Genbank
<i>Ruminococcus flavefaciens</i> YL228	Rúmen	SAMN02910280	3.36	15	BioSample/

Hungate1000*

<i>Ruminococcus flavefaciens</i> YRD2003	Rúmen bovino	NZ_FNZT00000000.1	3.56	43	Genbank
<i>Ruminococcus gnavus</i> AGR2154	Fezes de bezerro	JAGQ00000000.1	3.72	5	Genbank
<i>Ruminococcus</i> sp. FC2018	Rúmen bovino	JHXH00000000.1	2.40	32	Genbank
<i>Ruminococcus</i> sp. HUN007	Rúmen bovino	JOOA00000000.1	4.21	4	Genbank
<i>Ruminococcus</i> sp. NK3A76	Rúmen de ovelha	JMMA00000000.1	3.57	2	Genbank
<i>Ruminococcus</i> sp. YE71	Rúmen	SAMN02910447	4.07	43	BioSample/ Hungate1000*
<i>Ruminococcus</i> sp. ye78	Rúmen	SAMN02910446	4.14	53	BioSample/ Hungate1000*
<i>Sarcina</i> sp. DSM 11001	Rúmen bovino	FNFZ00000000.1	4.16		Genbank
<i>Selenomonas bovis</i> 8-14-1	Rúmen	JNKB00000000.1	2.61	115	Genbank
<i>Selenomonas ruminantium</i> AC2024	Rúmen bovino	NZ_JIAC00000000.1	2.84	NA	Genbank
<i>Selenomonas ruminantium</i> AB3002	Rúmen bovino	JNIO00000000.1	3.66	8	Genbank
<i>Selenomonas ruminantium</i> C3	Rúmen	SAMN02910323	2.77	18	BioSample/ Hungate1000*
<i>Selenomonas ruminantium</i> GACV-9	Rúmen	SAMN02910356	2.82	36	BioSample/ Hungate1000*
<i>Selenomonas ruminantium</i> L14	Estômago de camelo	NZ_FOJX01000010.1	3.30	24	Genbank
<i>Selenomonas ruminantium lactilytica</i> DSM 2872	Rúmen	PRJNA254915	3.00	22	BioSample/ Hungate1000*
<i>Selenomonas ruminantium lactilytica</i> TAM 6421	Rúmen de ovelha	AP012292.1	3.63	NA	Genbank
<i>Selenomonas ruminantium</i> z108	Rúmen de	NZ_FOQK00000000.1	3.12	52	Genbank

	ovelha				
<i>Selenomonas ruminantium</i> ATCC 12561	Rúmen bovino	AUJE00000000	3.38	71	Genbank
<i>Selenomonas</i> sp. AE3005	Rúmen	JMME00000000.1	2.82	25	Genbank
<i>Selenomonas</i> sp. FC4001	Rúmen	JNJF00000000.1	2.88	22	Genbank
<i>Selenomonas</i> sp. ND2010	Rúmen	JNKI00000000.1	2.86	63	Genbank
<i>Sharpea azabuensis</i> DSM 18934	Rúmen	JNKU00000000.1	2.41	91	Genbank
<i>Sharpea azabuensis</i> KH1P5	Rúmen de ovelha	NZ_FOTA00000000.1	3.12	52	Genbank
<i>Sharpea azabuensis</i> KH2P10	Rúmen bovino	FOMV00000000.1	2.33	72	Genbank
<i>Shigella sonnei</i> NLAE-ZL-G496	Rúmen	NZ_FNIS00000000.1	4.92	101	Genbank
<i>Slackia heliotrinireducens</i> DSM 20476	Rúmen de ovelha	NC_013165.1	3.17	NA	Genbank
<i>Staphylococcus epidermidis</i> AG42	Rúmen de ovelha	JNLI00000000.1	2.56	9	Genbank
<i>Staphylococcus epidermidis</i> NLAE-ZL-G239	Fezes de cabrito	FOPD00000000.1	2.48	30	Genbank
<i>Streptococcus equinus</i> 2B	Rúmen	JNKQ00000000.1	1.91	9	Genbank
<i>Streptococcus equinus</i> AG46	Rúmen de ovelha	SAMN02841210	1.93	1	BioSample/ Hungate1000*
<i>Streptococcus equinus</i> B315	Rúmen bovino	AUJD00000000.1	1.78	12	Genbank
<i>Streptococcus equinus</i> JB1	Rúmen bovino	AUZH00000000.1	1.95	45	Genbank
<i>Streptococcus equinus</i> HC5	Rúmen bovino	JPGC00000000.1	1.85	8	Genbank
<i>Streptococcus equinus</i> SN033	Fezes de cervo	ATWZ00000000.1	1.84	13	Genbank
<i>Streptococcus equinus</i> AR3	Rúmen de ovelha	PRJNA322972	2.03	16	BioSample/ Hungate1000*
<i>Streptococcus equinus</i> GA-1	Rúmen	SAMN02910449	1.79	11	BioSample/

Hungate1000*

<i>Streptococcus equinus</i> MPR1	Estômago de camelo	NZ_FOFQ00000000.1	1.87	23	Genbank
<i>Streptococcus equinus</i> MPR2	Estômago de camelo	NZ_FNEP00000000.1	1.87	23	Genbank
<i>Streptococcus equinus</i> PGA-7	Rúmen bovino	NZ_FNJX00000000.1	1.80	6	Genbank
<i>Streptococcus equinus</i> PR-5	Rúmen bovino	NZ_FNZW00000000.1	1.82	11	Genbank
<i>Streptococcus equinus</i> SB04	Rúmen bovino	NZ_FNJK00000000.1	1.87	21	Genbank
<i>Streptococcus equinus</i> SB20	Rúmen bovino	FOZA00000000.1	1.89	9	Genbank
<i>Streptococcus equinus</i> SI	Rúmen de ovelha	FOYC00000000.1	1.92	11	Genbank
<i>Streptococcus gallolyticus</i> LMG 15572	Rúmen de cabrito	FPBN00000000.1	2.37	24	Genbank
<i>Streptococcus gallolyticus</i> VTM 1R27	Rúmen	SAMN02910295	2.18	19	BioSample/ Hungate1000*
<i>Streptococcus gallolyticus</i> VTM 2R47	Rúmen de alce	NZ_FOGM00000000.1	2.18	33	Genbank
<i>Streptococcus gallolyticus</i> VTM3R42	Rúmen de alce	NZ_FNFJ00000000.1	2.18	14	Genbank
<i>Streptococcus henry</i> A4	Rúmen	SAMN02910293	2.42	63	BioSample/ Hungate1000*
<i>Succiniclasticum ruminis</i> DSM11005	Rúmen bovino	FMYW00000000.1	2.91	27	Genbank
<i>Succinivibrio dextrinosolvens</i> ACV-10	Rúmen bovino	FPAH00000000.1	2.97	94	Genbank
<i>Succinivibrio dextrinosolvens</i> H5	Ovine rumen	JNKL00000000.1	2.67	105	Genbank
<i>Treponema bryantii</i> NK4A124	Rúmen de ovelha	ATWV00000000.1	3.31	28	Genbank
<i>Treponema bryantii</i> XBD1002	Rúmen bovino	FORI00000000.1	3.20	24	Genbank
<i>Treponema sp.</i> JC4	Rúmen bovino	AJGU01000000.1	3.03	147	Genbank

<i>Treponema saccharophilum</i> DSM 2985	Rúmen bovino	AGRW00000000.1	3.45	56	Genbank
<i>Treponema</i> sp. c6a8	Rúmen bovino	JHVB00000000.1	2.86	23	Genbank
<i>Wolinella succinogenes</i> DSM 1740	Rúmen bovino	BX571656.1	2.11	NA	Genbank

*The genomes were obtained from Genome portal of the JGI genome (The Hungate1000 Project). NA: not applicable

Tabela S2: Banco de dados 1 (sequências nucleotídicas de 4 genes)

Peptídeo laço	Genes	Microrganismo produtor	Número de acesso	Referências
Microcin j25	mcjA	<i>Escherichia coli</i> AY25	AAD28494.1	Salomon e Farias (1992)
	mcjB	<i>Escherichia coli</i> AY25	AAD28495.1	Salomon e Farias (1992)
	mcjC	<i>Escherichia coli</i> AY25	AAD28496.1	Salomon e Farias (1992)
	mcjD	<i>Escherichia coli</i> AY25	AAD28494.1	Salomon e Farias (1992)

Tabela S3: Banco de dados 2 (sequências de aminoácidos de 20 genes relacionados a peptídeos laço)

Peptídeo laço	Genes	Microorganismo produtor	Número de acesso	Referências
Astexin I	<i>atxA1</i>	<i>Asticcacaulis excentricus</i> CB 48	PDB: 2LTI_A	Zimmermann et al. (2013)
Astexin I	<i>atxB1</i>	<i>Asticcacaulis excentricus</i> CB 48	E8RMD3	Zimmermann et al. (2013)
Astexin III	<i>atxA3</i>	<i>Asticcacaulis excentricus</i> CB 48	PDB: 2M8F_A	Zimmermann et al. (2013)
Astexin III	<i>atxD3</i>	<i>Asticcacaulis excentricus</i> CB 48	PDB: 5TXE_D	Zimmermann et al. (2013)
Capistruin	<i>capB</i>	<i>Burkholderia thailandensis</i>	-----	Knappe et al. (2008)
Caulonodin	gene A	<i>Caulobacter</i> sp. K31	PDB: 2MLJ_A	Zimmermann et al. (2014)
Caulosegnin 1	<i>csegA1</i>	<i>Caulobacter segnis</i>	PDB: 2LX6_A	Hegemann et al. (2013)
Lariat B (5 genes)	<i>larA/larB/larC/larD/larE</i>	<i>Rhodococcus jostii</i> K01-BD171	AB593691.1	Inokoshi et al. (2012)
lasso peptide StcB	<i>StcB</i>	<i>Stenotrophomonas acidaminiphila</i>	ALJ29424.1	Uniprot 2016
lasso peptide StcA	<i>StcA</i>	<i>Stenotrophomonas acidaminiphila</i>	ALJ29423.1	Uniprot 2016
Rp71,955	Gene B	<i>Streptomyces AA3891</i>	-----	Helynck et al. (1993)
Sphingopyxin I	<i>spiD</i>	<i>Sphingopyxis alaskensis</i> RB2256	PDB: 5JRL	Fage et al. (2016)
Streptomomicin	<i>StmA</i>	<i>Streptomonospora alba</i>	KIH99826.1	Metelev et al. (2015)
Streptomomicin	<i>StmD</i>	<i>Streptomonospora alba</i>	KII00663.1	Metelev et al. (2015)
Streptomomicin	<i>StmE</i>	<i>Streptomonospora alba</i>	KII00660.1	Metelev et al. (2015)
Xanthomonin I	<i>xgaA1</i>	<i>Xanthomonas gardneri</i> ATCC 19865	PDB: 2MFV_A	Hegemann et al. (2013)