

ISABELLE GONÇALVES DE OLIVEIRA PRADO

**ANÁLISE GENÔMICA E DETERMINAÇÃO DO PROTEOMA REFERÊNCIA
DE ISOLADOS CLÍNICOS DE *Actinobacillus pleuropneumoniae* SOROTIPO 8**

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

**VIÇOSA
MINAS GERAIS - BRASIL
2016**

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

T

P896a
2016
Prado, Isabelle Gonçalves de Oliveira, 1990-
Análise genômica e determinação do proteoma referência
de isolados clínicos de *Actinobacillus pleuropneumoniae*
sorotipo 8 / Isabelle Gonçalves de Oliveira Prado. – Viçosa,
MG, 2016.
viii, 56f. : il. (algumas color.) ; 29 cm.

Inclui apêndice.

Orientador: Denise Mara Soares Bazzolli.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.39-44.

1. *Actinobacillus pleuropneumoniae*. 2. Proteínas.
3. Genomas. 4. Genômica. I. Universidade Federal de Viçosa.
Departamento de Microbiologia. Programa de Pós-graduação em
Microbiologia Agrícola. II. Título.

CDD 22. ed. 579.34

ISABELLE GONÇALVES DE OLIVEIRA PRADO

**ANÁLISE GENÔMICA E DETERMINAÇÃO DO PROTEOMA REFERÊNCIA
DE ISOLADOS CLÍNICOS DE *Actinobacillus pleuropneumoniae* SOROTIPO 8**

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

APROVADA: 22 de fevereiro de 2016.

Mateus Ferreira Santana
(Coorientador)

Talles Eduardo Ferreira Maciel

Prof^a Denise Mara Soares Bazzolli
(Orientadora)

Aos meus pais, Júlio César do Prado e Denise Gonçalves de Oliveira Prado, pelo amor incondicional.

À minha irmã Danielle e vovó Diná, pela presença e amor.

Aos amigos e familiares.

DEDICO

AGRADECIMENTOS

Agradeço a Deus por ser minha fonte de sabedoria, pelas bênçãos a mim concedidas e por mais essa conquista.

À Universidade Federal de Viçosa (UFV) pelo acolhimento e por ter me proporcionado uma formação diferenciada.

Ao CNPq, CAPES e FAPEMIG que apoiaram a pesquisa e concederam bolsa de fomento durante minha trajetória acadêmica.

Ao Departamento de Microbiologia e ao NUBIOMOL pela assessoria e clareza nas informações.

Aos meus pais que sempre me incentivaram e me deram total apoio durante toda minha trajetória. Agradeço minha irmã Danielle pela segurança transmitida e solicitude. Vocês são meu porto seguro, meus amores e minha inspiração.

À minha vizinha Diná por ser meu maior exemplo de coragem, força e amor.

Aos meus familiares e amigos pela confiança, torcida e compreensão, principalmente por me ausentar em função do mestrado.

Aos mestres por acompanharem de perto minha evolução, auxiliando nas dificuldades e indicando o caminho.

À professora e orientadora Denise Mara Soares Bazzolli pela dedicação, compreensão e por acreditar no meu potencial.

Ao professor e coorientador Moysés Nascimento por ter me acolhido em seu laboratório e fazer desse ambiente de trabalho o mais agradável possível.

Aos colegas do laboratório de Microbiologia Agrícola, Estatística e Genética que sempre foram muito solícitos e grandes companheiros de trabalho.

À Monalessa Fábria Pereira, Pedro Marcus P. Vidigal, Juan Villada e Mateus F. Santana pelas reuniões, e-mails sem fim e por me fazerem acreditar que seria possível.

Aos membros da banca, professor Mateus Ferreira Santana e Talles Eduardo Ferreira Maciel, pela disponibilidade e por colaborarem com este trabalho.

A toda a equipe do BIOAGRO, que sempre de bom humor pelos corredores e que através de um simples bom dia, deixavam meu dia mais feliz e produtivo.

Às verdadeiras amizades construídas em Viçosa que levarei sempre comigo e também a todos que fizeram parte da minha vida nesses dois anos, tornando-a mais leve e divertida.

Enfim, obrigada a todos que de alguma forma contribuíram para meu crescimento e minha formação.

ÍNDICE

RESUMO	v
ABSTRACT	vii
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	3
3. MATERIAL E MÉTODOS	9
3.1. Análise dos genomas de <i>Actinobacillus pleuropneumoniae</i>	9
3.2. Proteoma de <i>A. pleuropneumoniae</i>	10
3.2.1. Montagem do proteoma referência <i>in silico</i> de <i>A. pleuropneumoniae</i> sorotipo 8	10
3.2.2. Análise comparativa entre proteomas de <i>A. pleuropneumoniae</i> de diferentes sorotipos	11
3.2.3. Análise de grupos de genes ortólogos de <i>A. pleuropneumoniae</i>	12
3.2.4. Análise de fatores de virulência e de resistência presentes no proteoma referência de <i>A. pleuropneumoniae</i> sorotipo 8.....	12
4. RESULTADOS.....	13
4.1. Características gerais dos genomas	13
4.1.1. Análise comparativa entre os genomas.....	14
4.1.2. Ilhas genômicas preditas em <i>A. pleuropneumoniae</i> sorotipo 8	17
4.1.3. Uso preferencial de códons.....	19
4.2. Caracterização do proteoma de <i>A. pleuropneumoniae</i> sorotipo 8.....	20
4.2.1. Análise de similaridade entre os proteomas dos diferentes sorotipos de <i>A. pleuropneumoniae</i>	29
4.2.2. Fatores de virulência e resistência a antibióticos do proteoma <i>in silico</i> de <i>A. pleuropneumoniae</i> sorotipo 8.....	30
5. DISCUSSÃO	32
6. CONCLUSÕES GERAIS	38
REFERÊNCIAS BIBLIOGRÁFICAS.....	39
MATERIAL SUPLEMENTAR.....	45

RESUMO

PRADO, Isabelle Gonçalves de Oliveira. M.Sc., Universidade Federal de Viçosa, fevereiro de 2016. **Análise genômica e determinação do proteoma referência de isolados clínicos de *Actinobacillus pleuropneumoniae* sorotipo 8.** Orientadora: Denise Mara Soares Bazzolli. Coorientadores: Mateus Ferreira Santana e Moysés Nascimento.

Actinobacillus pleuropneumoniae é o agente etiológico da pleuropneumonia suína, uma doença respiratória severa que resulta em significantes perdas econômicas no setor da suinocultura. Atualmente, *A. pleuropneumoniae* é classificado em 16 sorotipos diferentes que podem causar a doença com virulência distinta entre eles. Em Minas Gerais, Brasil, *A. pleuropneumoniae* é encontrado nas granjas e a maior distribuição é do sorotipo 8, sendo este até pouco tempo negligenciado em estudos de epidemiologia devido a problemas de identificação. Da mesma forma, atualmente é aceito que este sorotipo é também o de maior distribuição nos Estados Unidos, Canadá, Reino Unido além do Brasil. Até recentemente, não havia disponibilidade de genomas de *A. pleuropneumoniae* sorotipo 8 em banco de dados, por isso não existem informações genômicas específicas para este sorotipo, especialmente provenientes de isolados clínicos recentemente circulantes nas granjas. Neste contexto, somente a partir de 2015 os primeiros genomas foram disponibilizados para serem analisados, sendo assim sete genomas de isolados clínicos de *A. pleuropneumoniae* sorotipo 8 estão hoje disponíveis e estes foram analisados neste estudo. A partir das sequências genômicas dos isolados clínicos de *A. pleuropneumoniae* sorotipo 8 foi proposto um modelo de proteoma referência desse sorotipo com um total de 2.396 proteínas categorizadas nas diferentes classes de grupos de ortólogos. Na análise comparativa realizada entre genomas de *A. pleuropneumoniae* sorotipo 8 e demais sorotipos foi observado um padrão de conservação na organização do genoma entre esses. No entanto, foram identificadas algumas diferenças pontuais e dois segmentos genômicos diferenciais entre os genomas analisados com rearranjos e /ou inversões de 42,2 Kb e 59,6 Kb. O primeiro segmento foi encontrado nos genomas dos sorotipos 5 (L20), 7 (AP76) e em quatro dos isolados clínicos do sorotipo 8 investigados, sendo eles MV518, MV780, MV1022 e MV5651 contendo sequências de profago. O segundo segmento diferencial foi identificado apenas no genoma de *A. pleuropneumoniae* sorotipo 7 (AP76) e contém sequências como transposases caracterizando a presença de uma sequência de inserção no segmento. De acordo com a análise dos códons preferenciais de todos os sorotipos investigados, não foram observadas diferenças marcantes entre eles. Na análise

comparativa dos proteomas, a maioria das sequências do proteoma referência de *A. pleuropneumoniae* sorotipo 8 que apresentaram baixa similaridade com os demais sorotipos foram caracterizadas como hipotéticas, não tendo sua função conhecida. Nessa porção diferencial específica foram relatadas sequências codificadoras relacionadas à plasmídeos o que caracteriza possíveis eventos de transferência horizontal de genes (THG) ao longo da história evolutiva do genoma e fatores de resistência a antibióticos como cloranfenicol, sulfonamida e tetraciclina. Na análise de similaridade entre os proteomas houve maior similaridade das proteínas do proteoma referência do sorotipo 8 com as proteínas do sorotipo 6, o que pode estar associado à dificuldade em separar esses sorotipos nas técnicas de sorotipagem. Com as análises comparativas e a caracterização das diferenças entre os sorotipos e isolados será possível compreender os mecanismos de virulência dos isolados de *A. pleuropneumoniae* e identificar potenciais candidatos vacinais mais eficientes.

ABSTRACT

PRADO, Isabelle Gonçalves de Oliveira. M.Sc., Universidade Federal de Viçosa, February, 2016. **Genomic analysis and determination of reference proteome of clinic isolates of *Actinobacillus pleuropneumoniae* serotype 8.** Adviser: Denise Mara Soares Bazzolli. Co-advisers: Mateus Ferreira Santana and Moysés Nascimento.

Actinobacillus pleuropneumoniae is the causative agent of porcine pleuropneumonia, a severe respiratory illness that it results in significant economic losses in the swine industry. Currently, *A. pleuropneumoniae* is classified into 16 different serotypes that it can cause disease with distinct virulence between them. In Minas Gerais, Brazil, *A. pleuropneumoniae* is found on the farms and the largest distribution is of serotype 8, which, until shortly time, it was neglected in epidemiological studies because of the identification problems. Similarly, it is currently accepted that this serotype is also the most widely distributed in the United States of America, Canada, the United Kingdom as well as Brazil. Until recently, there was no availability of genomes of *A. pleuropneumoniae* serotype 8 in the database, because of this, there are no specific genomic information for this serotype, especially from of the clinical isolates that, recently, it are circulating on the farms. In this context, only from 2015, the first genomes were available to be analysed, so, today there are seven genomes of clinical isolates of *A. pleuropneumoniae* serotype 8 available and these were analyzed in this study. From the genomic sequences of clinical isolates of *A. pleuropneumoniae* serotype 8 was proposed a model of proteome which it is reference for this serotype with a total of 2.396 proteins categorized into different classes of orthologous groups. In a comparative analysis of genomes of *A. pleuropneumoniae* serotype 8 and other serotypes, it was observed a pattern of conservation in the organization of the genome between these. However, it were identified some specific differences and two genomic segments differentials between the genomes analyzed with rearrangements and/or reversals of 42.2 Kb and 59.6 Kb. The first segment was found in the genome of the serotype 5 (L20), 7 (AP76) and in four of the clinical isolates of serotype 8 that it were investigated: MV518, MV780, MV1022 and MV5651, which it containing prophage sequences. The second differential segment was identified only in the genome of *A. pleuropneumoniae* serotype 7 (AP76) and it contains sequences as transposases that it characterizes the presence of an insertion sequence in the segment. Regarding the use and the preference of codons between the serotypes investigated, there were no significant differences among them. In the comparative analysis of proteomes, the

sequences of the reference proteome of *A. pleuropneumoniae* serotype 8, which it showed low similarity with the remaining serotypes, it were identified as hypothetical, it does not having its known function. In this specific differential portion, coding sequences associated to plasmids were described, that it characterize possible events by horizontal transfers of genes (HTG) throughout the evolutionary history of the genome and antibiotic resistance factors, such as chloramphenicol, tetracycline and sulfonamide. In the similarity analysis between the proteomes, there was greater similarity of the proteome reference of serotype 8 with the proteins of the serotype 6, which it may be associated on difficulty to separate these serotypes in the serotyping techniques. With the comparative analysis and the characterization of the differences between the serotypes and the isolates, it will be possible to understand the virulence mechanisms of the isolates of *A. pleuropneumoniae* and to identify vaccine candidate potential more effective.

1. INTRODUÇÃO

Desde 1978, a carne suína vem se consolidando como importante fonte de proteína animal e atualmente é uma das mais consumidas no mundo. Nesse contexto, o bom desempenho do Brasil e demais países é garantido pelos avanços tecnológicos e organizacionais do sistema de criação de suínos. No entanto, esse sistema de criação intensivo tem tornado frequente a ocorrência de doenças respiratórias, acometendo grande parte da produção, causando prejuízos e perdas econômicas significativas na suinocultura. Uma das mais importantes doenças respiratórias de suínos, a pleuropneumonia suína, é causada pela bactéria *Actinobacillus pleuropneumoniae*, pertencente à família Pasteurellaceae. Atualmente, *A. pleuropneumoniae* é classificado em 16 sorotipos com base em propriedades antigênicas dos polissacarídeos da cápsula e separados em dois biotipos de acordo com a dependência do fator V de crescimento, a nicotinamida adenina dinucleotídeo (NAD), sendo o biotipo 1 β -NAD-dependente e o biotipo 2 β -NAD-independente, podendo sintetizá-lo a partir de precursores. Todos os sorotipos são capazes de causar a pleuropneumonia suína com padrões diferentes de virulência entre eles, e sua distribuição varia entre países e regiões. A patogênese da pleuropneumonia suína é complexa e envolve diferentes fatores de virulência produzidos pela bactéria.

Na maioria dos rebanhos, embora vários sorotipos tenham sido encontrados em um mesmo ambiente, é comum um sorotipo predominar. Trabalhos realizados no sudeste do Brasil permitiram a caracterização de isolados clínicos de *A. pleuropneumoniae* amostrados nas fazendas produtoras, com uma maior distribuição dos sorotipos 2, 7 e 8. Esses estudos apontaram que, embora exista uma distribuição variada dos diferentes sorotipos de *A. pleuropneumoniae* em Minas Gerais - Brasil, o sorotipo 8 é o mais encontrado nas propriedades suinícolas estudadas e dentro desse sorotipo, os isolados amostrados apresentaram diferenças no padrão de virulência e polimorfismo genético. *Actinobacillus pleuropneumoniae* sorotipo 8 por muito tempo foi negligenciado em estudos de identificação devido a falhas nas técnicas de sorotipagem e em função disso, estudos de genômica envolvendo este sorotipo são inexistentes. No entanto, estudos atuais têm apresentado uma ampla distribuição deste sorotipo em diversas regiões geográficas como no Reino Unido, América do Norte e Brasil. Recentemente, genomas correspondentes a *A. pleuropneumoniae* sorotipo 8 foram depositados em bancos de dados, entretanto, nenhuma análise descritiva e comparativa do sorotipo 8 com os demais sorotipos tem sido realizada.

Em estudos realizados pelo nosso grupo de pesquisa com o hospedeiro alternativo, a larva *Galleria mellonella*, diferentes padrões de virulência foram observados em isolados clínicos de *A. pleuropneumoniae* sorotipo 8. Com base nos resultados obtidos a partir desse trabalho, seis isolados com diferentes perfis fenotípicos foram selecionados para o sequenciamento, gerando informações para estudos comparativos. Embora o genoma de *A. pleuropneumoniae* seja relativamente conservado em termos de organização, podem existir variações dentro do mesmo sorotipo que podem influenciar na resposta imunológica do hospedeiro frente ao patógeno. Assim, com a genômica comparativa será possível entender a variabilidade genética da espécie *A. pleuropneumoniae* e relacionar as diferenças genéticas com padrões distintos de virulência entre os isolados do mesmo sorotipo. A relação estreita entre diferenças genômicas e variações na patogenicidade, epidemiologia, imunogenicidade e virulência entre os diferentes isolados reforçam a importância de se ampliar os conhecimentos sobre diferenças genéticas desses patógenos, possibilitando a descoberta de novos candidatos vacinais e de diagnóstico.

Este trabalho teve como objetivos a análise do genoma e montagem *in silico* do proteoma referência de *A. pleuropneumoniae* sorotipo 8 de forma representativa a partir dos sete isolados clínicos já sequenciados objetivando analisar as semelhanças compartilhadas com os demais sorotipos bem como determinar a parte genética diferencial encontrada no sorotipo 8.

2. REVISÃO DE LITERATURA

Na década de setenta, a carne suína se consolidou como a fonte de proteína animal mais importante no mundo, sendo a mais consumida desde então. Nesse contexto, o Brasil apresenta grande competitividade no mercado internacional da suinocultura, sendo atualmente considerado o quarto maior produtor e exportador em escala mundial (EMBRAPA, 2015). O bom desempenho do Brasil e demais países é garantido pelos avanços tecnológicos e organizacionais do sistema de criação que está relacionado entre outros fatores a um elevado número de animais confinados em um espaço limitado. Esse sistema de criação intensivo tem tornado frequente a ocorrência de doenças respiratórias, acometendo grande parte da produção e causando prejuízos e perdas econômicas significativas na suinocultura. Entre as doenças respiratórias que acometem suínos, a pleuropneumonia suína se destaca como uma doença importante, altamente contagiosa, frequentemente fatal e ocorre em diversos países nos quais a suinocultura é praticada (Chiers *et al.*, 2010).

A pleuropneumonia suína está associada a uma mortalidade elevada, diminuição no crescimento e na eficiência alimentar do animal e portanto, atraso no período de abate. Os custos de tratamento e a necessidade de se obter boas carcaças para o abate contribuem para que esta doença tenha um impacto econômico significativo. Os grandes danos causados ao produtor estão associados à diminuição da média diária do ganho de massa corporal do suíno, perdas por necessidade de abate e gastos com intervenções, como vacinação e medicação, além da morte precoce do animal (Krejci & Newberry, 2011). Outro problema encontrado se relaciona ao uso intensivo de antibióticos no controle das doenças respiratórias, que aumentam a chance da ocorrência de resistência bacteriana aos antimicrobianos comumente utilizados (Blaha *et al.*, 1994).

O agente etiológico da pleuropneumonia suína é a bactéria Gram negativa, cocobacilar, anaeróbia facultativa, capsulada, não formadora de esporos e suas colônias produzem hemólise em ágar sangue denominada de zona β -hemolítica (Sebunya & Saunders, 1983). Pertencente à família Pasteurellaceae, *Actinobacillus pleuropneumoniae* foi isolado pela primeira vez em 1957 nos Estados Unidos (Pattison *et al.*, 1957) e posteriormente, em 1964 na Argentina (Shope, 1964). O habitat natural de *A. pleuropneumoniae* é o trato respiratório de suínos e pode ser isolado a partir de diferentes regiões do sistema respiratório de organismos infectados tais como cavidades nasais, tonsilas e pulmões (Rycroft & Garside, 2000). *Actinobacillus pleuropneumoniae* é classificado em dois biotipos, de acordo com a dependência do fator V de

crescimento, a nicotinamida adenina dinucleotídeo (NAD), sendo o biotipo 1 β -NAD-dependente e o biotipo 2 β -NAD-independente, podendo sintetizá-lo a partir de precursores (Pohl *et al.*, 1983).

Atualmente, *Actinobacillus pleuropneumoniae* é classificado em 16 sorotipos com base em propriedades antigênicas dos polissacarídeos da cápsula (Sarkozi *et al.*, 2015). Todos os sorotipos são capazes de causar a pleuropneumonia suína com padrões diferentes de virulência entre eles, e sua distribuição varia entre países e regiões (Blackall *et al.*, 2002; O'Neill *et al.*, 2010; Gottschalk & Lacouture, 2014; Rossi *et al.*, 2013). Todos os sorotipos de *A. pleuropneumoniae* secretam diferentes combinações de quatro toxinas RTX (*Repeats in toxin*): ApxI, ApxII, ApxIII e ApxIV, sendo esta última secretada por todos os sorotipos (Chien *et al.*, 2009) e utilizada na diferenciação dos sorotipos por PCR multiplex (PCRm). Os sorotipos 1, 5, 9 e 11 são considerados mais virulentos, secretam as toxinas ApxI e ApxII e estão comumente associados aos surtos da doença com maiores índices de mortalidade; os sorotipos 2, 3, 4, 6, 8 e 15 apresentam virulência média e secretam ApxII e ApxIII; os sorotipos 7, 12, e 13 secretam ApxII, enquanto os sorotipos 10 e 14 produzem ApxI (Auger *et al.*, 2009; Frey, 2011). A correta sorotipagem de *A. pleuropneumoniae* é importante do ponto de vista epidemiológico e no controle da doença, pois a prevalência dos diferentes sorotipos e a virulência são fatores que podem variar entre países e regiões (Dubreuil *et al.*, 2000).

A patogênese da pleuropneumonia suína é complexa e envolve diferentes fatores de virulência produzidos pela bactéria (Chiers *et al.*, 2010). A virulência é multifatorial e está relacionada à combinação de fatores como as toxinas da família RTX, composição e estrutura de polissacarídeos da cápsula, ao lipopolissacarídeo de membrana externa (LPS), sideróforos de ferro, formação de biofilme, capacidade de se transformar naturalmente, presença de adesinas e diversos fatores envolvidos com a transferência horizontal de genes (THG) que acabam por promover o *fitness* da bactéria (Chiers *et al.*, 2010). Os elementos genéticos móveis presentes nas bactérias são caracterizados como contínuos desafios à estabilidade genômica, promovendo a evolução através da transferência horizontal de genes, mutação, recombinação e alteração da expressão gênica. Esse grupo engloba plasmídeos, bacteriófagos, elementos transponíveis, ilhas genômicas e muitos outros elementos genéticos especializados (Darmon & Leach, 2014).

A forma mais comum de adquirir novas funções é através da transferência de material genético a partir de organismos não relacionados por mecanismos de

transformação, conjugação e transdução. Esses mecanismos permitem aos organismos se adaptarem rapidamente às mudanças ambientais (Medini *et al.*, 2005). Análises filogenéticas revelaram que transferências horizontais de sequências de DNA geram um impacto substancial à longo prazo na evolução do genoma bacteriano e permite adaptação em ambientes dinâmicos (Popa & Dagan, 2011). Um desses mecanismos é a transformação natural que é um processo pelo qual DNA é captado do meio ambiente e transportado para o citoplasma como fita simples de DNA (Dubnau, 1999). Nesse contexto, alguns sorotipos de *A. pleuropneumoniae* apresentam competência natural, sendo, portanto, comum a ocorrência de transformação natural nesta espécie, e dessa forma a aquisição de genes por transferência horizontal e obtenção de mutantes em ensaios experimentais é facilitada (Bossé *et al.*, 2009; Bossé *et al.*, 2014).

Como fator limitante à transferência horizontal de genes temos o sistema CRISPR, *Clustered Regularly Interspaced Short Palindromic Repeats*, associado a um grupo de proteínas denominadas Cas, caracterizado como um sistema imune bacteriano específico que identifica e degrada elementos geneticamente móveis recentemente adquiridos (Popa & Dagan, 2011). Os tamanhos das repetições de CRISPR e dos espaçadores variam entre 23 a 47 pares de bases e 21 a 72 pares de bases respectivamente. Os genes *cas* codificam uma grande e heterogênea família de proteínas que possuem domínios funcionais típicos de nucleases, helicases, polimerases e proteínas de ligação a polinucleotídeos. A habilidade de prover defesa contra elementos genéticos invasores pode tornar o sistema CRISPR/Cas desejável em ambientes hostis e pode explicar a sua propensão de ser transferido entre organismos distantes (Horvath & Barrangou, 2010). Em *A. pleuropneumoniae* não existem informações sobre a organização e a funcionalidade do sistema CRISPR.

Como fator de virulência, temos também, as ilhas genômicas desempenhando um papel crítico na patogênese de doenças bacterianas. As ilhas genômicas podem codificar para muitas funções que podem estar envolvidas em simbiose ou patogênia, e podem ajudar na adaptação do organismo. Essas regiões podem conter genes de provável origem horizontal e são de grande interesse por estarem relacionadas na maioria das vezes a genes de virulência, resistência a antibióticos ou outras adaptações (Langille & Brinkman, 2009). Essas sequências muitas vezes apresentam composição diferente do genoma hospedeiro. As estratégias para avaliar essas diferenças vêm sendo estudadas e uma das mais simples se baseia no conteúdo GC e também em medições utilizando oligonucleotídeos de vários comprimentos. A partir de uma lista de ilhas genômicas é possível fazer as análises e estabelecer os limites da ilha e inspecionar os

genes em cada ilha genômica para determinar se há alguma pista sobre sua origem ou modo de transferência. Muitas ilhas podem ser definidas como prováveis profagos e, geralmente, identificadas por genes anotados com características virais. Estas regiões podem ser usadas em uma análise mais aprofundada com ferramentas de bioinformática e assim analisar suas sequências, porcentagem do conteúdo GC e tendências no uso de códons (Langille *et al.*, 2010).

Além dos fatores de virulência e das ilhas genômicas, a resistência a antimicrobianos é também caracterizada como fator de virulência e pode ser codificada tanto pelo cromossomo como por plasmídeos. A presença de plasmídeos relacionados à resistência em *A. pleuropneumoniae* foi reportada pela primeira vez em 1982 (Hirsh *et al.*, 1982) e desde então houve subsequentes relatos, como a presença de plasmídeos mobilizáveis conferindo resistência a trimetoprim (Bossé *et al.*, 2009; Bossé *et al.*, 2015). Além das diferenças no padrão de virulência entre os 16 sorotipos, são encontradas ainda diferenças entre os isolados do mesmo sorotipo (Pereira *et al.*, 2015a). Mesmo havendo um padrão de conservação no genoma de *A. pleuropneumoniae*, entre os sorotipos existem diferenças relacionadas à aquisição de DNA, mutações pontuais, inserções e deleções que originam variabilidade dentro da espécie (Xu *et al.*, 2010).

Na maioria dos rebanhos, embora vários sorotipos tenham sido demonstrados em um mesmo ambiente, é comum um sorotipo predominar (Chiers *et al.*, 2002). Trabalhos realizados por nosso grupo de trabalho na região sudeste do Brasil permitiram a caracterização de isolados clínicos de *A. pleuropneumoniae* amostrados nas fazendas produtoras, com uma maior distribuição dos sorotipos 2, 7 e 8 (Rossi *et al.*, 2013). Em Minas Gerais, o sorotipo 8 é o mais difundido nas propriedades suínícolas estudadas e dentro desse sorotipo, os isolados amostrados apresentaram diferenças no padrão de virulência e uma alta variabilidade (Rossi *et al.*, 2013; Pereira *et al.*, 2015a). Em estudos realizados com isolados clínicos de *A. pleuropneumoniae* sorotipo 8 foram observados um elevado polimorfismo genético, virulência e ocorrência de um elevado número de genes de resistência associados a fenótipos de multirresistência à drogas (Santos, 2013; Rossi *et al.*, 2013; Pereira *et al.*, 2015a). O sorotipo 8 foi caracterizado apresentando alta atividade citotóxica, sendo responsável por uma baixa taxa de mortalidade, porém alta taxa de morbidade (O'Neill *et al.*, 2010). Em estudos realizados pelo nosso grupo de pesquisa utilizando o hospedeiro alternativo, a larva *Galleria mellonella*, diferentes padrões de virulência foram observados em isolados clínicos de *A. pleuropneumoniae* sorotipo 8 (Pereira *et al.*, 2015a), sendo os isolados MV780 e MV1022 caracterizados

como de menor e maior virulência, respectivamente. Em uma análise mais aprofundada com esses dois isolados, foi observado que esta diferença está inicialmente relacionada com a capacidade do patógeno de proliferar e evadir o sistema imunológico do inseto (Pereira *et al.*, 2015a). Com base nos resultados obtidos a partir desse trabalho, seis isolados clínicos brasileiros com diferentes perfis fenotípicos foram selecionados para o sequenciamento genético (Pereira *et al.*, 2015b). Mesmo tendo sido relatada essas diferenças, nenhuma descrição do sorotipo 8 e nenhuma análise comparativa do genoma e do proteoma desse sorotipo foi realizada até o momento.

Trabalhos com *A. pleuropneumoniae* sorotipos 1, 5a, 5b, 7, 8 e 12 revelaram uma heterogeneidade pronunciada na estrutura cromossômica e *clusters* de genes variados envolvidos na produção de toxina, cápsula e biossíntese de LPS (Gouré *et al.*, 2009). A partir das diferenças dos componentes gênicos entre as linhagens mais e menos virulentas é possível identificar novos candidatos a fatores de virulência. Essas diferenças intraespecíficas na arquitetura do genoma podem influenciar o fenótipo bacteriano mesmo na presença do mesmo repertório gênico (Xu *et al.*, 2010).

A genômica comparativa é o primeiro passo na análise do pan-genoma que é representado pela coleção de todos os perfis funcionais encontrados nos genomas. Quanto mais genomas são adicionados nessa construção, existe uma tendência do conjunto de genes aumentar (Lukjancenko *et al.*, 2013). O genoma *core* é a coleção de genes que são conservados entre os genomas analisados e o tamanho desse genoma *core* tende a diminuir à medida que mais genomas são adicionados às análises. Esse estudo pode ser realizado com os genes e podem se estender às sequências codificadoras. Dessa forma, as diferenças entre os proteomas podem ser identificadas pelo perfil do proteoma acessório que inclui proteínas que estão presentes em vários, mas não em todos os genomas analisados ou são específicas de um genoma particular ou de um grupo de genomas. A proteína para ser específica deve estar presente na sequência query e ausente no conjunto de sequências dos organismos analisados (Lukjancenko *et al.*, 2013). A identificação e caracterização das proteínas em uma célula são de extrema importância para compreender os processos moleculares que medeiam a fisiologia celular. Com as melhorias contínuas em hardwares e métodos para analisar e estudar o proteoma é possível formular hipóteses testáveis com base nas informações biológicas pré-existentes, sendo a metodologia de análises de bioinformática na genômica comparativa de grande utilidade (Schmidt *et al.*, 2014; Duan *et al.*, 2014; Baig *et al.*, 2015).

Em um estudo prévio com diferentes genomas de sorotipos de *A. pleuropneumoniae* foi possível observar que a maior parte dos genes diferenciais estão relacionados à mecanismos de defesa, replicação, recombinação, sistemas de restrição, além de um grande número de transposases, recombinases e integrases provenientes de elementos transponíveis e profagos (Xu *et al.*, 2010). Análises de genômica comparativa usando sequências de genomas originados de 12 sorotipos mostraram que o pan genoma de *A. pleuropneumoniae* consiste de 3,303 *clusters* gênicos totais, sendo dividido em 1,709 genes do genoma core, 822 genes distribuídos (compartilhado) e 772 genes exclusivos de cada cepa (Xu *et al.*, 2010). Neste trabalho, nenhuma informação sobre *A. pleuropneumoniae* sorotipo 8 foi fornecida, uma vez que apenas recentemente este sorotipo foi reconhecido como altamente distribuído na América do Norte, Reino Unido e Brasil (O'Neill *et al.*, 2010; Gottschalk & Lacouture, 2014; Rossi *et al.*, 2013). Uma possível explicação está relacionada às reações cruzadas e consequentes erros de sorotipagem, como já observado em trabalho que relaciona o sorotipo 8 como sorotipo predominante dentro do grupo de reação cruzada 3/6/8/15 em análises com isolados de suínos doentes na América do Norte (Gottschalk & Lacouture, 2014). Além disso, somente recentemente genomas correspondentes a *A. pleuropneumoniae* sorotipo 8 (Pereira *et al.*, 2015a; Bossé *et al.*, 2016) foram depositados em bancos de dados, entretanto, nenhuma análise descritiva e comparativa correspondente ao genoma do sorotipo 8 foi realizada anteriormente.

A relação estreita entre diferenças genômicas e variações na patogenicidade, epidemiologia, imunogenicidade e virulência entre os diferentes isolados de *A. pleuropneumoniae* reforçam a importância de se ampliarem os conhecimentos sobre diferenças genéticas desses patógenos, possibilitando a descoberta de novos candidatos vacinais e de diagnóstico. O estudo de genômica comparativa com base em análises de transferência horizontal de genes, sistema CRISPR, genes de virulência, ilhas genômicas, Índice de Adaptação de Códon e frequência no uso dos códon podem ser uma ferramenta ímpar na detecção de diferenças minuciosas que podem se traduzir em variação no padrão de virulência entre os isolados do mesmo sorotipo.

A partir da montagem do proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8 de forma representativa a partir dos sete isolados clínicos será possível analisar as semelhanças compartilhadas com os demais sorotipos bem como determinar a parte diferencial encontrada no sorotipo 8.

3. MATERIAL E MÉTODOS

3.1. Análise dos genomas de *Actinobacillus pleuropneumoniae*

As análises foram realizadas a partir de genomas de *A. pleuropneumoniae* de diferentes sorotipos, sendo eles: 1, 2, 3, 4, 5b, 6, 7, 8, 9, 10, 11, 12 e 13, o que inclui linhagens referência e isolados clínicos. Os códigos de acesso dos genomas utilizados neste estudo estão dispostos na Tabela 1.

Tabela 1. Genomas dos diferentes sorotipos de *A. pleuropneumoniae* utilizados na análise.

Sorotipo	Linhagem	Código de acesso - NCBI	Referência
Sorotipo 1	4074	ADOD000000000.1	Xu <i>et al.</i> , 2010
Sorotipo 2	4226	ADXN000000000.1	Zhan <i>et al.</i> , 2010
Sorotipo 3	JL03	CP000687.1	Xu <i>et al.</i> , 2008
Sorotipo 4	M62	ADOF000000000.1	Xu <i>et al.</i> , 2010
Sorotipo 5b	L20	CP000569.1	Foote <i>et al.</i> , 2008
Sorotipo 6	Femo	ADOG000000000.1	Xu <i>et al.</i> , 2010
Sorotipo 7	AP76	CP001091.1	Linke <i>et al.</i> , 2008
Sorotipo 8	MV460	JSVG000000000.1	
Sorotipo 8	MV518	JSVZ000000000.1	
Sorotipo 8	MV597	JSVX000000000.1	Pereira <i>et al.</i> , 2015b
Sorotipo 8	MV780	JSVV000000000.1	
Sorotipo 8	MV1022	JSVF000000000.1	
Sorotipo 8	MV5651	JSVY000000000.1	
Sorotipo 8	MIDG2331	LN908249.1	Bossé <i>et al.</i> , 2016
Sorotipo 9	CVJ13261	ADOI000000000.1	
Sorotipo 10	D13039	ADOJ000000000.1	
Sorotipo 11	56153	ADOK000000000.1	Xu <i>et al.</i> , 2010
Sorotipo 12	1096	ADOL000000000.1	
Sorotipo 13	N273	ADOM000000000.1	

As informações disponíveis nesta tabela são provenientes do banco de dados NCBI.

As análises comparativas entre os genomas foram realizadas a partir do software *Progressive Mauve* v.2.3.1 (Darling *et al.*, 2004).

A identificação de possíveis ilhas genômicas foi realizada utilizando o programa *Island Viewer* (Langille & Brinkman, 2009) e o genoma de *A. pleuropneumoniae* JL03 do sorotipo 3 foi utilizado como referência no alinhamento com os genomas de *A. pleuropneumoniae* sorotipo 8 nesta análise. A escolha do genoma do sorotipo 3 se deu

pela qualidade da montagem e por ser o que compartilha maior número de sequências codificadoras com o sorotipo 8 em relação aos demais genomas que estão fechados, uma vez que o genoma do sorotipo 8, MIDG2331 não está disponibilizado no programa. A predição de ilhas genômicas foi baseada em três métodos: IslandPath-DIMOB, SIGI-HMM e *Island-Pick*, mas só foram detectadas pelos dois primeiros métodos (Figura S1). O primeiro identifica ilhas com base na tendência de dinucleotídeos e presença de genes de mobilidade associados (integrases, transposases). O método de SIGI-HMM identifica tendência no uso de códons com abordagem no modelo da cadeia de Markov. Já o método *Island Pick* requer vários genomas filogeneticamente relacionados para ser capaz de fazer uma previsão e depende assim dos genomas utilizados na comparação (Langille & Brinkman, 2009; Dhillon *et al.*, 2015).

A análise do uso preferencial de códons e conteúdo GC foi realizada a partir do programa EMBOSS (Rice, 2000) para os diferentes sorotipos de *A. pleuropneumoniae*. O índice de adaptação do códon (IAC ou CAI - *Codon Adaptation Index*) foi calculado de acordo com Sharp & Li, 1987. A adaptabilidade relativa de cada códon foi calculada pela razão entre o uso de cada códon e o códon mais abundante para um mesmo aminoácido. Este valor foi calculado de 0 a 1 (Sharp & Li, 1987). A determinação do uso de cada códon sinônimo foi calculada a partir do cálculo de RSCU (*Relative Synonymous Codon Usage*). O valor de RSCU calculado para cada códon foi o parâmetro usado para avaliar o tipo de seleção dos códons, sendo valores =1 característico de códons usados com igual frequência, valores > 1 de seleção positiva e < 1 seleção negativa. A contagem de aminoácidos e respectiva análise comparativa foram realizadas como proposto por Plotkin & Kudla, 2011.

A identificação de *locus* correspondentes a Repetições Palindrômicas Curtas Agrupadas e Regularmente Interespaçadas associados a proteínas Cas (CRISPR-Cas) foram preditas utilizando o programa *CRISPRFinder* (Grissa *et al.*, 2007).

3.2. Proteoma de *A. pleuropneumoniae*

3.2.1. Montagem do proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8

A montagem do proteoma referência de *A. pleuropneumoniae* sorotipo 8 foi baseada nos sete genomas de isolados clínicos desse sorotipo provenientes de granjas produtoras de suínos no Brasil (Pereira *et al.*, 2015b) e de um isolado clínico do Reino Unido (Bossé *et al.*, 2016) sequenciados previamente. Foi realizada também a

montagem do proteoma de *A. pleuropneumoniae* sorotipo 8 baseada apenas nos seis genomas de isolados clínicos brasileiros (Pereira *et al.*, 2015b) afim de realizar comparações entre os isolados do mesmo sorotipo, mas de regiões geográficas distintas.

Nestas análises foi utilizado o programa CD-HIT v. 4.6.1 (Weizhong & Godzik, 2006; Limin *et al.*, 2012) com parâmetro de 0,85 de identidade entre as sequências e valores padrões do programa.

Na anotação funcional do proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8 foram utilizados cinco bancos de dados, sendo eles COG (Tatusov *et al.*, 2000), CDD (Bauer *et al.*, 2011), PFAM (Finn *et al.*, 2010), SMART (Letunic *et al.*, 2006) e UNIPROT (The UniProt Consortium, 2015). A anotação teve como padrão os dados gerados pelo banco de dados UNIPROT e os demais bancos de dados foram utilizados para solucionar conflitos entre as informações geradas e garantir maior consistência na anotação. A análise teve como parâmetro um *E-value* $\leq 10^{-5}$.

3.2.2. Análise comparativa entre proteomas de *A. pleuropneumoniae* de diferentes sorotipos

Nesta análise foram utilizados 12 proteomas de diferentes sorotipos de *A. pleuropneumoniae* já depositados no banco de dados do UNIPROT (*The UniProt Consortium* UniProt, 2015) sendo estes sorotipos 1, 2, 3, 4, 5b, 6, 7, 9, 10, 11, 12 e 13. A análise foi realizada através do algoritmo *BLAST* (*Basic Local Alignment Search Tool*) contrastando os proteomas dos sorotipos analisados contra o proteoma de *A. pleuropneumoniae* sorotipo 8 baseado em sete isolados clínicos montado nesse trabalho. Os índices de similaridade (*greatest positive %*) e de identidade (*greatest identity %*) foram usados como critério para inferir sobre a conservação do proteoma dentro da espécie (Lukjancenکو *et al.*, 2013). A partir dessa análise fizemos grupamentos baseados no padrão de similaridade. O conjunto total de 2.174 sequências codificadoras de *A. pleuropneumoniae* sorotipo 8 MIDG2331 (Bossé *et al.*, 2016) extraído do banco de dados do NCBI foi contrastado com o proteoma referência predito de *A. pleuropneumoniae* sorotipo 8 baseado nos seis isolados clínicos do Brasil também montado neste trabalho com a finalidade de analisar as diferenças e similaridades intra-sorotípica.

3.2.3. Análise de grupos de genes ortólogos de *A. pleuropneumoniae*

A partir das sequências dos 12 proteomas de *A. pleuropneumoniae* e do proteoma referência sorotipo 8, baseado nos sete isolados clínicos foi montado um banco de dados contendo 28.002 sequências de proteínas correspondentes a todos os sorotipos da espécie. Através do programa CD-HIT v. 4.6.1 (Weizhong & Godzik, 2006; Limin *et al.*, 2012) com parâmetro de 0,70 de identidade entre as sequências e valores padrões do programa foi realizada uma análise para caracterizar o conjunto total de proteínas da espécie.

Os grupos de proteínas identificados pelo CD-HIT foram classificados como *Core*, compartilhado e diferencial. Adicionalmente, o proteoma referência de *A. pleuropneumoniae* sorotipo 8 foi comparado individualmente com o proteoma dos demais sorotipos usando o algoritmo *BLAST*.

3.2.4. Análise de fatores de virulência e de resistência presentes no proteoma referência de *A. pleuropneumoniae* sorotipo 8

Os fatores de virulência presentes no proteoma referência de *A. pleuropneumoniae* sorotipo 8 foram investigados a partir do portal PATRIC (*PathoSystems Resource Integration Center*) (www.patricbrc.org) e na base de dados de interação patógeno hospedeiro, PHI base (www.PHI-base.org), sendo os parâmetros identidade superior a 30%, *E-value* $\leq 10^{-5}$ e proteínas com fenótipo não afetado desconsideradas na análise final. A partir do banco de dados ARDB (*Antibiotic Resistance Genes Database*) (<http://ardb.cbc.umd.edu/>) foram pesquisadas sequências de genes envolvidos na resistência a antibióticos (Liu & Pop, 2009) seguindo parâmetros do programa, sendo *E-value* menor igual a 10^{-5} e identidade acima de 80%.

4. RESULTADOS

4.1. Características gerais dos genomas

O tamanho dos genomas de *A. pleuropneumoniae* variou de 2.185.499 pb a 2.345.435 pb, sendo a média de 2.265.454 pb e o número de genes variou de 2.082 a 2.300 genes. Os valores do conteúdo GC foram muito próximos entre os sorotipos variando de 41,00% a 41,30% (Tabela 2). Os genomas de *A. pleuropneumoniae* sorotipos 3 (JL03), 5 (L20), 7 (AP76) e 8 (MIDG2331) estão fechados. Já os genomas dos demais sorotipos encontram-se em contigs (Tabela 2).

Tabela 2. Organização dos genomas dos diferentes sorotipos de *A. pleuropneumoniae*.

Sorotipo	Cepa	Tamanho (pb)	Genes	GC (%)	Locus CRISPR
1	4074	2.254.524	2.158	41,20	1
2	4226	2.314.083	2.228	41,20	1
3	JL03	2.242.062	2.115	41,20	1
4	M62	2.260.565	2.186	41,20	1
5b	L20	2.274.482	2.168	41,30	1
6	Femo	2.302.700	2.300	41,00	1
7	AP76	2.345.435	2.234	41,21	1
8	MV460	2.213.381	2.116	41,10	1
8	MV518	2.275.540	2.189	41,10	1
8	MV597	2.219.395	2.115	41,10	1
8	MV780	2.274.000	2.179	41,10	1
8	MV1022	2.262.828	2.179	41,10	1
8	MV5651	2.264.279	2.179	41,10	1
8	MIDG2331	2.337.633	2.235	41,10	1
9	CVJ13261	2.256.417	2.163	41,20	1
10	D13039	2.266.276	2.155	41,20	1
11	56153	2.257.884	2.154	41,20	1
12	1096	2.185.499	2.082	41,20	1
13	N273	2.236.660	2.148	41,20	1

Na análise realizada no programa CRISPRFinder foi observada a presença de um locus CRISPR (Tabela 2) em todos os genomas de *A. pleuropneumoniae* e independente do sorotipo, a organização e o tipo de CRISPR foi o mesmo, sendo do tipo II, modificando apenas os espaçadores (tamanho e composição). As proteínas Cas associadas encontradas pertencem à família Cas1, Cas3, Csy1, Csy2, Csy3 e Csy4. As

repetições diretas (DR – Direct Repeats) são conservadas com um tamanho de vinte e oito pares de bases.

4.1.1. Análise comparativa entre os genomas

No alinhamento entre os contigs representativos dos genomas dos seis isolados clínicos brasileiros de *A. pleuropneumoniae* sorotipo 8 e genomas fechados dos sorotipos 3 (JL03), 5b (L20), 7 (AP76) e 8 (MIDG2331) foi verificada uma conservação na estrutura do genoma. Os genomas de *A. pleuropneumoniae* compartilham praticamente os mesmos blocos, denominados LCBs – “*Selecting Locally Collinear Blocks*” (Figura 1). Embora tenha sido observada uma conservação foi possível constatar regiões de aquisição/ perda de material genético e rearranjos (Figura 1). Entre essas variações, observamos algumas diferenças pontuais e duas regiões se destacam pelo tamanho e conteúdo de proteínas. Na primeira observação, foi constatada a presença de um bloco diferencial e conservado presente apenas nos genomas dos sorotipos 5 (L20), 7 (AP76) e em quatro dos isolados clínicos do sorotipo 8, sendo eles MV518, MV780, MV1022 e MV5651 (Figura 1). Neste bloco composto por aproximadamente 42.206 pb e com conteúdo GC de 40,37% estão presentes sequências como integrase ([WP_005620278.1](#)), pirofosfatase ([WP_005620280.1](#)), a proteína de recombinação RdgC ([WP_011848390.1](#)), DNA metiltransferase ([WP_011848391.1](#)), antirepressor ([WP_011848395.1](#), [WP_011848414.1](#)), DNA metilase ([WP_011848397.1](#)), endodeoxiribonuclease RuvA ([WP_011848399.1](#)), terminase ([WP_043880767.1](#)), peptidase S14 ([WP_043877971.1](#)) e diversas proteínas de fago ([WP_011848407.1](#), [WP_011848405.1](#), [WP_011848408.1](#), [WP_011848412.1](#)). Já no segundo segmento genômico diferencial foi constatada uma região de inversão e rearranjo de um bloco de aproximadamente 59.649 pb e conteúdo GC de 40,91% (Figura 1). Este segmento foi encontrado apenas no sorotipo 7 e foram relatadas sequências correspondentes a genes que codificam para enzimas carboxilases ([WP_005617934.1](#), [WP_005602033.1](#)), oxidoredutases ([WP_005598646.1](#), [WP_005602054.1](#), [WP_012478542.1](#)), redutase ([WP_005598644.1](#)), fatores de virulência envolvidos na captação de ferro ([WP_005602070.1](#)), integrase ([WP_005617888.1](#)) e transposases ([WP_005599960.1](#)). Foram observadas ainda variações pontuais entre os diferentes genomas. Entre essas diferenças encontramos uma pequena região de aproximadamente 13.000 pb presente nos sete genomas de *A. pleuropneumoniae* sorotipo 8 e as sequências codificadoras encontradas foram: ligase de tRNA-glutamato ([WP_005608501.1](#)), tRNA-Ala, preproteína translocase

([WP_005612726.1](#)), regulador transcricional da família Rha ([WP_039768145.1](#)), antirepressor ([WP_058230489.1](#)), proteína de utilização de propanediol ([WP_039768152.1](#)), proteína da família de prevenção da morte de hospedeiro (Phd) ([WP_005598318.1](#)), toxina YoeB ([WP_005605064.1](#)), quinase de tetracil dissacarídeo ([WP_005598320.1](#), [WP_005608502.1](#)) e nove proteínas hipotéticas. Outra região de aproximadamente 13.180 pb foi encontrada presente nos genomas de *A. pleuropneumoniae* sorotipo 5 (L20), sorotipo 7 (AP76) e em três isolados clínicos sorotipo 8 (MIDG2331, MV1022 e MV518). As sequências correspondentes a genes presentes nessa região codificam para operon de *flpC* ([WP_039709641.1](#)), operon de *flpB* ([WP_039709034.1](#), [WP_011848427.1](#)), proteína de fímbria ([WP_058230512.1](#), [WP_039709035.1](#), [WP_005611759.1](#), [WP_011848428.1](#)), RNA helicases ATP dependente – RhlB ([WP_039709036.1](#)), proteína de ligação a ATP ([WP_005611764.1](#)), substrato de ligação ao transportador ABC de ferro ([WP_005611765.1](#)) e três proteínas hipotéticas.

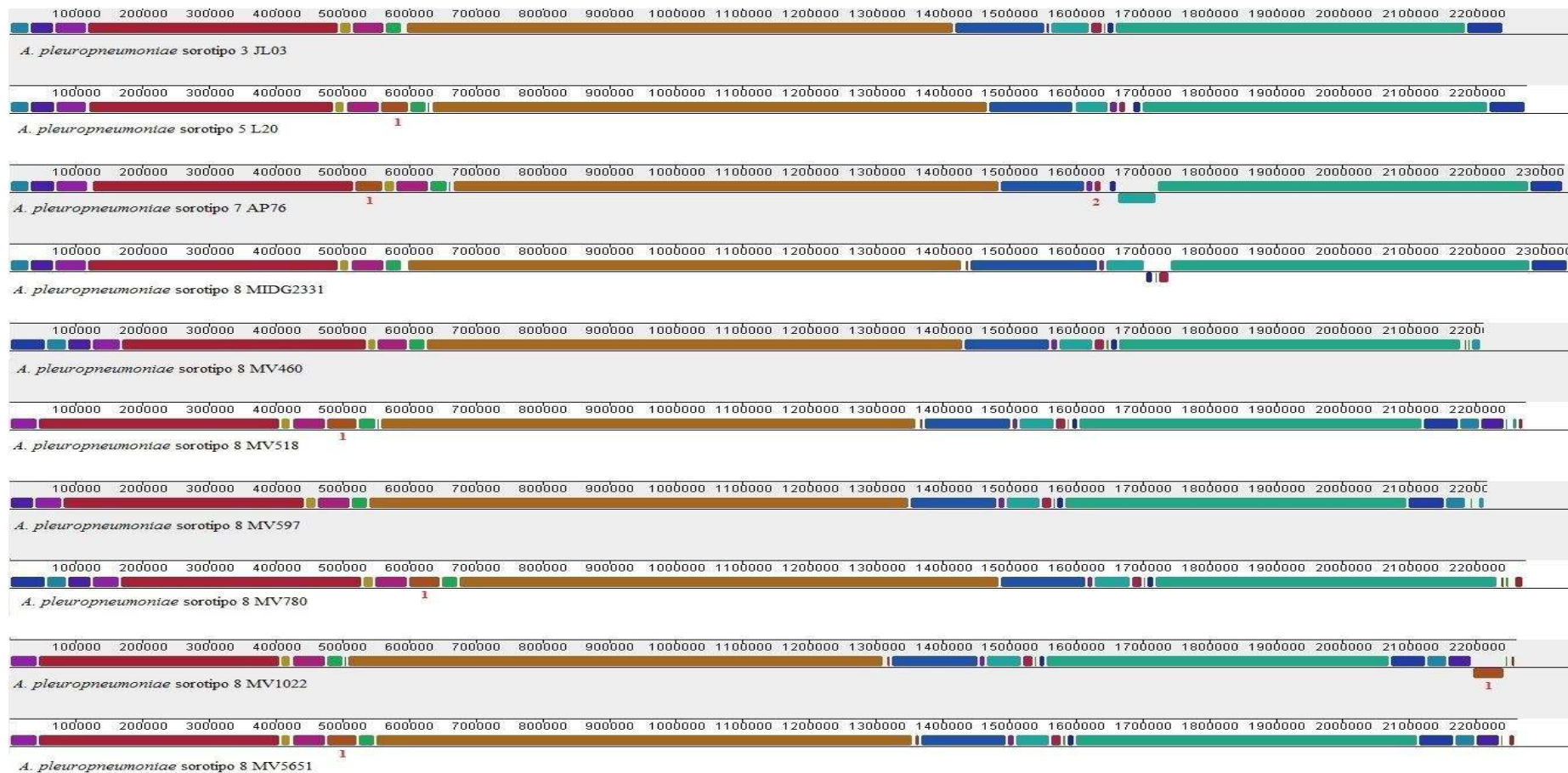


Figura 1. Análise comparativa entre genomas de *A. pleuropneumoniae*. O alinhamento foi realizado utilizando o programa Mauve. Os genomas representados correspondem a *A. pleuropneumoniae* sorotipo 3 JL03; *A. pleuropneumoniae* sorotipo 5 L20; *A. pleuropneumoniae* sorotipo 7 AP76; *A. pleuropneumoniae* sorotipo 8 MIDG2331; *A. pleuropneumoniae* sorotipo 8 MV460; *A. pleuropneumoniae* sorotipo 8 MV518; *A. pleuropneumoniae* sorotipo 8 MV597; *A. pleuropneumoniae* sorotipo 8 MV780; *A. pleuropneumoniae* sorotipo 8 MV1022 e *A. pleuropneumoniae* sorotipo 8 MV5651. As barras horizontais representam o tamanho do genoma (kb). A região identificada em 1 representa a aquisição e perda de informação genômica e a região 2 representa evento de recombinação.

4.1.2. Ilhas genômicas preditas em *A. pleuropneumoniae* sorotipo 8

A busca e predição de ilhas genômicas foi realizada com os genomas de *A. pleuropneumoniae* sorotipo 8. Antes da análise nos genomas não fechados dos seis isolados de *A. pleuropneumoniae* sorotipo 8 foi realizado o alinhamento no programa *Island Viewer* com o genoma completo JL03 do sorotipo 3. O propósito dessa análise foi identificar segmentos de DNA provavelmente oriundos de transferência horizontal gênica e no programa foram caracterizados como ilhas.

A partir dessa análise foi possível identificar diferentes segmentos de DNA sugestivos de ilhas genômicas, mas confirmativos de transferência horizontal de genes. Estes variaram em número e tamanho quando se investigou cada genoma dos isolados clínicos de *A. pleuropneumoniae* sorotipo 8 separadamente. De todos os elementos preditos, cinco foram conservados entre os genomas dos seis isolados clínicos brasileiros de *A. pleuropneumoniae* e o isolado clínico de *A. pleuropneumoniae* MIDG2331 (Tabela 3). Esses segmentos sugestivos de transferência horizontal tiveram tamanhos variando de 5kb a 13,7kb sendo o tamanho entre os isolados próximos com sequências codificadoras equivalentes. Nas sugestivas ilhas genômicas identificadas, encontram-se proteínas de ligação ao DNA, enzimas de restrição, reguladores transcricionais, hemolisinas, proteínas de cápsula e de membrana, proteínas relacionadas ao antígeno O como Wzz e Wzy, transposases e diversas proteínas caracterizadas como hipotéticas. Embora o conjunto de sequências codificadoras presentes nas ilhas sejam similares entre os genomas investigados, nos isolados MV518, MV780, MV1022 e MV5651 identificamos uma região diferencial com sequências de fago como relatado no alinhamento Mauve (Figura 1) e diferenças pontuais entre os diferentes isolados foram detectadas.

Tabela 3. Relação das ilhas genômicas conservadas nos genomas de *A. pleuropneumoniae* sorotipo 8 - MIDG2331 e demais genomas não fechados dos seis isolados clínicos de *A. pleuropneumoniae* sorotipo 8, sendo eles MV460, MV518, MV597, MV780, MV1022 e MV5651.

Sorotipo/Isolado	Ilha 1		Ilha 2		Ilha 3		Ilha 4		Ilha 5	
	Posição	Tamanho (pb)	Posição	Tamanho (pb)	Posição	Tamanho (pb)	Posição	Tamanho (pb)	Posição	Tamanho (pb)
S8 / 460	375130-380266	5136	1085623-1097335	11712	1670482-1678396	7914	1790527-1798179	7652	1800742-1812414	11672
S8 / 518	238782-243919	5137	1011580-1025314	13734	1607319-1615212	7893	1725203-1732855	7652	1735418-1747091	11673
S8 / 597	291214-296351	5137	1010357-1024090	13733	1596929-1604843	7914	1712058-1719711	7653	1722274-1733948	11674
S8 / 780	363786-368912	5126	1139156-1150868	11712	1725225-1733139	7914	1845259-1852911	7652	1855474-1867146	11672
S8 / 1022	245834-250958	5124	964777-978511	13734	1560009-1567924	7915	1682040-1689788	7748	1692351-1704024	11673
S8 / 5651	238892-244028	5136	1001464-1013187	11723	1600432-1608346	7914	1719752-1727405	7653	1729968-1741643	11675
S8 / 2331	322541 - 327503	4962	1055434-1068953	13519	1722576-1730491	7915	1837169-1843924	6755	1846610-1858285	11675

4.1.3. Uso preferencial de códons

Em relação à análise dos códons foi observado elevado padrão de conservação no uso de códons entre todos os sorotipos de *A. pleuropneumoniae* investigados, o que inclui os isolados clínicos do sorotipo 8 estudados neste trabalho (Tabela S1; Figura S2). Na Figura 2 temos representado o uso de códons por *A. pleuropneumoniae*. Os códons com maiores valores de RSCU resultam em maior seleção positiva para seus respectivos aminoácidos (Figura 2A).

Na Figura 2B temos representado a utilização de aminoácidos de *A. pleuropneumoniae*. Não observamos diferenças significativas nas proporções do uso de aminoácidos entre os diferentes isolados e nem entre os sorotipos (Tabela S2; Figuras S2). Os aminoácidos mais utilizados foram leucina (L - 10,55%), alanina (A - 8,71%), isoleucina (I - 6,83%) e valina (V - 6,82%), ao passo que cisteína (C - 1,02%) e triptofano (W - 1,17%) foram os mais raramente utilizados.

O Índice de Adaptação de Códon (CAI) foi calculado para caracterizar a eficiência de tradução dos genes, prever o nível de expressão e demonstrar a probabilidade de uma ORF ser codificadora de proteína. Observamos que a distribuição dos valores para os diferentes sorotipos foram similares e o valor médio do conjunto de dados totais de cada sorotipo variou de 0,705 a 0,710 (Tabela S3). Esses valores podem ser analisados no gráfico boxplot (Figura 3).

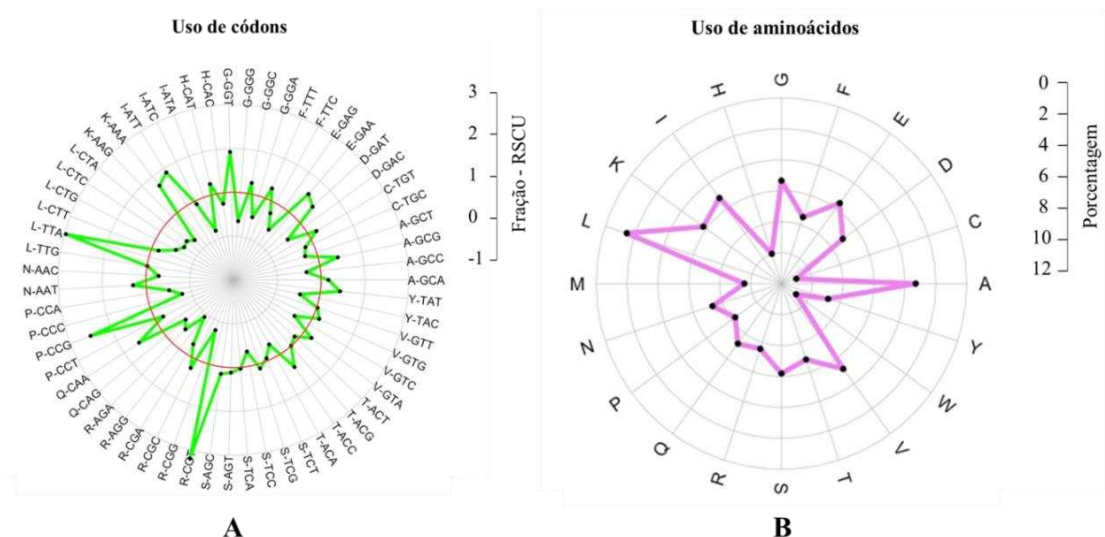


Figura 2. Uso de códons e respectivos aminoácidos de *A. pleuropneumoniae*. **(A)** Tendência no uso de códons representado em mapa circular. Metionina, triptofano e códons de parada foram omitidos. Códons sinônimos para um aminoácido usado com igual frequência tem valor de RSCU = 1, indicado pela linha circular vermelha. **(B)** Porcentagem do uso de aminoácidos representados em mapa circular.

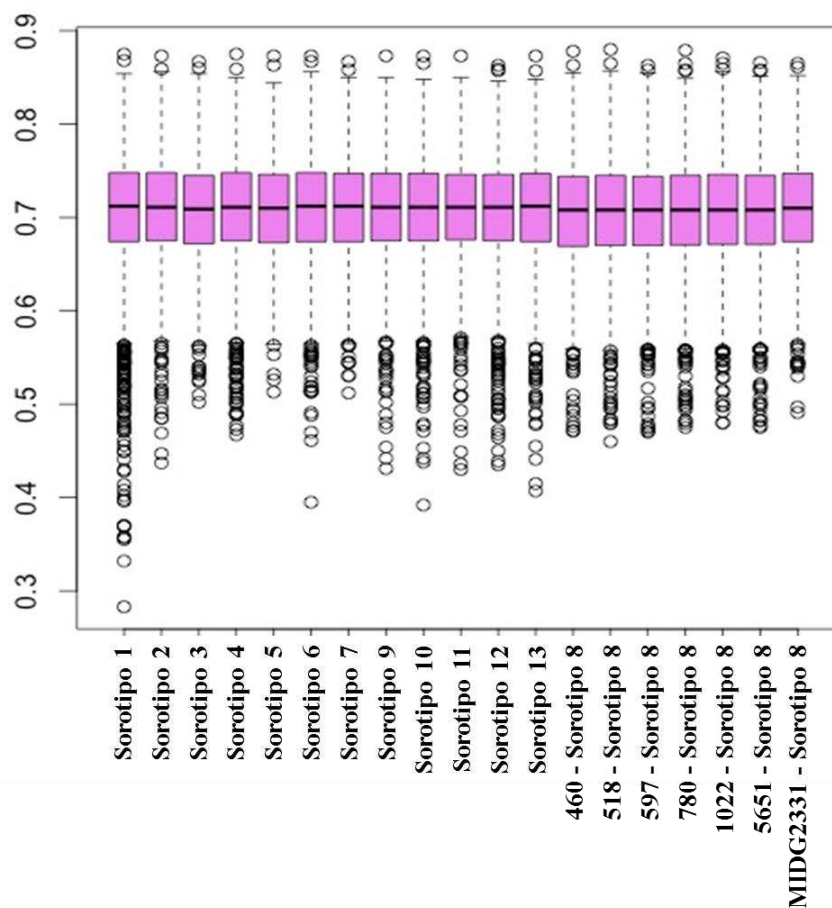


Figura 3: Índice de Adaptação de Códon – CAI obtido para *A. pleuropneumoniae*. Os círculos representam valores fora da mediana.

4.2. Caracterização do proteoma de *A. pleuropneumoniae* sorotipo 8

Previamente à análise do proteoma de *A. pleuropneumoniae* sorotipo 8, os proteomas de todos os sorotipos investigados neste estudo foram analisados. A partir dos *clusters* obtidos do conjunto total de sequências codificadoras da espécie foi realizada a caracterização geral das sequências codificadoras com a distinção da porção *Core*, compartilhada e diferencial. Do total, 1.737 *clusters* foram caracterizados como região *Core*, estando presente nos treze proteomas analisados; 756 como *clusters* de proteínas compartilhadas e 491 *clusters* correspondentes a proteínas diferenciais de cada sorotipo.

Em relação ao proteoma total de cada sorotipo, a porção *Core* equivale em média a 82,49% do proteoma (Tabela 4) evidenciando uma conservação entre os diferentes sorotipos.

Tabela 4. Características dos proteomas dos diferentes sorotipos de *A. pleuropneumoniae* utilizados na análise.

Sorotipo	Referência	Proteínas Totais	CORE		COMPARTILHADO		DIFERENCIAL		
			Proteínas	% do Core	Proteínas	% do Comp.	Proteínas	% do Dif.	
Sorotipo 1	Xu <i>et al.</i> , 2010	2176	1765	81.11	404	18.57	7	0.32	
Sorotipo 2	Zhan <i>et al.</i> , 2010	2064	1774	85.95	275	13.32	15	0.73	
Sorotipo 3	Xu <i>et al.</i> , 2008	2026	1756	86.67	260	12.83	10	0.49	
Sorotipo 4	Xu <i>et al.</i> , 2010	2219	1790	80.67	325	14.65	104	4.69	
Sorotipo 5b	Foote <i>et al.</i> , 2008	2004	1765	88.07	208	10.38	31	1.55	
Sorotipo 6	Xu <i>et al.</i> , 2010	2211	1768	79.96	384	17.37	59	2.67	
Sorotipo 7	Linke <i>et al.</i> , 2008	2113	1774	83.96	327	15.48	12	0.57	
REFERÊNCIA 8									
	* Isolados: MIDG2331, 460, 518, 597, 780, 1022, 5651	Pereira <i>et al.</i> , 2015	2396	1801	75.17	436	18.20	159	6.64
Sorotipo 9			2197	1779	80.97	416	18.93	2	0.09
Sorotipo 10			2170	1774	81.75	321	14.79	75	3.46
Sorotipo 11	Xu <i>et al.</i> , 2010		2184	1767	80.91	414	18.96	3	0.14
Sorotipo 12			2081	1771	85.10	294	14.13	16	0.77
Sorotipo 13			2145	1760	82.05	381	17.76	4	0.19

* O proteoma referência de *A. pleuropneumoniae* sorotipo 8 foi montado a partir de sequências codificadoras do genoma de sete isolados clínicos sequenciados previamente.

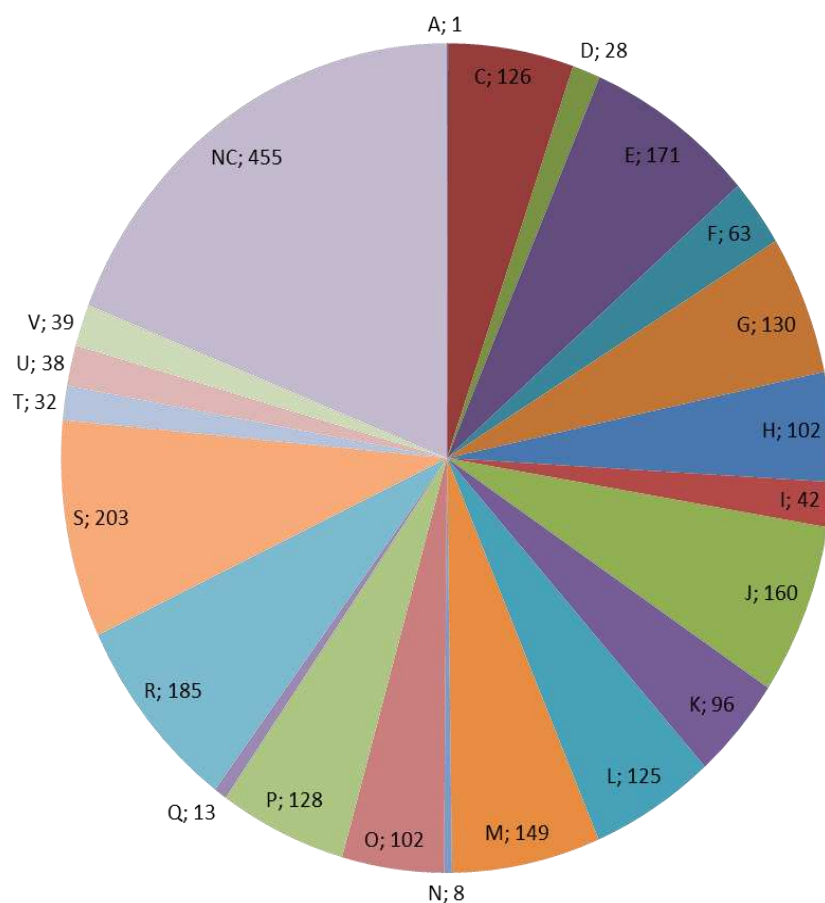
Na anotação de proteínas relacionadas como diferenciais foi observada uma variação na proporção presente entre os sorotipos, o que pode ser parcialmente explicado pelas informações disponíveis nos banco de dados de determinados sorotipos não estarem revisadas, gerando assim, redundância nas análises.

O proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8 gerado a partir dos sete isolados clínicos já sequenciados corresponde a um total de 2.396 sequências codificadoras. Destas, 1.801 (75,17%) consideradas *core*, 436 proteínas compartilhadas (18,20%) com outros sorotipos, mas não com todos e 159 sequências (6,63%) preditas como diferenciais do proteoma referência sorotipo 8.

Entre as 2.396 sequências codificadoras do proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8, um total de 1.941 (81,01%) foram categorizadas na base de dados COG. Entre essas, 1.553 proteínas (80,01%) representam categorias funcionais conhecidas (excluindo S – “*Function unknown*” e R “*General function prediction only*”). No gráfico dos resultados da base de dados COG do proteoma referência *in silico* sorotipo 8 tem representada a distribuição funcional das proteínas nas respectivas categorias (Figura 4).

A partir da distinção das regiões *core*, compartilhada e diferencial do proteoma referência de *A. pleuropneumoniae* sorotipo 8 foi realizada a análise dos *clusters* de grupos de ortólogos. Do total de 1.801 sequências compoendo a porção *core*, 1.685 proteínas foram afiliadas às categorias da base de dados COG. Entre essas, 1.358 proteínas (80,59%) representam categorias funcionais conhecidas (Figura 4). Observamos que a maioria das sequências caracterizadas como *core* estão relacionadas ao metabolismo e transporte de aminoácidos; processos de tradução, estrutura ribossomal e biogênese; biogênese da parede, membrana e envelope celular; produção e conservação de energia, entre outras atividades caracterizadas como essenciais à sobrevivência do patógeno.

Número de Proteínas



DESCRIÇÃO DAS CLASSES COG	CORE	COMP	DIF	TOTAL
A [A] Modificação e processamento de RNA	1	0	0	1
C [C] Conversão e produção de energia	117	9	0	126
D [D] Controle do ciclo e divisão celular, particionamento do cromossomo	25	3	0	28
E [E] Metabolismo e transporte de aminoácidos	158	10	3	171
F [F] Metabolismo e transporte de nucleotídeos	60	2	1	63
G [G] Metabolismo e transporte de carboidratos	117	11	2	130
H [H] Metabolismo e transporte de coenzimas	91	10	1	102
I [I] Metabolismo e transporte e de lipídeos	39	3	0	42
J [J] Tradução, estrutura ribossomal e biogênese	153	7	0	160
K [K] Transcrição	77	10	9	96
L [L] Replicação, recombinação e reparo	95	20	10	125
M [M] Biogênese da parede celular, membrana, envelope	127	21	1	149
N [N] Motilidade celular	6	1	1	8
O [O] Modificação pós traducional, turnover de proteínas e chaperonas	96	6	0	102
P [P] Metabolismo e transporte de íons inorgânicos	104	24	0	128
Q [Q] Biossíntese de metabólitos secundários, transporte e catabolismo	8	5	0	13
R [R] Predição de funções gerais	158	23	4	185
S [S] Função desconhecida	169	31	3	203
T [T] Mecanismos de transdução de sinal	31	1	0	32
U [U] Tráfego intracelular, secreção e transporte vesicular	33	5	0	38
V [V] Mecanismos de defesa	20	18	1	39
NC [NC] Proteínas não categorizadas em COG	116	216	123	455
TOTAL de proteínas afiliadas	1685	220	36	1941
TOTAL de proteínas do sorotipo 8	1801	436	159	2396

Figura 4. Análise de Grupos de Ortólogos (COG) das proteínas que compõe o proteoma referência de *A. pleuropneumoniae* sorotipo 8. As seqüências codificadoras foram caracterizadas como pertencentes ao grupo *core*, compartilhadas e diferenciais com base em análise realizada a partir do conjunto total de proteínas. Comp, Compartilhado; Dif, Diferencial.

Em relação à porção compartilhada, do total de 436 proteínas, 220 foram afiliadas às categorias da base de dados COG, sendo que 166 (75,45%) representam categorias funcionais conhecidas. Observamos uma maioria de proteínas relacionadas ao metabolismo e transporte de íons inorgânicos; biogênese do envelope celular; replicação, recombinação e reparo do DNA; e metabolismo em geral.

A porção diferencial do proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8 apresentou 159 sequências codificadoras, sendo 92 caracterizadas como hipotéticas e apenas 36 sequências foram afiliadas às categorias da base de dados COG. Entre as 36 sequências codificadoras, 29 (80,56%) representam categorias funcionais conhecidas, sendo que dezenove foram relacionadas a processo de transcrição e mecanismos de replicação, recombinação e reparo. Essas sequências codificadoras diferenciais (Tabela 5) estão relacionadas aos processos de regulação e mecanismos de transferência horizontal de genes como plasmídeos e fagos. Foram relatadas nessa porção proteínas de resistência a antibióticos como tetraciclina, florfenicol e sulfonamida, reguladores transcricionais como LysR, proteína de reparo, transposons como transposon gamma-delta resolvase, metilases, proteínas envolvidas na replicação, na biossíntese de lipooligossacarídeos, proteínas envolvidas em transporte como acetiltransferase e simporte de sódio e glutamato.

Tabela 5. Proteínas diferenciais encontradas no proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8.

Identificação	UNIPROT	DESCRIÇÃO DA PROTEÍNA	COG
1	A0A0S4QBQ6	Acetyltransferase	R
2	A0A0S4QBH9	AntA/AntB antirepressor	K
3	A0A0S4QBY2	Antibiotic biosynthesis monooxygenase	R
4	D9PEA9	bacteriophage CII protein	-
5	NC	COG1061: DNA or RNA helicase of superfamily II	K
6	B3H0W4	cro repressor	K
7	Q4W1V9	Dihydropteroate synthase type-2	H
8	A0A0S4QAY3	DNA primase TraC	L
9	A0A0S4QBD3	DNA repair protein	L
10	A0A0S4QDK6	DNA topoisomerase III	L
11	A0A0S4QB64	DNA-binding transcriptional regulator LysR	K
12	E0F444	filamentation induced by cAMP protein Fic	S
13	A0A0S4QBY0	filamentation induced by cAMP protein Fic	S
14	A0A0G3Y3F6	florfenicol/chloramphenicol resistance protein FloR	G
15	A3N393	Gram-positive signal peptide protein: YSIRK family	-
16	A0A0S4QB36	integrase/recombinase XerC	L
17	D9PEB1	LexA family repressor/S24 family protease	K
18	E0EKB4	LysR family transcriptional regulator	K
19	E0E8L3	membrane protein	-
20	Q3HYC8	MobC	-
21	A0A0S4QDG2	N-6 DNA Methylase	V
22	A0A0S4Q8V1	Phage X family protein	-
23	A0A0S4QB66	Plasmid protein of unknown function (plasmid_RAQPRD)	-
24	E0E833	Prepilin peptidase dependent protein D (ApfA)	-
25	A0A0S4QBU2	Putative helicase	-
26	A0A0S4Q9A0	putative HTH-type transcriptional regulator	K
27	A0A0S4Q9B2	recombination protein F	L
28	Q3HYC5	Rep	L
29	A0A0S4QBB8	replicative DNA helicase	L
30	A0A0S4QBA4	sodium/glutamate symporter	E
31	Q3B8S2	TetH	G
32	Q3B8S3	tetracycline repressor protein TetR GN=tetR	K
33	A0A0S4QDE8	Transposon gamma-delta resolvase	L
34	A0A0S4QB51	TraU protein	-
35	A0A0S4QBF4	TraX protein	-
36	A0A0S4QA99	Type II secretion system protein D	N

Esta análise foi baseada no proteoma *in silico* de *Actinobacillus pleuropneumoniae* sorotipo 8. ^aUNIPROT é um centro que coleta informações funcionais de proteínas, sendo os dados derivados dos bancos EMBL-Bank/GenBank/DDBJ. ^bCOG agrupa proteínas com base em relações filogenéticas. NC, não categorizado.

Tabela 5. Cont. Proteínas diferenciais encontradas no proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8.

Identificação	UNIPROT	DESCRIÇÃO DA PROTEÍNA	COG
37	A3MZ09		-
38	D9P9R5		-
39	E0FQZ0		-
40	E0F7J4		-
41	NC		-
42	NC		-
43	NC		-
44	NC		-
45	NC		-
46	NC		-
47	NC		-
48	NC		-
49	E0FQZ0		-
50	NC	hypothetical protein	-
51	NC		-
52	D9PEA3		-
53	NC		-
54	E0E6Z6		-
55	NC		-
56	D9PEJ6		-
57	NC		-
58	NC		-
59	NC		-
60	A9HSL6		-
61	A0A0S4QAU0		-
62	Q4W2Q8		-
63	NC		-
64	A0A0S4Q8W0		-
65	A0A0S4Q9Q3		-
66	A0A0S4Q8T1		-
67	A0A0S4Q971		S
68	A0A0S4Q8U2		-
69	A0A0S4Q8S2		-
70	A0A0S4QBE1	hypothetical protein MIDG2331	-
71	A0A0S4Q920		-
72	A0A0S4Q8R1		-
73	A0A0S4Q8P5		-
74	A0A0S4QB97		-
75	A0A0S4QAU7		-

Esta análise foi baseada no proteoma *in silico* de *Actinobacillus pleuropneumoniae* sorotipo 8. ^aUNIPROT é um centro que coleta informações funcionais de proteínas, sendo os dados derivados dos bancos EMBL-Bank/GenBank/DDBJ. ^bCOG agrupa proteínas com base em relações filogenéticas. NC, não categorizado.

Tabela 5. Cont. Proteínas diferenciais encontradas no proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8.

Identificação	UNIPROT	DESCRIÇÃO DA PROTEÍNA	COG
76	A0A0S4QAW4		-
77	A0A0S4QB60		-
78	A0A0S4QB27		-
79	A0A0S4QB44		-
80	A0A0S4QDP9		-
81	A0A0S4QB87		-
82	A0A0S4QBA8		-
83	A0A0S4QDC9		-
84	A0A0S4QAY4		-
85	A0A0S4QBS1		-
86	A0A0S4QCA7		-
87	A0A0S4QAY1		-
88	A0A0S4QBV7		R
89	A0A0S4QB74		-
90	A0A0S4QB40		-
91	A0A0S4QB59		-
92	A0A0S4QDW8		-
93	A0A0S4QB99		-
94	A0A0S4QDG8		-
95	A0A0S4QDH9	hypothetical protein MIDG2331	-
96	A0A0S4QB00		-
97	A0A0S4QDE3		-
98	A0A0S4QAZ9		-
99	A0A0S4QBT4		-
100	A0A0S4QCB9		-
101	A0A0S4QAZ7		-
102	A0A0S4QBX0		-
103	A0A0S4QB90		-
104	A0A0S4QB75		-
105	A0A0S4QDY1		-
106	A0A0S4QBB6		-
107	A0A0S4QDI1		-
108	A0A0S4QDJ4		-
109	A0A0S4QB88		-
110	A0A0S4QDZ8		-
111	A0A0S4QBG6		-
112	A0A0S4QB32		-
113	A0A0S4QDH3		-
114	A0A0S4QBW0		-

Esta análise foi baseada no proteoma *in silico* de *Actinobacillus pleuropneumoniae* sorotipo 8. ^aUNIPROT é um centro que coleta informações funcionais de proteínas, sendo os dados derivados dos bancos EMBL-Bank/GenBank/DDBJ. ^bCOG agrupa proteínas com base em relações filogenéticas. NC, não categorizado.

Tabela 5. Cont. Proteínas diferenciais encontradas no proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8.

Identificação	UNIPROT	DESCRIÇÃO DA PROTEÍNA	COG
115	A0A0S4QCE2		-
116	A0A0S4QB19		-
117	A0A0S4QBA2		-
118	A0A0S4QBZ5		-
119	A0A0S4QCM0		-
120	A0A0S4QBW1		R
121	A0A0S4QC22	hypothetical protein MIDG2331	-
122	A0A0S4QCV2		-
123	A0A0S4QDA7		-
124	A0A0S4QCL8		-
125	A0A0S4QDL4		-
126	A0A0S4QC39		-
127	A0A0S4QCC7		-
128	A0A0S4QD57		-
129	E0FJK9		[EC:1.1.1.1]
130	E0E8I6	[EC:1.17.4.1]	F
131	E0EL87	[EC:2.4.1.25]	-
132	E0F667	[EC:2.5.1.55]	M
133	D9PAP1	[EC:2.7.1.- 2.7.1.29]	-
134	NC	[EC:2.7.7.6]	-
135	A0A0S4QDI4	[EC:3.4.21.-]	-
136	D9PEA6	[EC:3.6.4.12]	L
137	NC		-
138	NC		-
139	NC		-
140	E0F7I2		K
141	E0EKX5		-
142	NC		-
143	NC		-
144	NC		-
145	B0BSI2	NC	E
146	NC		-
147	Q3HYC5		-
148	A0A0S4Q9S4		E
149	E0EK47		-
150	NC		-
151	NC		-
152	NC		-
153	E0FEU9		-

Esta análise foi baseada no proteoma *in silico* de *Actinobacillus pleuropneumoniae* sorotipo 8. ^aUNIPROT é um centro que coleta informações funcionais de proteínas, sendo os dados derivados dos bancos EMBL-Bank/GenBank/DDBJ. ^bCOG agrupa proteínas com base em relações filogenéticas. NC, não categorizado.

Tabela 5. Cont. Proteínas diferenciais encontradas no proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8.

Identificação	UNIPROT	DESCRIÇÃO DA PROTEÍNA	COG
154	NC		-
155	NC		-
156	D9PEA7		-
157	NC	NC	L
158	E0E991		-
159	D9P9W5		-

Esta análise foi baseada no proteoma *in silico* de *Actinobacillus pleuropneumoniae* sorotipo 8. ^aUNIPROT é um centro que coleta informações funcionais de proteínas, sendo os dados derivados dos bancos EMBL-Bank/GenBank/DDBJ. ^bCOG agrupa proteínas com base em relações filogenéticas. NC, não categorizado.

4.2.1. Análise de similaridade entre os proteomas dos diferentes sorotipos de *A. pleuropneumoniae*

A partir do algoritmo *BLAST* foi realizado o alinhamento entre o proteoma referência de *A. pleuropneumoniae* sorotipo 8 contra o banco de dados contendo todos os demais proteomas oriundos de genomas de *A. pleuropneumoniae* utilizados neste estudo (Tabela 4) gerando grupamentos baseado no padrão de similaridade (Tabela S4). Do total de 2.396 proteínas, 2.196 (91,65%) apresentaram o padrão de similaridade maior que 95%, revelando assim alta conservação do proteoma predito do sorotipo 8 em relação aos demais. Nessa análise, 90 proteínas identificadas no proteoma de *A. pleuropneumoniae* sorotipo 8 não apresentaram correspondência com nenhuma sequência codificadora dos outros proteomas.

Com base na análise dos resultados do *BLAST* (Figura 5), foram obtidos três grupos principais de similaridade relacionados com padrões de virulência. Houve um maior compartilhamento das sequências codificadoras do proteoma predito do sorotipo 8 com as sequências do sorotipo 6, seguido pelo sorotipo 3.

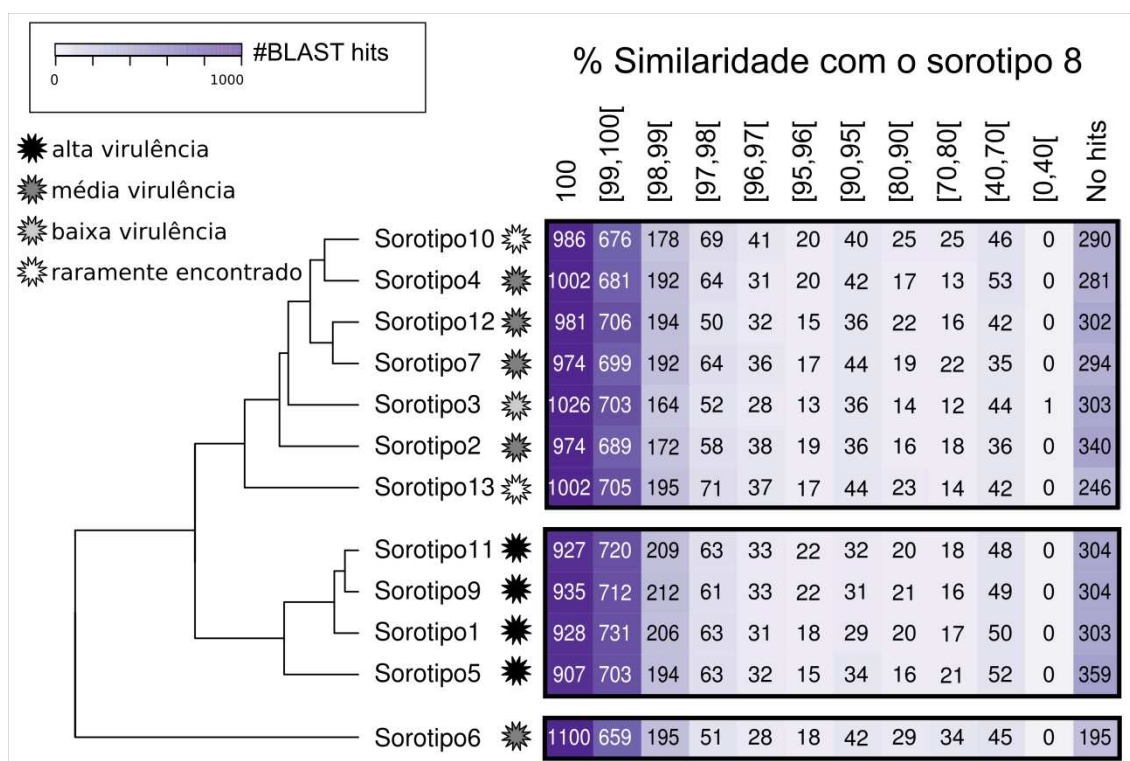


Figura 5. Análise de similaridade entre proteoma referência de *A. pleuropneumoniae* sorotipo 8 e os demais sorotipos. Valores de similaridade indicados em porcentagem. Os proteomas foram agrupados de acordo com a similaridade em relação ao sorotipo 8. Padrão de virulência relacionado como de alta, média e baixa virulência e os raramente encontrados associados a surtos de doenças.

Na análise realizada no algoritmo *BLAST* contrastando as 2.186 seqüências codificadoras do proteoma predito de *A. pleuropneumoniae* sorotipo 8 montado a partir dos seis isolados clínicos brasileiros contra o proteoma do isolado MIDG2331 contendo 2.174 seqüências e vice-versa, encontramos um total de 2.101 correspondências entre os proteomas. Entre as seqüências codificadoras do proteoma predito que não tiveram correspondência com o proteoma do isolado MIDG2331, algumas se encontram na porção considerada diferencial (Tabela 5) e as demais proteínas são proteínas consideradas hipotéticas, proteínas de fago, integrases, proteínas de recombinação, transposases, MobA, MobC e terminases, seqüências de plasmídeos, reguladores transcricionais, e proteína de choque frio.

4.2.2. Fatores de virulência e resistência a antibióticos do proteoma *in silico* de *A. pleuropneumoniae* sorotipo 8

Na análise do proteoma referência de *A. pleuropneumoniae* na base de dados de resistência à antibióticos – ARDB, obtivemos cinco seqüências que apresentaram resultados acima do valor de corte. Entre essas seqüências temos proteínas de resistência à estreptomicina, cloranfenicol, sulfonamida e tetraciclina, sendo as três

últimas presentes na porção diferencial do proteoma referência de *A. pleuropneumoniae* sorotipo 8.

Na análise realizada no PHI base foi feita uma relação das proteínas do proteoma referência de *A. pleuropneumoniae* sorotipo 8 que afetam a interação do patógeno com o hospedeiro. Ao todo, 174 proteínas tiveram um correspondente no banco de dados sendo relacionadas quanto ao nível de patogenicidade e virulência. Observamos que a maior parte das proteínas identificadas nessa análise, totalizando 134 (77,00%) se encontram na porção *core* do proteoma, 38 proteínas (21,85%) na porção compartilhada e apenas duas (1,15%) como diferenciais, sendo prepilina peptidase e proteína de ligação LysR.

Na plataforma PATRIC foi realizada a busca por fatores de virulência nos sete isolados clínicos de *A. pleuropneumoniae* sorotipo 8. Ao todo, cinquenta e quatro sequências codificadoras diferentes foram consideradas fatores de virulência e duas como fatores de resistência a antibióticos, sendo essas de resistência à tetraciclina e à sulfanamida, presentes na porção diferencial.

5. DISCUSSÃO

Actinobacillus pleuropneumoniae pertence à família Pasteurellaceae, ordem Pasteurellales, representada por um grupo diversificado de bactérias comensais e patogênicas para animais e humanos, dentro da classe Gammaproteobacteria e filo Proteobacteria. Os principais gêneros dessa família são *Actinobacillus*, *Haemophilus*, e *Pasteurella* e a distinção destes grupos se baseia no conteúdo GC e em diversas características fenotípicas (Naushad *et al.*, 2015). A caracterização de *A. pleuropneumoniae* nos diferentes sorotipos é baseada em técnicas de sorotipagem e até recentemente o sorotipo 8 foi negligenciado em estudos de identificação devido a falhas nas técnicas empregadas na sorotipagem e na genotipagem (Gottschalk & Lacouture, 2014).

Na análise comparativa entre genomas de *A. pleuropneumoniae* realizada no programa Mauve foram constatadas variações pontuais entre os genomas dos sorotipos. A presença de sequências codificadoras de bacteriófagos e uma pequena variação no conteúdo GC dessa região são indícios da presença de profago. Essa alteração pode ter ocorrido por perda ou ganho em processo de transferência horizontal de genes. Já os processos de inversão e rearranjo no sorotipo 7 em relação aos demais pode ter sido ocasionado pela presença de elementos de inserção, indicando processo de integração de elementos móveis. Mesmo observando um padrão de conservação no genoma de *A. pleuropneumoniae* entre os sorotipos no alinhamento das sequências, existem diferenças relacionadas à aquisição de DNA, inserções, deleções e mutações pontuais que podem vir a originar grande variabilidade. Essa observação já foi relatada em trabalho com *A. pleuropneumoniae* (Xu *et al.*, 2010). Além disso, esse microrganismo é capaz de realizar transformação natural e possui diferentes níveis de competência entre os sorotipos e até mesmo entre os isolados (Bossé *et al.*, 2009, Bossé *et al.*, 2014). As demais variações pontuais observadas entre os genomas dos diferentes isolados e sorotipos, embora sejam pequenas e o conteúdo de proteínas não revelar características relevantes na análise, podem ser uma fonte de variação entre os isolados que acabam resultando em um padrão diferenciado da manifestação da doença. Qualquer alteração, inserção, deleção ou rearranjo que possam vir a ocorrer em um genoma entre os diferentes sorotipos e seus respectivos isolados, resultam em variações globais no estudo da patogênese do microrganismo e podem alterar a expressão de genes adjacentes, gerando um impacto substancial na expressão gênica (Chandler & Mahillon, 2002; Ooka *et al.*, 2009).

Na análise do conteúdo GC, do uso de códons e uso de aminoácidos entre os diferentes sorotipos de *A. pleuropneumoniae* não foram observadas diferenças marcantes. Isso reforça a ideia de que os sorotipos compartilham um conjunto de sequências codificadoras altamente conservadas como já evidenciado pela porção *core* do proteoma e que não trazem diferenças no repertório geral. Em relação ao uso de aminoácidos é importante observar que entre os aminoácidos mais utilizados estão os aminoácidos de cadeia ramificada, sendo eles leucina, isoleucina e valina. Esses aminoácidos de cadeia ramificada já foram relacionados como sendo requeridos na sobrevivência e virulência de *A. pleuropneumoniae* no suíno e a capacidade de sintetizar esses aminoácidos é crítica para patógenos do trato respiratório (Subashchandrabose *et al.*, 2009). Na maioria dos genomas, os códons sinônimos não são usados com frequências iguais e isso tem um efeito marcante na expressão gênica e função celular (Plotkin & Kudla, 2011). O valor médio do CAI, Índice de Adaptação de Códon, próximo a 0,7 indica uma aproximada adequação da utilização de códons dos mRNAs, sendo este valor correlacionado positivamente com o nível de expressão gênica (Sharp & Li, 1987).

A presença do *locus* CRISPR em *A. pleuropneumoniae* foi confirmada nas análises realizadas *in silico*. Os processos de transferência horizontal de genes promovem e aumentam a variabilidade entre os genomas e tem como fator limitante o sistema CRISPR, *Clustered Regularly Interspaced Short Palindromic Repeats*. Este sistema é caracterizado como um sistema imune bacteriano que identifica e degrada plasmídeos estrangeiros, além dos fagos (Popa & Dagan, 2011). Essa é uma estratégia dos microrganismos para reconhecer e fazer a distinção entre DNA exógeno e DNA próprio mantendo assim a integridade genética. CRISPR provê imunidade adquirida (Horvath & Barrangou, 2010) e em combinação com proteínas *Cas* formam o sistema CRISPR/Cas.

A presença de um conjunto de sequências codificadoras diferenciadas entre os isolados, como relatado na predição de ilhas genômicas, pode ser a explicação das diferenças observadas em ensaios experimentais realizados com o organismo modelo *Galleria mellonella* em que foram relatados graus diferentes de patogenicidade a partir de isolados clínicos de diferentes origens e sorotipos (Pereira *et al.* 2015a). As sequências codificadoras diferenciais e a presença de proteínas hipotéticas relacionadas entre os diferentes sorotipos acabam sendo a fonte de variação no tamanho e número das ilhas preditas nos diferentes isolados. Foram ainda identificadas nas ilhas genômicas dos isolados MV460, MV1022 e MIDG2331 proteínas relacionadas ao transposon

Tn10, que codifica o gene *tetC*, que confere resistência à tetraciclina, representando diferenças entre os isolados analisados. O transposon Tn10 já foi caracterizado em *Escherichia coli* e os níveis de resistência a tetraciclina foi afetada em mutantes para TetC (Braus *et al.*, 1984).

Na análise dos *clusters* a partir do conjunto de proteínas compartilhadas entre os sorotipos, verificou-se que o sorotipo 8 compartilha um alto número de proteínas com os sorotipos caracterizados como de média virulência, sendo esses sorotipos o 2, 4, 6, 7 e 12. Já foi relatado que essa característica de média virulência favorece na persistência do patógeno no meio, sendo um fator importante na patogênese da pleuropneumonia suína (Chiers *et al.*, 2010). Através do alinhamento com o algoritmo *BLASTP* foi possível avaliar *in silico* o perfil das proteínas de *A. pleuropneumoniae* de forma comparativa entre os sorotipos e avaliar o perfil de similaridade entre as proteínas do proteoma referência de *A. pleuropneumoniae* sorotipo 8 em relação aos demais sorotipos e foi observado um grande compartilhamento de proteínas do sorotipo 8 com o sorotipo 6, seguido do sorotipo 3. Como já relatado, em análises de sorotipagem, alguns grupos podem sofrer reações cruzadas e serem caracterizados de forma equivocada. Os sorotipos 3, 6 e 8 de *A. pleuropneumoniae* em trabalhos de sorotipagem realizados na América do Norte, constituem um mesmo grupo e a identificação final de determinado sorotipo dentro deste grupo utilizando anti-soros é extremamente difícil. Nesse trabalho, relataram que cinquenta por cento das cepas analisadas apresentaram reações cruzadas entre os sorotipos 6 e 8 por coaglutinação (Gottschalk & Lacouture, 2014). Diante dessas informações é possível estabelecer estratégias que potencializem as técnicas de sorotipagem a fim de evitar que sorotipos que compartilham um alto repertório de proteínas similares sejam agrupados de forma inespecífica, como o que vem ocorrendo entre os sorotipos 6 e 8. O padrão de agrupamento por similaridade foi compatível com a classificação dos sorotipos nas três categorias de virulência (Xu *et al.*, 2010), sendo de alta, média e baixa virulência.

Na análise COG do proteoma predito de *A. pleuropneumoniae* sorotipo 8, a região *core* é caracterizada por genes que codificam funções metabólicas e genes essenciais à sobrevivência e crescimento, relacionados aos genes *housekeeping*. Já a porção acessória, que compreende a porção compartilhada e a diferencial é caracterizada por genes que conferem benefícios ao microrganismo em certas condições ambientais e resulta em adaptação e diversificação. A grande maioria das sequências codificadoras relacionadas aos mecanismos de defesa se encontram na porção compartilhada. Essas sequências compartilhadas por alguns sorotipos acabam

resultando nos diferentes grupamentos observados em processos de sorotipagem com perfil de virulência em comum.

A porção diferencial, como observado nas análises *in silico*, tem uma forte relação com processos de transferência horizontal de genes, contendo proteínas comuns a plasmídeos e fagos. Essa região pode resultar em adaptações importantes, influenciando na interação diferenciada do patógeno com o hospedeiro, além de ter um papel importante na diferenciação de sorotipos, mecanismos de virulência e na replicação do DNA. Em relação à porção diferencial, poucas sequências codificadoras diferenciais foram afiliadas às categorias COG. Como essa parte do proteoma não foi estudada de forma criteriosa, temos uma grande relação de proteínas caracterizadas como hipotéticas, não sendo categorizadas na análise COG. Nas análises contrastando as sequências codificadoras do proteoma *in silico* de *A. pleuropneumoniae* sorotipo 8 montado a partir dos seis isolados clínicos brasileiros contra as sequências do isolado do Reino Unido, MIDG2331, do mesmo sorotipo, existe uma porção capaz de diferenciar os isolados. As sequências que os diferem se relacionam a prováveis elementos genéticos móveis que contribuem para essa variação entre isolados do mesmo sorotipo. A diversidade em sequência, organização e comportamento fazem desses elementos genéticos móveis importantes objetos de estudo por contribuírem com uma modificação na expressão dos genes. Essa diversidade acaba levando à classificação desses elementos como Tn, que seriam os transposons, ISs como as sequências de inserção e os ICEs que são os elementos conjugativos integrativos. A fronteira entre esses elementos é cada vez mais ambígua (Siguier *et al.*, 2014).

Na porção diferencial do proteoma referência sorotipo 8, além das sequências relacionadas à plasmídeos e elementos móveis característicos de processos de transferência horizontal de genes, encontramos proteínas de resistência à tetraciclina, florfenicol e sulfanamida. Embora a porção diferencial do sorotipo 8 seja relativamente baixa (6.64%) é necessária grande atenção, uma vez que esse conjunto de sequências codificadoras podem estar relacionadas à fatores de virulência, de resistência e fatores de competência. Isso acaba por resultar em um padrão diferenciado de resposta quando é levado em consideração o padrão clínico da doença, além de resultar em um escape no controle da doença e gerar variabilidade intra-específica. Em estudos prévios já foi caracterizada a presença de plasmídeos mobilizáveis similares em *A. pleuropneumoniae* e em outros membros da família Pasteurellaceae, conferindo resistência a florfenicol e cloranfenicol (Bossé *et al.*, 2015). A capacidade de mobilização deste entre as espécies foi confirmada e é discutida a importância de se ter o controle de susceptibilidade a

florfenicol em *Actinobacillus* e em outros patógenos Gram Negativos que coexistem no trato respiratório de suínos. Análises com a sequência codificadora de FloR (*chloramphenicol/florfenicol efflux MFS transporter FloR*) a partir do *BLASTP* realizadas com os isolados apresentados nesse estudo apresentaram baixa identidade com o banco de dados disponibilizado, indicando um padrão diferenciado do que se tem descrito na literatura (Bossé *et al.*, 2015). Por ser uma sequência relacionada à resistência ao antibiótico e associada à plasmídeo, não está presente em todos os sorotipos de *A. pleuropneumoniae* e, portanto, é um fator diferencial importante entre os sorotipos dentro da mesma espécie e até mesmo entre isolados. Em trabalho desenvolvido utilizando o organismo modelo, *Galleria mellonella*, avaliou-se a interação desse inseto com o microrganismo *A. pleuropneumoniae* e constatou-se resposta imune celular diferenciada nos hospedeiros em tratamento com diferentes isolados do sorotipo 8. Essa variação se deu em função da presença de diferentes fatores de virulência presentes nos microrganismos que são capazes de desencadear respostas imunes diferenciadas com o objetivo comum de eliminar o patógeno (Pereira *et al.* 2015a). Essa observação reforça a importância da porção diferencial presente entre os isolados de *A. pleuropneumoniae* como explicação para as variações na manifestação clínica da doença, a interação do hospedeiro com o patógeno e o perfil de virulência do mesmo.

A virulência de *A. pleuropneumoniae* está relacionada a diversos fatores como um processo multifatorial. Os sorotipos secretam diferentes combinações das quatro toxinas RTX: ApxI, ApxII, ApxIII e ApxIV. Embora todos os sorotipos causem virtualmente as mesmas lesões, os de alta virulência causam maior taxa de mortalidade (Frey, 2011). A partir das análises realizadas no PHI base, entre as diversas proteínas relatadas, duas proteínas relacionadas à virulência, prepilina peptidase e LysR, estão presentes na porção diferencial do proteoma referência de *A. pleuropneumoniae* sorotipo 8. A proteína prepilina peptidase já foi relatada em processos de transferência genética, virulência e persistência ambiental de uma grande variedade de agentes patogênicos bacterianos gram-negativos (LaPointe & Taylor, 2000), além de ter papel de aderência e secreção em *A. pleuropneumoniae* (Xu *et al.*, 2008), e atividade de clivagem na biogênese do pili tipo IV (Tfp) em outros membros da família Pasteurellaceae (Boekema *et al.*, 2004). Já a proteína regulatória, LysR, constitui uma família de reguladores transcricionais que regulam um conjunto diversificado de genes, incluindo aqueles envolvidos na virulência, metabolismo, *quorum sensing* e motilidade (Maddocks *et al.*, 2008). Este regulador também foi relacionado em processos de

regulação de genes que codificam para urease em bactérias patogênicas (Bossé *et al.*, 2001). Esses dados são importantes para melhor avaliar o controle e uso de antibióticos no tratamento da pleuropneumonia. Mesmo diante desses resultados não é possível concluir que *A. pleuropneumoniae* sorotipo 8 é susceptível a antibióticos, pois os genes de resistência também são frequentemente encontrados em plasmídeos.

Estabelecer as relações entre os genomas bacterianos é importante devido a processos de transferência horizontal de genes, pressões variadas de seleção, aquisição e perda de genes, divergência dos genes, mesmo quando se trata de linhagens intimamente relacionadas (Rokicki *et al.*, 2014). A importância de descrever o conjunto total de proteínas de patógenos bacterianos está relacionada com a compreensão da patogênese, geração de diagnósticos moleculares e identificação de novos agentes vacinais para controle.

6. CONCLUSÕES GERAIS

A análise comparativa realizada entre o proteoma referência de *A. pleuropneumoniae* sorotipo 8 montado de forma representativa a partir dos sete isolados clínicos e o conjunto total de proteínas dos diferentes sorotipos da espécie sugere que esses proteomas sejam conservados e exista um mesmo padrão no uso de códons e de aminoácidos. As diferenças fenotípicas entre os sorotipos que podem estar associadas com a porção diferencial do proteoma podem ser explicadas por informações recebidas na transferência horizontal gênica (THG), uma vez que na porção diferencial foram encontradas proteínas relacionadas aos plasmídeos e fagos, proteínas de recombinação, bem como presença de sequências de inserção, e estas fazem parte de ilhas genômicas preditas. Na análise comparativa dos proteomas dos diferentes sorotipos observamos a formação de três grandes grupos com padrão de virulência distintos, sendo o sorotipo 8 de virulência média. A maior similaridade entre o conjunto de proteínas do sorotipo 8 com o sorotipo 6 resulta muitas vezes em um fator crítico nas análises de sorotipagem, sendo provável fonte dos erros na identificação. A variabilidade genômica observada entre linhagens de uma mesma espécie revelou ser necessário estudar as características do microrganismo de forma mais representativa. A partir da investigação do proteoma de *A. pleuropneumoniae* será possível fornecer informações para um banco de dados robusto que poderá ser utilizado na avaliação de isolados/sorotipos recentemente identificados ou emergentes. Com a caracterização das diferenças entre os isolados do mesmo sorotipo e entre os sorotipos, considerando a porção diferencial e os mecanismos de virulência e resistência a antibióticos, será possível lançar estratégias no diagnóstico da doença e desenvolvimento de candidatos vacinais utilizando proteínas alvo específicas identificadas por análises de bioinformática.

REFERÊNCIAS BIBLIOGRÁFICAS

- Auger, E., Deslandes, V., Ramjeet, M., Contreras, I., Nash, J. H. E., Harel, J., Gottschalk, M., Olivier, M., Jacques, M., 2009, Host-pathogen interactions of *Actinobacillus pleuropneumoniae* with porcine lung and tracheal epithelial cells. *Infection and immunity* 77 (4), 1426–41.
- Baig, A., Weinert, L. A., Peters, S. E., Howell, K. J., Chaudhuri, R. R., Wang, J., Holden, M. T. G., Parkhill, J., Langford, P. R. and BradPit consortium, 2015, Whole genome investigation of a divergent clade of the pathogen *Streptococcus suis*. *Frontiers in Microbiology* 6, 1191.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., Horvath, P., 2007, CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712.
- Bauer, A. M.; Lu, S.; Anderson, J. B.; Chitsaz, F.; Derbyshire, M. K.; DeWeese-Scott, C.; Fong, J. H.; & other authors, 2011, CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research* 39, D225–9.
- Blackall, P. J., Klaasen, H. L. B. M., Van Den Bosch, H., Kuhnert, P., Frey, J., 2002, Proposal of a new serovar of *Actinobacillus pleuropneumoniae*: serovar 15. *Veterinary Microbiology* 84, 47–52.
- Blaha T., von Hamell M.L. & Beilage E, 1994, Slaughter checks-integrated quality assurance systems, meat quality, antibiotics. In: Proceedings of the 13th Congress of the International Pig Veterinary Society, Bangkok, Thailand p.424.
- Boekema, B. K. H. L., Putten, J. P. M. Van, Stockhofe-Zurwieden, N., Smith, H. E., 2004, Host cell contact-induced transcription of the type IV fimbria gene cluster of *Actinobacillus pleuropneumoniae*, *Infection and Immunity*, 72, 691–700.
- Bossé, J.T.; Gilmour, H. D.; MacInnes, J. I., 2001, Novel genes affecting urease activity in *Actinobacillus pleuropneumoniae*. *Journal of Bacteriology*, 183: 1242–1247.
- Bossé, J. T., Sinha, S., Schippers, T., Kroll, J. S., Redfield, R. J., Langford, P. R., 2009, Natural competence in strains of *Actinobacillus pleuropneumoniae*. *FEMS Microbiology Letters*, 298(1), 124–30.
- Bossé, J. T., Soares-Bazzolli, D. M., Li, Y., Wren, B. W., Tucker, A. W., Maskell, D. J., Rycroft, A. N., Langford, P. R. & on behalf of the BRaDPIT Consortium, 2014, The generation of successive unmarked mutations and chromosomal insertion of heterologous genes in *Actinobacillus pleuropneumoniae* using natural transformation. *PLoS ONE* 9, e111252.
- Bossé, J. T., Li, Y., Walker, S., Atherton, T., Crespo, R. F., Williamson, S. M., Rogers, J., Chaudhuri, R. R., Weinert, L. A. & other authors.. Langford, P.R., 2015, Identification of *dfrA14* in two distinct plasmids conferring trimethoprim resistance in *Actinobacillus pleuropneumoniae*. *Journal of Antimicrobial Chemotherapy* 6–11.
- Bossé, J. T., Chaudhuri, R. R., Li, Y., Leanse, L. G., Crespo, F., Coupland, P., Holden, M. T. G., Bazzolli, D. Ms., Maskell, D. J. & other authors., 2016, Complete Genome

- Sequence of MIDG2331, a Genetically Tractable Serovar 8 Clinical Isolate of *Actinobacillus pleuropneumoniae*. American Society for Microbiology 4, 4–5.
- Braus, G., Argast, M., Beck, C. F., 1984, Identification of additional genes on transposon Tn10: *tetC* and *tetD*. Journal of Bacteriology, nov. 160(2):504–509.
- Chandler, M.; Mahillon, J., 2002, Insertion sequences revisited. In: Craig N, & other authors. Mobile DNA II. American Society for Microbiology Press; Washington, DC. pp. 305–336.
- Chien, M. S., Chan, Y. Y., Chen, Z. W., Wu, C. M., Liao, J. W., Chen, T. H., Lee, W. C., Yeh, K. S., Hsuan, S. L., 2009, *Actinobacillus pleuropneumoniae* serotype 10 derived ApxI induces apoptosis in porcine alveolar macrophages. Veterinary Microbiology 135, 327–333.
- Chiers, K., Donné, E., Overbeke, I. V., Ducatelle, R., Haesebrouck, F., 2002, *Actinobacillus pleuropneumoniae* infections in closed swine herds: infection patterns and serological profiles. Veterinary Microbiology 85: 343–352.
- Chiers, K., De Waele, T., Pasmans, F., Ducatelle, R., Haesebrouck, F., 2010, Virulence factors of *Actinobacillus pleuropneumoniae* involved in colonization, persistence and induction of lesions in its porcine host. Veterinary Research 41, 1–16.
- Darling, A.C., Mau, B., Blattner, F.R., Perna, N.T., 2004, Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research 14, 1394–1403.
- Darmon, E. & Leach, D. R., 2014, Bacterial genome instability, Microbiology and Molecular Biology Reviews 78(1):1–39.
- Dhillon, B. K., Laird, M. R., Shay, J. a, Winsor, G. L., Lo, R., Nizam, F., Pereira, S. K., Waglechner, N., McArthur, A. G. & other authors, 2015, IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. Nucleic Acids Research 43, W104–8.
- Duan, C., Xiong, X., Qi, Y., Gong, W., Jiao, J. & Wen, B., 2014, Genomic and comparative genomic analyses of *Rickettsia heilongjiangensis* provide insight into its evolution and pathogenesis. Infection, genetics and evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases. Elsevier B.V 26, 274–82.
- Dubnau, D., 1999, DNA uptake in bacteria. Annual Review of Microbiology 53, 217–244.
- Dubreuil, J.D., Jacques, M., Mittal, K.R., Gottschalk, M., 2000, *Actinobacillus pleuropneumoniae* surface polysaccharides: their role in diagnosis and immunogenicity. Animal Health Research Reviews 1, 73 - 93.
- EMBRAPA, Empresa Brasileira de Pesquisa Agropecuária. Embrapa Suínos e Aves. Disponível em: < <https://www.embrapa.br/suinos-e-aves> > Acesso em: 21 de dezembro de 2015.
- Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G. & other authors, 2010, The Pfam protein families database. Nucleic Acids Research 38, D211–22.

- Foote, S. J., Bossé, J. T., Bouevitch, A. B., Langford, P. R., Young, N. M. & Nash, J. H. E., 2008, The complete genome sequence of *Actinobacillus pleuropneumoniae* L20 (serotype 5b). *Journal of bacteriology* 190, 1495–6.
- Frey, J., 2011, The role of RTX toxins in host specificity of animal pathogenic *Pasteurellaceae*. *Veterinary Microbiology* 153, 51-58.
- Gottschalk, M., Lacouture, S., 2014, *Actinobacillus pleuropneumoniae* serotypes 3, 6, 8 and 15 isolated from diseased pigs in North America. *The Veterinary Record* 174 (18), 452.
- Gouré, J., Findlay, W.A., Deslandes, V., Bouevitch, A., Foote, S. J., MacInnes, J. I., Coulton, J.W., Nash, J. H., Jacques, M., 2009, Microarray-based comparative genomic profiling of reference strains and selected Canadian field isolates of *Actinobacillus pleuropneumoniae*. *BMC Genomics* 10:88.
- Grissa, I., Vergnaud, G., Pourcel, C., 2007, *CRISPRFinder*: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research* 35, 52-57.
- Hirsh, D.C., Martin, L.D., Libal, M.C., 1982, Plasmid-mediated antimicrobial resistance in *Haemophilus pleuropneumoniae*. *American Journal of Veterinary Research* 43, 269–272.
- Krejci, R., Newberry, J., 2011, Pleuropneumonia in pigs - its importance and prevention. *International Pigs Topics* 26, 15-17.
- Horvath, P., Barrangou, R., 2010, CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* 327: 167–170.
- Langille, M.G., Brinkman, F.S., 2009, IslandViewer: an integrated interface for computational identification and visualization of genomic islands. *Bioinformatics* 25, 664–665.
- Langille, M. G. I., Hsiao, W. W. L. & Brinkman, F. S. L., 2010, Detecting genomic islands using bioinformatics approaches. *Nature Reviews Microbiology* 8, 373–82. Nature Publishing Group.
- Lapointe, C. F. & Taylor, R. K., 2000, The Type 4 Prepilin Peptidases Comprise a Novel Family of Aspartic Acid Proteases, 275, 1502–1510.
- Letunic, I., Copley, R. R., Pils, B., Pinkert, S., Schultz, J. & Bork, P., 2006, SMART 5: domains in the context of genomes and networks. *Nucleic Acids Research* 34, D257–60.
- Limin F., Beifang N., Zhengwei Z., Sitao W. and Weizhong L., 2012, CD-HIT: accelerated for clustering the next generation sequencing data. *Bioinformatics*, 28 (23): 3150-3152.
- Linke B., Buettner F., Martinez-Arias R., Goesmann A., Baltés N., Tegetmeyer H., Singh M., Gerlach G.F., 2008, Genome and proteome analysis of *A. pleuropneumoniae* serotype 7, EMBL/GenBank/DDBJ databases.
- Liu, B., Pop, M., 2009, ARDB – Antibiotic resistance genes database. *Nucleic Acids Research*. 37, D443–447.

- Lukjancenko, O., Thomsen, M. C., Voldby Larsen, M. & Ussery, D. W., 2013, PanFunPro: PAN-genome analysis based on FUNctional PROfiles. F1000Research.
- Maddocks SE, Oyston PCF, 2008, Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* 154: 3609–3623.
- Medini, D., Donati, C., Tettelin, H., Masignani, V., Rappuoli, R., 2005, The microbial pan-genome, *Current Opinion in Genetics & Development* 15, 589- 594.
- Morioka, A., Asai, T., Nitta, H., Yamamoto, K., Ogikubo, Y., Takahashi, T., Suzuki, S., 2008. Recent trends in antimicrobial susceptibility and the presence of the tetracycline resistance gene in *Actinobacillus pleuropneumoniae* isolates in Japan. *Journal of Veterinary Medical Science* 70, 1261–1264.
- Naushad, S.; Adeolu, M.; Goel, N.; Khadka, B.; Al-Dahwi, A. & Gupta, R. S., 2015, Phylogenomic and molecular demarcation of the core members of the polyphyletic Pasteurellaceae genera *Actinobacillus*, *Haemophilus*, and *Pasteurella*. *International Journal of Genomics* 2015, 198560. Hindawi Publishing Corporation.
- O'Neill, C., Jones, S. C., Bossé, J. T., Watson, C. M., Williamson, S. M., Rycroft, A. N., Kroll, J. S., Hartley, H. M., Langford, P.R., 2010, Prevalence of *Actinobacillus pleuropneumoniae* serovars in England and Wales. *Veterinary Research* 167, 661-662.
- Ooka, T.; Ogura, Y.; Asadulghani, M.; Ohnishi, M.; Nakayama, K.; Terajima, J.; Watanabe, H.; Hayashi, T., 2009, Inference of the impact of insertion sequence (IS) elements on bacterial genome diversification through analysis of small size structural polymorphisms in *Escherichia coli* O157 genomes. *Genome Research* 19:1809–16.
- Pattison, I. H., Howell, D. G., Elliot J., 1957, A *Haemophilus*-like organism isolated from pig lung and the associated pneumonic lesions. *Journal of Comparative Pathology* 67:320–329.
- Pereira, M. F., Rossi, C. C., de Queiroz, M. V., Martins, G. F., Isaac, C., Bossé, J. T., Li, Y., Wren, B. W., Terra, V. S., Cuccui J., Langford, P. R., Bazzolli, D. M. S., 2015a, *Galleria mellonella* is an effective model to study *Actinobacillus pleuropneumoniae* infection. *Microbiology (Reading, England)* 161, 387–400.
- Pereira, M. F., Rossi, C.C., De Carvalho, F.M., De Almeida, L.G. P., Souza, R. C., De Vasconcelos, A.T.R., Bazzolli, D.M. S., 2015b, Draft Genome Sequences of Six *Actinobacillus pleuropneumoniae* Serotype 8 Brazilian Clinical Isolates: Insight into New Applications. *Genome Announcements* v. 3, p. e01585-14.
- Plotkin, J. B. & Kudla, G., 2011, Synonymous but not the same: the causes and consequences of codon bias. *Nature reviews Genetics* 12, 32–42. Nature Publishing Group.
- Pohl, S., Bertschinger, H. U., Frederiksen, W., Mannheim, W., 1983, Transfer of *Haemophilus pleuropneumoniae* and the *Pasteurella haemolytica*-like organism causing porcine necrotic pleuropneumonia to the genus *Actinobacillus* (*Actinobacillus pleuropneumoniae* comb. nov.) on the basis of phenotypic and deoxyribonucleic acid relatedness. *International Journal of Systematic Bacteriology* 33, 510-514.
- Popa, O., Dagan, T., 2011, Trends and barriers to lateral gene transfer in prokaryotes. *Current Opinion in Microbiology* 14:615–623.

- Rice, P., 2000, The European Molecular Biology Open Software Suite EMBOSS : The European Molecular Biology Open Software Suite 16, 2–3.
- Rokicki, J., Knox, D., Dowell, R. D. & Copley, S. D., 2014, CodaChrome: a tool for the visualization of proteome conservation across all fully sequenced bacterial genomes. *BMC genomics* 15, 65.
- Rossi, C. C., Vicente, A. M., Guimarães, W. V., Araújo, E. F., Queiroz, M. V., Bazzolli, D. M. S., 2013, Face to face with *Actinobacillus pleuropneumoniae*: Landscape of the distribution of clinical isolates in Southeastern Brazil. *African Journal of Microbiology Research* 7, 2916–2924.
- Rycroft, A. N., Garside, L. H., 2000, *Actinobacillus* species and their role in animal disease. *Veterinary Journal* 159:18–36.
- Santos, C. M. A., 2013, Perfil de resistência a antimicrobianos em isolados clínicos de *Actinobacillus pleuropneumoniae* do Estado de Minas Gerais. Dissertação (Mestrado em Microbiologia Agrícola) – Universidade Federal de Viçosa, Viçosa, MG, 90p.
- Schmidt, A., Forne, I. & Imhof, A., 2014, Bioinformatic analysis of proteomics data. *BMC systems biology* 8 Suppl 2, S3. BioMed Central Ltd.
- Sebunya, T. N., Saunders, J. R., 1983, *Haemophilus pleuropneumoniae* infection in swine: a review. *Journal of the American Veterinary Medical Association* 182, 1331-1337.
- Sharp, P. M., and Li, W. H., 1987, The codon adaptation index a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Research* 15:1281-1295.
- Shope, R. E., 1964, Porcine contagious pleuropneumonia. I. Experimental Transmission, Etiology, and Pathology. *Journal of Experimental Medicine* 119, 357-68.
- Siguier P, Gournayre E, Chandler M., 2014, Bacterial insertion sequences: their genomic impact and diversity. *FEMS microbiology reviews* 38(5):865–91.
- Subashchandrabose, S., LeVeque, R. M., Wagner, T. K., Kirkwood, R. N., Kiupel, M., Mulks, M. H., 2009, Branched-chain amino acids are required for the survival and virulence of *Actinobacillus pleuropneumoniae* in swine. *Infection and Immunity*, 77(11):4925–33.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V., 2000, The COG database : a tool for genome-scale analysis of protein functions and evolution 28, 33–36.
- The UniProt Consortium UniProt, 2015, A hub for protein information, *Nucleic Acids Research* 43: D204-D212.
- Vanni, M., Merenda, M., Barigazzi, G., Garbarino, C., Luppi, A., Tognetti, R., Intorre, L., 2012, Antimicrobial resistance of *Actinobacillus pleuropneumoniae* isolated from swine. *Veterinary microbiology* 156, 172–7. Elsevier B.V.
- Weizhong Li & Adam Godzik, 2006, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, *Bioinformatics*, 22:1658-9.

Xu, Z., Zhou, Y., Li, L., Zhou, R., Xiao, S., Wan, Y., Zhang, S., Wang, K., Li, W. & other authors, 2008, Genome biology of *Actinobacillus pleuropneumoniae* JL03, an isolate of serotype 3 prevalent in China. PloS one 3, e1450.

Xu, Z., Chen, X., Li, L., Li, T., Wang, S., Chen, H., Zhou, R., 2010, Comparative Genomic Characterization of *Actinobacillus pleuropneumoniae*. Journal of Bacteriology 192, 5625 - 5635.

Zhan B, Angen Ø, Hedegaard J, Bendixen C, Panitz F, 2010, Draft genome sequences of *Actinobacillus pleuropneumoniae* serotypes 2 and 6. Journal of Bacteriology 192: 5846–5847.

MATERIAL SUPLEMENTAR

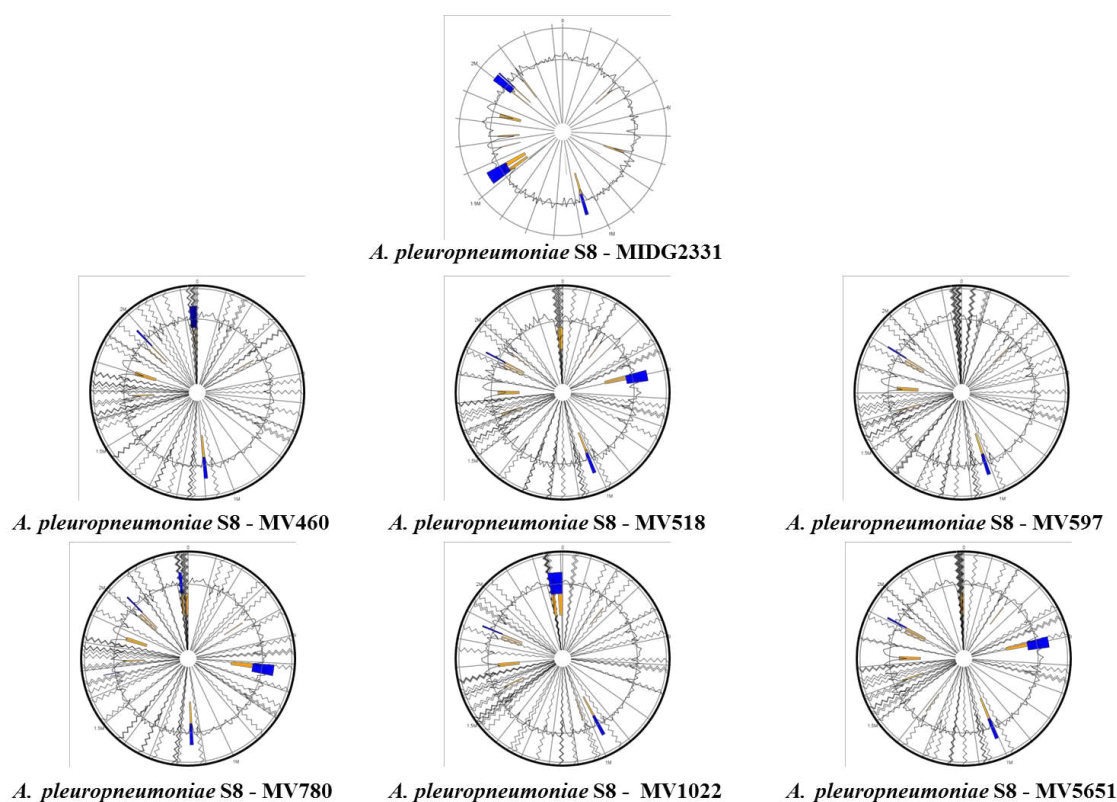


Figura S1. Representação das ilhas genômicas nos genomas de *A. pleuropneumoniae* sorotipo 8. Na primeira figura temos representada a predição de ilhas genômicas do genoma completo MIDG2331. Os demais isolados que possuem genoma não fechado foram tratados no programa com o genoma referência completo de *A. pleuropneumoniae* sorotipo 3 str. JL03, indicado pelo círculo externo. Temos representados os isolados MV460, MV518, MV597, MV780, MV1022 e MV5651. Métodos de predição: Azul, predição por *IslandPath-DIMOB*; Amarelo: *SIGI-HMM*. As linhas tortuosas radiais em cinza nas representações das ilhas genômicas dos isolados indicam os *contigs*.

Tabela S1. Resultado do RSCU para cada um dos códons dos diferentes sorotipos de *A. pleuropneumoniae*.

#Codon	AA	S 1	S 2	S 3	S 4	S 5	S 6	S 7	S 9	S 10	S 11	S 12	S 13	SS- 460	SS- 518	SS- 597	SS- 780	SS- 1022	SS- 5651	S 8- MIDG2331	Média	Desv. padrão	
GCA	A	0.29	0.29	0.29	0.29	0.29	0.30	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.30	0.30	0.30	0.30	0.29	0.00
GCC	A	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.00
GCG	A	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.00
GCT	A	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.00
TGC	C	0.42	0.41	0.40	0.42	0.40	0.41	0.40	0.41	0.41	0.40	0.41	0.41	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.01
TGT	C	0.58	0.59	0.60	0.58	0.60	0.59	0.60	0.59	0.59	0.60	0.59	0.59	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.01
GAC	D	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.00
GAT	D	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.00
GAA	E	0.80	0.80	0.80	0.80	0.81	0.80	0.80	0.81	0.80	0.81	0.81	0.80	0.81	0.80	0.81	0.80	0.80	0.80	0.80	0.80	0.80	0.00
GAG	E	0.20	0.20	0.20	0.20	0.19	0.20	0.20	0.19	0.20	0.20	0.19	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.00
TTC	F	0.37	0.37	0.37	0.37	0.38	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.00
TTT	F	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.00
GGA	G	0.13	0.12	0.12	0.12	0.12	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.00
GGC	G	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.00
GGG	G	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.00
GGT	G	0.47	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.00
CAC	H	0.38	0.38	0.38	0.38	0.38	0.37	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.37	0.38	0.38	0.38	0.38	0.37	0.38	0.00
CAT	H	0.62	0.62	0.62	0.62	0.62	0.63	0.63	0.62	0.62	0.62	0.62	0.62	0.63	0.63	0.63	0.62	0.62	0.62	0.63	0.63	0.62	0.00
ATA	I	0.07	0.06	0.06	0.06	0.06	0.07	0.07	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.00
ATC	I	0.31	0.30	0.31	0.30	0.31	0.30	0.31	0.31	0.31	0.31	0.31	0.30	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.00
ATT	I	0.63	0.64	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.00
AAA	K	0.86	0.86	0.87	0.86	0.87	0.86	0.87	0.87	0.86	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.00

AA, aminoácido; S_, sorotipo.

Tabela S1. Cont. Resultado do RSCU para cada um dos códons dos diferentes sorotipos de *A. pleuropneumoniae*.

#Codon	AA	S 1	S 2	S 3	S 4	S 5	S 6	S 7	S 9	S 10	S 11	S 12	S 13	SS- 460	SS- 518	SS- 597	SS- 780	SS- 1022	SS- 5651	S 8- MIDG2331	Média	Desv. padrão	
AAG	K	0.14	0.14	0.13	0.14	0.13	0.14	0.14	0.13	0.14	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.00
CIA	L	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.00
CTC	L	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.00
CTG	L	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.00
CTT	L	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.00
TTA	L	0.49	0.50	0.50	0.49	0.50	0.50	0.49	0.49	0.49	0.50	0.50	0.49	0.50	0.49	0.50	0.49	0.50	0.49	0.50	0.49	0.49	0.00
TTG	L	0.17	0.16	0.16	0.17	0.16	0.16	0.16	0.16	0.17	0.16	0.16	0.16	0.16	0.17	0.16	0.17	0.16	0.16	0.17	0.16	0.16	0.00
ATG	M	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
AAC	N	0.36	0.35	0.36	0.36	0.36	0.35	0.36	0.36	0.36	0.35	0.36	0.36	0.36	0.35	0.36	0.35	0.35	0.35	0.35	0.35	0.35	0.00
AAT	N	0.64	0.65	0.65	0.64	0.64	0.65	0.65	0.65	0.65	0.64	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.00
CCA	P	0.12	0.12	0.12	0.12	0.12	0.13	0.12	0.13	0.12	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.00
CCC	P	0.06	0.05	0.05	0.06	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.00
CCG	P	0.62	0.63	0.63	0.62	0.63	0.61	0.62	0.62	0.63	0.62	0.63	0.63	0.63	0.62	0.63	0.63	0.62	0.62	0.62	0.61	0.62	0.01
CCT	P	0.20	0.20	0.20	0.20	0.20	0.21	0.21	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.21	0.20	0.00
CAA	Q	0.79	0.79	0.80	0.79	0.80	0.79	0.80	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.80	0.79	0.79	0.79	0.79	0.80	0.79	0.00
CAG	Q	0.21	0.21	0.20	0.21	0.20	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.00
AGA	R	0.10	0.09	0.09	0.10	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.09	0.10	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.00
AGG	R	0.02	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
CGA	R	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.00
CGC	R	0.20	0.21	0.21	0.20	0.21	0.20	0.21	0.20	0.20	0.20	0.20	0.20	0.20	0.21	0.20	0.20	0.21	0.20	0.20	0.20	0.20	0.00
CGG	R	0.05	0.04	0.03	0.05	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.04	0.00

AA, aminoácido; S_i, sorotipo.

Tabela S1. Cont. Resultado do RSCU para cada um dos códons dos diferentes sorotipos de *A. pleuropneumoniae*.

#Codon	AA	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_9	S_10	S_11	S_12	S_13	S_1460	S_1518	S_1597	S_1780	S_11022	S_15651	S_18MIDG2331	Média	Desv. padrão	
CGC	R	0.20	0.21	0.21	0.20	0.21	0.20	0.21	0.20	0.20	0.20	0.20	0.20	0.20	0.21	0.20	0.21	0.20	0.20	0.20	0.20	0.20	0.00
CGT	R	0.50	0.53	0.54	0.52	0.54	0.52	0.53	0.52	0.52	0.53	0.53	0.52	0.54	0.53	0.54	0.53	0.54	0.54	0.53	0.54	0.53	0.01
AGC	S	0.20	0.19	0.19	0.20	0.20	0.19	0.19	0.20	0.20	0.20	0.20	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.00
AGT	S	0.18	0.19	0.19	0.18	0.19	0.19	0.19	0.19	0.18	0.19	0.19	0.19	0.18	0.18	0.18	0.18	0.19	0.18	0.19	0.19	0.19	0.00
TCA	S	0.17	0.17	0.17	0.17	0.17	0.18	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.00
TCC	S	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.00	
TCG	S	0.19	0.18	0.19	0.18	0.19	0.18	0.18	0.18	0.19	0.18	0.19	0.18	0.19	0.18	0.19	0.18	0.18	0.18	0.18	0.18	0.18	0.00
TCT	S	0.15	0.16	0.15	0.15	0.15	0.16	0.15	0.15	0.15	0.15	0.15	0.15	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.00	
ACA	T	0.17	0.17	0.17	0.17	0.17	0.18	0.18	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.00	
ACC	T	0.35	0.36	0.36	0.35	0.36	0.35	0.35	0.35	0.36	0.35	0.36	0.36	0.36	0.35	0.36	0.36	0.36	0.36	0.35	0.35	0.00	
ACG	T	0.26	0.25	0.25	0.25	0.26	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.00	
ACT	T	0.22	0.22	0.22	0.22	0.22	0.23	0.22	0.23	0.22	0.23	0.22	0.22	0.23	0.23	0.23	0.23	0.23	0.22	0.23	0.23	0.00	
GTA	V	0.30	0.31	0.31	0.30	0.31	0.31	0.31	0.31	0.30	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.30	0.31	0.31	0.31	0.00	
GTC	V	0.16	0.15	0.16	0.15	0.16	0.15	0.15	0.15	0.16	0.15	0.16	0.15	0.16	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.00	
GTG	V	0.29	0.29	0.29	0.29	0.29	0.28	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.00	
GTT	V	0.26	0.25	0.25	0.26	0.25	0.26	0.25	0.26	0.25	0.25	0.25	0.26	0.25	0.26	0.25	0.26	0.26	0.26	0.26	0.26	0.00	
TGG	W	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	
TAC	Y	0.28	0.28	0.27	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.27	0.00	
TAT	Y	0.72	0.73	0.73	0.72	0.72	0.73	0.72	0.72	0.72	0.72	0.72	0.72	0.73	0.72	0.73	0.72	0.72	0.72	0.72	0.73	0.00	
TAA	*	0.66	0.70	0.79	0.68	0.79	0.72	0.78	0.70	0.69	0.71	0.69	0.72	0.72	0.74	0.72	0.72	0.74	0.73	0.78	0.78	0.04	
TAG	*	0.18	0.17	0.15	0.16	0.15	0.17	0.15	0.15	0.18	0.16	0.17	0.16	0.19	0.18	0.19	0.18	0.18	0.18	0.16	0.16	0.01	
TGA	*	0.16	0.13	0.06	0.17	0.06	0.11	0.07	0.15	0.12	0.13	0.14	0.13	0.10	0.09	0.09	0.10	0.08	0.09	0.06	0.06	0.03	

AA, aminoácido; S_ , sorotipo.

Tabela S2. Uso de aminoácidos dos diferentes sorotipos de *A. pleuropneumoniae*.

AA	S 1	S 2	S 3	S 4	S 5	S 6	S 7	S 9	S 10	S 11	S 12	S 13	SS- 460	SS- 518	SS- 597	SS- 780	SS- 1022	SS- 5651	S_8- MIDG2331	Média	Desvio-padrão
A	8.62	8.70	8.71	8.69	8.76	8.61	8.68	8.73	8.66	8.74	8.69	8.72	8.70	8.71	8.72	8.70	8.71	8.68	8.67	8.69	0.04
C	1.05	1.02	0.99	1.04	0.98	1.00	1.00	1.03	1.03	1.01	1.04	1.02	1.01	1.02	1.01	1.01	1.01	1.02	0.98	1.01	0.02
D	4.86	4.89	4.88	4.90	4.96	4.92	4.96	4.93	4.88	4.92	4.87	4.91	4.88	4.91	4.88	4.91	4.90	4.95	4.96	4.91	0.03
E	6.32	6.40	6.45	6.37	6.43	6.41	6.42	6.40	6.41	6.40	6.39	6.41	6.40	6.41	6.41	6.43	6.42	6.45	6.39	6.41	0.03
F	4.49	4.50	4.55	4.48	4.47	4.48	4.48	4.48	4.50	4.47	4.50	4.48	4.54	4.52	4.54	4.51	4.53	4.49	4.49	4.50	0.02
G	6.72	6.66	6.65	6.69	6.77	6.69	6.72	6.74	6.65	6.72	6.67	6.70	6.64	6.61	6.64	6.65	6.64	6.64	6.72	6.68	0.04
H	2.03	2.03	2.06	2.02	2.03	2.04	2.04	2.03	2.04	2.03	2.06	2.03	2.04	2.03	2.04	2.03	2.02	2.02	2.04	2.04	0.01
I	6.79	6.81	6.84	6.78	6.76	6.84	6.78	6.80	6.81	6.82	6.80	6.81	6.84	6.82	6.84	6.83	6.83	6.81	6.81	6.81	0.02
K	6.29	6.22	6.24	6.24	6.24	6.35	6.29	6.21	6.26	6.22	6.20	6.22	6.21	6.25	6.20	6.25	6.24	6.29	6.28	6.25	0.04
L	10.43	10.50	10.56	10.44	10.47	10.41	10.43	10.43	10.54	10.45	10.51	10.46	10.60	10.55	10.60	10.53	10.55	10.49	10.49	10.50	0.06
M	2.36	2.37	2.39	2.36	2.37	2.37	2.37	2.36	2.41	2.37	2.40	2.38	2.41	2.39	2.40	2.40	2.41	2.37	2.37	2.38	0.02
N	4.70	4.68	4.64	4.73	4.69	4.78	4.74	4.71	4.67	4.70	4.69	4.70	4.66	4.70	4.67	4.70	4.67	4.74	4.74	4.70	0.03
P	3.71	3.72	3.74	3.69	3.72	3.68	3.69	3.70	3.72	3.71	3.73	3.70	3.72	3.70	3.72	3.70	3.70	3.69	3.69	3.71	0.02
Q	4.77	4.79	4.83	4.79	4.80	4.76	4.78	4.77	4.79	4.78	4.78	4.79	4.80	4.82	4.79	4.80	4.81	4.80	4.78	4.79	0.02
R	4.70	4.50	4.44	4.57	4.45	4.47	4.45	4.54	4.52	4.52	4.56	4.52	4.44	4.45	4.44	4.43	4.43	4.46	4.40	4.49	0.07
S	5.87	5.84	5.77	5.86	5.76	5.86	5.81	5.81	5.81	5.80	5.82	5.84	5.80	5.82	5.79	5.81	5.81	5.78	5.82	5.81	0.03
T	5.19	5.21	5.14	5.21	5.24	5.25	5.25	5.23	5.17	5.23	5.19	5.19	5.14	5.14	5.15	5.16	5.17	5.16	5.25	5.19	0.04
V	6.74	6.82	6.77	6.79	6.81	6.73	6.75	6.77	6.79	6.78	6.77	6.77	6.83	6.79	6.83	6.81	6.82	6.80	6.77	6.79	0.03
W	1.17	1.17	1.16	1.18	1.14	1.17	1.17	1.17	1.18	1.17	1.18	1.18	1.17	1.17	1.17	1.17	1.17	1.17	1.16	1.17	0.01
Y	3.20	3.17	3.19	3.17	3.14	3.20	3.18	3.17	3.17	3.15	3.17	3.17	3.18	3.19	3.18	3.18	3.17	3.18	3.19	3.18	0.01

AA, aminoácido; S_n, sorotipo.

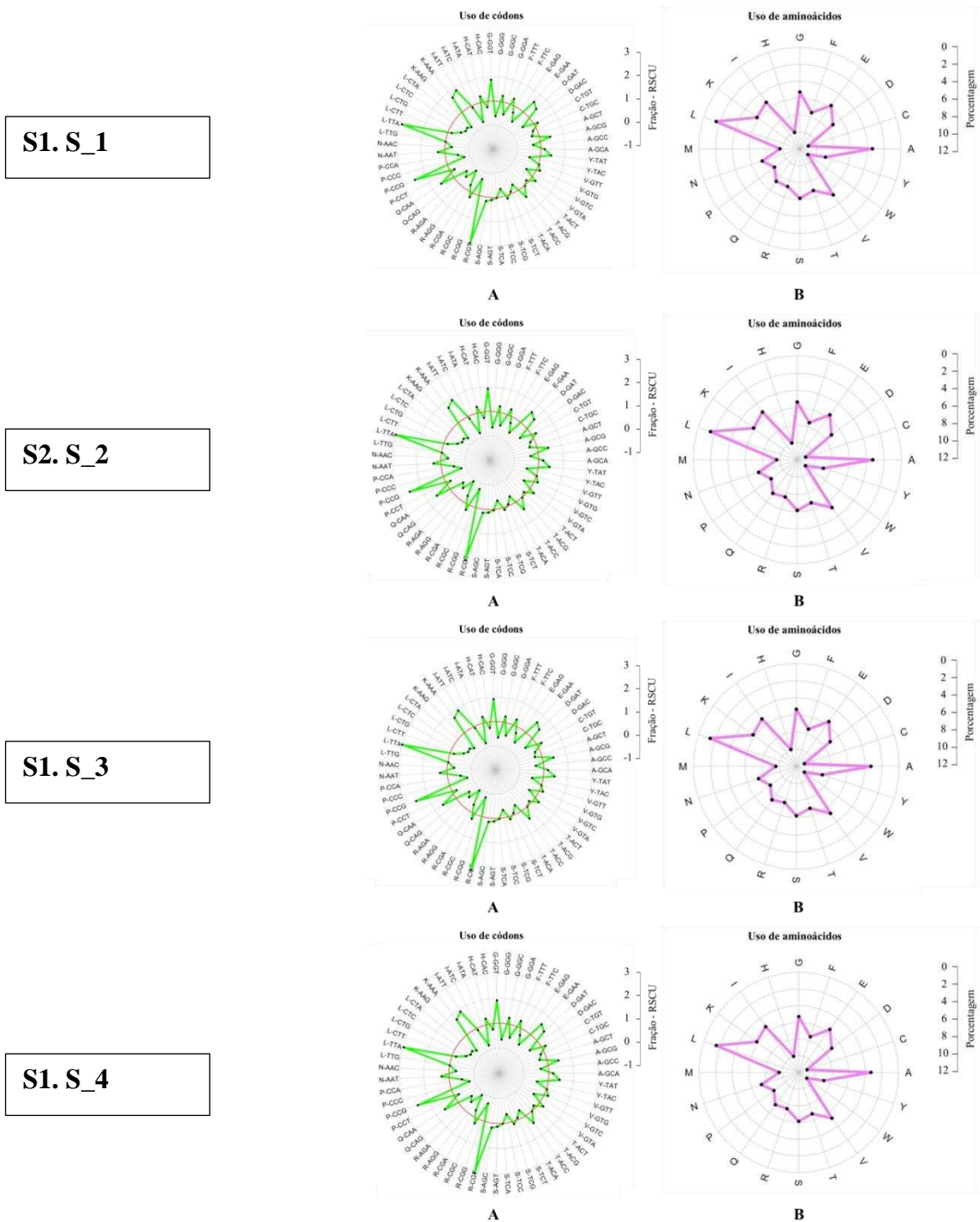
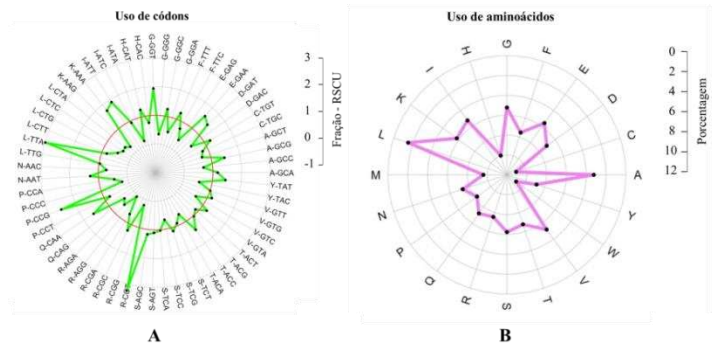
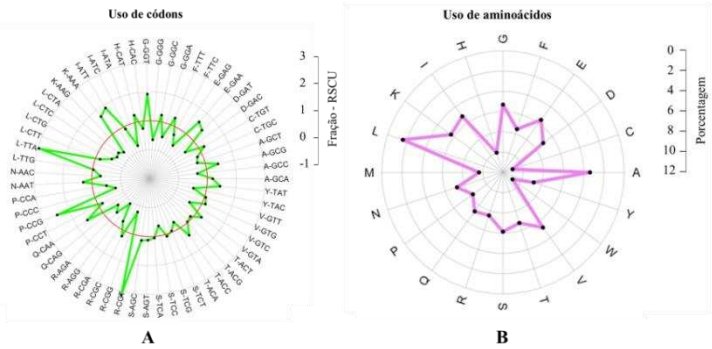


Figura S2. Uso de códons e uso de aminoácidos dos diferentes sorotipos de *A. pleuropneumoniae*. (A) Tendência no uso de códons representado em mapa circular. Metionina, triptofano e códons de parada foram omitidos. Códons sinônimos para um aminoácido usado com igual frequência tem valor de RSCU igual a 1, caracterizado pela linha circular vermelha. (B) Porcentagem do uso de aminoácidos representados em mapa circular. **S1.S_1**: sorotipo 1; **S1.S_2**: sor. 2; **S1.S_3**: sor. 3; **S1.S_4**: sor. 4; **S1.S_5**: sor. 5; **S1.S_6**: sor. 6; **S1.S_7**: sor. 7; **S1.S_9**: sor. 9; **S1.S_10**: sor. 10; **S1.S_11**: sor. 11; **S1.S_12**: sor. 12; **S1.S_13**: sor. 13; **S1.S_8-MV460**: isolado MV460 sor. 8; **S1.S_8-MV518**: isolado MV518 sor. 8; **S1.S_8-MV597**: isolado MV597 sor. 8; **S1.S_8-MV780**: isolado MV780 sor. 8; **S1.S_8-MV1022**: isolado MV1022 sor. 8; **S1.S_8-MV5651**: isolado MV5651 sor. 8; e **S1.S_8-MIDG2331**: isolado MIDG2331 sor. 8.

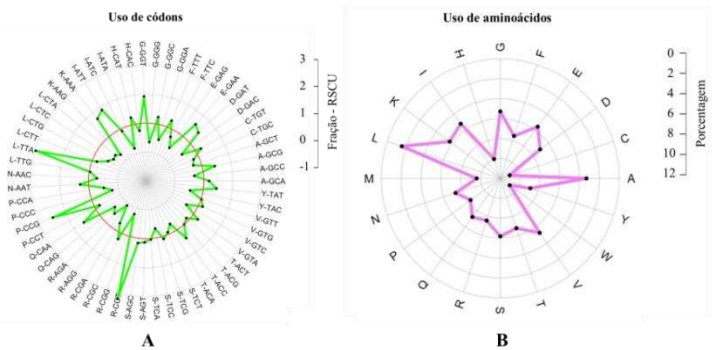
S1.S_5



S1.S_6



S1.S_7



S1.S_9

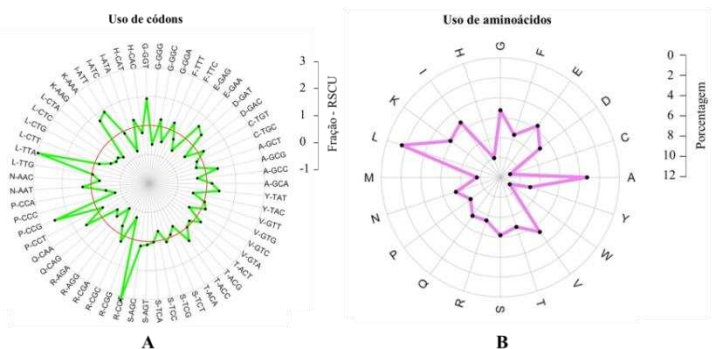
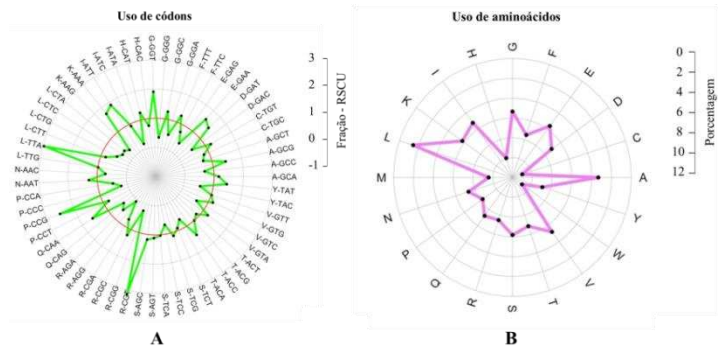
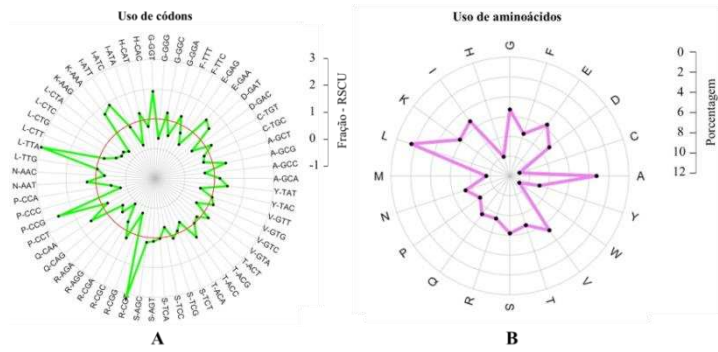


Figura S2. Uso de códons e uso de aminoácidos dos diferentes sorotipos de *A. pleuropneumoniae*. **(A)** Tendência no uso de códons representado em mapa circular. Metionina, triptofano e códons de parada foram omitidos. Códons sinônimos para um aminoácido usado com igual frequência tem valor de RSCU igual a 1, caracterizado pela linha circular vermelha. **(B)** Porcentagem do uso de aminoácidos representados em mapa circular. **S1.S_1:** sorotipo 1; **S1.S_2:** sor. 2; **S1.S_3:** sor. 3; **S1.S_4:** sor. 4; **S1.S_5:** sor. 5 **S1.S_6:** sor. 6; **S1.S_7:** sor. 7; **S1.S_9:** sor. 9; **S1.S_10:** sor. 10; **S1.S_11:** sor. 11; **S1.S_12:** sor. 12; **S1.S_13:** sor. 13; **S1.S_8-MV460:** isolado MV460 sor. 8; **S1.S_8-MV518:**isolado MV518 sor. 8; **S1.S_8-MV597:** isolado MV597 sor. 8; **S1.S_8-MV780:** isolado MV780 sor. 8; **S1.S_8-MV1022:** isolado MV1022 sor. 8; **S1.S_8-MV5651:** isolado MV5651 sor. 8; e **S1.S_8-MIDG2331:** isolado MIDG2331 sor. 8.

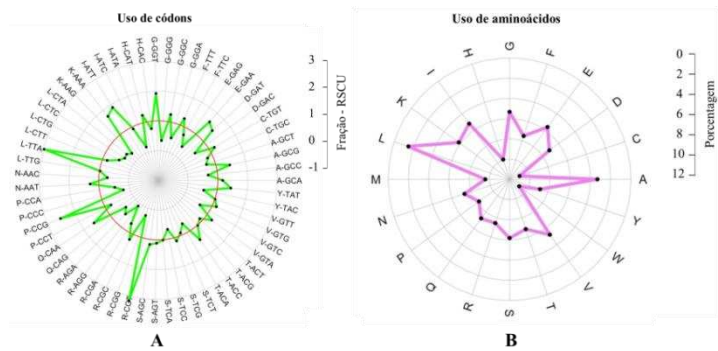
S1.S_10



S1.S_11



S1.S_12



S1.S_13

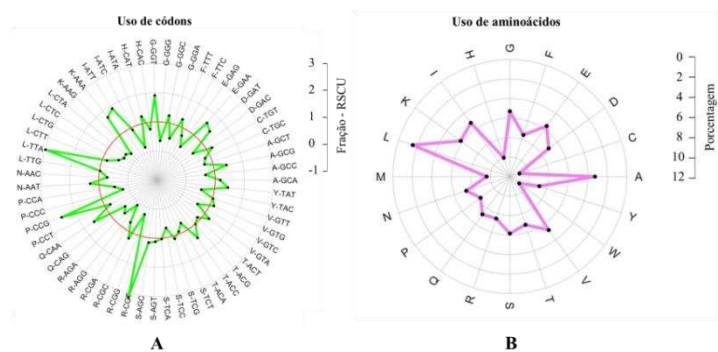
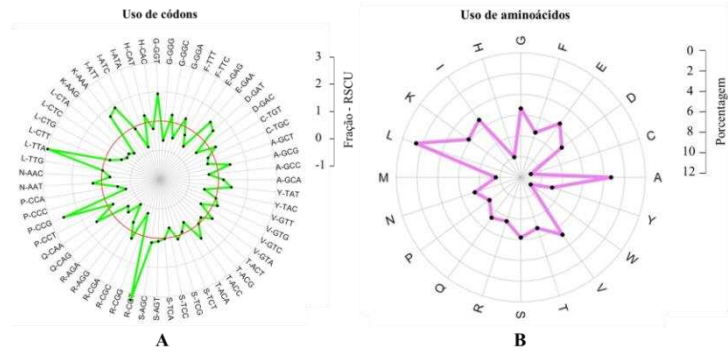
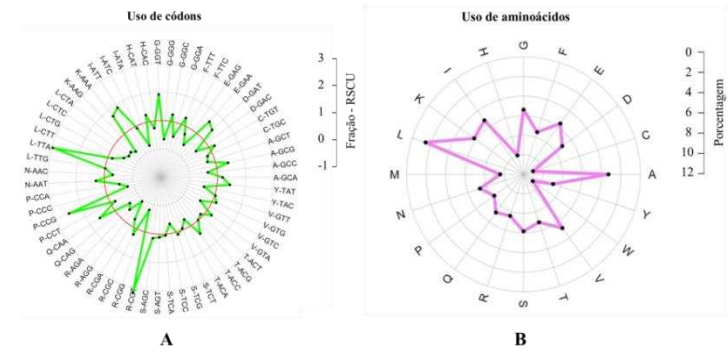


Figura S2. Uso de códons e uso de aminoácidos dos diferentes sorotipos de *A. pleuropneumoniae*. (A) Tendência no uso de códons representado em mapa circular. Metionina, triptofano e códons de parada foram omitidos. Códons sinônimos para um aminoácido usado com igual frequência tem valor de RSCU igual a 1, caracterizado pela linha circular vermelha. (B) Porcentagem do uso de aminoácidos representados em mapa circular. S1.S_1: sorotipo 1; S1.S_2: sor. 2; S1.S_3: sor. 3; S1.S_4: sor. 4; S1.S_5: sor. 5 S1.S_6: sor. 6; S1.S_7: sor. 7; S1.S_9: sor. 9; S1.S_10: sor. 10; S1.S_11: sor. 11; S1.S_12: sor. 12; S1.S_13: sor. 13; S1.S_8-MV460: isolado MV460 sor. 8; S1.S_8-MV518: isolado MV518 sor. 8; S1.S_8-MV597: isolado MV597 sor. 8; S1.S_8-MV780: isolado MV780 sor. 8; S1.S_8-MV1022: isolado MV1022 sor. 8; S1.S_8-MV5651: isolado MV5651 sor. 8; e S1.S_8-MIDG2331: isolado MIDG2331 sor. 8.

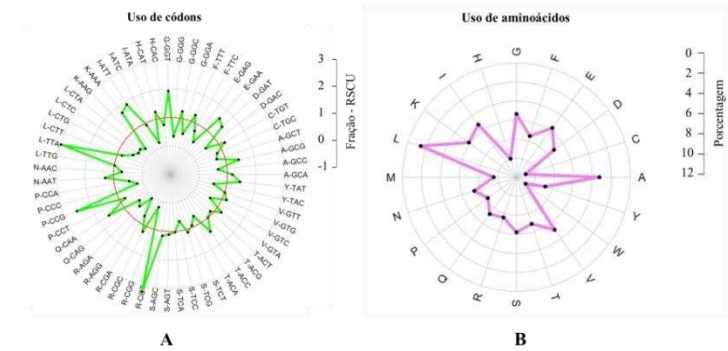
S1.S_8-MV460



S1.S_8-MV518



S1.S_8-MV597



S1.S_8-MV780

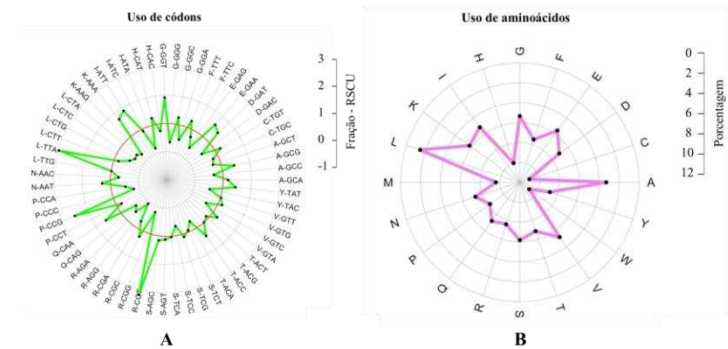
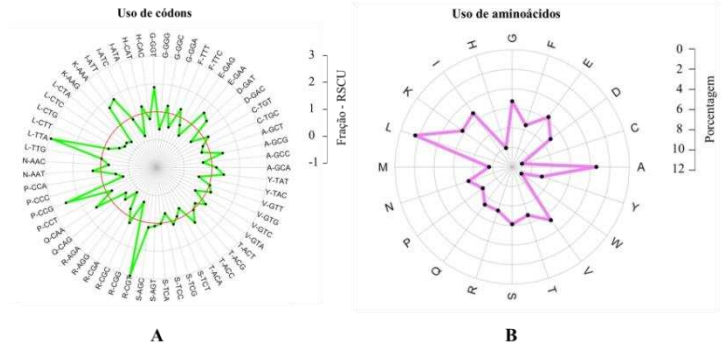
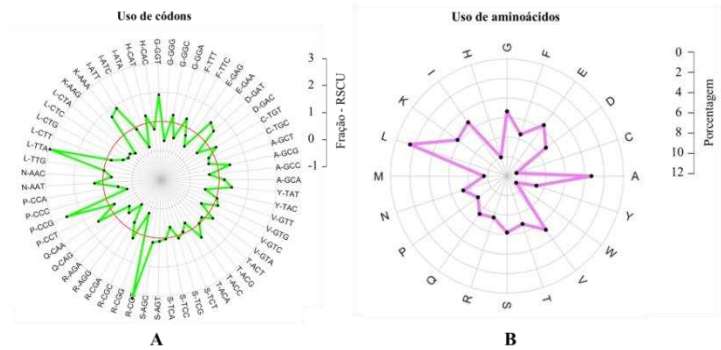


Figura S2. Uso de códons e uso de aminoácidos dos diferentes sorotipos de *A. pleuropneumoniae*. (A) Tendência no uso de códons representado em mapa circular. Metionina, triptofano e códons de parada foram omitidos. Códons sinônimos para um aminoácido usado com igual frequência tem valor de RSCU igual a 1, caracterizado pela linha circular vermelha. (B) Porcentagem do uso de aminoácidos representados em mapa circular. S1.S_1: sorotipo 1; S1.S_2: sor. 2; S1.S_3: sor. 3; S1.S_4: sor. 4; S1.S_5: sor. 5; S1.S_6: sor. 6; S1.S_7: sor. 7; S1.S_9: sor. 9; S1.S_10: sor. 10; S1.S_11: sor. 11; S1.S_12: sor. 12; S1.S_13: sor. 13; S1.S_8-MV460: isolado MV460 sor. 8; S1.S_8-MV518: isolado MV518 sor. 8; S1.S_8-MV597: isolado MV597 sor. 8; S1.S_8-MV780: isolado MV780 sor. 8; S1.S_8-MV1022: isolado MV1022 sor. 8; S1.S_8-MV5651: isolado MV5651 sor. 8; e S1.S_8-MIDG2331: isolado MIDG2331 sor. 8.

S1.S_8-MV1022



S1.S_8-MV5651



S1.S_8-MIDG2331

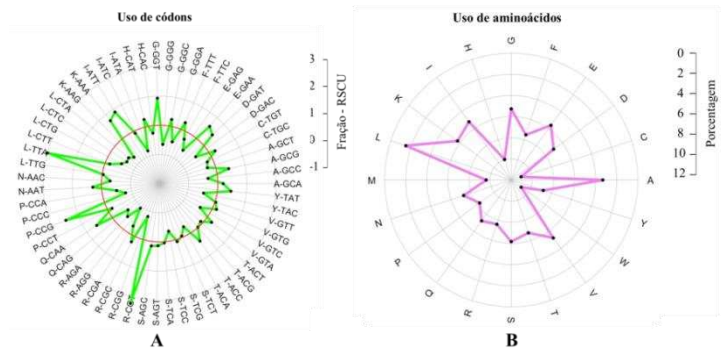


Figura S2. Uso de códons e uso de aminoácidos dos diferentes sorotipos de *A. pleuropneumoniae*. (A) Tendência no uso de códons representado em mapa circular. Metionina, triptofano e códons de parada foram omitidos. Códons sinônimos para um aminoácido usado com igual frequência tem valor de RSCU igual a 1, caracterizado pela linha circular vermelha. (B) Porcentagem do uso de aminoácidos representados em mapa circular. S1.S_1: sorotipo 1; S1.S_2: sor. 2; S1.S_3: sor. 3; S1.S_4: sor. 4; S1.S_5: sor. 5; S1.S_6: sor. 6; S1.S_7: sor. 7; S1.S_9: sor. 9; S1.S_10: sor. 10; S1.S_11: sor. 11; S1.S_12: sor. 12; S1.S_13: sor. 13; S1.S_8-MV460: isolado MV460 sor. 8; S1.S_8-MV518: isolado MV518 sor. 8; S1.S_8-MV597: isolado MV597 sor. 8; S1.S_8-MV780: isolado MV780 sor. 8; S1.S_8-MV1022: isolado MV1022 sor. 8; S1.S_8-MV5651: isolado MV5651 sor. 8; e S1.S_8-MIDG2331: isolado MIDG2331 sor. 8.

Tabela S3. Estatística do Índice de Adaptação de códon com valores de média e desvio padrão representados.

SOROTIPO	MÍNIMO	1st QUARTIL	MEDIANA	MÉDIA	3rd QUARTIL	MÁXIMO	NA's
Sorotipo1	0.283	0.674	0.712	0.705	0.748	0.875	260
Sorotipo2	0.437	0.675	0.711	0.709	0.748	0.873	383
Sorotipo3	0.502	0.672	0.709	0.707	0.745	0.867	412
Sorotipo4	0.467	0.675	0.711	0.709	0.748	0.875	297
Sorotipo5	0.513	0.673	0.710	0.709	0.746	0.873	435
Sorotipo6	0.395	0.674	0.712	0.709	0.748	0.873	224
Sorotipo7	0.512	0.674	0.712	0.710	0.747	0.867	317
Sorotipo9	0.431	0.675	0.711	0.709	0.747	0.873	326
Sorotipo10	0.392	0.675	0.711	0.708	0.747	0.873	340
Sorotipo11	0.430	0.676	0.711	0.709	0.746	0.873	336
Sorotipo12	0.435	0.675	0.711	0.707	0.746	0.863	403
Sorotipo13	0.407	0.674	0.712	0.708	0.747	0.873	346
Sorotipo8-460	0.471	0.669	0.708	0.705	0.744	0.878	52
Sorotipo8-518	0.460	0.670	0.708	0.706	0.745	0.880	178
Sorotipo8-597	0.470	0.670	0.708	0.706	0.744	0.863	44
Sorotipo8-780	0.475	0.671	0.708	0.706	0.745	0.879	--
Sorotipo8-1022	0.480	0.671	0.708	0.706	0.746	0.871	230
Sorotipo8-5651	0.475	0.671	0.708	0.706	0.745	0.866	161
Sorotipo8-							
MIDG2331	0.491	0.674	0.710	0.709	0.747	0.865	274
MÉDIA	---	---	0.710	0.708	---	---	---
DESVIO							
PADRÃO	---	---	0.002	0.002	---	---	---

Tabela S4. Similaridade entre as 2.396 seqüências codificadoras do proteoma referência *in silico* de *A. pleuropneumoniae* sorotipo 8 e o conjunto de proteínas dos demais sorotipos. Noventa proteínas do proteoma *in silico* de *A. pleuropneumoniae* sorotipo 8 não apresentaram nenhuma correspondência com o banco de dados.

Intervalo de similaridade (%)	Nº de proteínas
No hits	90
[40-70[27
[70-80[26
[80-90[30
[90-95[27
[95-98[59
[98-99[81
[99-99,5[134
[99,5 a 100[319
[100]	1603
TOTAL	2.396