UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

Dissertação

Análise evolutiva baseada no genoma de *Pasteurella multocida* proveniente de isolados veterinários

Discente: Raquel Enma Hurtado Castillo
Orientador: Prof. Dr. Vasco Ariston de Carvalho Azevedo
Coorientadora: Profa. Dra. Flávia Figueira Aburjaile

BELO HORIZONTE
2019

Raquel Enma Hurtado Castillo

Dissertação apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Bioinformática.

Orientador: Prof. Dr. Vasco Ariston de Carvalho Azevedo
Coorientadora: Profa. Dra. Flávia Figueira Aburjaile

BELO HORIZONTE

2019

"Não há céu sem tempestades, nem caminhos sem acidentes. Não tenha medo da vida, tenha medo de não vivê-la intensamente".

Augusto Cury

AGRADECIMENTOS

Agradeço a Universidade Federal de Minas Gerais e aos professores com os quais tive o prazer de ter aulas. Sou grata pela excelente formação concedida nestes anos.

Agradeço ao professor Vasco Azevedo pela oportunidade e confiança, por acreditar em minhas capacidades e nunca duvidar de mim.

Agradeço a Flávia Aburjaile pela coorientação neste trabalho, pela confiança, dedicação e acompanhamento durante a minha formação.

À minha mãe Edma, meu Pai Florencio, minha tia Rosana, meus irmãos Nora e Ronal, agradeço infinitamente, vocês foram o meu melhor apoio e fortaleza durante este tempo. Mesmo não presente por perto, foram minha maior inspiração.

Aos meus amigos da moradia que abriram as portas em seus corações para mim: Michel, Victor, Jonas, Gilson, Zé Carlos e Dayane.

Agradeço também aos meus amigos e companheiros de discussões acadêmicas e ajuda, vocês somaram na realização deste trabalho Dennis, Gustavo, Marcus, Álvaro, Diego e Tetsu.

Aos meus amigos e colegas do LGCM que me acompanharam de perto e me deram suporte, em especial a: Alessandra, Ana Carolina, Brenda, Luiz, Rose, Stephany, Thiago, Izabel, Rodrigo, Vinicius, Alfonso e Anne. Sou muito grata por fazerem os meus dias melhores e por aguentarem o meu jeito peruano.

Sou imensamente agradecida pelo acolhimento neste país maravilhoso que me abriu as portas. Tenho uma grande admiração pela cultura e as pessoas lindas que encontrei durante o meu caminho, que me acrescentaram muito profissionalmente e pessoalmente.

RESUMO

Pasteurella multocida é um patógeno comensal e oportunista que causa várias doenças como pneumonia, rinite atrófica, cólera aviária e septicemia hemorrágica em diferentes hospedeiros. Essas doenças causam um alto índice de morbidade e mortalidade, resultando em grandes prejuízos econômicos a nível mundial. No Peru, *P. multocida* é o principal agente causador de pneumonias, possuindo alta taxa de mortalidade em neonatos e filhotes de alpacas na época das geladas, o qual representa uma grande problemática, uma vez que a produção de alpacas é um dos principais recursos econômicos da região dos Andes. Nesse contexto, é essencial o conhecimento de características genômicas da espécie *P. multocida* para o desenvolvimento de estratégias futuras para controle e prevenção dessas doenças. Por esse motivo, foi sequenciada a linhagem UNMSM isolada de uma alpaca com pneumonia e 22 linhagens de *P. multocida* foram recuperadas da base de dados *National Center for Biotechnology Information* (*NCBI*). Com o objetivo de decifrar a diversidade genética e o processo adaptativo de *P. multocida*, este trabalho realizou análises filogenômicas e pangenômicas de isolados de distintas doenças e hospedeiros dessa espécie. Análises de seleção positiva, predição de ilhas genômicas e de genes componentes do genoma acessório foram realizadas nos grupos fenotípicos estabelecidos. Como resultado, comprovou-se a patogenicidade da linhagem UNMSM devido à presença dos principais genes de virulência da espécie e da presença de ilhas de patogenicidade. Entretanto, a análise filogenética demostrou que a linhagem UNMSM diverge das demais linhagens do grupo que causam a pneumonia. A análise da diversidade baseada na filogenia de nucleotídeos, aminoácidos e presença e ausência de genes demonstrou a concordância do agrupamento de linhagens aos grupos fenotípicos estabelecidos. A análise da diversidade por filogenia baseada na presença e ausência de genes, mostrou uma maior variabilidade no mesmo grupo. Em relação aos estudos de pangenoma, os genes acessórios codificam proteínas associadas ao metabolismo e transporte de carboidratos, aminoácidos, biogênese e membrana de parede celular e proteínas hipotéticas. A presença desses grupos funcionais como parte do genoma acessório sugere a importância no processo adaptativo, indicando a fixação no genoma através da pressão seletiva. Análises de seleção positiva no genoma identificaram genes submetidos a seleção, relevantes na interação com hospedeiro, mas não estão implicados no processo de adaptação em patogenias. Finalmente, a presença de ilhas genômicas e determinadas linhagens teriam um papel crucial na preferência por determinado hospedeiro e/ou doença. Análises genômicas e transcricionais poderão fornecer novas informações acerca do papel essencial de eventos de transferência horizontal e polimorfismos de nucleotídeo único envolvidos na diversificação e compreensão dos mecanismos de adaptação de *P. multocida*.

Palavras-chave: *Pasteurella multocida*, adaptação evolutiva, diversidade genômica, pangenoma, filogenômica e genômica comparativa.

ABSTRACT

Pasteurella multocida is a commensal and opportunistic pathogen that causes several diseases such as snuffles, pneumonia, atrophic rhinitis, fowl cholera and hemorrhagic septicemia in different hosts. These diseases cause a high rate of morbidity and mortality, causing important economic loss worldwide. In Peru, *P. multocida* is a mean agent of pneumonic infection, with a high mortality rate in neonatal alpacas in the seasons of icy. This represents a major problem since the production of alpacas is one of the main economic resources from the Andes region. In this context, knowledge of genomic characteristics from *P. multocida* species is essential for the development of future strategies for the control and prevention of these diseases. For this reason, the UNMSM strain isolated from an alpaca with pneumonia was sequenced and 22 *P. multocida* strains were retrieved from the NCBI database. In this case, to decipher the genetic diversity and adaptive process of *P. multocida*, this work aimed to perform phylogenomic and pan genomic analyses of isolates from distinct diseases and hosts. Analysis of positive selection, prediction of genomic islands and accessory genome were performed in the phenotype established phenotypic groups. As a result, the pathogenicity of the UNMSM strain was evidenced due to the presence of the main virulence genes of the species and the presence of pathogenicity islands. However, phylogenetic analysis has shown that the UNMSM strain diverges from other lineages in the pneumonia group. Analysis of the diversity based on the nucleotide/aminoacid phylogeny and presence and absence of genes show the agreement of the clustering of lineages from established phenotypic groups. Analysis of diversity by phylogeny based on the presence and absence of genes showed greater variability in the same group. Notably, the accessory genes encode proteins associated with metabolism and transport of carbohydrates, aminoacids, cell wall and membrane biogenesis, and hypothetical proteins. The presence of these functional groups as part of the accessory genome suggest the importance in the adaptation process, suggesting their fixation in the genome through selective pressure. Positive selection analysis identified genes under selection, relevant in host interaction but are not involved in the pathogen adaptation process. In conclusion, the presence of genomic islands and determinate strains would play a crucial role in host/disease predilection. Transcriptional and genomic analyzes may provide new information about the essential role of horizontal transfer events and single nucleotide polymorphisms involved in the diversification and understanding of the mechanisms of *P. multocida* adaptation.

Keywords: *Pasteurella multocida*, evolution, diversity genomics, pangenome, phylogenomic, comparative genomics.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS

API	Analytical Profile Index
BRD	Bovine Respiratory Disease
BRIG	Blast Ring Image Generator
BLASTp	Basic Local Alignment Search Tool – protein
BOX-PCR	BOX- A1R element
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CC	Clonal Complex
CCT	CGView Comparison Tool
CDSs	Coding sequences
CG-SNPs	Complete genomics Single Nucleotide Polymorphisms
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
COG	Cluster of Orthologous Genes
DNA	Deoxyribonucleic acid
ERIC-PCR	Enterobacterial Repetitive Intergenic Consensus Polymerase Chain Reaction
EUA	Estados Unidos da América
FAPEMIG	Fundação de Amparo à Pesquisa do Estado de Minas Gerais
FMV	Facultad de Medicina Veterinaria
ICB	Instituto de Ciências Biológicas
GEIs	Genomic Islands
G+C	Guanine + Cytosine
GIPSy	Genomic Island Prediction Software
HGT	Horizontal Gene Transfer
HS	Hemorrhagic Septicemia
ICEs	Integrative and Conjugative Elements
LGCM	Laboratório de Genética Celular e Molecular
LGT	Lateral Gene Transfer
LPS	Lipopolysaccharide
MCL	Markov Clustering Algorithm
MGEs	Mobile Genetic Elements
MIs	Metabolic Islands
MLST	Multilocus Sequence Typing

MDR	Multidrug Resistant
NCBI	National Center of Biotechnology Information
PAIs	Pathogenicity Islands
PATRIC	Pathosystems Resource Integration Center
PFGE	Pulsed-field gel electrophoresis
RAPD-PCR	Ribotyping, random amplification of polymorphic DNA
RAST	Rapid Annotation using Subsystem Technology
REA	Restriction Endonuclease Analysis
RIRDC	Rural Industries Research and Development Corporation
RIs	Resistance Islands
SNPs	Single Nucleotide Polymorphisms
STs	Sequences Types
SIs	Symbiotic Islands
TA	Toxin-antitoxin
UNMSM	Universidad Nacional Mayor de San Marcos
UNA	Universidad del Altiplano
UFMG	Universidade Federal de Minas Gerais
UniProt	Universal Protein Resource
VFDB	Virulence Factor Database
VFs	Virulence Factors
WGS	Whole Genome Sequencing
wgMLST	Whole-genome Multilocus Sequence Typing

APRESENTAÇÃO

Colaborações

Este trabalho foi realizado no Laboratório de Genética Celular e Molecular (LGCM) do Instituto de Ciências Biológicas (ICB), Universidade Federal de Minas Gerais (UFMG) supervisionado pelo Prof. Dr. Vasco Azevedo e Profa. Dra. Flávia Aburjaile. Este trabalho foi realizado em colaboração com o Prof. Dr Lenin Maturrano, Pesquisador e Professor da *Universidad Nacional Mayor de San Marcos* (*UNMSM*), Seção de Biologia Molecular e Genética da Faculdade de Medicina Veterinária (FMV), localizado no Peru.

Este trabalho teve o apoio financeiro das seguintes agências brasileiras: Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Além disso, o projeto contou com o apoio financeiro do Programa Nacional de Inovação para a Competitividade e Produtividade (*Innovate* Peru), visando o controle e prevenção de infeções de âmbito veterinário na região dos Andes.

Prefácio

Durante o desenvolvimento da dissertação do mestrado, iniciou-se uma colaboração entre a equipe de pesquisa do Peru da Secção de Biologia e Genética Molecular FMV-UNMSM, e o Laboratório de Genética Celular e Molecular, ICB-UFMG. Em ambos laboratórios, visava-se o estudo genômico de bactérias de interesse público e veterinário, para o desenvolvimento de estratégias de controle e prevenção em diferentes animais de produção. Durante esse período foram desenvolvidos projetos em conjunto, dos quais resultaram na publicação de três artigos científicos.

O primeiro artigo foi elaborado como parte do desenvolvimento do projeto: "Vacinologia reversa para o desenvolvimento de nova geração de vacinas para controle e/ou prevenção de pneumonia em alpacas com pasteurelose", subvencionado pelo Programa *Innovate* Peru. Este projeto apresenta relevância devida a que a produção de camelídeos sul-americanos ser um dos principais recursos econômicos para as comunidades andinas no Peru, isto pelo consumo de carne e a demanda de venda da fibra da alpaca. Porém a produção desses animais diminuiu principalmente por infeções pneumônicas, que são uma das principais causas da mortalidade em neonatos e filhotes de alpacas. *Pasteurella multocida* é uma bactéria Gram-negativa, membro da microbiota normal do trato superior desses animais. No entanto, devido a fatores ambientais e de manejo, os animais apresentam uma susceptibilidade imunitária, a qual permite o desenvolvimento de infecções bacterianas, como pneumonia (Rosadio *et al*., 2011). *P. multocida* é o principal agente patogênico em animais domésticos e selvagens, causando pneumonias (em animais ruminantes) e septicemia hemorrágica (bovinos), rinite atrófica (porcos), cólera aviária e mais usualmente infeções na pele (humanos) (Wilson and Ho, 2013).

Nesse contexto, redigiu-se uma revisão sistemática de *P. multocida*, sendo o primeiro artigo de revisão a abranger estudos envolvendo áreas, como a genômica e epidemiologia, para responder questões a respeito da diversidade genômica e adaptação evolutiva. Essa revisão intitulada: "*Insights of Pasteurella multocida: genomic diversity, epidemiology and evolutionary adaptation*" foi recentemente submetida ao periódico *International Journal of Medical Microbiology*, e encontra-se no **Capitulo 1 - REVISÃO DA LITERATURA** deste manuscrito.

Tratanando-se de problemas relacionados às infecções pneumônicas em alpacas estão as medidas do controle e prevenção não eficientes, o que causa o uso de antibióticos genéricos, e sem um controle da susceptibilidade e resistência a antimicrobianos. Além disso, não se tem o registro adequado em relação ao local de isolamento e detecção da doença ao longo dos anos. Tendo em vista esse cenário, em 2014, um estudo de campo foi realizado pela *Universidad del*

Altiplano (*UNA*), no estado *de Puno,* Peru. Foram coletadas 24 linhagens de *P. multocida de* 46 animais com evidência clínica de pneumonia procedentes do centro experimental *la Raya*. Todos esses isolados de *P.* m*ultocida* pertencem ao sorogrupo A e genótipo L6. Por meio de técnicas moleculares, como ERIC-PCR e BOX-PCR, acredita-se que o surto ocorreu procedente de uma infeção comum, devido à homogeneidade genética encontrada na maioria das linhagens (Rocío *et al*., 2017). Dentre os isolados obtidos neste estudo, foi selecionada a linhagem UNMSM. A partir dos dados genômicos foi possível identificar importantes genes de virulência e 75 genes relacionados a fagos. Esse resultado nos permitiu inferir sobre a virulência da linhagem e a plasticidade genômica da mesma. O resultado deste trabalho foi apresentado no artigo publicado no periódico *Journal of Genomics*: "*Draft Genome Sequence of a Virulent Strain of Pasteurella Multocida Isolated From Alpaca*" o qual está no **Capítulo 2 – RESULTADOS** desta dissertação.

Além de provocar doenças em alpacas, com alto índice de virulência e patogenicidade, como demonstrado no artigo anterior, *P. multocida* é um patógeno veterinário que afeta a produção de bovinos, porcos, aves e coelhos. O tratamento com antibióticos é a estratégia mais utilizada contra essa bactéria, porém apresenta uso excessivo e indiscriminado. Além do uso dos antibióticos, vem sendo utilizadas vacinas composta por toxinas, bacterinas, bactérias atenuadas, proteínas e peptídeos recombinantes, que na maioria dos casos não confere completa imunidade e a maioria não são comercialmente disponíveis (Wilson and Ho, 2013). Com isso, ainda não é suficiente os esforços para combater os diferentes processos infeciosos, por isso é de suma importância conhecer características em nível genômico que contribuem no tropismo e patogenicidade desse microrganismo, para a busca e melhora das estratégias de prevenção e controle. Com o contínuo aumento de genomas sequenciados é possível acessar a informação genômica de linhagens procedentes das diversas doenças e hospedeiros.

Em meio a essas questões, surge a ideia de realizar análises pangenômica, predição de elementos genéticos móveis e de genes submetidos à pressão seletiva, que possam estar envolvidos na diversidade intraespecífica e na adaptação patogênica, permitindo encontrar diferenças genômicas entre linhagens procedentes de diversos fenótipos. A contribuição deste trabalho permitiu conhecer os recursos genômicos que promovem a adaptabilidade aos diferentes nichos e patogênese. Esses resultados conferem dados importante para novas estratégias de controle e tratamento de diversas doenças acometidas por essa espécie. Os resultados deste trabalho foram apresentados no artigo publicado no periódico *GENE "Pan-genomic approach shows insight of genetic divergence and pathogenic adaptation of Pasteurella multocida*", e exibido no **Capítulo 2 - RESULTADOS**. Além disso, foi realizada

a predição de sequências tipo e de grupos clonais que permitiram agregar valor as análises filogenômicas. O resultado permitiu a identificação de linhagens ancestrais associadas à capacidade de infecção de um hospedeiro e/ou doença, resultados que corroboram com as análises populacionais. Os detalhes desses resultados estão apresentados na seção de **RESULTADOS ADICIONAIS**.

Ao fim deste manuscrito, temos também uma seção composta por discussão, conclusão e perspectivas deste trabalho de dissertação ao longo destes últimos dois anos, reunindo os resultados obtidos em nossos estudos e em trabalhos científicos prévios para a compreensão de *P. multocida* no contexto evolutivo, epidemiológico e genômico.

Objetivos

Objetivo geral

Analisar a diversidade genômica e adaptação evolutiva de 23 linhagens de *Pasteurella multocida* isoladas de diferentes hospedeiros e doenças.

Objetivos específicos

- Isolar e caracterizar microbiologicamente a linhagem UNMSM de *P. multocida* procedente do estado do Cusco, Peru;
- Sequenciar, montar, anotar o primeiro isolado pertencente a linhagem UNMSM *P. multocida*, proveniente de um pulmão de alpaca com pneumonia;
- Analisar a diversidade intraespecífica através da relação filogenética baseada em substituição nucleotídica/aminoacídica e presença e ausência de genes;
- Identificar e caracterizar a presença de genes e de ilhas genômicas possivelmente associadas a capacidade de infectar um determinado hospedeiro e/ou causar uma doença específica;
- Identificar genes submetidos a pressão seletiva que estejam envolvidos com a doença ou adaptação ao hospedeiro.

CAPÍTULO 1: REVISÃO DA LITERATURA

Artigo de revisão intitulado "***Insights of Pasteurella multocida: genomic diversity, epidemiology and evolutionary adaptation*".** O artigo foi submetido ao periódico *International Journal of Medical Microbiology* (*Elsevier*) em janeiro de 2019.

Está revisão descreve a história genômica desde o primeiro genoma isolado de *P. multocida* proveniente de um caso de cólera aviária reportado no ano de 2011, até os últimos estudos de genômica comparativa em nível populacional, com o aumento crescente de genomas sequenciados provenientes de origem veterinária. Os resultados das análises comparativas e pangenômicas, revelaram a espécie com um genoma flexível, devido à presença de uma grande variedade de elementos genéticos móveis, tais como ICEs, GEIs, ISs e plasmídeos. A identificação de ilhas genômicas exclusivas e a presença de genes de resistência a antibiótico, indica a fixação destes no genoma, pelo aumento do *fitness* da bactéria em um nicho específico. Adicionalmente, análises epidemiológicas baseadas na tipificação de sequências multilocus, serotipagem, genotipagem e polimorfismo de nucleotídeo único permitiu explorar um pouco mais a relação de genótipos e grupos filogenéticos associados à capacidade de causar uma determinada doença e/ou predileção por um hospedeiro. Além disso, encontrou-se elementos genéticos móveis presentes em determinados grupos genéticos ou adquiridos por convergência evolutiva que atuariam na capacidade de causar uma doença ou infectar um hospedeiro. A análise de seleção de diversificação encontrou proteínas expostas na superfície sujeitas ás pressões de seleção relacionadas a sua capacidade de adaptação ao nicho. Portanto, é essencial explorar o conteúdo genético e os *Single Nucleotide Polymorphisms* (*SNPs*) desse patógeno que desempenham papéis cruciais na virulência bacteriana, na diversificação da população e na adaptação aos nichos hospedeiros. Esse artigo de revisão é o primeiro estudo descrevendo a relação genômica e epidemiológica para uma melhor compreensão da diversidade e evolução de *P. multocida*.

REVIEW ARTICLE

Insights of *Pasteurella multocida*: genomic diversity, epidemiology and evolutionary adaptation

Raquel Hurtado[1,2], Lenin Maturrano[2], Vasco Azevedo[1]† and Flávia Aburjaile[1,3]†*

[1] Laboratory of Cellular and Molecular Genetics, Federal University of Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brazil

[2] Laboratory of Molecular Biology and Genetics, Veterinary Medicine Faculty, San Marcos University, Lima, Peru

[3] Laboratory of Plant Genetics and Biotechnology, Federal University of Pernambuco, Recife, 50670-901, Pernambuco, Brazil

†Contributed equally

*Corresponding author: faburjaile@gmail.com

+55 (31) 98797-5210

Laboratory of Plant Genetics and Biotechnology

Federal University of Pernambuco, Recife, 50670-901, Pernambuco, Brazil

Abstract

Pasteurella multocida is an important veterinary pathogen that harbors the ability to infect a wide range of animals in a broad spectrum of diseases. *P. multocida* is a complex microorganism concerning genomic flexibility, host adaptation and pathogenesis. Epidemiological analysis based in multilocus sequence typing, serotyping, genotyping, association to virulence genes and single nucleotide polymorphisms (SNPs) allow to asses intraspecies diversity, phylogenetic relationships and strain-specific associated with host predilection or disease. A high number of sequenced genomes to draw us to a more accurate interpretation genomic and epidemiological to possibly conclude that certain lineages can having host predilection or produce disease. Comparative analysis and pan-genomic approaches have shown a flexible genome for hosting mobile genetic elements and therefore show a significant variation on the genes content. Moreover, it was possible to find lineage-specific the mobile genetic elements (MGEs) from the same niche, showing acquisition probable due to an evolutionary convergence event or by a genetic group with the capacity of infect or produce disease. Furthermore, diversification selection analysis exhibits proteins exposed on the surface subject to

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

selection pressures with an interstrain heterogeneity and it would be related to their ability to adapt. Therefore, it is essential to explore the genes content and SNPs of this pathogen that would play crucial roles in bacterial virulence, population diversification and niche adaptation. This article is the first study describing the genomic and epidemiology relationship for a better understanding of the diversity and evolution of *P. multocida.*

Highlights

- *P. multocida* can infect a wide range of hosts and cause different diseases.
- *P.* multocida presents genomic flexibility in host adaptation and pathogenesis.
- SNPs have roles in population diversification and niche adaptation.
- Integration of genomic and epidemiological data is essential to cognize evolution.

Keywords: genomics; epidemiology; host predilection; phenotypes associated strains; virulence.

1. Introduction

Pasteurella multocida was isolated for the first attempts by Pasteur in 1881, from a fowl cholera outbreak and to develop an attenuated bacterial used as a vaccine, which can stimulate an immune response in chickens (Pasteur, 1881, 1880). Belonging to the genus *Pasteurella*, are mostly commensal and opportunistic pathogens bacterial found in the oral cavity, nasopharyngeal, and upper respiratory tract of a wide range of domestic and wild animals throughout the world.

P. multocida is a Gram-negative rod-shaped, non-flagellated (4). *P. multocida* isolates are aerobic or facultative anaerobic and grow well at 37°C on 5% sheep's blood in different medium agar, however no growth on MacConkey agar. Although *P. multocida* grows on blood agar, it is not hemolytic. Most clinical isolates are catalase, oxidase, indole, and ornithine decarboxylase positive but Urease negative (Boerlin et al., 2000). Identification of *P. multocida* is performed using biochemical strips (such as API 20E/20NE) (Collins et al., 1981). This microorganism is the primary agent causing of fowl cholera, predominantly in poultry and a variety of avian species (Christensen et al., 2008; Shane and Stern, 2003), hemorrhagic septicemia in ungulates animals, especially in cattle and buffalo (Rimler and Wilson, 1994), atrophic rhinitis in pigs (Chanter and Rutter, 1990; Davies et al., 2003), rhinitis (snuffles) and pneumonia in rabbits (DiGiacomo et al., 1991; Tayeb et al., 2004). *P. multocida* as secondary agent is implicated in the pneumonic infection (Dabo et al., 2007; Hansen et al., 2010; Rosadio et al., 2011) in host as cattle (Dabo et al., 2007), sheep (Davies et al., 2003), pigs (Chanter and Rutter, 1990; Davies et al., 2003; Hansen et al., 2010), rabbit, goat (Wilkie et al., 2012) and alpaca (Hurtado et al., 2017; Wilkie et al., 2012). *P. multocida* is an zoonotic agent, being

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

68 the most predominant pathogen isolated in human infections, and their transmission occurs through
69 animal bites or contact with nasal secretions (Wilson and Ho, 2013). The infection associated with
70 *Pasteurella* are generically called "Pasteurellosis" that can be presented as an acute or chronic
71 infection and that can lead to significative mortality (Harper et al., 2006; Wilson and Ho, 2013). More
72 information about the pathogenesis by specific disease were described previously (Wilkie et al., 2012;
73 Wilson and Ho, 2013).
74 *P. multocida* was subdivided into three subspecies: *P. multocida* subspecies *multocida*, *P. multocida*
75 subspecies *septica*, and *P. multocida* subspecies *gallicida (Mutters et al., 1985)* classified by
76 phylogenetic relations (Kuhnert et al., 2000), phenotypes (Mutters et al., 1985) and niche adaptation
77 (Holst et al., 1992; Muhairwa et al., 2000). The *P. multocida* strains are classified based in the
78 serotyping, which recognize 5 types capsular polysaccharide antigens, distinguishing: A (Hyaluronic
79 acid), B (arabinose, mannose and galactose), D (heparin), E (uncharacterized), or F(chondroitin)
80 serogroups (CARTER, 1955). The strains also are classified on 16 serotypes based on the recognizing
81 of lipopolysaccharide (LPS) antigens using the Heddleston scheme (Heddleston et al., 1972). For a
82 simplified typing of LPS antigens and accuracy is replaced Heddleston serotyping scheme by a
83 multiplex PCR targeting the genes encoding the structure of the LPS outer core biosynthesis loci
84 (Harper et al., 2006) which, recognize and classify the eight distinct genetic *loci* in L1 through L8
85 genotype corresponding at Heddleston serovar: L1(1, 14); L2 (1,2, 5); L3 (3,4); L4 (6,7); L5(9);
86 L6(10,11,12,15); L7 (8,13) and L8 (16). In this review, we discuss phylogenomic and epidemiological
87 strategies to describe the relationship between some *Pasteurella multocida* strains and its capacity of
88 host predilection or cause a disease. We also summarize comprehensive genomic analyses, such as
89 pan-genomic, pathogenomics and evolutionary forces (acting on genes and SNPs) for a better
90 understanding of the diversity and evolutionary adaptation of *P. multocida* (Figure 1). A timeline of
91 key genomic and pathogenesis research of *P. multocida* is given in Figure 1 in the attempt to
92 understand the genetic diversity and evolution of specie (Figure 2).
93
94 **Figure 1.** Workflow comparative genomic studies of *P. multocida* isolates. *P. multocida* can infect a
95 wide range of host. The resource of genetic diversity is mediated mainly by point mutations and
96 acquisition of mobile elements as IS, transposons, GEIs and plasmids. An increase of genome
97 sequencing of diverse pathogenic phenotype isolates allows a fine knowledge of diversity and
98 evolutionary adaptation. Studies of comparative genomic are reported including phylogeny
99 construction, pangenome analysis, epidemiology and positive selection.
100 **Figure 2.** Timeline showing key genomic analyses and pathogenesis studies in the attempt to decipher
101 the genetic diversity and evolution of *P. multocida*.

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

2. Epidemiology: multilocus sequence typing, serotyping, genotyping and virulence-associated genes

P. multocida is genetically diverse according to numerous studies (Blackall et al., 1998; Davies et al., 2004, 2003; García-Alvarez et al., 2017; Hotchkiss et al., 2011; Marois et al., 2009; Shivachandra et al., 2005; Subaaharan et al., 2010). Numerous studies try to decipher the genetic diversity, pathogenic association, epidemiological origin and dissemination of *P. multocida* isolates. Main issues in discussing is to identify genetic differences among strains that could be associated with the capacity of hosts predilection and cause diverse diseases (Davies et al., 2004; Gunawardana et al., 2000; Shivachandra et al., 2005). The bacterial pathogenesis result of complex interactions between host factors (e.g., species, age, immune status) and the presence of bacterial virulence factors (e.g., LPS, capsule, adhesins, virulence genes) (Boyce et al., 2010). A variety of gene based strategies are applied to know the genetic diversity, genetic relationships and epidemiological depth among *P. multocida* strains (García-Alvarez et al., 2017; Massacci et al., 2018; Peng et al., 2018a), such as 16S rRNA, restriction endonuclease analysis (REA), ribotyping, random amplification of polymorphic DNA (RAPD)-PCR, pulsed-field gel electrophoresis (PFGE) and multilocus sequence type analysis (MLST) (Mühldorfer et al., 2011; Sthitmatee et al., 2010).

Being MLST analyses a strategies more popularity to explore global epidemiology, population structure and surveillance in long-term studies (Mühldorfer et al., 2011; Wang et al., 2013a). MLST scheme presents two databases operating on BIGSdb with separated housekeeping genes. RIRDC (Rural Industries Research and Development Corporation) MLST database (Subaaharan et al., 2010) is the largest concerning isolates compared to the multi-host MLST database for *P. multocida* (Davies et al., 2004). MLST analyses in bacteria as *Campylobacter* (Zautner et al., 2011) and *S. aureus* (Herron-Olson et al., 2007) has identified clonal complexes in correlation to host specialization (Smith et al., 2005; Smyth et al., 2009), these findings could form the basis for similar work in *P. multocida*. Among epidemiological marker genes that allow to show distribution and prevalence of virulence genotypes and correlated with the pathogenic mechanism of *P. multocida* isolates to cause a disease, are the capsule biosynthesis genes (cap A, B, D, E and F), genogroups of L1 through to L8 and 14 virulence-associated genes (Ewers et al., 2006) of which, *toxA*, *tbpA*, *hgbB*, and *pfhA* genes present a greater variability of frequency and association among hosts and disease (Bethe et al., 2009; Ewers et al., 2006; Katsuda et al., 2013; Li et al., 2018). Majority of epidemiological analyses are presented to continuation in attempt to find genotypes associated to pathogenic phenotype isolates.

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

2.1. Fowl cholera and epidemiology

Countable epidemiological analysis of *P. multocida* isolates from fowl cholera outbreaks in avian species has been reported in America, Australia, India and China (Amonsin et al., 2002; Blackall et al., 1998; Gunawardana et al., 2000; Kardos and Kiss, 2005; Li et al., 2018; Muhairwa et al., 2000; Shivachandra et al., 2005; Wang et al., 2013a). In 2010, Subaaharan developed the prediction of sequences types (STs) to species *P. multocida*, of an analysis of 63 isolates from Australian poultry in 8 fowl cholera outbreaks, it was reported 26 STs (Subaaharan et al., 2010). These results show great diversity and representativeness of STs in samples from Australian origin. In Chine, frequently occur outbreaks of fowl cholera, an outbreak was reported between 2011- 2012, the MLST analysis of 40 strains classified all strains with the sequence type, ST129 (Table 1) (Wang et al., 2013a). MLST analysis of 45 isolates of an outbreak of fowl cholera occurring during 2016-2017 in southwest China reported three different STs (ST129, ST8 and ST342), being ST129, the dominant sequence type (Table 1) (Li et al., 2018). To isolates from outbreaks of fowl cholera in Chine, it was identified ST129 as the genotype of major prevalence in avian *P. multocida* strain (Table 1) (Li et al., 2018; Peng et al., 2018a; Wang et al., 2013a), suggesting that this allelic variant have a predisposition to cause fowl cholera in avian, however, ST129 was identified in bovine (Hotchkiss et al., 2011), pig and another animals, being an ST that is shared across host species worldwide (Li et al., 2018; Peng et al., 2018a; Wang et al., 2013a). Another way could be explained by a limited number of isolates, limited geographic distribution or the same strain origin (Gunawardana et al., 2000; Hotchkiss et al., 2011). Likewise, it was also reported ST specific to host avian with fowl cholera, as a new ST342 (Hotchkiss et al., 2011; Li et al., 2018; Subaaharan et al., 2010).

Table 1. Sequence types (STs) of isolates of *P. multocida* obtained from all known studies.

Multihost (#/%)	RIRDC ST (#/%)	CC	Number of isolates	Host	Disease/Source	Country/Continent	Reference
	ST 1-28 (Except 3, 6); ST8 (12 isolates)		63	Poultry	Fowl cholera	Australia	(Subaaharan et al., 2010)
	ST13, ST79 and ST80 (91/128); ST122, ST51, ST81, ST86, ST123, ST125, ST137, ST9.	CC13 (82.03%)	128	Bovine	Respiratory isolates of healthy and clinically affected animals	UK, USA, Scotland and France	(Hotchkiss et al., 2011)
	ST122		12	Cattle, Buffalo, Elephant, Bison	Hemorrhagic septicemia	Asia	
	ST132, ST95, ST98, ST99, ST102, ST124		28	Ovine	Pneumonia	NZ, Spain	
	ST50, ST73, ST74, ST196		15	Porcine	Bronchopneumonia	UK	
	ST13		9	Avian		Asia/unknown	
	ST13 (21/29)	CC13	29	Porcine	Bronchopneumonia	Denmark	(Pors et al., 2011a)
	ST149 (1/29)	CC13					
	ST74 (3/29)	CC74					
	ST148 (1/29)	CC74					
	ST50	CC50					
	ST150	CC50					
	ST146	CC50					
	ST129		40	Poultry	Fowl cholera: 21 strains from the liver ,3 eggs, 2 oviduct,1 eye of a chicken. 8 strains of the liver, 2 heart of duck. 2 strains from liver goose, and one from liver pigeon	China, Jiangsu province	(Wang et al., 2013b)
	ST131, ST132, ST137, ST173, ST174, ST175		116	Ruminants	Pneumonia	Western part of Europe	(Bisgaard et al., 2013)
ST47	ST322		10	Pigs	Hemorrhagic septicemia	Spain	(Cardoso-Toset et al., 2013)
	ST122 (100%)		20	Buffalo, cattle	Hemorrhagic septicemia	Pakistan and Thailand, South Asia	(Moustafa et al., 2013)

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

	ST122		58	Ruminants	Hemorrhagic septicemia	Indonesia, ND, Sri Lanka	
	ST122 (50)		64	Bovine, pig and buffalo	Hemorrhagic septicemia	Vietnam, south and central Africa, India, Philippines, Malaysia, Iraq and Burma	(Petersen et al., 2014)
	ST63 (1)						
	ST147 (2)						
	ST162 (2)						
	ST151 (9)				Nasal swabs or tonsils of apparently healthy pigs		
ST10	ST74		38	Rabbit	Pneumonia (lungs), pyometra (Uterus), skin abscesses, mastitis (mammary gland), septicemia (liver or spleen)	Spain, Portugal	(García-Alvarez et al., 2015)
ST11	ST50						
ST12	ST9						
ST25	ST9						
ST50	ST203 (1) ST321 (1) and ST323 (1)		12	Ovine	Pneumonia	Spain	(García-Alvarez et al., 2017)
ST19	ST131 (1)						
ST19, ST52	STT320 (2)						
ST19, ST52, ST56	ST324 (4)						
ST20	ST95 (2)						
ST12, ST25	ST9 (4)		13	Pig	Pneumonia, septicemia and arthritis		
ST3	ST13 (2)						
ST11	ST50 (2)						
ST62	ST27 (2)						
ST10	ST74 (2)						
ST47	ST322 (1)						
	ST129(83.7%)	CC129	45	Poultry	Fowl cholera	China	(Li et al., 2018)
	ST8 (4.44%)	CC8					
	ST342 (15.55%)	CC129					

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

	23/52 (ST9, ST24, ST27, ST50, ST74, ST204, ST206); 29/52 (ST293, ST294, ST295, ST296, ST297, ST298, ST299, ST300, ST301, ST302, ST303, ST304, ST305, ST310, ST311, ST312, ST313)	CC ST9, CC ST50, CC ST74	52	Rabbit	Respiratory disease, metritis, mastitis, skin with abscesses and septicemia	Italian, Denmark and other countries	(Massacci et al., 2018)
	ST8 (2)	ST8	16	Avian	Fowl cholera		(Peng et al., 2018a)
	ST9 (1)	ST9					
	ST27 (1)	ST74					
	ST53 (1)	ND					
	ST60 (2)	ST158					
	ST129 (7) (43.75%)	ST129					
	ST156 (1)	ST156					
	ST159 (1)	ND					
ST12	ST9 (52.94%)		17	Rabbit		France	
ST11	ST50 (5.88%)						
	ST27, ST204, ST298						
	ST80(5%)	ST13	20	Bovine			
	ST122 (60.00%)	ST122					
	ST79 (30%)	ST13					
	ST65(5%)	ND					
	ST7, ST9, ST13(23.21%), ST27, ST50 (37.5%), ST74(21.43%), ST122, ST287		56	Porcine			
ST3 (7/40)	ST13		40	Porcine	Lung of respiratory disease	Chine	(Peng et al., 2018b)
ST10 (13/40)	ST74						
ST11 (18/40)	ST50						
ST16 (2/40)							

160

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

2.2. Hemorrhagic septicemia and epidemiology

Hemorrhagic septicemia (HS) is a disease of high economic importance, reported with a high prevalence and mortality in Asia and Africa regions (Benkirane and De Alwis, 2002; Hotchkiss et al., 2011; Moustafa et al., 2013). There are reports in Central and South America countries and Southern and Eastern European countries (De Alwis, 1999; Shivachandra et al., 2011). In two studies of 12 and 20 isolates of hemorrhagic septicemia cases, a unique genotype, ST122 was reported being the last case collected from different climate zones and outbreaks in different years from Pakistan and Thailand regions (Table 1) (Hotchkiss et al., 2011; Moustafa et al., 2013). Of an outbreak of hemorrhagic septicemia in pigs in Spain, 10 isolates belonged to ST322 (typing ST47 by multi-host MLST) (Table 1) (Cardoso-Toset et al., 2013) grouping in the cluster with other STs associated to HS in cattle indicating a close genetic relationship among genotype related to disease (García-Alvarez et al., 2017; Hotchkiss et al., 2011). MLST analysis of 55 isolates of *P. multocida* associated with hemorrhagic septicemia in mainly bovines and also in pigs and buffalo from south and central Africa and Asia, reporting the genotypes, ST122 (90.9%), ST63 (1.8%), ST147 (3.6%) and ST162 (1.8%), forming a clonal complex (Petersen et al., 2014). Of 20 isolates, 12 belonged to ST122 genotype (Peng et al., 2018a). Interestingly, ST122 is reported in an isolated of bovine, buffalo, and a pig with hemorrhagic septicemia with wide geographic distribution (Cardoso-Toset et al., 2013; García-Alvarez et al., 2017; Peng et al., 2018a; Petersen et al., 2014), suggesting that a clonal population of *P. multocida* is associated with HS but not with a host (Cardoso-Toset et al., 2013; García-Alvarez et al., 2017; Peng et al., 2018a; Petersen et al., 2014). Likewise, some sequence types were identified in another host as ST8 (Li et al., 2018; Peng et al., 2018a), ST129 in birds (Li et al., 2018; Wang et al., 2013a), also some genotype would be specific of septicemia hemorrhagic bovine as the ST322 (Hotchkiss et al., 2011).

2.3 Respiratory diseases and epidemiology

2.3.1. Rabbit pneumonia

In the MLST analysis, 15 isolates of *P. multocida* from spanish rabbit with pneumonia were reported in major frequency to ST50 (typing ST11 by multi-host MLST) and another STs as ST74 and ST9 (García-Alvarez et al., 2015). Of an epidemiologic study of 39 *P. multocida* isolates from Italian rabbit with a variety of clinical manifestation, 19 STs were reported mainly distributed into three clonal complexes (CC) that include the 82% of all strains, being three sequence types more prevalent ST9, ST50 and ST74 (Massacci et al., 2018). A summary of epidemiological study based on MLST, associated with 17 *P. multocida* isolates from France, show five genotypes (ST9, ST27, ST50, ST204, and ST298), being the most prevalent, ST9 (52.94%). Genotypes ST50, ST74, ST13

and ST9 have been identified different host and not exclusively in rabbit (García-Alvarez et al., 2017; Hotchkiss et al., 2011; Massacci et al., 2018; Peng et al., 2018a; Pors et al., 2011a). Interestingly, the predominant MLST genotype isolates from pneumonia rabbit is ST9. However, the analyzes were poorly represented by the number of samples (García-Alvarez et al., 2015; Massacci et al., 2018). Likewise, they have identified unique STs in rabbit as ST37, ST204, ST298 among others (García-Alvarez et al., 2015; Massacci et al., 2018).

2.3.2. Bovine pneumonia

Bovine respiratory disease (BRD) in cattle include pneumonia and shipping fever diseases. In a study of respiratory cases in cattle with 128 isolates from the UK, USA, Scotland and France collected between 1984 -2008 were reported 91 isolates belonging to three STs (ST13, ST79 and ST80), 105 bovine respiratory isolates belonging to CC13, which include cases from a wide range of countries, years and presentations (Table 1) (Hotchkiss et al., 2011) that corresponds with last analyses (Peng et al., 2018a). The CC13 is associated strongly with bovine respiratory isolates, but not exclusively, due to another host belonging to this CC, according to the database, considering then a virulent clone with some degree of specificity for respiratory disease (Bisgaard et al., 2013; Hotchkiss et al., 2011; Peng et al., 2018b; Pors et al., 2011a). However, there is some unique genotype ST; find in a specific disease (Hotchkiss et al., 2011) for example, ST74 in pneumonia bovine, pig and rabbit (García-Alvarez et al., 2017; Hotchkiss et al., 2011; Massacci et al., 2018; Peng et al., 2018a; Pors et al., 2011b). However, homogeneity of STs could be the result of getting samples of a limited geographical distribution (Hotchkiss et al., 2011; Peng et al., 2018a).

2.3.3. Ovine Pneumonia

An MLST analysis of 43 *P. multocida* isolates mostly from pneumonia ovine cases in Spain, reported ST95, ST131, ST203, ST320, ST324, ST321, and ST323 genotypes (Table 1) (García-Alvarez et al., 2017) and 28 isolates pneumonia ovine from Spain and New Zeland reported ST132, ST95, ST98, ST99, ST102 and ST124 genotypes (Table 1) (Hotchkiss et al., 2011), all STs with exceptions ST132 have been exclusively detected from ovine pneumonia cases (Bisgaard et al., 2013; García-Alvarez et al., 2017; Hotchkiss et al., 2011). Also, they identified six STs (ST131, ST132, ST137, ST173, ST174 and ST175) isolated from pneumonia in ruminants on western of Europe (Bisgaard et al., 2013), one of this, ST131 was reported on ovine (García-Alvarez et al., 2017). The detection of these STs almost exclusively from sheep and goat might indicate the existence of genetic groups of *P. multocida* adapted to this host, or at least to small ruminant hosts and with European origin (Hotchkiss et al., 2011).

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

2.3.4. Pigs pneumonia

The 29 isolates with bronchopneumonia in Danish pig herds, as the major representants of genetic diversity, were submitted to MLST analysis. The results report three clonal complex, CC13, CC74 and CC50 (Table 1) (Pors et al., 2011a), being these three CC associated to pneumonia cases (García-Alvarez et al., 2015; Hotchkiss et al., 2011; Pors et al., 2011b). MLST analysis of 48 isolates of pig with pneumonia, septicemia and arthritis from Spain reported ST9, ST13, ST27, ST50, ST74 and ST322 genotypes (Table 1) (García-Alvarez et al., 2017). All STs identified among the porcine isolates, except ST322, had been identified previously in *P. multocida* isolates from pig, avian, bovine, rabbit, lion, goat, and human, associated to different clinical backgrounds and from diverse geographical areas (Bethe et al., 2009; Blackall et al., 2000; Davies et al., 2003; Fussing et al., 1999; Pors et al., 2011b). MLST analysis of 40 *P. multocida* isolates in China randomly selected, reported three genotypes, ST13, ST74, and ST50 (typing ST3, ST10 and ST11 by multi-host MLST) (Table 1) (Peng et al., 2018b). To 56 pig isolates were identified prevalently the ST50 (37.50%), ST13 (23.21%), and ST74 (19.64%) genotypes (Peng et al., 2018a). The number of cases of pneumonia or respiratory associated disease concludes the presence in major frequency of determinate complex clonal as CC13, CC50 and CC74, CC13 in rabbit (García-Alvarez et al., 2015; Massacci et al., 2018), cattle (Hotchkiss et al., 2011; Peng et al., 2018a) and pig (Hotchkiss et al., 2011; Peng et al., 2018b). CC50 and CC74 were presents in rabbits (García-Alvarez et al., 2015; Massacci et al., 2018) and pigs (Hotchkiss et al., 2011).

These findings suggest that several evolutionary lines belong to same group of STs or be independent and may adapt to specific niches (phenotype disease/host), and it supports the idea that exists some genetic groups that may be truly pathogenic (virulent clone) rather than opportunistic (Bisgaard et al., 2013; Davies et al., 2004, 2003; Hotchkiss et al., 2011; Pors et al., 2011b). This proposes that some lineages have increased the capacity to acquired virulence factor or present compensatory mutations which could lead to the emergence of dominant lineages (Casali et al., 2014; Davies et al., 2003; Enright and Spratt, 1998; Ford et al., 2013; King et al., 2002). For example the ST131 S. Typhimurium, an emergence lineage was associated to the presence of multiple drug resistance and evidence a partial genome degradation, their presence could support some degree of human adaptation, however other lineages also have the ability to cause invasive disease indicating that others host factors are drives as susceptibility and not by bacterial genotype (Klemm and Dougan, 2016).

2.4. Virulence genes and epidemiology

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

Epidemiological analysis based on the profile and distribution of virulence factor allows correlating the presence of them to the pathogenic potential of *P. multocida* lineages (Christensen and Bisgaard, 2010; Dabo et al., 2007; Davies et al., 2004; Ewers et al., 2006; Hata et al., 2010; Sheppard et al., 2010). This strategy allows explaining the association of virulence genes in isolates with the capacity to cause the same clinical case or having host predilection. Analysis of the diversity of virulence-associated genes proven of isolates from various clinically healthy and diseased hosts record gene *tbpA* (31.5%) that encode a transferrin binding protein firstly detected in ruminant strains (Ewers et al., 2006) but lately, reported the *tbpA* (82%) gene in poultry, swine, and rabbit isolates (García-Alvarez et al., 2015; Sarangi et al., 2014). A gene *pfhA* (37.0%) that encode a filamentous hemagglutinin was positively associated to diseased bovine (Ewers et al., 2006), but lately, *pfhA* (95.3%) gene was identified in high prevalence irrespectively in host species (Sarangi et al., 2014). Additionally, the *pfhA, nanB,* and *nanH* genes were regularly distributed among all the serogroups (Sarangi et al., 2014). In relation to respiratory diseases cases, in a study of isolates of *P. multocida* from clinically healthy and diseases calves reported four virulence-associated genes (*pfhA, tadD, tbpA* and *HAS*) associated with the serogroup A, in agreement with previous studies, *pfhA, tbpA* and *hgbB* genes were associated with respiratory bovine diseases (Ewers et al., 2006; Katsuda et al., 2013). In a study of isolates from the lungs of swine, it was identified the gene *HAS* (63.5%). Also, *pfhA, tadD, HAS* and *hgbA* genes were more prevalent in serogroup A than serogroup D, while *hsf-1, hgbB* and *plpB* were more prevalent in serogroup D (Peng et al., 2018a)(Peng et al., 2018a).

In a study of *P. multocida* isolates from pneumonia rabbit cases, it was identified serotype D associated to *hgbB* (88.9%) and serotype A and *hgbB* + *pfhA* negatively associated with pneumonia (García-Alvarez et al., 2015) in accordance with isolates from rabbit affected by various diseases, it was reported *pfhA* (33.33%), *hgbB* (53.84%) and *pfhA* +*hgbB* (10.25%) genes (Massacci et al., 2018). The *pfhA, tadD, HAS, hgbA* and *nanH* genes were more prevalent in LPS genotype L3 than genotype L6, while and *nanB* were more prevalent in LPS genotype L6 (Peng et al., 2018b). The gene *toxA* encode a dermonecrotoxin, present in 12.5% of isolates proven all host and 66.7% from swine, similarly was detected in 44.1% of swine isolates (Peng et al., 2018b; Sarangi et al., 2014). *ToxA* was initially detected in serogroup D associated with atrophic rhinitis in pigs (Davies et al., 2003; Gardner et al., 1994). Likewise, *toxA* (2.6%) gene have shown a low prevalence in isolates from pneumonia swine (Djordjevic et al., 1998a; Peng et al., 2018a) and absent in pneumonia rabbit, serogroup D (García-Alvarez et al., 2015). To isolates of *P. multocida* from fowl cholera cases, they were found the *ompH, oma87, sodC, hgbA, hgbB, exBD-tonB* and *nanB* genes in 100% of the isolates, and it was identified *sodA* and *nanH* (96%), *ptfA* (92%), and *pfhA* (60%) genes. Gene *toxA* was not identify

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

(Furian et al., 2013). Analysis based on the profile and distribution of virulence-associated genes show their association to pathogenicity but not distinguish clearly among them.

2.5. Serotyping and epidemiology

The *P. multocida* strains are classified based in the serotyping and genotyping, in 5 serogroups, 16 serovars and 8 LPS genogroups. Fowl cholera cases have been frequently designated in serogroups A , B, D and F (Davies et al., 2003; Wilson and Ho, 2013); and serovars 1, 3 and 4 (Table 2) (Rhoades and Rimler, 1990; Wilson and Ho, 2013) that correspond to LPS genogroups L1, L3 (Li et al., 2018) and lately reported L2, L5, and L6 (Table 2) (Peng et al., 2018a). Hemorrhagic septicemia is caused by *P. multocida* mainly in bovine and buffalo species caused by specific serotypes B:2 and E:2 and in lower frequency in various other serotypes (De Alwis, 1999), and bovine respiratory diseases have been reported in the highest frequency to serotype A:3 (Table 2) (Dabo et al., 2007, 1997; Wilson and Ho, 2013). *P. multocida* isolates from bovine with diverse diseases belong in major frequency to serogroups A and B; and LPS genogroups L1, L2, L3 (Table 2) (Peng et al., 2018a). Isolates from rabbit have been classified in serotypes A, D and F (Chengappa et al., 1982; Dabo et al., 1999; Jaglic et al., 2004), and in the serovars 1, 3, 4, 11, 12 and 15 (Brogden, 1980; Chengappa et al., 1982; DiGiacomo et al., 1993).

Last studies of *P. multocida* isolates from Pasteurellosis cases in rabbits belonged to serogroups A, D and F and corresponded to serovars 3 and 4 or L3 and serovars 6 or L6 (Table 2) (García-Alvarez et al., 2015; Massacci et al., 2018; Peng et al., 2018a; Townsend et al., 2001). *P. multocida* strains isolated from respiratory diseases in pig were frequently designated to serogroups A and D in China (Djordjevic et al., 1998b; Liu et al., 2017; Pijoan et al., 1983; Tang et al., 2009), and in low frequency to serogroups B and F (Table 2) and in the serovars 3, 4, and 12 (Jamaludin et al., 2005; Lainson et al., 2002; Pijoan et al., 1983), in addition, this was classified in three LPS genogroups, in high-frequency L3 and L6 and low frequency L2 (Table 1) (Peng et al., 2018b). To porcine atrophic rhinitis cases often belong in the highest frequency to serogroups D and A (Davies et al., 2003; Gardner et al., 1994; Sakano et al., 1992). Analysis based on the profile and distribution of serotypes, serogroups and, LPS genogroups shows the distribution of specific genotypes associated with phenotypes corroborated with the last study at the genomic level (Peng et al., 2018a).

Preliminary genetic diversity and epidemiological analyses have identified that determined lineages could have the predisposition to infect a host or cause a disease. However MLST approach does not differentiate among members of the same clone, by consequence, it presents a limited discriminations power (Blackall et al., 1998). Another strategies have identified the profile, distribution and correlation of virulence genes and serotyping with the pathogenic potential of *P.*

330 *multocida* lineages (Davies et al., 2004; Ewers et al., 2006; Shirzad Aski and Tabatabaei, 2016; Tang
331 et al., 2009). Although the reports of virulence-associated genes, capsule and lipopolysaccharide are
332 correlated to pathogenic phenotypes, they could not discriminate between the diversity of phenotypes
333 strains. That is how the massive whole-genome sequencing (WGS) data analysis could allow an
334 approach more representative of diversity on genes content and nucleotide sequence variability. An
335 approach more representative of nucleotide diversity is the single nucleotide polymorphism analysis
336 (SNPs) at a whole-genome level (Achtman, 2008; Pearson et al., 2009).
337
338 **Table 2.** Serotyping LPS genogroups of isolates of *P. multocida* obtained from all known studies.

Diseases/Hosts	Serotypes (%/N°)	Serovars	LPS genogroups (%)	Reference
Fowl cholera	A, F and B, D			(Davies et al., 2003)
		1, 3 and 4	L1(1) and L3 (3,4)	(Hirsh et al., 1985)
	A (85.7%) n=20			(Ewers et al., 2006)
	A (93.75%) n=16		(L1, L2, L3, L5, and L6), L1 (56.25%), L3 (25.0%)	(Peng et al., 2018a)
	A (100%) n= 45		L1 (95.6%) and L3 (4.4%)	(Li et al., 2018)
Hemorrhagic septicemia/bovine and buffalo	B and E	2	L2 (2)	(De Alwis, 1999)
Respiratory diseases and health / bovine	A	3	L3(3)	(Dabo et al., 2007, 1997)
	A (93.7%) and D (6.3%) n=338			(Katsuda et al., 2013)
	type A (92.3%) n=91			(Ewers et al., 2006)
Variety disease bovine	A (40.00%) and B (60.00%) n=20		(L1, L2, L3), L2 (60.00%) and L3 (35.00%) B:L2 (60.00%) and A: L3 (35.00%)	(Peng et al., 2018a)
Rabbit	A, D and F			(Chengappa et al., 1982; Dabo et al., 1999; Jaglic et al., 2004)
		1, 3, 4, 11, 12 and 15		(Brogden, 1980; Chengappa et al., 1982)
	A (93.75%) and D n=17		L3 (76.47%) and L6 (23.53)	(Peng et al., 2018a)
	A: (76.47%)		L3 (76.47%)	(Townsend et al., 2001)
	A:3 or A:4			(Harper et al., 2013)
	A (47.0%), D (28.0%), and F (25.0%) n=100			(García-Alvarez et al., 2015)
	A (51.28%), D (23.07%) and F (25.64%)	3 (56.41%) or 6 (43.59%)		(Massacci et al., 2018)
Respiratory diseases and health /pig	A and D			(Djordjevic et al., 1998a; Pijoan et al., 1983)
		3, 4, and 12		(Jamaludin et al., 2005; Lainson et al., 2002; Pijoan et al., 1983)
	A (39.5%) and D (54.9%) and untypeable (5.2%) n=233			(Tang et al., 2009)
	A (49.3%) and D (47.6%) n=296			(Liu et al., 2017)
	A 100% n=111			(Pors et al., 2011b)
	A (48.21), B, D (42.865) and F, n=56		(L2, L3, L6), among them, L6 (66.07%) and L3 (32.14%), concluding A: L3 (76.47%) n=56	(Peng et al., 2018a)
	A (49.6%), D (46.1%), and untypeable (4.3%); n=115		L3 (22.6%) and L6 (77.4%) n=115	(Peng et al., 2018b)
Atrophic rhinitis	A, D			(Davies et al., 2003; Gardner et al., 1994; Sakano et al., 1992)

3. Phylogenomics

Phylogenetic analysis seeks to infer the evolutionary history and relationship among organisms from nucleic acids or amino acid sequences, genes or genomes. The increasing availability of whole

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

genome sequences allows an appropriate genome-scale phylogenetic inference of organisms called phylogenomic (Eisen and Fraser, 2003; Maiden et al., 1998; O'Brien and Stanyon, 1999; Sicheritz-Ponten, 2001). Phylogenomic methods are based on different approaches for the construction of phylogenies using a whole phylogenetic signature, such as sequence variability, gene content (Gu and Zhang, 2004; Huson and Steel, 2004; Snel et al., 1999), concatenated sequences (Brown et al., 2001; Rokas et al., 2003) and consensus and supertree approach with the integration of multiple gene tree. Phylogenomics constructions are the basis of important studies as pathogen transmission mapping, geographical distribution, identification of emerging strains and as well as gene-to-phenotype association studies (Klemm and Dougan, 2016). Construction of the phylogenetic trees SNP-based variant calling is currently the most widely used technique. SNP-based methods have the power to distinguish between closely related genomes that differ by only a few SNPs (Klemm and Dougan, 2016). The common prediction of phylogenies encompasses the vertically transmitted phylogenetic signal. However, one of the challenges could represent infer the horizontal evolutive way through of a phylogenetic representation (Comas et al., 2007; Jain et al., 1999; Kurland et al., 2003). Phylogenomic construction based on content genes will be the best way to infer the horizontal evolution. To the date, *P. multocida* presents an increase of available sequencing genomes from isolates sampled across age lineages, regions and diseases (Bentley and Parkhill, 2015) however still is insufficient the representativeness of samples. (Bentley and Parkhill, 2015). The continue increases of sequenced genomes of *P. multocida* would allow a possible approximation to an evolutionary reality with the analysis of the largest number of polymorphisms.

3.1. Phylogeny in *P. multocida*

First pan-genomic analysis of the nine *P. multocida* strains (complete genomes Pm70, 36950 and incomplete genomes X73, VP161, Anand1P, Anand1G, P903, M1404 and P3480) based on 7,931 SNPs at shared positions shows a little or no correlation between the phylogenetic relatedness of the strains and the country of isolation, the serogroup, serovar, the disease or the host predilection (Boyce et al., 2012). Phylogenetic analysis based on similarity of common genes among six complete *P. multocida* strains (3480, HN06, p1059, Pm70, 36950 and X73) shows only subtle nucleotide differences among them (Wilson and Ho, 2013). Phylogenetic analysis using 7,892 CG-SNPs show the clustering of 12 *P. multocida* HS strains, clearly separated of non-HS strains (Pm70, X73, VP161, P1059 and Anand1P) (Moustafa et al., 2015). This finding could show a correlation between lineage and HS disease. However, there was no clear correlation with another type of diseases or hosts. Additionally, phylogenetic relatedness of only the HS group using 722 CG-SNPs, show separate

clustering among Asiatic strains (Thai and Pakistani strains) and M1404 (the North American isolate) revealing a clear genetic relationship by geographic local (Moustafa et al., 2015).

Later, a core genome SNP phylogeny of 33 *P. multocida* genomes from the diverse host, present four major phylogenetic groups (Cao et al., 2017). Group IV includes all serotype-B strains that includes HS bovine strains from different countries, except the swine host. Group II shows a close relationship between isolates of poultry and rabbit host, probably derived from a recent common ancestor. Phylogenetic group I, contain ATCC 43137 strain isolated of pig and 36950, P1062 (isolates of pneumonia in cattle) and HB03 isolated of respiratory diseases in swine (Cao et al., 2017). Concluding that with the highest number of sequenced genomes, it is possible to find genetic variability but not relate phylogenetic groups to pathogenic phenotypes among the *P. multocida* isolates (Cao et al., 2017), considering a limited available number of genomes and the lack of information on the origin of samples. Phylogenetic-based analysis on SNPs across the WGSs and the SNPs within all single-copy within 1806 genes among 109 *P. multocida* genomes show a corresponding phylogenetic relationship with the combining of LPS: MLST genotype (Peng et al., 2018a). The distribution of LPS: MLST genotypes in the tree allow clustering some isolates cases with a same MLST genotype and host but was not analyzed by disease association. A last study among 114 isolates of *P. multocida* from different host species show the phylogenetic relationship and classification of MLST genotype of these strains intimately related to the host and geographical location (Zhu, 2019). The MLST and SNP analysis found some important genetic groups likely relate to strain-specific phenotypes (Christensen and Bisgaard, 2010; Dabo et al., 2007; Davies et al., 2004; Ewers et al., 2006; Hata et al., 2010; Sheppard et al., 2010) .

A similar result of phylogenetic-based analysis on 1688 core proteins of 23 *P. multocida* isolates show the clustering of Pneumonia group isolates from various countries except for 3480 and HN07 strains, and a group from snuffles of rabbit strains. Remains the avian strains clustering in two groups, except strain Pm70 (Hurtado et al., 2018). What leads one to think that specific genetic groups may be related to the capacity of cause a disease or host predilection. Phylogenomic analysis based on the presence/absence genes, could represent the better way of the horizontal evolution on bacteria. The clustering of mobile genetic element included on accessory genome could explain the association to pathogenic phenotype-specific strain (Comas et al., 2007; Hurtado et al., 2018; Medini et al., 2005; Vernikos et al., 2015).

The phylogeny based on the accessory genome shows a major diversity at genes level on *P. multocida* strains but no stronger correlation between gene content, and host predilection or disease manifestation (Hurtado et al., 2018). This likely reflects that the differences associated to pathogenic phenotypes isolates could be defined for acquisition or loss genes events that could explain by

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

evolutionary convergence (Hurtado et al., 2018) or by a genetic group with the capacity of cause disease or having host predilection. It will be of interest to discover whether this trend continues with more *P. multocida* sequenced genomes.

4. Comparative genomics

Since the first complete genome sequencing of *Haemophilus influenza*, nowadays exist a massive number of genome sequenced product of high-throughput next-generation sequencing technologies. The existence of a large amount of genome from the same species or lineages has allowed a study of diversity and evolution at a high scale and level of resolution (Loman and Pallen, 2015).

Comparative genomic strategies allow to identifying similarities and differences at gene and sequence variation among isolates enabling their clustering them. When these differences are linking to information of isolates origin as time, space and clinical data, it can bring an understanding of species evolution, geographical spread and disease associations (Bentley and Parkhill, 2015). (Bentley and Parkhill, 2015; Mira et al., 2010; Vernikos et al., 2015).

Increasing the number of the sequences of representative lineages of the same species appears the enormous variability of gene content at the intraspecific level on bacterial (McInerney et al., 2017). Being that the concept of pangenome refered at all set of gene families that are found in the multiple genomes from specie bacterial, and subdivided in core genome, genes involved in essential processes and a group of accessory genes (indispensable and unique genes) involved in adaptive process. The first study in this field revealed the detection of the difference in gene content in 8 strains of *Streptococcus agalactiae* (Tettelin et al., 2005). Some prokaryotic species have extensive (or open) pangenomes while others have genomes that manifest very few gene content differences (closed pangenomes) (McInerney et al., 2017).

The dominant source of gene-containing variability is the cause of acquisition and loss of genes by horizontal transfer mechanisms, allied to differential gene losses, with gene duplications also playing a role, albeit a lesser one (Treangen and Rocha, 2011). Accessory genes associated with a specific phenotype could usually have functions that confer a selective advantage (Frost et al., 2005). As a consequence of the advantageous acquisition of these genes, the organism can migrate to new niches and maintain itself through natural selection, increasing its fitness (McInerney et al., 2017).

4.1. Comparative genomics in *P. multocida*

In the attempt to decipher the diversity and evolutionary adaptation of *P. multocida* strains, phylogenetic trees based on MLST, 16S RNA gene, core and whole genomes were constructed. Complementary, comprehensive genome-wide comparisons analysis based on strategies of pan-

445 genomics, pathogenomics and evolutionary selection forces prediction of genes and SNPs associated
446 with *P. multocida* strain-specific phenotypes, facilitate a better understanding of diversity and
447 evolutionary adaption. The genomic history in *P. multocida* begin with the first complete genome
448 sequencing, Pm70 strain, belonging to a representative clonal group from a fowl cholera case reported
449 in 2001, then provide the first highlights of genomic information with identification of a variety of
450 virulence factors, such as two genes encoding filamentous hemagglutinin and 50 genes with a
451 potential role in iron acquisition and metabolism (May et al., 2001).

452 Nowadays, the sequencing of new genomes from diverse phenotypes presents 189 all sequenced
453 genomes available of *P. multocida*, of which 51 are complete and 138 are incomplete in the NCBI
454 database. All available *P. multocida* genomes are between 2.30Mb in length with approximately G+C
455 content 40.40% (Table 1). Comparative analysis between complete genomes (Pm70 and 36950
456 strains) and incomplete genome (X73, VP161, Anand1P, Anand1G, P903, M1404 and P3480 strains)
457 allowed the first pan-genomic analysis resulting with a core of 1110 genes, but using only complete
458 genomes it resulted in 1,780 core genes and a pan-genome with more than 2,800 genes (Boyce et al.,
459 2012). The 36950 strain presents a unique large integrative conjugative element, ICEPmu1 (Michael
460 et al., 2012) of 82kb containing 88 genes of which 12 genes encoding resistance to antibiotics (Boyce
461 et al., 2012). To M1404 strain, presents 60 unique phage-related genes. The presence of this mobile
462 genetic element, as ICEPmu1 from another pathogen bacterial (*Histophilus somni* 2336 and
463 *Mannheimia haemolytica* PHL213) which share the same host niche with *P. multocida*. Another
464 unique region present in Pm70 strain is a homolog cluster present in another member of
465 Pasteurellaceae family (*H. influenzae* R2846, *H. somni* 2336, and *Gallibacterium anatis* UMN179)
466 (Wilson and Ho, 2013). The result of genetic exchange will be facilitated share the same host niche
467 and present homolog regions for recombination by to be members of the Pasteurellaceae family.

468 The comparison between HN06 and 3480 strains, HN06 presents a unique region of 18kbp carrying
469 14 genes that include a *toxA* gene and several phage-related genes. Two shared regions of 4.8-kbp
470 and 16kbp between 3480 and HN06. To 3480 strain, a unique region of 37-kbp and HN06 a unique
471 region of 33kb no found in any of the other strains. Pm70 strains compared with 36950 strain, present
472 a region with a site-specific recombinase, a restriction/modification system and two unique phage-
473 like genes (Wilson and Ho, 2013). First pathogenomics analysis mediated by the comparison between
474 virulent avian *P. multocida* genomes (P1059 and/or X73) versus an avirulent strain Pm70, allow the
475 identification of 336 genes which include 61 genes with unknown function, the differential presence
476 of this genes could be associated with their capacity of virulence (Johnson et al., 2013). There was
477 identified a region shared among P1059 and F218 strains of 8.5 kb containing six genes involved in
478 the transport and modification of citrate and a region shared among X73 and F216 strains of 12.78kb

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

containing genes encode galactitol-specific phosphotransferases, reported uniquely in isolates of fowl cholera, but absents in 21 isolates from other different hosts and diseases (Hurtado et al., 2018). A region shared among P1059 and/or X73 and absent in Pm70 strains contained ten CDSs involved in the transport and utilization of L-fucose. Overall, the presence of unique genes/systems related to metabolism and adhesion could provide to avian strains additional resources for increased fitness leading to higher virulence (Johnson et al., 2013).

Additionally, comparison with a new avian virulent GX-Pm strain compared with P1059, X73, Pm70 strains, show the presence of unique GEIs containing drug resistance gene such as 2 streptomycin resistance genes, 1 chloramphenicol resistance and 1 sulfonamide resistance genes, and also several function-unknown proteins in prophage regions (Yu et al., 2016). Furthermore, 22 genes were predicted shared between all virulence strains (P1059, X73, GX-Pm) which include a periplasmic lysozyme inhibitor and flavohemoglobin synthase. Also, 475 genes were included in at least some of virulent strains, among these were predicted Type I restriction-modification System, an antitoxin protein HigA, SprT and Opacity-associated protein (Opa) that contribute on the process of adhesion to the host cells (Yu et al., 2016). The potential of virulence-associated genes shared by virulent strains was corroborated with new analysis among virulent strains P1059, X73 or HB02 versus avirulent Pm70 (Peng et al., 2017). The comparison among them shows the variation of distribution of known virulence genes, like *hgbA, hgbB, ptfA, tadD, flp1, pfhB1, pfhB2, nanB* and *nanH*. This analysis also allowed to characterize virulence-related genes present in the subspecies septica (HB02 strain) (Peng et al., 2017). Another diversity variant is the presence of pseudogenes and the duplication of genes (paralogous presence). Of the comparison between 36950 and Pm70 strains shows the presence of 30 non-functional genes (pseudogenes) contained in 36950. An example is the presence of two pseudogenes Afu2 and Afu3 that conventionally encode iron-binding proteins. Another particularity is the presence of putative functional paralogues from pseudogenes as example *PfhB*1 and *PfhB2, Hsf1* and Hsf2 (Boyce et al., 2012).

Another characteristic is the variation of distribution of the paralogous genes among the strains (Pm70, P1059, X73, 36950, HN06 3480). OmpH2 is more highly conserved than the *ompH1* gene. *OmpH3* gene is present in all sequenced strains except for X73 strain. The *pfhB1* gene is present in strains Pm70, P1059, X73, and 3480 strains. *PfhB2* gene is present in strains Pm70, P1059, X73, and 36950. The *pfhB3* and *pfhB4* genes are absent in strain Pm70. A *pfhB3* gene is present in strains P1059, X73, and 36950 strains. A *pfhB4* gene present in strains P1059, HN06, and 3480 (Johnson et al., 2013). The occurrence of pseudogenes and paralogous genes could be due to the likely significant redundancy of function and the loss of these functional genes that will have a little impact on the phenotype of the strains. It is too probable that exist an incomplete distribute number of genes as a

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

consequence of genetic variation rather than the complete absence of these genes (Garcia et al., 2011; Johnson et al., 2013).

Comparative genomics analysis among high virulent strains (PmCQ2, Pm36950, and PmHN06) with low virulent strain (PmCQ6) and avirulent strains (Pm3480 and Pm70) show the presence of a Transposase IS200 which contains three proteins ModA, ModB, and ModC in only the three high virulent strains. The virulent PmCQ2 strain shows the presence of unique elements as IS605, ModB, nonspecific tight adherence protein D and include recombinase, phage-related genes, phage N-6-adenine-methyltransferase, phage terminase, and prophage integrase. The ModB encodes a class of zinc finger-containing transcription factor with a broad range of functions (Du et al., 2016). Virulence genes and insertion sequences present in high virulent strains might be responsible for the virulence differences among high virulent, low virulent and avirulent strains. Also, these results stand out the virulence factors among strains from bovine host (Du et al., 2016).

The first genomic analysis of all HS-associated strains (M1404, Pakistani and Thai) compared to strains not associated with HS (Pm70, 3480, 36950 and HN06) show two unique intact prophages of approximately 34 kb and 15kb in length present on all HS strains (Moustafa et al., 2015). Overall, the HS strains share 96 genes dispersed around the genome and that are absent from the other genomes analyzed. Also, the twelve Asian HS strains share an incomplete phage of approximately 44 kb region with 59 genes absent from the American HS strain, M1404 (Moustafa et al., 2015). This region includes 44 hypothetical proteins and includes an IgA FC receptor with antiphagocytic function and a protein that encodes cell filamentation protein Fic, component of the toxin-antitoxin module; these results were conferred with an additional Razi_0001 HS strain and compared with 19 isolates from others phenotypes (Hurtado et al., 2018).

The seven Pakistani strains share a questionable phage of approximately 50 kb region containing 39 genes that include phage elements as well as hypothetical proteins but it is not present in other genomes including 9 Thailand strains (Moustafa et al., 2015). Additionally, the strains 36950, TX1 and BUKK share 35 unique genes that include a range of antimicrobial resistance genes, designated as a new integrative conjugative element, ICEPmu2 (Moustafa et al., 2015). Additionally, accessory genome analysis between Razi_0001, V1, Lslm, THA, and PMTB2.1 identified 127 CDSs exclusively shared by the HS strains that include 110 hypothetical proteins and virulence-related genes such as TonB-dependent heme receptor A, XRE family transcriptional regulator and ArsR family transcriptional regulator (Hurtado et al., 2018).

News pan-genomic analysis on 33 *P. multocida* strains reveal the composition of a pangenome with 4036 genes, 1602 core genes (39%), 1,364 dispensable (33%) genes, and 1,070 strain-specific genes (27%) (Cao, 2017). Likewise, pan-genomic analysis with 23 strains show similar proportions

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

with a pan-genome of 3,585 genes, 1,688 (47.08%) core genes, 1,200 (33.47%) were part of the dispensable genome, and 697 (19.44%) were unique genes. The result shows 52.91% of variable genome content, additionally suggesting that *P. multocida* species present an open pan-genome. Also, it shows that content of accessory genes is related to carbohydrate transport and metabolism (9.54%), cell wall/membrane/envelope biogenesis (5.55%), transcription (4.85%), replication, recombination and repair (3.08%), inorganic ion transport and metabolism (4.6%), and unknown function (49%) (Hurtado et al., 2018). Last studies of total repertoire genes on 109 *P. multocida* strains show a composition of 1,806 core genes, 1,841 dispensable genes, and 609 strain-specific genes. Likewise, complementary to the studies of the association of accessory genome no found presence of specific genes to a predilection to host (Peng et al., 2018a), On the contrary, an association of accessory genome to specific diseases would be expected (Hurtado et al., 2018; Johnson et al., 2013; Moustafa et al., 2015). All comparative genomic analysis shows the presence of prophages, ICEs, insertions sequences with the content of virulence genes from horizontal transfer events, which allow present to *P. multocida* with a dynamic genome (Boyce et al., 2012; Hurtado et al., 2018). The important intraspecific diversity is conferred with pan-genomic analyses based on the presence and absence genes from 179 genomes of *P. multocida* (as shown in Figure 3).

4.1.1 A comparative genomics approach to identify genes under diversification and positive selection pressure

Genomic comparison analysis between five complete genomes of *P. multocida* (Pm70, HB03, HN06, 3480, and 36950 strains) show a total of 500 putative mutant protein, which are categorized into 10 groups as outer membrane proteins, capsule and fimbria, and involved in processes such as carbohydrate, energy, nucleic acid and amino acid metabolisms, in addition to transport, and drug resistance (Okay, 2015). These proteins were described in previous studies highlighting their functional importance in response to the host environment, adaptive responses and stress response, host-pathogen interaction and pathogenicity (Boyce et al., 2010; Guo et al., 2012; Hatfaludi et al., 2010; Michael et al., 2012; Paustian et al., 2002). Likewise, various mutations were detected in hypothetical proteins, non-coding regions, some of which are on repetitive DNA regions (Nakagawa et al., 1986; Okay and Kurt Kızıldoğan, 2015), and the effects of mutational variants in these regions are still unknown.

SNPs analysis in three *P. multocida* genomes (X73, P1059 and Pm70) has indicated a higher dN/dS ratio within proteins localized on the outer membrane or secreted such as PfhB2, HgbA, HemR, a secreted effector protein, an iron-regulated outer membrane protein, TadD-F (pilus assembly proteins), RcpB-C (pilus assembly proteins), and PlpP proteins. This event suggests that they are

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

under diversifying selection due to interactions with the host immune system, although further analyses would be required to confirm this observation (Johnson et al., 2013). There were reported known virulence genes *bcbAB*, *fcbC*, *lipA*, *bexDCA*, *ctrCD*, *lgtA*, *lgtC*, *lic2A* involved in the biogenesis of surface polysaccharides, *hsf* encoding autotransporter adhesin, and *fhaB* encoding filamentous haemagglutinin, are under diversifying selection in 33 strains of *P. multocida* (Cao et al., 2017).

Later, a screening based on genome-wide positive selection found a total of 35 genes subject to strong selection pressure, preferably located in extracellular space, because of the interactions with host immune systems (Chen et al., 2006; Fitzpatrick et al., 2005; Johnson et al., 2013) or with a role in the bacterial colonization and adaptation to host niches (Lin et al., 2002). Among the Positively Selected Genes were described *omp47*, *ompP5* and *ompA*, associated with bacterial evasion of host immune responses. *hgbA* gene, TonB-dependent receptor family protein involved in heme and iron acquisition and icsA gene that encodes an autotransporter related to the Type V secretion system (Cao et al., 2017). Likewise, three genes (*rffG*, *ntpA*, *wzz*,) involved in the biogenesis of LPS, gene *ftsW* coding a transporter of lipid-linked peptidoglycan precursors, *yedZ* encoding a sulfoxide reductase heme-binding subunit involved in protein repair were also subject to positive selection. The function associated with the metabolism of protein and DNA on the adaptation process still unknown (Cao et al., 2017). Also, it was identified two genes *gabR*, encoding an HTH-type transcriptional regulator with significant evidence for positive selection. Besides, the gene *qseC*, which encodes a sensor kinase in the two-component *qseBC* quorum- sensing system, was also positively selected in *P. multocida (Cao et al., 2017).* These findings indicated that a positive selection of genes involved in transcriptional control might contribute to bacterial adaptation to various niches. Bacterial two-component regulatory systems and transcription factors are known to enable rapid adaptation to environmental conditions (Cases et al., 2003).

There is still a challenge in knowing in depth the genomic diversity of the species *P. multocida*, this because there are few isolated strains, most of which are representative strains of independent populations. The availability of a major number of *P. multocida* genomes and new strategies of genomic comparison will contribute to a fine knowledge of diversity and evolutionary adaptation of species.

5. Genetic divergence

The resource of genomic plasticity and the evolutionary forces in bacterial are the key points for the enormous diversity and adaptive capacity at new and changing environments. The generation of genetic variation is mediated by mutation rate, recombination, HGT and the evolutionary forces

which address the mutational fixation are natural selection, migration, effective populations size and genetic drift (Arber, 2000). Among the major resource of variation genetic is the horizontal transfer of mobile genetic element, that allow an evolution more significative and saltational, contributing more greatly to diversification than do point mutations that are long-term (Burrus and Waldor, 2004; Levin and Bergstrom, 2000). The new genomic integration generate a genomic restructuration and could be fixed by natural selection (Brussow et al., 2004; Levin and Bergstrom, 2000).

5.1. Mobile genetic elements

The MGEs are DNA regions that encode proteins that allow mobilizing of DNA within and without genome bacterial. MGEs are transferred by three mechanisms of horizontal: transformation, conjugation and transduction (Frost et al., 2005). (Brussow et al., 2004; Levin and Bergstrom, 2000)(Brussow et al., 2004; Levin and Bergstrom, 2000)(Brussow et al., 2004; Levin and Bergstrom, 2000)(Brussow et al., 2004; Levin and Bergstrom, 2000)MGEs often carry fitness genes associated with entry, invasion and survival in a host cell (Welch et al., 2002), considered the major drivers of bacterial evolution and provide sources with adaptive traits that enhance the fitness of bacteria in new and changing niches (Wiedenbeck and Cohan, 2011). The mobile genetic elements more known are plasmids, prophages, insertion sequences, transposons and ICEs. Nowadays, a populational study of 114 *P. multocida* isolates predicted a total of 280 GEIs, 123 of which were PAIs, 27 RGIs, 97 symbiosis islands, 4 metabolic islands, 7 pathogenicity and metabolic islands, 1 pathogenicity and symbiosis island and 21 pathogenicity and resistance islands. Likewise a total of 324 phages. From accessory genomes 42.3 % of functional genes was present in the GEIs. Reporting prevalently proteins with role on the replication and repair, related to defense mechanisms and cell membranes (Zhu, 2019). The presence of numerous GEIs from accessory genome accomplish an important role on the diversification and adaptation of *P. multocida*. The presence of MGEs in *P. multocida* will be detailed at continuation.

5.1.1. Plasmids

Most of these plasmids confer adaptive advantages and increase the survival of specific conditions, in the majority of cases carrying multiple antibiotic resistance genes (Heuer and Smalla, 2012). Nowadays, antibiotic treatment is an effective strategy against *P. multocida*. However, there is limited information regarding the antibiotic susceptibility and resistance, with excessive and unjustified use of antimicrobials and high prevalence of multidrug resistance containing in plasmid from *P. multocida* (Furian et al., 2016; Oh et al., 2018). A total of 19 plasmids from *P. multocida* have been

submitted in the NCBI database, that being compost by resistance genes. The plasmids are classified on conjugative and small or mobilization plasmids (Frost et al., 2005).

Plasmids have also been detected in the majority of porcine *P. multocida* strains examined (Gardner et al., 1994; Schwarz and Chaslus-Dancla, 2001). The complete nucleotide sequence of a plasmid carrying resistance genes will be descriptive, founding streptomycin and sulphonamide resistance genes (Table 3) (Wright et al., 1997). The pCCK1900 plasmid present florfenicol resistance genes (Corinna Kehrenberg et al., 2008), there were reported multi-resistance isolates carrying on multiple plasmids harboring ampicillin, streptomycin, tetracycline and sulphonamide resistance gene (Table 3) (San Millan et al., 2009).

In the outbreaks from fowl cholera of different countries, isolates resistant to streptomycin, trimethoprim, lincomycin, tetracycline (Diallo et al., 1995) lincomycin, sulfadiazine and tetracycline (Jonas et al., 2001), ampicillin, amikacin and tetracycline were reported (Shivachandra et al., 2004); but without correlation with their presence on plasmids. Likewise, there were reported plasmids containing streptomycin, sulfonamide, tetracycline, and kanamycin resistance genes (Hirsh et al., 1985). The complete nucleotide sequence of a plasmid carrying resistance genes was pJR1 plasmid harboring sulfonamide, tetracycline, chloramphenicol resistance genes (Table 3); pJR2 plasmid harbored streptomycin and spectinomycin, ampicillin and carbenicillin resistance genes (Table 3)(Wu et al., 2015a; Zhao et al., 2012). Pvm111 plasmid harbored sulfonamide, streptomycin and tetracycline resistance gene (Kehrenberg et al., 2003). Also, it was identified a cryptic plasmid, pLEM isolated from avian (Table 3)(McGee and Bejcek, 2001).

Plasmids isolated from cattle strains of *P. multocida* code for streptomycin, sulphonamide, tetracycline (Hirsh et al., 1985; Silver et al., 1979), or ampicillin resistance (Livrelli et al., 1988) or are phenotypically cryptic (Price et al., 1993). The complete nucleotide sequence of Ppmt1 and Pmth1 plasmid presented tetracycline resistance gene (Table 3) (Kehrenberg et al., 2001, 1998), non-conjugative plasmids encoding resistance to ampicillin in both cases from pneumonia (Rosenau et al., 1991). The pCCK411 plasmid present kanamycin resistance gene (Table 3) (Yamashita et al., 2014). The pCCK381 plasmid probably results in from inter-plasmid recombination carrying chloramphenicol and florfenicol resistance genes (Table 3)(Kehrenberg and Schwarz, 2005). Small plasmids were isolated from hemorrhagic septicemia outbreaks, which do not carry any virulence gene or antibiotic resistance, with no apparent role in pathogenicity (Ameen, MahmoodJamal et al., 2005). Also, it was reported a pCCK647 plasmid containing spectinomycin/streptomycin resistance gene (Table 3)(Kehrenberg et al., 2005). In the lambs lung were detected the pB1018 plasmid that presented tetracycline resistance gene (Table 3) (Ares-Arroyo et al., 2018).

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

The last studies show resistance or susceptibility for several antibiotics of class aminoglycosides (Rigobelo et al., 2013; WANG et al., 2017; Yeh et al., 2017), penicillin (Ferreira et al., 2012; Furian et al., 2016; Rigobelo et al., 2013), tetracyclines (Oh et al., 2018; Rigobelo et al., 2013) and florfenicol (Ferreira et al., 2012; Oh et al., 2018). However, there are isolated with resistance to macrolides (Ferreira et al., 2012; Yeh et al., 2017) and sulphonamides (Furian et al., 2016; Oh et al., 2018). Also, there are isolates susceptible to quinolones (Ferreira et al., 2012) and Cephalosporins (Rigobelo et al., 2013) being considered the most effective drugs (Ferreira et al., 2012; Kehrenberg et al., 2001).

Different studies describe resistance to different antibiotics class, however, it is not clear if the transference of them will be chromosomic, although already the presence of multidrug resistance and spread would show their acquisition by plasmid transfer (Corinna Kehrenberg et al., 2008; Tang et al., 2009). Even so, the presence of plasmids has not been reported in most sequenced genomes; this would show the lack of strategies to complement the genomic information of the bacterium.

Table 3. Complete Plasmids and ICEs identified in *P. multocida*.

Antibiotic class						Aminoglycosides									Su *	Macrolides			Penicillins			Tetracyclines					Cm* or/and Ff *	
Nome plasmid	**Accession**	**Host / Tissue / Disease**	**Size**	**Resistance genes**	**References**	**aphA1**	**strB**	**strA**	**aadB**	**aadA25**	**aadA1**	**aphA3**	***aadA14***	**aph(3'')-Ib**	**sul2**	**erm(42)**	**msr(E)**	**mph(E)**	**bla OXA−2**	**BlaROB-1**	**blaP1**	**tetH**	**tetR**	**tetO**	**tetB**	**tetG**	**catB2**	**floR**
no information	no information	Calf/lung	4.3kb	b-lactamase resistant to ampicillin, ticarcillin, and narrow-spectrum cephalosporins	(Rosenau et al., 1991)																							
pIG1	NC_001774.1	Pig	5.36 kb	sulphonamide and streptomycin resistance	(Wright et al., 1997)																							
pPMT1	no information	Cattle/pneumonia	4.828 kb	tet(H)-tetR resistance	(Kehrenberg et al., 1998)																	X	X					
pMHT1	no information	no information	4.4 kb	tet(H), tetR	(Kehrenberg et al., 2001)																	X	X					
pLEM	no information	Avian	3.745 kb	no information	(McGee and Bejcek, 2001)																							
pJR1	NC_004771.1	Fowl cholera	6792 bp	sulfonamide resistant (sul2), tetracycline resistance protein (tetG), chloramphenicol acetyltransferase gene (catB2)	(Wu et al., 2015b, 2003)										X											X	X	
pJR2	NC_004772.1	Fowl cholera	5252 bp	aminoglycoside adenylyltransferase (aadA1), beta-lactamase (blaP1)	(Wu et al., 2015b, 2003)						X										X							
Pvm111	no information	Fowl cholera	no information	sulphonamide (sul2), streptomycin (strA-strB) and tetracycline resistance [tetR-tet(H)]	(Kehrenberg et al., 2003)		X	X							X							X	X					
pCCK381	NC_006994.1	Bovine respiratory disease	10.8 kb	florfenicol , chloramphenicol resistance (floR)	(Kehrenberg et al., 2005)																							X
pCCK647	NC_006868.1	Calves/peritonitis	5.198 kb	spectinomycin/streptomycin resistance gene, designated aadA14	(Kehrenberg et al., 2005)								X															
pCCK1900	NC_011378.1	Pig	10.226 kb	no information	(C. Kehrenberg et al., 2008)																							X
pB1000		Pig	4.613 kb	ampicillin resistance gene (blaROB-1)	(San Millan et al., 2009)															X								

Single Nucleotide Polymorphisms (SNPs)

Mobile Genetic Elements (MGEs)

pB1005	NC_012215.1	Pig	no information	streptomycin and sulphonamide (sul2 and strA)	(San Millan et al., 2009)			X							X														
pB1006	NC_012216.1	Pig	no information	tetracycline resistance genes tet(O)	(San Millan et al., 2009)																			X					
pB1001		Pig	5.128 kb	tet(B)	(San Millan et al., 2009)																				X				
p9956	no information	Pig	no information	tet(H)	(San Millan et al., 2009)																	X							
p1002	no information	Pig	no information	ampicillin resistance gene (blaROB-1)	(San Millan et al., 2009)															X									
p1003	no information	Pig	no information	sul2 and strA	(San Millan et al., 2009)			X							X														
ICE*Pmu1*,	CP003022	Bovine respiratory disease	85 kb	streptomycin/spectinomycin (aadA25), streptomycin (strA and strB), gentamicin (aadB), kanamycin/neomycin (aphA1), tetracycline [tetR-tet(H)], chloramphenicol/florfenicol (floR), sulphonamides (sul2), tilmicosin/ clindamycin [erm(42)] or tilmicosin/tulathromycin [msr(E)-mph(E)]	(Michael et al., 2012)	X	X	X	X	X					X	X	X	X	X			X					X
pCCK411	NC_016973.1	Cattle		kanamycin resistance gene	(Yamashita et al., 2014)							X																	
ICEPmu3358	no information	Bovine respiratory disease	no information	kanamycin/neomycin aphA1,streptomycin strB, streptomycin, strA, sul2, chloramphenicol/florfenicol floR, tilmicosin/ clindamycin erm(42), gentamicin aadB, aadA25, bla OXA−2, tilmicosin/tulathromycin msr(E), mph(E), tetracycline tetH	(Beker et al., 2018)	X	X	X	X	X					X	X	X	X	X			X					X
ICEPmu3361	no information	Bovine respiratory disease	no information	aadB, aadA25, bla OXA−2, msr(E), mph(E), tetH	(Beker et al., 2018)				X	X							X	X	X			X						

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

ICEPmu12591	no information	Bovine respiratory disease	no information	aphA1, strB, strA, sul2, tetH	(Beker et al., 2018)	X	X	X							X							X						X
ICEPmu12601	no information	Bovine respiratory disease	no information	aphA1, strB, strA, sul2, tetH	(Beker et al., 2018)	X	X	X							X							X						X
pOV	NC_019381.1		13,551 kb	blaROB-1 , sul2, and strAB genes	(Li et al., 2018)		X	X							X					X								
pB1018	NC_019981.1	Lambs/lungs		tetracycline resistance gene tetH	(Ares-Arroyo et al., 2018)																	X						
	NC_017035.1	no information	5360 bp	aminoglycoside O-phosphotransferase APH(3")-Ib. sulfonamide-resistant dihydropteroate synthase Sul2	NCBI									X	X													
	NZ_CP020346	Rabbit/lung	325.255 kb	no information	NCBI																							
	NZ_CP020404.1	no information	3742 bp	no information	NCBI																							
pS298D	NZ_CM009574.1	no information	29.269 kb	no information	NCBI																							
	CP020348, CP020349	no information	34.596 kb	no information	NCBI																							

Abbreviation of antibiotics class*: Su, sulfonamides Cm, chloramphenicol; Ff, florfenicol.

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

5.1.2. Integrative and conjugative elements (ICEs)

ICEs are self-transmissible elements of DNA able, like conjugative plasmid, and capable of integrating into the chromosome and its replication occurs together with chromosomal replication (Burrus et al., 2002; Wozniak and Waldor, 2010). The presence of ICEs will be related to confer fitness of the host and its survival under specific conditions (Burrus and Waldor, 2004).

The strain 36950 isolated from a case of BRD, present an ICE called ICE*Pmu1*, element of 85 kb length that carries 88 genes, including 12 different resistance genes which confer resistance to streptomycin/spectinomycin, streptomycin, gentamicin, kanamycin/neomycin, tetracycline, chloramphenicol/florfenicol, sulphonamide, tilmicosin/clindamycin or tilmicosin/tulathromycin, additionally contain genes involved in conjugative transfer and its chromosomal excision/integration (Table 3) (Michael et al., 2012). ICEpmcn07 was identified in HN07 strain, containing set of genes to encoding a putative type IV secretion system (T4SS), their presence may also be associated with the virulence variability as ICEhin1056, with which present high similarity (Peng et al., 2017). ICEPmu3358, ICEPmu12591, ICEPmu1260 and ICEPmu3361 were sequenced that include clustering of the antibiotic class aminoglycosides, florfenicol, macrolides, sulfonamides, beta-lactams and tetracyclines (Table 3) (Beker et al., 2018).

5.1.3. ISs and Transposons

Insertion sequences and transposons present the ability to move and carrying the information necessary for transposition (Burrus et al., 2002). Typically, an IS consist of a gene encoded a transposase enzyme and it is flanked by a specific inverted repeat. Transposons are constituted by two IS flanking and generally containing genes that confer antibiotics resistance.

The role of the transposons is important for the dissemination of multiple antibiotic resistance containing in conjugative plasmids (Burrus et al., 2002; Burrus and Waldor, 2004). Such is the case of mapping of plasmid pPMT1 and pMHT1 from *P. multocida* containing integrated transposon Tn5706. This transposon of size 4,378-bp consists of an internal *tet*(H)-*tetR* resistance gene region which is flanked by almost identical insertion elements, IS*1596* and IS1597. The nucleotide sequence of the 4,828bp region of plasmid pPMT1, including the entire transposon, Tn5706 and its adjacent flanking regions (Kehrenberg et al., 1998).

5.1.4. Genomic islands

Genomic islands (GEIs) are common large genomic region transferred and inserted into bacterial genome (Juhas et al., 2009). GEIs are elements that drive at the genome evolution and adaptation of

731 bacteria (Dobrindt et al., 2004). Some GEIs are stable and not easily transferable, while others are
732 highly mobile (Gyles and Boerlin, 2014).
733 GEIs are composed by clusters of genes, which allows classifying by the function of genes
734 encoding virulence factors (Dobrindt et al., 2000; Gal-Mor and Finlay, 2006); responsible for
735 antibiotic resistance (Krizova and Nemec, 2010); related to metabolism (Tumapa et al., 2008); for
736 sustaining a bacterium symbiotic relationship (Barcellos et al., 2007).
737 Several studies of *P. multocida* report the presence of complete, incomplete and questionable
738 genomic islands that were previously discussed (Hurtado et al., 2018; Moustafa et al., 2015; Okay
739 and Kurt Kızıldoğan, 2015).
740
741 **Conclusion**
742 This is the first review article that related genomic and epidemiology to answer evolutionary
743 adaptation of this important veterinary pathogen. The presence of a diverse number of mobile genetic
744 elements shows a dynamic genome of *P. multocida,* fixed by selective pressure, such as the presence
745 of resistance genes on multiple plasmids. The diversity of mobile genetic elements and single
746 nucleotide polymorphism reveals that the strains are in a constant diversification process,
747 evolutionary convergence and selection pressures by adaptation to niche. A union of all information
748 correlated to the availability of genome data from *P. multocida* would allow the elaboration of new
749 strategies to control, combat and elucidate the pathogenomics with *P. multocida*.
750
751 **Competing interests**
752 The authors declare that they have no conflict of interest.
753
754 **Funding**
755 This work was supported by the Programa Nacional de Innovación para la Competitividad y
756 Productividad (Innóvate Perú), under the contract 133-FINCyT-IB-2013, as part of the project
757 "Reverse Vaccinology: Developing a new generation vaccine for the control and/or prevention
758 pasteurellosis pneumonia in alpacas". In addition, this study was financed in part by the Coordenação
759 de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) and Conselho Nacional de
760 Desenvolvimento Científico e Tecnológico (CNPq).
761
762 **References**
763 Achtman, M., 2008. Evolution, Population Structure, and Phylogeography of Genetically
764 Monomorphic Bacterial Pathogens. Annu. Rev. Microbiol.

765 https://doi.org/10.1146/annurev.micro.62.081307.162832
766 Ameen, MahmoodJamal, H., Chua, K.H., Frederick, D., Mahmood Ameen, A., Salmah, I., 2005.
767 Plasmid DNA Analysis of Pasteurella multocida Serotype B isolated from Haemorrhagic
768 Septicaemia outbreaks in Malaysia. Malays. J. Microbiol.
769 Amonsin, A., Wellehan, J.F.X., Li, L.L., Laber, J., Kapur, V., 2002. DNA fingerprinting of
770 Pasteurella multocida recovered from avian sources. J. Clin. Microbiol.
771 https://doi.org/10.1128/JCM.40.8.3025-3031.2002
772 Arber, W., 2000. Genetic variation: Molecular mechanisms and impact on microbial evolution.
773 FEMS Microbiol. Rev. https://doi.org/10.1016/S0168-6445(99)00032-7
774 Ares-Arroyo, M., Bernabe-Balas, C., Santos-Lopez, A., Baquero, M.R., Prasad, K.N., Cid, D.,
775 Martin-Espada, C., Millan, A.S., Gonzalez-Zorn, B., 2018. PCR-based analysis of ColE1
776 plasmids in clinical isolates and metagenomic samples reveals their importance as gene capture
777 platforms. Front. Microbiol. https://doi.org/10.3389/fmicb.2018.00469
778 Barcellos, F.G., Menna, P., Batista, J.S.D.S., Hungria, M., 2007. Evidence of horizontal transfer of
779 symbiotic genes from a Bradyrhizobium japonicum inoculant strain to indigenous diazotrophs
780 Sinorhizobium (Ensifer) fredii and Bradyrhizobium elkanii in a Brazilian savannah soil. Appl.
781 Environ. Microbiol. https://doi.org/10.1128/AEM.01823-06
782 Beker, M., Rose, S., Lykkebo, C.A., Douthwaite, S., 2018. Integrative and Conjugative Elements
783 (ICEs) in Pasteurellaceae Species and Their Detection by Multiplex PCR. Front. Microbiol. 9,
784 1329. https://doi.org/10.3389/fmicb.2018.01329
785 Benkirane, A., De Alwis, M.D.L., 2002. Haemorrhagic septicaemia, its significance, prevention and
786 control in Asia. Vet. Med. (Praha). https://doi.org/10.17221/5830-VETMED
787 Bentley, S.D., Parkhill, J., 2015. Genomic perspectives on the evolution and spread of bacterial
788 pathogens. Proc. R. Soc. B Biol. Sci. https://doi.org/10.1098/rspb.2015.0488
789 Bethe, A., Wieler, L.H., Selbitz, H.J., Ewers, C., 2009. Genetic diversity of porcine Pasteurella
790 multocida strains from the respiratory tract of healthy and diseased swine. Vet. Microbiol.
791 https://doi.org/10.1016/j.vetmic.2009.04.027
792 Bisgaard, M., Petersen, A., Christensen, H., 2013. Multilocus sequence analysis of Pasteurella
793 multocida demonstrates a type species under development. Microbiol. (United Kingdom).
794 https://doi.org/10.1099/mic.0.063461-0
795 Blackall, P.J., Fegan, N., Chew, G.T.I., Hampson, D.J., 1998. Population structure and diversity of
796 avian isolates of Pasteurella multocida from Australia. Microbiology.
797 https://doi.org/10.1099/00221287-144-2-279
798 Blackall, P.J., Fegan, N., Pahoff, J.L., Storie, G.J., McIntosh, G.B., Cameron, R.D.A., O'Boyle, D.,

Frost, A.J., Bará, M.R., Marr, G., Holder, J., 2000. The molecular epidemiology of four outbreaks of porcine pasteurellosis. Vet. Microbiol. https://doi.org/10.1016/S0378-1135(99)00192-3

Boerlin, P., Siegrist, H.H., Burnens, A.P., Kuhnert, P., Mendez, P., Prétat, G., Lienhard, R., Nicolet, J., 2000. Molecular identification and epidemiological tracing of Pasteurella multocida meningitis in a baby. J. Clin. Microbiol.

Boyce, J.D., Harper, M., Wilkie, I.W., Adler, B., 2010. Pasteurella, in: Pathogenesis of Bacterial Infections in Animals: Fourth Edition. https://doi.org/10.1002/9780470958209.ch17

Boyce, J.D., Seemann, T., Adler, B., Harper, M., 2012. Pathogenomics of Pasteurella multocida. Curr. Top. Microbiol. Immunol. 361, 23–38. https://doi.org/10.1007/82-2012-203

Brogden, K.A., 1980. Physiological and serological characteristics of 48 Pasteurella multocida cultures from rabbits. J. Clin. Microbiol.

Brown, J.R., Douady, C.J., Italia, M.J., Marshall, W.E., Stanhope, M.J., 2001. Universal trees based on large combined protein sequence data sets. Nat. Genet. https://doi.org/10.1038/90129

Brussow, H., Canchaya, C., Hardt, W.-D., 2004. Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. Microbiol. Mol. Biol. Rev. https://doi.org/10.1128/MMBR.68.3.560-602.2004

Burrus, V., Pavlovic, G., Decaris, B., Guédon, G., 2002. Conjugative transposons: The tip of the iceberg. Mol. Microbiol. https://doi.org/10.1046/j.1365-2958.2002.03191.x

Burrus, V., Waldor, M.K., 2004. Shaping bacterial genomes with integrative and conjugative elements. Res. Microbiol. https://doi.org/10.1016/j.resmic.2004.01.012

Cao, P., Guo, D., Liu, J., Jiang, Q., Xu, Z., Qu, L., 2017. Genome-wide analyses reveal genes subject to positive selection in Pasteurella multocida. Front. Microbiol. 8. https://doi.org/10.3389/fmicb.2017.00961

Cardoso-Toset, F., Gómez-Laguna, J., Callejo, M., Vela, A.I., Carrasco, L., Fernández-Garayzábal, J.F., Maldonado, A., Luque, I., 2013. Septicaemic pasteurellosis in free-range pigs associated with an unusual biovar 13 of Pasteurella multocida. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2013.08.005

CARTER, G.R., 1955. Studies on Pasteurella multocida. I. A hemagglutination test for the identification of serological types. Am. J. Vet. Res. 16, 481–484.

Casali, N., Nikolayevskyy, V., Balabanova, Y., Harris, S.R., Ignatyeva, O., Kontsevaya, I., Corander, J., Bryant, J., Parkhill, J., Nejentsev, S., Horstmann, R.D., Brown, T., Drobniewski, F., 2014. Evolution and transmission of drug-resistant tuberculosis in a Russian population. Nat. Genet. https://doi.org/10.1038/ng.2878

Cases, I., De Lorenzo, V., Ouzounis, C.A., 2003. Transcription regulation and environmental adaptation in bacteria. Trends Microbiol. https://doi.org/10.1016/S0966-842X(03)00103-3

Chanter, N., Rutter, J.M., 1990. Colonisation by Pasteurella multocida in atrophic rhinitis of pigs and immunity to the osteolytic toxin. Vet. Microbiol. https://doi.org/10.1016/0378-1135(90)90082-7

Chen, S.L., Hung, C.-S., Xu, J., Reigstad, C.S., Magrini, V., Sabo, A., Blasiar, D., Bieri, T., Meyer, R.R., Ozersky, P., Armstrong, J.R., Fulton, R.S., Latreille, J.P., Spieth, J., Hooton, T.M., Mardis, E.R., Hultgren, S.J., Gordon, J.I., 2006. Identification of genes subject to positive selection in uropathogenic strains of Escherichia coli: A comparative genomics approach. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.0600938103

Chengappa, M.M., Myers, R.C., Carter, G.R., 1982. Capsular and somatic types of Pasteurella multocida from rabbits. Can. J. Comp. Med. https://doi.org/10.4271/2005-01-1413

Christensen, H., Bisgaard, M., 2010. Molecular classification and its impact on diagnostics and understanding the phylogeny and epidemiology of selected members of Pasteurellaceoe of veterinary importance. Berl. Munch. Tierarztl. Wochenschr. https://doi.org/10.2376/0005-9366-123-20

Christensen, J.P., Bojesen, A.M., Bisgaard, M., 2008. Fowl cholera, in: Poultry Diseases. https://doi.org/10.1016/B978-0-7020-2862-5.50015-5

Collins, M.T., Weaver, N., Ellis, R.P., 1981. Identification of Pasteurella multocida and Pasteurella haemolytica by APSI 20E, Minitek, and Oxi/Ferm systems. J. Clin. Microbiol. 13, 433–437.

Comas, I., Moya, A., González-Candelas, F., 2007. From phylogenetics to phylogenomics: The evolutionary relationships of insect endosymbiotic γ-proteobacteria as a test case. Syst. Biol. https://doi.org/10.1080/10635150601109759

Dabo, S.M., Confer, a W., Montelongo, M., Lu, Y.S., 1999. Characterization of rabbit Pasteurella multocida isolates by use of whole-cell, outer-membrane, and polymerase chain reaction typing. Lab. Anim. Sci.

Dabo, S.M., Confer, A.W., Murphy, G.L., 1997. Outer membrane proteins of bovine Pasteurella multocida serogroup A isolates. Vet. Microbiol. https://doi.org/10.1016/S0378-1135(96)01274-6

Dabo, S.M., Taylor, J.D., Confer, a W., 2007. Pasteurella multocida and bovine respiratory disease. Anim. Health Res. Rev. https://doi.org/10.1017/S1466252307001399

Davies, R.L., MacCorquodale, R., Baillie, S., Caffrey, B., 2003. Characterization and comparison of Pasteurella multocida strains associated with porcine pneumonia and atrophic rhinitis. J. Med. Microbiol. 52, 59–67. https://doi.org/10.1099/jmm.0.05019-0

867 Davies, R.L., MacCorquodale, R., Reilly, S., 2004. Characterisation of bovine strains of Pasteurella
868 multocida and comparison with isolates of avian, ovine and porcine origin. Vet. Microbiol.
869 https://doi.org/10.1016/j.vetmic.2003.11.013
870 De Alwis, M.D.L., 1999. Haemorrhogic septicemia. Aust. Cent. Int. Agric. Res. 129–141.
871 Diallo, I.S., Bensink, J.C., Frost, A.J., Spradbrow, P.B., 1995. Molecular studies on avian strains of
872 Pasteurella multocida in Australia. Vet. Microbiol. https://doi.org/10.1016/0378-
873 1135(95)00099-V
874 DiGiacomo, R.F., Deeb, B.J., Brodie, S.J., Zimmerman, T.E., Veltkamp, E.R., Chrisp, C.E., 1993.
875 Toxin production by Pasteurella multocida isolated from rabbits with atrophic rhinitis. Am. J.
876 Vet. Res.
877 DiGiacomo, R.F., Xu, Y.M., Allen, V., Hinton, M.H., Pearson, G.R., 1991. Naturally acquired
878 Pasteurella multocida infection in rabbits: clinicopathological aspects. Can. J. Vet. Res.
879 Djordjevic, S.P., Eamens, G.J., Ha, H., Walker, M.J., Chin, J.C., 1998a. Demonstration that
880 Australian Pasteurella multocida isolates from sporadic outbreaks of porcine pneumonia are
881 non-toxigenic (toxA-) and display heterogeneous DNA restriction endonuclease profiles
882 compared with toxigenic isolates from herds with progressive atrophic rhinitis. J. Med.
883 Microbiol. 47, 679–688. https://doi.org/10.1099/00222615-47-8-679
884 Djordjevic, S.P., Eamens, G.J., Ha, H., Walker, M.J., Chin, J.C., 1998b. Demonstration that
885 Australian Pasteurella multocida isolates from sporadic outbreaks of porcine pneumonia are
886 non-toxigenic (toxA-) and display heterogeneous DNA restriction endonuclease profiles
887 compared with toxigenic isolates from herds with progressiv. J. Med. Microbiol.
888 https://doi.org/10.1099/00222615-47-8-679
889 Dobrindt, U., Hochhut, B., Hentschel, U., Hacker, J., 2004. Genomic islands in pathogenic and
890 environmental microorganisms. Nat. Rev. Microbiol. https://doi.org/10.1038/nrmicro884
891 Dobrindt, U., Janke, B., Piechaczek, K., Nagy, G., Ziebuhr, W., Fischer, G., Schierhorn, A., Hecker,
892 M., Blum-Oehler, G., Hacker, J., 2000. Toxin genes on pathogenicity islands: Impact for
893 microbial evolution. Int. J. Med. Microbiol. https://doi.org/10.1016/S1438-4221(00)80028-4
894 Du, H., Fang, R., Pan, T., Li, T., Li, N., He, Q., Wu, R., Peng, Y., Zhou, Z., 2016. Comparative
895 Genomics Analysis of Two Different Virulent Bovine *Pasteurella multocida* Isolates. Int. J.
896 Genomics 2016, 1–14. https://doi.org/10.1155/2016/4512493
897 Eisen, J.A., Fraser, C.M., 2003. Phylogenomics: Intersection of evolution and genomics. Science
898 (80-.). https://doi.org/10.1126/science.1086292
899 Enright, M.C., Spratt, B.G., 1998. A multilocus sequence typing scheme for Streptococcus
900 pneumoniae: Identification of clones associated with serious invasive disease. Microbiology.

https://doi.org/10.1099/00221287-144-11-3049
Ewers, C., Lübke-Becker, A., Bethe, A., Kießling, S., Filter, M., Wieler, L.H., 2006. Virulence genotype of Pasteurella multocida strains isolated from different hosts with various disease status. Vet. Microbiol. 114, 304–317. https://doi.org/10.1016/j.vetmic.2005.12.012
Ferreira, T.S.P., Felizardo, M.R., Sena de Gobbi, D.D., Gomes, C.R., Nogueira Filsner, P.H. de L., Moreno, M., Paixão, R., Pereira, J. de J., Micke Moreno, A., 2012. Virulence Genes and Antimicrobial Resistance Profiles of *Pasteurella multocida* Strains Isolated from Rabbits in Brazil. Sci. World J. 2012, 1–6. https://doi.org/10.1100/2012/685028
Fitzpatrick, D.A., Creevey, C.J., McInerney, J.O., 2005. Evidence of positive Darwinian selection in putative meningococcal vaccine antigens. J. Mol. Evol. https://doi.org/10.1007/s00239-004-0290-6
Ford, C.B., Shah, R.R., Maeda, M.K., Gagneux, S., Murray, M.B., Cohen, T., Johnston, J.C., Gardy, J., Lipsitch, M., Fortune, S.M., 2013. Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. Nat. Genet. https://doi.org/10.1038/ng.2656
Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: The agents of open source evolution. Nat. Rev. Microbiol. https://doi.org/10.1038/nrmicro1235
Furian, T.Q., Borges, K.A., Laviniki, V., da Silveira Rocha, S.L., de Almeida, C.N., do Nascimento, V.P., Salle, C.T.P., de Souza Moraes, H.L., 2016. Virulence genes and antimicrobial resistance of Pasteurella multocida isolated from poultry and swine. Brazilian J. Microbiol. https://doi.org/10.1016/j.bjm.2015.11.014
Furian, T.Q., Borges, K.A., Rocha, S.L.S., Rodrigues, E.E., do Nascimento, V.P., Salle, C.T.P., Moraes, H.L.S., 2013. Detection of virulence-associated genes of Pasteurella multocida isolated from cases of fowl cholera by multiplex-PCR. Pesqui. Vet. Bras. https://doi.org/10.1590/S0100-736X2013000200007
Fussing, V., Nielsen, J.P., Bisgaard, M., Meyling, A., 1999. Development of a typing system for epidemiological studies of porcine toxin-producing Pasteurella multocida ssp. multocida in Denmark. Vet. Microbiol. https://doi.org/10.1016/S0378-1135(98)00288-0
Gal-Mor, O., Finlay, B.B., 2006. Pathogenicity islands: A molecular toolbox for bacterial virulence. Cell. Microbiol. https://doi.org/10.1111/j.1462-5822.2006.00794.x
García-Alvarez, A., Chaves, F., Fernández, A., Sanz, C., Borobia, M., Cid, D., 2015. An ST11 clone of Pasteurella multocida, widely spread among farmed rabbits in the Iberian Peninsula, demonstrates respiratory niche association. Infect. Genet. Evol. https://doi.org/10.1016/j.meegid.2015.07.018

García-Alvarez, A., Vela, A.I., San Martín, E., Chaves, F., Fernández-Garayzábal, J.F., Lucas, D., Cid, D., 2017. Characterization of Pasteurella multocida associated with ovine pneumonia using multi-locus sequence typing (MLST) and virulence-associated gene profile analysis and comparison with porcine isolates. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2017.04.015

Garcia, N., Fernandez-Garayzabal, J.F., Goyache, J., Dominguez, L., Vela, A.I., 2011. Associations between biovar and virulence factor genes in Pasteurella multocida isolates from pigs in Spain. Vet. Rec. 169, 362–362. https://doi.org/10.1136/vr.d4869

Gardner, I.A., Kasten, R., Eamens, G.J., Snipes, K.P., Anderson, R.J., 1994. Molecular Fingerprinting of Pasteurella Multocida Associated with Progressive Atrophic Rhinitis in Swine Herds. J. Vet. Diagnostic Investig. https://doi.org/10.1177/104063879400600407

Gu, X., Zhang, H., 2004. Genome phylogenetic analysis based on extended gene contents. Mol. Biol. Evol. https://doi.org/10.1093/molbev/msh138

Gunawardana, G.A., Townsend, K.M., Frost, A.J., 2000. Molecular characterisation of avian Pasteurella multocida isolates from Australia and Vietnam by REP-PCR and PFGE. Vet. Microbiol. https://doi.org/10.1016/S0378-1135(99)00191-1

Guo, D., Lu, Y., Zhang, A., Liu, J., Yuan, D., Jiang, Q., Lin, H., Si, C., Qu, L., 2012. Identification of genes transcribed by Pasteurella multocida in rabbit livers through the selective capture of transcribed sequences. FEMS Microbiol. Lett. https://doi.org/10.1111/j.1574-6968.2012.02559.x

Gyles, C., Boerlin, P., 2014. Horizontally Transferred Genetic Elements and Their Role in Pathogenesis of Bacterial Disease. Vet. Pathol. https://doi.org/10.1177/0300985813511131

Hansen, M.S., Pors, S.E., Jensen, H.E., Bille-Hansen, V., Bisgaard, M., Flachs, E.M., Nielsen, O.L., 2010. An investigation of the pathology and pathogens associated with porcine respiratory disease complex in Denmark. J. Comp. Pathol. https://doi.org/10.1016/j.jcpa.2010.01.012

Harper, M., Boyce, J.D., Adler, B., 2006. Pasteurella multocida pathogenesis: 125 Years after Pasteur. FEMS Microbiol. Lett. https://doi.org/10.1111/j.1574-6968.2006.00442.x

Harper, M., St.Michael, F., John, M., Vinogradov, E., Steen, J.A., van Dorsten, L., Steen, J.A., Turni, C., Blackall, P.J., Adler, B., Cox, A.D., Boycea, J.D., 2013. Pasteurella multocida heddleston serovar 3 and 4 strains share a common lipopolysaccharide biosynthesis locus but display both inter- and intrastrain lipopolysaccharide heterogeneity. J. Bacteriol. https://doi.org/10.1128/JB.00779-13

Hata, E., Katsuda, K., Kobayashi, H., Uchida, I., Tanaka, K., Eguchi, M., 2010. Genetic variation among Staphylococcus aureus strains from bovine milk and their relevance to methicillin-resistant isolates from humans. J. Clin. Microbiol. https://doi.org/10.1128/JCM.01940-09

969 Hatfaludi, T., Al-Hasani, K., Boyce, J.D., Adler, B., 2010. Outer membrane proteins of Pasteurella
970 multocida. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2010.01.027
971 Heddleston, K.L., Gallagher, J.E., Rebers, P.A., 1972. Fowl Cholera: Gel Diffusion Precipitin Test
972 for Serotyping Pasteurella multocida from Avian Species. Avian Dis. 16, 925.
973 https://doi.org/10.2307/1588773
974 Herron-Olson, L., Fitzgerald, J.R., Musser, J.M., Kapur, V., 2007. Molecular correlates of host
975 specialization in Staphylococcus aureus. PLoS One.
976 https://doi.org/10.1371/journal.pone.0001120
977 Heuer, H., Smalla, K., 2012. Plasmids foster diversification and adaptation of bacterial populations
978 in soil. FEMS Microbiol. Rev. https://doi.org/10.1111/j.1574-6976.2012.00337.x
979 Hirsh, D.C., Martin, L.D., Rhoades, K.R., 1985. Resistance plasmids of Pasteurella multocida
980 isolated from turkeys. Am. J. Vet. Res.
981 Holst, E., Rollof, J., Larsson, L., Nielsen, J.P., 1992. Characterization and distribution of Pasteurella
982 species recovered from infected humans. J. Clin. Microbiol.
983 https://doi.org/10.1016/j.jcpa.2003.10.002
984 Hotchkiss, E.J., Hodgson, J.C., Lainson, F.A., Zadoks, R.N., 2011. Multilocus sequence typing of a
985 global collection of Pasteurella multocida isolates from cattle and other host species
986 demonstrates niche association. BMC Microbiol. https://doi.org/10.1186/1471-2180-11-115
987 Hurtado, R., Carhuaricra, D., Soares, S., Viana, M.V.C., Azevedo, V., Maturrano, L., Aburjaile, F.,
988 2018. Pan-genomic approach shows insight of genetic divergence and pathogenic-adaptation
989 of Pasteurella multocida. Gene. https://doi.org/10.1016/j.gene.2018.05.084
990 Hurtado, R.E., Aburjaile, F., Mariano, D., Canário, M.V., Benevides, L., Fernandez, D.A., Allasi,
991 N.O., Rimac, R., Juscamayta, J.E., Maximiliano, J.E., Rosadio, R.H., Azevedo, V., Maturrano,
992 L., 2017. Draft Genome Sequence of a Virulent Strain of *Pasteurella Multocida* Isolated From
993 Alpaca. J. Genomics 5, 68–70. https://doi.org/10.7150/jgen.19297
994 Huson, D.H., Steel, M., 2004. Phylogenetic trees based on gene content. Bioinformatics.
995 https://doi.org/10.1093/bioinformatics/bth198
996 Jaglic, Z., Kucerova, Z., Nedbalcova, K., Hlozek, P., Bartos, M., 2004. Identification of Pasteurella
997 multocida serogroup F isolates in rabbits. J. Vet. Med. Ser. B Infect. Dis. Vet. Public Heal.
998 https://doi.org/10.1111/j.1439-0450.2004.00807.x
999 Jain, R., Rivera, M.C., Lake, J.A., 1999. Horizontal gene transfer among genomes: The complexity
1000 hypothesis. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.96.7.3801
1001 Jamaludin, R., Blackall, P.J., Hansen, M.F., Humphrey, S., Styles, M., 2005. Phenotypic and
1002 genotypic characterisation of Pasteurella multocida isolated from pigs at slaughter in New

Zealand. N. Z. Vet. J. https://doi.org/10.1080/00480169.2005.36505

Johnson, T.J., Abrahante, J.E., Hunter, S.S., Hauglund, M., Tatum, F.M., Maheswaran, S.K., Briggs, R.E., 2013. Comparative genome analysis of an avirulent and two virulent strains of avian Pasteurella multocida reveals candidate genes involved in fitness and pathogenicity. BMC Microbiol. 13, 1. https://doi.org/10.1186/1471-2180-13-106

Jonas, M., Morishita, T.Y., Angrick, E.J., Jahja, J., 2001. Characterization of nine Pasteurella multocida isolates from avian cholera outbreaks in Indonesia. Avian Dis. https://doi.org/10.2307/1593009

Juhas, M., Van Der Meer, J.R., Gaillard, M., Harding, R.M., Hood, D.W., Crook, D.W., 2009. Genomic islands: Tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol. Rev. https://doi.org/10.1111/j.1574-6976.2008.00136.x

Kardos, G., Kiss, I., 2005. Molecular epidemiology investigation of outbreaks of fowl cholera in geographically related poultry flocks. J. Clin. Microbiol. https://doi.org/10.1128/JCM.43.6.2959-2961.2005

Katsuda, K., Hoshinoo, K., Ueno, Y., Kohmoto, M., Mikami, O., 2013. Virulence genes and antimicrobial susceptibility in Pasteurella multocida isolates from calves. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2013.09.029

Kehrenberg, C., Catry, B., Haesebrouck, F., de Kruif, A., Schwarz, S., 2005. Novel spectinomycin/streptomycin resistance gene, aadA14, from Pasteurella multocida. Antimicrob. Agents Chemother. 49, 3046–9. https://doi.org/10.1128/AAC.49.7.3046-3049.2005

Kehrenberg, C., Salmon, S.A., Watts, J.L., Schwarz, S., 2001. Tetracycline resistance genes in isolates of Pasteurella multocida, Mannheimia haemolytica, Mannheimia glucosida and Mannheimia varigena from bovine and swine respiratory disease: intergeneric spread of the tet(H) plasmid pMHT1. J. Antimicrob. Chemother. 48, 631–640. https://doi.org/10.1093/jac/48.5.631

Kehrenberg, C., Schwarz, S., 2005. Plasmid-borne florfenicol resistance in Pasteurella multocida. J. Antimicrob. Chemother. 55, 773–775. https://doi.org/10.1093/jac/dki102

Kehrenberg, C., Tham, N.T.T., Schwarz, S., 2003. New plasmid-borne antibiotic resistance gene cluster in Pasteurella multocida. Antimicrob. Agents Chemother. 47, 2978–80. https://doi.org/10.1128/AAC.47.9.2978-2980.2003

Kehrenberg, C., Wallmann, J., Schwarz, S., 2008. Molecular analysis of florfenicol-resistant Pasteurella multocida isolates in Germany. J. Antimicrob. Chemother. https://doi.org/10.1093/jac/dkn359

Kehrenberg, C., Wallmann, J., Schwarz, S., 2008. Molecular analysis of florfenicol-resistant

Pasteurella multocida isolates in Germany. J. Antimicrob. Chemother. 62, 951–955. https://doi.org/10.1093/jac/dkn359

Kehrenberg, C., Werckenthin, C., Schwarz, S., 1998. Tn5706, a transposon-like element from pasteurella multocida mediating tetracycline resistance. Antimicrob. Agents Chemother. 42, 2116–2118.

King, S.J., Leigh, J.A., Heath, P.J., Luque, I., Tarradas, C., Dowson, C.G., Whatmore, A.M., 2002. Development of a multilocus sequence typing scheme for the pig pathogen Streptococcus suis: Identification of virulent clones and potential capsular serotype exchange. J. Clin. Microbiol. https://doi.org/10.1128/JCM.40.10.3671-3680.2002

Klemm, E., Dougan, G., 2016. Advances in Understanding Bacterial Pathogenesis Gained from Whole-Genome Sequencing and Phylogenetics. Cell Host Microbe. https://doi.org/10.1016/j.chom.2016.04.015

Krizova, L., Nemec, A., 2010. A 63 kb genomic resistance island found in a multidrug-resistant Acinetobacter baumannii isolate of European clone I from 1977. J. Antimicrob. Chemother. https://doi.org/10.1093/jac/dkq223

Kuhnert, P., Boerlin, P., Emler, S., Krawinkler, M., Frey, J., 2000. Phylogenetic analysis of Pasteurella multocida subspecies and molecular identification of feline P. multocida subsp. septica by 16S rRNA gene sequencing. Int. J. Med. Microbiol. https://doi.org/10.1016/S1438-4221(00)80008-9

Kurland, C.G., Canback, B., Berg, O.G., 2003. Horizontal gene transfer: A critical view. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.1632870100

Lainson, F.A., Aitchison, K.D., Donachie, W., Thomson, J.R., 2002. Typing of Pasteurella multocida isolated from pigs with and without porcine dermatitis and nephropathy syndrome. J. Clin. Microbiol. https://doi.org/10.1128/JCM.40.2.588-593.2002

Levin, B.R., Bergstrom, C.T., 2000. Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. Proc. Natl. Acad. Sci. U. S. A. 97, 6981–5.

Li, Z., Cheng, F., Lan, S., Guo, J., Liu, W., Li, X., Luo, Z., Zhang, M., Wu, J., Shi, Y., 2018. Investigation of genetic diversity and epidemiological characteristics of Pasteurella multocida isolates from poultry in southwest China by population structure, multi-locus sequence typing and virulence-associated gene profile analysis. J. Vet. Med. Sci. https://doi.org/10.1292/jvms.18-0049

Lin, J., Huang, S., Zhang, Q., 2002. Outer membrane proteins: Key players for bacterial adaptation in host niches. Microbes Infect. https://doi.org/10.1016/S1286-4579(02)01545-9

Liu, H., Zhao, Z., Xi, X., Xue, Q., Long, T., Xue, Y., 2017. Occurrence of Pasteurella multocida among pigs with respiratory disease in China between 2011 and 2015. Ir. Vet. J. https://doi.org/10.1186/s13620-016-0080-7

Livrelli, V.O., Darfeuille-Richaud, A., Rich, C.D., Joly, B.H., Martel, J.L., 1988. Genetic determinant of the ROB-1 β-lactamase in bovine and porcine Pasteurella strains. Antimicrob. Agents Chemother. https://doi.org/10.1128/AAC.32.8.1282

Loman, N.J., Pallen, M.J., 2015. Twenty years of bacterial genome sequencing. Nat. Rev. Microbiol. https://doi.org/10.1038/nrmicro3565

Maiden, M.C.J., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M., Spratt, B.G., 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.95.6.3140

Marois, C., Fablet, C., Gaillot, O., Morvan, H., Madec, F., Kobisch, M., 2009. Molecular diversity of porcine and human isolates of Pasteurella multocida. J. Appl. Microbiol. https://doi.org/10.1111/j.1365-2672.2009.04360.x

Massacci, F.R., Magistrali, C.F., Cucco, L., Curcio, L., Bano, L., Mangili, P.M., Scoccia, E., Bisgaard, M., Aalbæk, B., Christensen, H., 2018. Characterization of Pasteurella multocida involved in rabbit infections. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2017.11.023

May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whittam, T.S., Kapur, V., 2001. Complete genomic sequence of Pasteurella multocida , Pm70. Proc.Natl.Acad.Sci.U.S.A 98, 3460–3465. https://doi.org/10.1073/pnas.051634598

McGee, J.E., Bejcek, B.E., 2001. A cryptic plasmid from Pasteurella multocida has a predicted protein nearly identical to a transport protein from Actinobacillus actinomycetemcomitans. Plasmid. https://doi.org/10.1006/plas.2001.1521

McInerney, J.O., McNally, A., O'Connell, M.J., 2017. Why prokaryotes have pangenomes. Nat. Microbiol. https://doi.org/10.1038/nmicrobiol.2017.40

Medini, D., Donati, C., Tettelin, H., Masignani, V., Rappuoli, R., 2005. The microbial pan-genome. Curr. Opin. Genet. Dev. https://doi.org/10.1016/j.gde.2005.09.006

Michael, G.B., Kadlec, K., Sweeney, M.T., Brzuszkiewicz, E., Liesegang, H., Daniel, R., Murray, R.W., Watts, J.L., Schwarz, S., 2012. ICEPmu1, an integrative conjugative element (ICE) of pasteurella multocida: Structure and transfer. J. Antimicrob. Chemother. 67, 91–100. https://doi.org/10.1093/jac/dkr411

Mira, A., Martín-Cuadrado, A.B., D'Auria, G., Rodríguez-Valera, F., 2010. The bacterial pan-genome: A new paradigm in microbiology. Int. Microbiol.

https://doi.org/10.2436/20.1501.01.110
Moustafa, A.M., Bennett, M.D., Edwards, J., Azim, K., Mesaik, M.A., Choudhary, M.I., Pathanasophon, P., Worarach, A., Ali, Q., Abubakar, M., Anjum, R., 2013. Molecular typing of haemorrhagic septicaemia-associated Pasteurella multocida isolates from Pakistan and Thailand using multilocus sequence typing and pulsed-field gel electrophoresis. Res. Vet. Sci. https://doi.org/10.1016/j.rvsc.2013.07.003
Moustafa, A.M., Seemann, T., Gladman, S., Adler, B., Harper, M., Boyce, J.D., Bennett, M.D., 2015. Comparative genomic analysis of Asian haemorrhagic septicaemia-associated strains of Pasteurella multocida identifies more than 90 haemorrhagic septicaemia-specific genes. PLoS One 10, 1–16. https://doi.org/10.1371/journal.pone.0130296
Muhairwa, A.P., Christensen, J.P., Bisgaard, M., 2000. Investigations on the carrier rate of Pasteurella multocida in healthy commercial poultry flocks and flocks affected by fowl cholera. Avian Pathol. https://doi.org/10.1080/03079450094162
Mühldorfer, K., Schwarz, S., Fickel, J., Wibbelt, G., Speck, S., 2011. Genetic diversity of Pasteurella species isolated from European vespertilionid bats. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2010.10.002
Mutters, R., Ihm, P., Pohl, S., Frederiksen, W., Mannheim, W., 1985. Reclassification of the Genus Pasteurella on the Basis of DNA Homology with Proposals for the New Species Pasteurella-Dagmatis New-Species and Pasteurella-Canis Pasteurella-Stomatis New-Species Pasteurella-Anatis New-Species Pasteurella-Langaa New-Species. Int. J. Syst. Bacteriol.
Nakagawa, M., Nakayama, K., Saito, M., Takayama, S., Watarai, S., Farm, F., 1986. Bacteriological and multocida Serological Infection Studies on Pasteurella in Rabbits animals and occasionally from the conjunctiva , lung and heart . rapid increase of the infection rate was observed at two to three months of age . Serum by use of the tu 35, 463–469.
O'Brien, S.J., Stanyon, R., 1999. Ancestral primate viewed. Nature. https://doi.org/10.1038/46450
Oh, Y.-H., Moon, D.-C., Lee, Y.J., Hyun, B.-H., Lim, S.-K., 2018. Antimicrobial resistance of *Pasteurella multocida* strains isolated from pigs between 2010 and 2016. Vet. Rec. Open 5, e000293. https://doi.org/10.1136/vetreco-2018-000293
Okay, S., Kurt Kızıldoğan, A., 2015. Comparative genome analysis of five Pasteurella multocida strains to decipher the diversification in pathogenicity and host specialization. Gene 567, 58–72. https://doi.org/10.1016/j.gene.2015.04.063
Pasteur, L., 1881. Sur les virus-vaccins du choléra des poules et du charbon. CR Trav Congr Int Dir Agron. Sess. Versailles 161–162.
Pasteur, L., 1880. De l'att´enuation du virus du chol´era des poules. C R Acad Sci 9, 673–680.

Paustian, M.L., May, B.J., Cao, D., Boley, D., 2002. Transcriptional Response of Pasteurella multocida to Defined Iron Sources 184, 6714–6720. https://doi.org/10.1128/JB.184.23.6714

Pearson, T., Okinaka, R.T., Foster, J.T., Keim, P., 2009. Phylogenetic understanding of clonal populations in an era of whole genome sequencing. Infect. Genet. Evol. https://doi.org/10.1016/j.meegid.2009.05.014

Peng, Z., Liang, W., Wang, F., Xu, Z., Xie, Z., Lian, Z., Hua, L., Zhou, R., Chen, H., Wu, B., 2018a. Genetic and Phylogenetic Characteristics of Pasteurella multocida Isolates From Different Host Species. Front. Microbiol. 9, 1–13. https://doi.org/10.3389/fmicb.2018.01408

Peng, Z., Liang, W., Wang, Y., Liu, W., Zhang, H., Yu, T., Zhang, A., Chen, H., Wu, B., 2017. Experimental pathogenicity and complete genome characterization of a pig origin Pasteurella multocida serogroup F isolate HN07. Vet. Microbiol. 198, 23–33. https://doi.org/10.1016/j.vetmic.2016.11.028

Peng, Z., Wang, H., Liang, W., Chen, Y., Tang, X., Chen, H., Wu, B., 2018b. A capsule/lipopolysaccharide/MLST genotype D/L6/ST11 of Pasteurella multocida is likely to be strongly associated with swine respiratory disease in China. Arch. Microbiol. https://doi.org/10.1007/s00203-017-1421-y

Petersen, A., Bisgaard, M., Townsend, K., Christensen, H., 2014. MLST typing of Pasteurella multocida associated with haemorrhagic septicaemia and development of a real-time PCR specific for haemorrhagic septicaemia associated isolates. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2014.02.022

Pijoan, C., Morrison, R.B., Hilley, H.D., 1983. Serotyping of Pasteurella multocida isolated from swine lungs collected at slaughter. J. Clin. Microbiol.

Pors, S.E., Hansen, M.S., Bisgaard, M., Jensen, H.E., 2011a. Occurrence and associated lesions of Pasteurella multocida in porcine bronchopneumonia. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2011.01.005

Pors, S.E., Hansen, M.S., Christensen, H., Jensen, H.E., Petersen, A., Bisgaard, M., 2011b. Genetic diversity and associated pathology of Pasteurella multocida isolated from porcine pneumonia. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2011.02.050

Price, S.B., Freeman, M.D., MacEwen, M.W., 1993. Molecular analysis of a cryptic plasmid isolated from avian strains of Pasteurella multocida. Vet. Microbiol. https://doi.org/10.1016/0378-1135(93)90180-F

Rhoades, K.R., Rimler, R.B., 1990. Pasteurella multocida Colonization and Invasion in Experimentally Exposed Turkey Poults. Avian Dis. https://doi.org/10.2307/1591423

Rigobelo, E.C., Joseph Blackall, P., Pariz Maluta, R., Antonio de Ávila, F., 2013. Identification and

antimicrobial susceptibility patterns of Pasteurella multocida isolated from chickens and Japanese quails in Brazil. Brazilian J. Microbiol. https://doi.org/10.1590/S1517-83822013000100023

Rimler, R.B., Wilson, M.A., 1994. Re-examination of Pasteurella multocida serotypes that caused haemorrhagic septicaemia in North America. Vet. Rec. https://doi.org/10.1136/vr.134.10.256

Rokas, A., Williams, B.I., King, N., Carroll, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature. https://doi.org/10.1038/nature02053

Rosadio, R., Cirilo, E., Manchego, A., Rivera, H., 2011. Respiratory syncytial and parainfluenza type 3 viruses coexisting with Pasteurella multocida and Mannheimia hemolytica in acute pneumonias of neonatal alpacas. Small Rumin. Res. https://doi.org/10.1016/j.smallrumres.2011.02.001

Rosenau, A., Labigne, A., Escande, F., Courcoux, P., Philippon, A., 1991. Plasmid-mediated ROB-1 β-lactamase in Pasteurella multocida from a human specimen. Antimicrob. Agents Chemother. 35, 2419–2422. https://doi.org/10.1128/AAC.35.11.2419

Sakano, T., Taneda, A., Okada, M., Ono, M., Hayashi, Y., Sato, S., 1992. Toxigenic type A Pasteurella multocida as a causative agent of nasal turbinate atrophy in swine. J. Vet. Med. Sci. https://doi.org/10.1292/jvms.54.403

San Millan, A., Escudero, J.A., Gutierrez, B., Hidalgo, L., Garcia, N., Llagostera, M., Dominguez, L., Gonzalez-Zorn, B., 2009. Multiresistance in Pasteurella multocida is mediated by coexistence of small plasmids. Antimicrob. Agents Chemother. 53, 3399–3404. https://doi.org/10.1128/AAC.01522-08

Sarangi, L.N., Priyadarshini, A., Kumar, S., Thomas, P., Gupta, S.K., Nagaleekar, V.K., Singh, V.P., 2014. Virulence genotyping of pasteurella multocida isolated from multiple hosts from India. Sci. World J. 2014. https://doi.org/10.1155/2014/814109

Schwarz, S., Chaslus-Dancla, E., 2001. Use of antimicrobials in veterinary medicine and mechanisms of resistance, in: Veterinary Research. https://doi.org/10.1051/vetres:2001120

Shane, S.M., Stern, N.J., 2003. Diseases of Poultry, Saif, Y. M. Barnes, H. J. Glisson, J. R. Fadly, A. M. McDougald, L. R. Swayne, D. E. https://doi.org/10.1354/vp.41-4-450

Sheppard, S.K., Colles, F., Richardson, J., Cody, A.J., Elson, R., Lawson, A., Brick, G., Meldrum, R., Little, C.L., Owen, R.J., Maiden, M.C.J., McCarthy, N.D., 2010. Host association of campylobacter genotypes transcends geographic variations. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.00124-10

Shirzad Aski, H., Tabatabaei, M., 2016. Occurrence of virulence-associated genes in Pasteurella multocida isolates obtained from different hosts. Microb. Pathog.

https://doi.org/10.1016/j.micpath.2016.04.008
Shivachandra, S.B., Kumar, A.A., Biswas, A., Ramakrishnan, M.A., Singh, V.P., Srivastava, S.K., 2004. Antibiotic Sensitivity Patterns among Indian Strains of Avian Pasteurella multocida. Trop. Anim. Health Prod. https://doi.org/10.1023/B:TROP.0000045950.35070.7f
Shivachandra, S.B., Kumar, A.A., Gautam, R., Saxena, M.K., Chaudhuri, P., Srivastava, S.K., 2005. Detection of multiple strains of Pasteurella multocida in fowl cholera outbreaks by polymerase chain reaction-based typing. Avian Pathol. https://doi.org/10.1080/03079450500367963
Shivachandra, S.B., Viswas, K.N., Kumar, A.A., 2011. A review of hemorrhagic septicemia in cattle and buffalo. Anim. Health Res. Rev. https://doi.org/10.1017/S146625231100003X
Sicheritz-Ponten, T., 2001. A phylogenomic approach to microbial evolution. Nucleic Acids Res. https://doi.org/10.1093/nar/29.2.545
Silver, R.P., Leming, B., Garon, C.F., Hjerpe, C.A., 1979. R-plasmids in Pasteurella multocida. Plasmid. https://doi.org/10.1016/0147-619X(79)90033-7
Smith, E.M., Green, L.E., Medley, G.F., Bird, H.E., Fox, L.K., Schukken, Y.H., Kruze, J. V., Bradley, A.J., Zadoks, R.N., Dowson, C.G., 2005. Multilocus sequence typing of intercontinental bovine Staphylococcus aureus isolates. J. Clin. Microbiol. https://doi.org/10.1128/JCM.43.9.4737-4743.2005
Smyth, D.S., Feil, E.J., Meaney, W.J., Hartigan, P.J., Tollersrud, T., Fitzgerald, J.R., Enright, M.C., Smyth, C.J., 2009. Molecular genetic typing reveals further insights into the diversity of animal-associated Staphylococcus aureus. J. Med. Microbiol. https://doi.org/10.1099/jmm.0.009837-0
Snel, B., Bork, P., Huynen, M.A., 1999. Genome phylogeny based on gene content. Nat. Genet. https://doi.org/10.1038/5052
Sthitmatee, N., Kataoka, Y., Sawada, T., 2010. Molecular epidemiology of Japanese avian Pasteurella multocida strains by the single-enzyme amplified fragment length polymorphism and pulsed-field gel electrophoresis. J. Vet. Med. Sci. https://doi.org/http://dx.doi.org/10.1292/jvms.10-0181
Subaaharan, S., Blackall, L.L., Blackall, P.J., 2010. Development of a multi-locus sequence typing scheme for avian isolates of Pasteurella multocida. Vet. Microbiol. https://doi.org/10.1016/j.vetmic.2010.01.017
Tang, X., Zhao, Z., Hu, J., Wu, B., Cai, X., He, Q., Chen, H., 2009. Isolation, antimicrobial resistance, and virulence genes of Pasteurella multocida strains from swine in China. J. Clin. Microbiol. https://doi.org/10.1128/JCM.02029-08

Tayeb, A.B. El, Morishita, T.Y., Angrick, E.J., 2004. Evaluation of *Pasteurella Multocida* Isolated from Rabbits by Capsular Typing, Somatic Serotyping, and Restriction Endonuclease Analysis. J. Vet. Diagnostic Investig. https://doi.org/10.1177/104063870401600205

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S. V., Crabtree, J., Jones, A.L., Durkin, A.S., DeBoy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J.B., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M., 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: Implications for the microbial "pan-genome." Proc. Natl. Acad. Sci. 102, 13950–13955. https://doi.org/10.1073/pnas.0506758102

Townsend, K.M., Boyce, J.D., Chung, J.Y., Frost, A.J., Adler, B., 2001. Genetic organization of Pasteurella multocida cap loci and development of a multiplex capsular PCR typing system. J. Clin. Microbiol. https://doi.org/10.1128/JCM.39.3.924-929.2001

Treangen, T.J., Rocha, E.P.C., 2011. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. PLoS Genet. https://doi.org/10.1371/journal.pgen.1001284

Tumapa, S., Holden, M.T.G., Vesaratchavest, M., Wuthiekanun, V., Limmathurotsakul, D., Chierakul, W., Feil, E.J., Currie, B.J., Day, N.P.J., Nierman, W.C., Peacock, S.J., 2008. Burkholderia pseudomallei genome plasticity associated with genomic island variation. BMC Genomics. https://doi.org/10.1186/1471-2164-9-190

Vernikos, G., Medini, D., Riley, D.R., Tettelin, H., 2015. Ten years of pan-genome analyses. Curr. Opin. Microbiol. https://doi.org/10.1016/j.mib.2014.11.016

Wang, Y., Zhu, J., Lu, C., Wu, B., Liu, D., Hang, W., Liu, H., Liu, X., 2013a. Evidence of circulation of an epidemic strain of Pasteurella multocida in Jiangsu, China by multi-locus sequence typing (MLST). Infect. Genet. Evol. 20, 34–38. https://doi.org/10.1016/j.meegid.2013.07.027

Wang, Y., Zhu, J., Lu, C., Wu, B., Liu, D., Hang, W., Liu, H., Liu, X., 2013b. Evidence of circulation of an epidemic strain of Pasteurella multocida in Jiangsu, China by multi-locus sequence typing (MLST). Infect. Genet. Evol. https://doi.org/10.1016/j.meegid.2013.07.027

WANG, Z., KONG, L.-C., JIA, B.-Y., LIU, S.-M., JIANG, X.-Y., MA, H.-X., 2017. Aminoglycoside susceptibility of <i>Pasteurella multocida</i> isolates from bovine respiratory infections in China and mutations in ribosomal protein S5 associated with

high-level induced spectinomycin resistance. J. Vet. Med. Sci. 79, 1678–1681. https://doi.org/10.1292/jvms.17-0219

Welch, R.A., Burland, V., Plunkett, G., Redford, P., Roesch, P., Rasko, D., Buckles, E.L., Liou, S.-R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C., Perna, N.T., Mobley, H.L.T., Donnenberg, M.S., Blattner, F.R., 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.252529799

Wiedenbeck, J., Cohan, F.M., 2011. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. FEMS Microbiol. Rev. https://doi.org/10.1111/j.1574-6976.2011.00292.x

Wilkie, I.W., Harper, M., Boyce, J.D., Adler, B., 2012. Pasteurella multocida: Diseases and pathogenesis. Curr. Top. Microbiol. Immunol. 361, 1–22. https://doi.org/10.1007/82-2012-216

Wilson, B.A., Ho, M., 2013. Pasteurella multocida: From Zoonosis to cellular microbiology. Clin. Microbiol. Rev. 26, 631–655. https://doi.org/10.1128/CMR.00024-13

Wozniak, R.A.F., Waldor, M.K., 2010. Integrative and conjugative elements: Mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat. Rev. Microbiol. https://doi.org/10.1038/nrmicro2382

Wright, C.L., Strugnell, R.A., Hodgson, A.L.M., 1997. Characterization of a Pasteurella multocida plasmid and its use to express recombinant proteins in P. multocida. Plasmid 37, 65–79. https://doi.org/10.1006/plas.1996.1276

Wu, J.-R., Shieh, H.K., Shien, J.-H., Gong, S.-R., Chang, P.-C., 2015a. Molecular characterization of plasmids with antimicrobial resistant genes in avian isolates of Pasteurella multocida. Avian Dis. https://doi.org/10.1637/z7035

Wu, J.-R., Shieh, H.K., Shien, J.-H., Gong, S.-R., Chang, P.-C., 2015b. Molecular characterization of plasmids with antimicrobial resistant genes in avian isolates of Pasteurella multocida. Avian Dis. 47, 1384–1392. https://doi.org/10.1637/z7035

Wu, J.-R., Shieh, H.K., Shien, J.-H., Gong, S.-R., Chang, P.-C., 2003. Molecular Characterization of Plasmids with Antimicrobial Resistant Genes in Avian Isolates of Pasteurella multocida. Avian Dis. https://doi.org/10.1637/z7035

Yamashita, A., Sekizuka, T., Kuroda, M., 2014. Characterization of Antimicrobial Resistance Dissemination across Plasmid Communities Classified by Network Analysis. Pathog. (Basel, Switzerland) 3, 356–76. https://doi.org/10.3390/pathogens3020356

Yeh, J.-C., Lo, D.-Y., Chang, S.-K., Chou, C.-C., Kuo, H.-C., 2017. Antimicrobial susceptibility, serotypes and genotypes of Pasteurella multocida isolates associated with swine pneumonia in

Taiwan. Vet. Rec. 181, 323. https://doi.org/10.1136/vr.104023
Yu, C., Sizhu, S., Luo, Q., Xu, X., Fu, L., Zhang, A., 2016. Genome sequencing of a virulent avian Pasteurella multocida strain GX-Pm reveals the candidate genes involved in the pathogenesis. Res. Vet. Sci. 105, 23–27. https://doi.org/10.1016/j.rvsc.2016.01.013
Zautner, A.E., Herrmann, S., Corso, J., Tareen, A.M., Alter, T., Groß, U., 2011. Epidemiological association of different Campylobacter jejuni groups with metabolism-associated genetic markers. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.02403-10
Zhao, Y., Wu, J., Yang, J., Sun, S., Xiao, J., Yu, J., 2012. PGAP: Pan-genomes analysis pipeline. Bioinformatics. https://doi.org/10.1093/bioinformatics/btr655

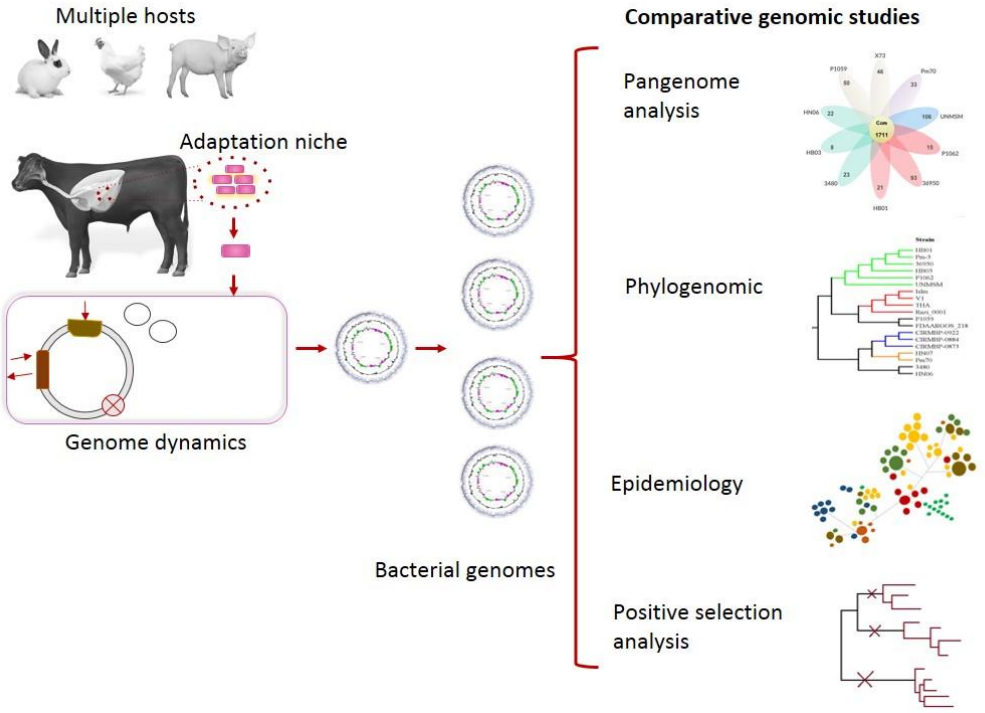


Figure 1. Workflow comparative genomic studies of *P. multocida* isolates. *P. multocida* can infect a wide range of host. The resource of genetic diversity is mediated mainly by point mutations and acquisition of mobile elements as IS, transposons, GEIs and plasmids. An increase of genome sequencing of diverse pathogenic phenotype isolates allows a fine knowledge of diversity and evolutionary adaptation. Studies of comparative genomic are reported including phylogeny construction, pangenome analysis, epidemiology and positive selection.

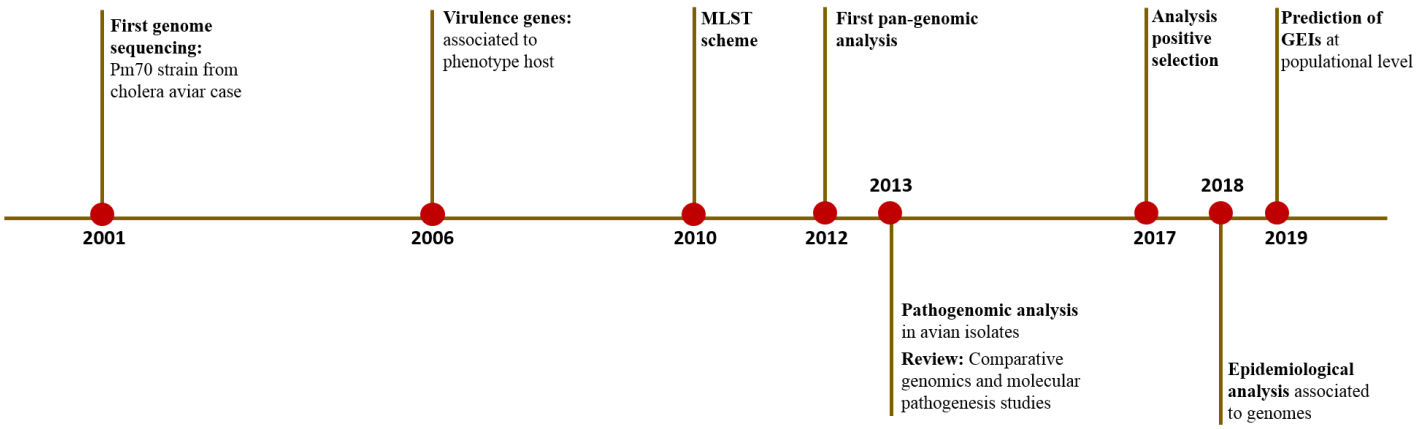


Figure 2. Timeline showing key genomic analyses and pathogenesis studies in the attempt to decipher the genetic diversity and evolution of *P. multocida*.

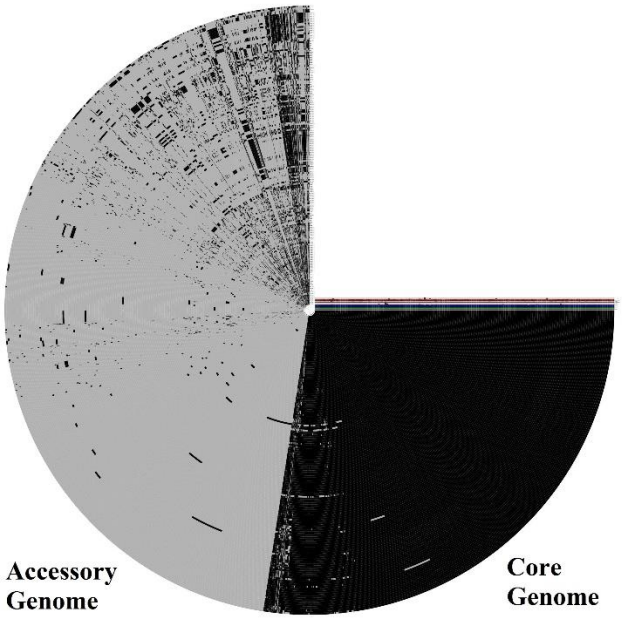


Figure 3. Pan-genome visualization of 179 *P. multocida* strains was obtained using Anvi'o tool. The proteins were predicted with interactive tool. The protein similarity calculation by DIAMOND and clustered with MCL algorithm. The dark bars indicate the occurrence of gene cluster that are organized across genomes by ring. Dark region indicate the core genome and the color variations, as dark and light show accessory genome. The dendrogram represents the hierarchical clustering of genomes based on the occurrence of gene clusters.

Single Nucleotide Polymorphisms (SNPs)
Mobile Genetic Elements (MGEs)

CAPÍTULO 2: RESULTADOS

O trabalho intitulado **"*Draft Genome Sequence of a Virulent Strain of Pasteurella multocida*"** foi publicado no periódico Journal of Genomics (Ivyspring) em junho de 2017, e teve como objetivo a apresentação e submissão do primeiro genoma de *P. multocida* isolado de um caso de uma alpaca com pneumonia na base de dados do *National Center of Biotechnology Information* (*NCBI*). O desenvolvimento desse trabalho nasceu do projeto "Vacinologia reversa para o desenvolvimento de nova geração de vacinas para controle e/ou prevenção de pneumonia em alpacas com pasteurelose", incentivado pelo governo do Peru, que considera de grande relevância econômica combater as diferentes problemáticas na produção de alpacas. Entre as principais causas de mortalidade em neonatos e filhotes de alpacas estão as infeções pneumônicas, nas quais a *P. multocida* é um dos agentes mais frequentemente reportados. Tendo em vista esse cenário, em 2014, um estudo de campo foi realizado na *Universidade del Altiplano* (*UNA*), no estado *de Puno* no Peru, pela equipe de Biologia Molecular e Genética da FMV, UNMSM. O projeto coletou 24 linhagens de *P. multocida* de 46 animais com evidência clínica de pneumonia procedente do centro experimental la Raya. Das análises de diversidade genética baseada na técnica molecular ERIC-PCR e BOX-PCR, reportou a homogeneidade genética em 22 linhagens isoladas do surto. Dentro do grupo clonal, foi selecionada a linhagem de *P. multocida* UNMSM. Dos resultados do sequenciamento e análises genômicas foram identificados importantes genes de virulência e 75 genes relacionados a fagos. Por conseguinte, nos permitiu inferir a natureza patogênica da linhagem de acordo com o caso clínico do qual foi isolada e pela presença de genes de virulência e genes associados a virulência, assim como pela plasticidade genômica da mesma.

IVYSPRING INTERNATIONAL PUBLISHER

Journal of Genomics

2017; 5: 68-70. doi: 10.7150/jgen.19297

Short Research Paper

Draft Genome Sequence of a Virulent Strain of *Pasteurella Multocida* Isolated From Alpaca

Raquel Enma Hurtado[1], Flavia Aburjaile[2], Diego Mariano[2], Marcus Vinicius Canário[2], Leandro Benevides[2], Daniel Antonio Fernandez[1], Nataly Olivia Allasi[1], Rocio Rimac[1], Julio Eduardo Juscamayta[1], Jorge Enrique Maximiliano[1], Raul Hector Rosadio[1], Vasco Azevedo[2] and Lenin Maturrano[1]✉

1. Laboratory of Molecular Biology and Genetics, Veterinary Medicine Faculty, San Marcos University, Lima, Peru;
2. Laboratory of Cellular and Molecular Genetics, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

✉ Corresponding author: Tel.: +551 956533581. E-mail: amaturranoh@unmsm.edu.pe (L. Maturrano)

Received: 2017.01.22; Accepted: 2017.05.14; Published: 2017.06.28

Abstract

Pasteurella multocida is one of the most frequently isolated bacteria in acute pneumonia cases, being responsible for high mortality rates in Peruvian young alpacas, with consequent social and economic costs. Here we report the genome sequence of *P. multocida* strain UNMSM, isolated from the lung of an alpaca diagnosed with pneumonia, in Peru. The genome consists of 2,439,814 base pairs assembled into 82 contigs and 2,252 protein encoding genes, revealing the presence of known virulence-associated genes (*ompH*, *ompA*, *tonB*, *tbpA*, *nanA*, *nanB*, *nanH*, *sodA*, *sodC*, *plpB* and *toxA*). Further analysis could provide insights about bacterial pathogenesis and control strategies of this disease in Peruvian alpacas.

Key words: Alpaca, genome, pasteurellosis, pneumonia.

Introduction

Pasteurella multocida is a commensal bacteria from the upper respiratory tract [1], which affects a wide range of hosts [2, 3]. This bacteria is the primary agent of many infections such as; avian cholera, hemorrhagic septicemia in ungulates, atrophic rhinitis in pigs and snuffles in rabbits [3], and acts as a secondary agent in infectious pneumonia, including cases of acute or chronic pneumonia in different hosts such as swine, calves, sheep, bovine and alpaca [3–5].

In Peru, alpaca raising represents an important economic activity for the High Andean population. However, acute pneumonia causes high mortality rates in young alpacas, in which *P. multocida* has been principally isolated [6]. *P. multocida* has a large number of virulence factors that play a role in pathogenesis, including capsule, lipopolysaccharide, fimbriae, adhesins, toxins, outer membrane proteins, iron regulated and iron acquisition proteins, acquisition proteins, hyaluronidase and sialidase [6]. In this study, we announce the draft genome of *P. multocida* strain UNMSM isolated from an alpaca lung affected with pneumonia.

Pasteurella multocida strain UNMSM is a gram negative, short rod shaped bacteria, oxidase and catalase positive and nonhemolytic, with approximate measures of 0.3 – 0.6 µm in width and 0.8 – 2.0 µm in length (Figure 1). Genome sequencing was performed using Illumina Hiseq sequencing platform. The paired-end library contained inserts of an average size of 100 bp. *De novo* assembly was performed using Edena v3.131028 and SIMBA v1.4 software [7], which produced 82 contigs, with a N50 value of 70,838, 2.4 Mb of size and mean depth coverage ~400-fold. The genome was annotated using the Rapid Annotations using Subsystems Technology (RAST) [8], following by manual curation of the predicted CDSs (Coding Sequences). The genome presents GC content around 40.2%. A total of 2,434

genes were predicted, of which 2,252 were protein-coding genes, eight rRNA genes (four 16S rRNA, one 23S rRNA and three 5S rRNA genes) and 55 were tRNA genes. In table 1, a few summary statistics of the genome are presented.

Table 1. Genome statistics of *P. multocida* strain UNMSM.

Attribute	Value	% of Total*
Genome size (bp)	2,439,814	100.00%
DNA coding region(bp)	2,066,315	84.69%
DNA G+C content (bp)	979,694	40.15%
Total genes	2,434	100.00%
Pseudogenes	115	4.72%
Genes assigned to COGs	1,921	78.92%
Genes with Pfam domains	1,183	48.60%
Genes with signal peptides	173	7.10%
Genes with transmembrane helices	490	20.13%
CRISPR repeats	1	-

* The total is based on either the size of the genome in base pairs or the total number of genes in the annotated genome.

The Cluster of Orthologous Genes (COG) [8] assignments were done using CGView Comparison Tool (CCT), and the presence of protein domains was predicted by Pfam. The signal peptides were identified with the SignalP 4.0 software [9], transmembrane helices were classified by method of Krogh and collaborators method [10], and the CRISPR motif was detected with a web tool described by Grissa and collaborators [11] (Table 1). The functional analysis obtained from Rapid Annotation using Subsystem Technology (RAST) revealed 395 collections of functionally related protein families (Figure 2). The genome has 54 genes responsible for iron acquisition and metabolism; 42 genes involved in virulence, disease and defense; 74 genes involved in membrane transport, and 75 genes related to phages, prophages and transposable elements. Analysis of genes showed that this strain belongs to capsular type A, and the presence virulence associated genes, including outer membrane proteins as *ompH* (UR07_03150), and *ompA* (UR07_05380); hyaluronidases as *nanA* (UR07_06210), *nanB* (UR07_07035) and *nanH* (UR07_07095); iron acquisition related genes, *tonB* (UR07_07610) and *tbpA* (UR07_03090); genes involved in sialic acid metabolism as *sodA* (UR07_00380) and *sodC* (UR07_00680); a lipoprotein, *plpB* (UR07_09985) and *toxA* (UR07_03905), a dermonecrotic toxin [5–12].

This work provides a better understanding of *P. multocida* UNMSM, the first genome isolated from alpaca. Future studies *in vitro* and *in vivo* associated with virulence genes predicted in *P. multocida*, will improve understanding of their pathogenicity and provide an assessment of new targets for the use of vaccines and drugs, which may be important for the control of pneumonia case numbers in alpacas. The genome projects of *P. multocida* strain UNMSM have been deposited in GenBank under the following accession numbers, LGRE00000000.

Acknowledgements

This work was conducted by Laboratory of Cellular and Molecular Genetics in Federal University of Minas Gerais and Veterinary Medicine Faculty in National University of San Marcos and present work was financially supported by the Programa Nacional de Innovación para la Competitividad y Productividad (Innóvate Perú), under the contract 133-FINCyT-IB-2013, as part of the project "Reverse Vaccinology: Developing a new generation vaccine for the control and/or prevention pasteurellosis pneumonia in alpacas".

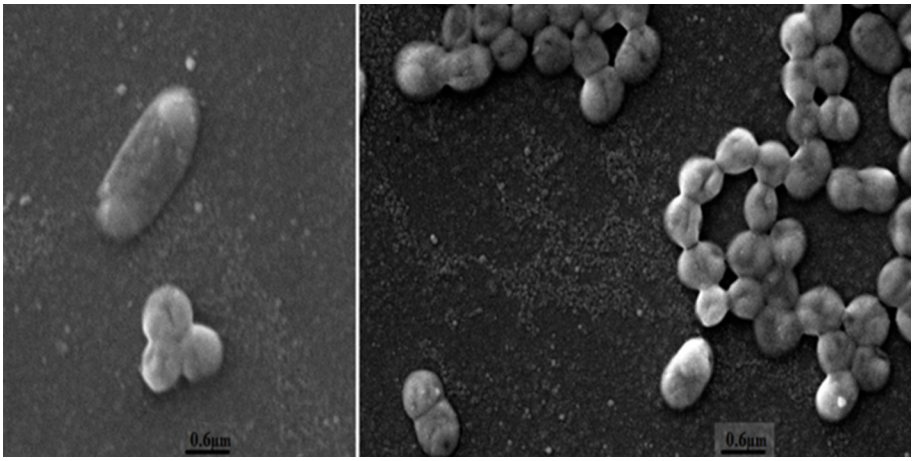


Figure 1. Scanning electron micrograph of cells of *P. multocida* strain UNMSM. The scale bar represents 0.6 μm.

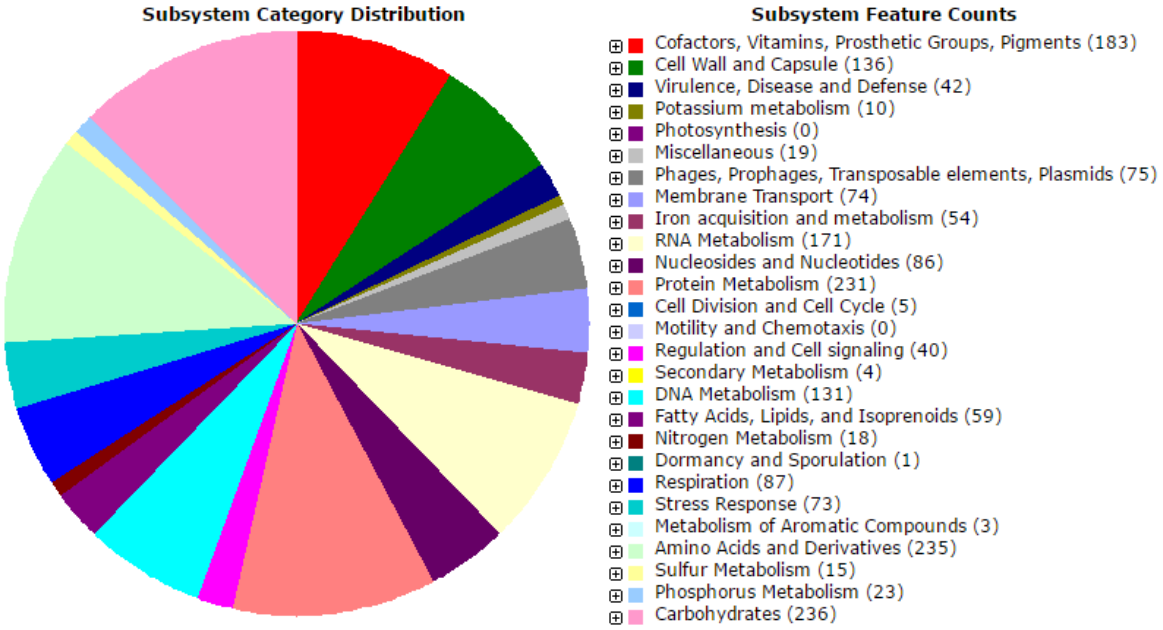


Figure 2. The overview of subsystem category coverage *P. multocida* UNMSM genome based on RAST server.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Timoney JF. Guillespie JA, Scott FW, Barlough JE. The genus Pasteurella. In. Hagan and Bruner's Microbiology and Infectious Diseases of Domestic Animals, 8th ed. Ithaca: Cornell University Press; 1989: 104–116.
2. Christensen H, Bisgaard M. The genus Pasteurella. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, eds. The Prokaryotes, 3rd ed. Berlin: Springer New York; 2006: 1062–1090.
3. Wilkie IW, Harper M, Boyce JD, Adler B. Pasteurella multocida: Diseases and Pathogenesis. In: Aktories K, Orth JHC, Adler B, eds. Pasteurella Multocida. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012: 1–22.
4. Rosadio R, Cirilo E, Manchego A, Rivera H. Respiratory syncytial and parainfluenza type 3 viruses coexisting with Pasteurella multocida and Mannheimia hemolytic in acute pneumonia of neonatal alpacas. Small Rumin. Res. 2011; 97: 110–6.
5. Ameghino E, De Martíni J. Mortalidad de crías de alpacas. Bol. Divulg. Inst. Vet. Investig. Trop. Altura IVITA UNMSM. 1991;71–80.
6. Ramírez A. Enfermedades infecciosas en camélidos sudamericanos. Lima, Perú: XII Reunión científica anual APPA; 1989.
7. Hernandez D, Francois P, Farinelli L, Osteras M, Schrenzel J. De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. Genome Res. 2008; 18: 802–9.
8. Grant JR, Arantes AS, Stothard P. Comparing thousands of circular genomes using the CGView Comparison Tool. BMC Genomics. 2012;13: 202.
9. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods. 2011; 8: 785–6.
10. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 2001; 305: 567–80.
11. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007; 35: W52–7.
12. Katsuda K, Hoshinoo K, Ueno Y, Kohmoto M, Mikami O. Virulence genes and antimicrobial susceptibility in Pasteurella multocida isolates from calves. Vet. Microbiol. 2013; 167: 737–41.

O artigo científico intitulado **"*Pan-genomic approach shows insight of genetic divergence and pathogenic adaptation of Pasteurella multocida*"** foi publicado no periódico *GENE* (*Elsevier*) em maio de 2018. O material suplementar encontra-se na seção **"ANEXOS: material complementar"** desse manuscrito.

P. multocida é um patógeno veterinário capaz de causar diversas doenças em um amplo espectro de animais de produção (bovinos, porcos, aves, alpacas e coelhos). As prévias análises genômicas comparativas e epidemiológicas mostraram a presença de elementos genéticos móveis específicos e grupos genéticos associados à capacidade de causar uma doença específica ou infectar um hospedeiro. Nesse contexto, surgiu a idéia de realizar análises pangenômicas, predição de elementos genéticos móveis e de genes submetidos à pressão seletiva, que possam estar envolvidos na diversidade intraespecífica e na adaptação patogênica da espécie *P. multocida*. Para isso, realizou-se um estudo em 23 linhagens provenientes de isolados veterinários com genomas disponíveis no NCBI. As análises pangenômicas mostraram que essa espécie possui um pangenoma aberto com um constante e gradual aumento de genes. Das análises funcionais do genoma acessório e predição de ilhas genômicas exclusivas por grupo fenotípico, resultou na composição em maior porcentagem de genes relacionados ao metabolismo e transporte de carboidratos e íon inorgânico, biogêneses da membrana e da parede celular, transcrição, replicação, recombinação e reparo, além de proteínas de função desconhecida. Genes associados a estas categorias funcionais, estão envolvidas nos mecanismos de adaptação entre as linhagens. A contribuição deste trabalho permitiu conhecer os recursos genômicos que promovem a adaptabilidade aos diferentes nichos e patogêneses. As análises de seleção positiva por grupo filogenético encontraram a pressão seletiva em dois genes no subgrupo sorotipo B, mas sem relação funcional com o grupo fenotípico de procedência. Baseado nestes resultados, é de suma importância conhecer as características do conteúdo genômico que contribuem no tropismo e diversificação desse microrganismo, para a busca e melhoria das estratégias de prevenção e controle.

Gene 670 (2018) 193–206

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Pan-genomic approach shows insight of genetic divergence and pathogenic-adaptation of *Pasteurella multocida*

Raquel Hurtado[a], Dennis Carhuaricra[b], Siomar Soares[c], Marcus Vinicius Canário Viana[a], Vasco Azevedo[a], Lenin Maturrano[b,1], Flávia Aburjaile[a,d,*,1]

[a] Laboratory of Cellular and Molecular Genetics, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[b] Laboratory of Molecular Biology and Genetics, Veterinary Medicine Faculty, San Marcos University, Lima, Peru
[c] Department of Microbiology, Immunology and Parasitology, Institute of Biological and Natural Sciences, Federal University of Triângulo Mineiro, Uberaba, Minas Gerais, Brazil
[d] Laboratory of Plant Genetics and Biotechnology, Federal University of Pernambuco, Recife, Pernambuco, Brazil

ARTICLE INFO

Keywords:
Genes
Genomic islands
Pan-genome analysis
Virulence factors

ABSTRACT

Pasteurella multocida is a gram-negative, non-motile bacterial pathogen, which is associated with chronic and acute infections as snuffles, pneumonia, atrophic rhinitis, fowl cholera and hemorrhagic septicemia. These diseases affect a wide range of domestic animals, leading to significant morbidity and mortality and causing significant economic losses worldwide. Due to the interest in deciphering the genetic diversity and process adaptive between *P. multocida* strains, this work aimed was to perform a pan-genome analysis to evidence horizontal gene transfer and positive selection among 23 *P. multocida* strains isolated from distinct diseases and hosts. The results revealed an open pan-genome containing 3585 genes and an accessory genome presenting 1200 genes. The phylogenomic analysis based on the presence/absence of genes and islands exhibit high levels of plasticity, which reflects a high intraspecific diversity and a possible adaptive mechanism responsible for the specific disease manifestation between the established groups (pneumonia, fowl cholera, hemorrhagic septicemia and snuffles). Additionally, we identified differences in accessory genes among groups, which are involved in sugar metabolism and transport systems, virulence-related genes and a high concentration of hypothetical proteins. However, there was no specific indispensable functional mechanism to decisively correlate the presence of genes and their adaptation to a specific host/disease. Also, positive selection was found only for two genes from sub-group hemorrhagic septicemia, serotype B. This comprehensive comparative genome analysis will provide new insights of horizontal gene transfers that play an essential role in the diversification and adaptation mechanism into *P. multocida* species to a specific disease.

1. Introduction

Pasteurella multocida is a gram-negative bacteria, commensal and opportunistic pathogen. It causes various diseases in a wide range of hosts, including fowl cholera; hemorrhagic septicemia in ungulates; atrophic rhinitis in pigs; snuffles in rabbits; pneumonia in pig, cattle, sheep, and alpaca; and, it is also associated with infection in humans transmitted from cats and dog bites. These diseases in animals are highly crucial for the animal production and lead to significant economic losses worldwide (Wilkie et al., 2012; Hurtado et al., 2017).

P. multocida has been classified into five serogroups capsules (A, B, D, E, and F), due to differences in capsular antigens, and 16 Heddleston serotypes based on lipopolysaccharide (LPS) antigens (Carter, 1955; Harper et al., 2015).

Studies of *P. multocida* attempting to elucidate its pathogenesis have described several virulence factors that include capsule,

Abbreviations: BRIG, BLAST Ring Image Generator; CDSs, Coding sequences; COG, Clusters of Orthologous Groups database; GC-content, Guanine-cytosine content; GEIs, Genomic islands; GIPSy, Genomic Island Prediction Software; HGT, Horizontal gene transfer; HS, Hemorrhagic septicemia; HTH, Helix–turn–helix motif; Kb, Kilobase(s); KDO, 3-deoxy-D-manno-oct-2-ulosonic acid; ICEPmu1, Integrative conjugative element (ICE) of *Pasteurella multocida*; LPS, Lipopolysaccharide; Mb, Megabase(s); MCL, Markov Clustering Algorithm; MIs, Metabolic islands; NCBI, National Center for Biotechnology Information Database; NSS, Neighbor Similarity Score; PAIs, Pathogenicity islands; PHASTER, PHAge Search Tool Enhanced Release; PHI, Pairwise Homoplasy Index; pident, Percent identity; PTS, Phosphoenolpyruvate transferase system; RIs, Resistance islands; RNA, Ribonucleic acid; rRNA, ribosomal RNA; SIs, Symbiotic islands; SNPs, Single Nucleotide Polymorphisms; VFDB, Virulence Factor Database; VFs, Virulence factors; TCA, Tricarboxylic acid

* Corresponding author at: Laboratory of Cellular and Molecular Genetics, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.
E-mail address: faburjaile@gmail.com (F. Aburjaile).
[1] Contributed equally.

https://doi.org/10.1016/j.gene.2018.05.084
Received 9 March 2018; Received in revised form 6 May 2018; Accepted 21 May 2018
Available online 24 May 2018
0378-1119/

Table 1
General information of 23 *P. multocida* strains.

N°	Strain Name	Host	Disease	Serogroup/Serotype	Accession	Complete (C)/Draft (# de contigs)	Prokka annotation		
							Size (bp)	CDS	Genes
1	3480	Swine	Pneumonia	A	PRJNA32177	C	2,378,127	2216	2272
2	36,950	Bovine	Pneumonia	A:3	PRJNA70829	C	2,349,518	2152	2207
3	HB01	Bovine	Pneumonia	A	PRJNA210009	C	2,416,068	2260	2317
4	HB03	Swine	Pneumonia	A	PRJNA83131	C	2,307,684	2107	2162
5	HN06	Swine	Atrophic Rhinitis	D	PRJNA80883	C	2,402,218	2252	2309
6	P1059	Avian	Fowl cholera	A:3	PRJNA152827	C	2,307,912	2119	2170
7	P1062	Bovine	Pneumonia	A:3	PRJNA201507	C	2,699,012	2412	2467
8	Pm70	Avian	Avirulent	F:3	PRJNA39	C	2,257,487	2031	2089
9	UNMSM	Alpaca	Pneumonia	A:6	PRJNA276484	82	2,439,814	2401	2458
10	X73	Avian	Fowl cholera	A:1	PRJNA155371	C	2,266,799	2078	2130
11	Pm-3	Bovine	Pneumonia	A	PRJNA312874	C	2,415,558	2289	2346
12	Razi_0001	Bovine	Hemorrhagic septicemia	B	PRJNA352138	C	2,360,663	2230	2292
13	HN07	Swine	Pneumonia	F	PRJNA231703	C	2,325,026	2120	2175
14	CIRMBP-0884	Rabbit	Snuffles	A	PRJNA355236	C	2,137,782	2391	2446
15	FDAARGOS_218	Turkey	Fowl Cholera	A:3	PRJNA231221	C	2,331,741	2137	2194
16	FDAARGOS_216	Avian	Fowl Cholera	A:1	PRJNA231221	C	2,295,161	2093	2151
17	PMTB2.1	Buffalo	Hemorrhagic septicemia	B:2	PRJNA172191	8	2,315,202	2134	2191
18	THA	Buffalo	Hemorrhagic septicemia	B:2	PRJNA244069	29	2,299,006	2131	2184
19	Islm	Buffalo	Hemorrhagic septicemia	B:2	PRJNA244093	32	2,349,312	2187	2240
20	V1	Buffalo	Hemorrhagic septicemia	B:2	PRJNA244097	27	2,348,747	2187	2240
21	CIRMBP-0922	Rabbit	Snuffles	A	PRJNA355236	14	2,275,582	2121	2171
22	Razi0002	Avian	Fowl cholera	A:1	PRJNA360234	C	2,289,036	2193	2251
23	CIRMBP-0873	Rabbit	Snuffles	A	PRJNA355236	C	2,418,212	2393	2454

lipopolysaccharide, fimbriae and adhesins, toxins, iron-regulated and iron acquisition proteins, sialic acid metabolism proteins, hyaluronidase and outer membrane proteins (Ewers et al., 2006; Harper et al., 2006). The first complete genome to be analyzed was *P. multocida* Pm70 strain. It showed a gene set involved in virulence, such as iron uptake, TCA cycle, glycolysis, glyconeogenesis, oxidative pentose phosphate and Entner Doudoroff pathways (May et al., 2001).

Phylogenetic analysis based on 16S rRNA genes and core genome of *P. multocida* strains 3480, HN06, P1059, Pm70, 36,950 and ×73 did not find any specific relationship by hosts or disease patterns (Wilson & Ho, 2013). Nevertheless, the comparison of these genomes revealed a unique region (18 kb) that includes the *toxA* gene (the toxin responsible for atrophic rhinitis) and phage-related genes in HN06 strain, also unique regions of 37 kb and 33 kb in strains 3480 and HN06, respectively. Additionally, there are two regions of 4.8 kb and 16 kb, shared among 3480 and HN06 strains. These regions are composed of phage-related genes, whose functions are still to be determined (Wilson & Ho, 2013).

The genetic diversity analysis of *P. multocida* strains Pm70, HB03, HN06, 3480, and 36,950 identified amino acids variation in proteins mainly involved in the outer membrane and distinct metabolic processes, playing a role in the pathogenicity and host adaptation (Okay & Kurt Kizildoğan, 2015). Previous comparative analysis of complete and draft genomes of Pm70, 36,950, ×73, VP161, Anand1P, Anand1G, P903, M1404 and P3480 strains identified accessory regions among them. Strain 36,950 had a region of 92 kb denominated integrative conjugative element (ICEPmu1), which contains 88 genes encoding antibiotic resistance proteins (Boyce et al., 2012; Michael et al., 2012). Within these genes, 35 encoded proteins sharing a significant level of identity with BUKK and TX1 HS strains (Moustafa et al., 2015), and Pm70 strain contains 15 unique genes. Also, analysis based on 7931 SNPs concluded that there is no clear correlation between phylogenetic relatedness and host predilection, disease, country of isolation and serogroup (Boyce et al., 2012).

The comparative analysis of three avian isolated *P. multocida* strains (Pm70, ×73 and P1059) identified 61 genes shared among 2 pathogenic strains (×73 and P1059) and absent in one non-pathogenic strain (Pm70), most of them classified as hypothetical protein and a region containing ten predicted proteins involved in the transport and utilization of L-fructose. In strain P1059, 148 unique genes were predicted, including a region involved in citrate transport and modification. In strain ×73, 127 unique genes were predicted, where five of them were involved in the galactitol-specific phosphotransferase and utilization system; and six genes involved in sugar D-allose transport and utilization systems. Such genomic differences in comparison with additional strains give light in understanding the fitness enhancement in the avian host and its role in adaptation (Christensen et al., 2004).

Comparative genomic analysis of *P. multocida* associated the 12 strains (incomplete genomes) that cause hemorrhagic septicemia (HS) with complete genomes of the strains 36,950 (pneumonia), Pm70 (avirulent), 3480 (pneumonia), HN06 (atrophic rhinitis) and M1404 (HS). They identified the presence of 4 phage regions in the septicemia group, 2 of them being intact phages and several dispersed genes uniquely present in all HS-associated strains. Overall, the HS strains share 96 genes that are absent in the other analyzed genomes, including a capsule biosynthesis locus. An incomplete prophage of 44 kb was shared with only 12 Asian hemorrhagic septicemia strains, and a probable prophage of 50 kb was shared among seven Pakistani strains (Moustafa et al., 2015). Considering that, the presence of different phage elements in different sets of strains may contain important virulence genes and impact on the virulence (Boyce et al., 2012). Also, phylogenetic analysis indicated that HS strains from North America, Pakistani and Thai isolates form clades based on the isolation country (Moustafa et al., 2015).

There are currently 136 complete and draft genomes from *P. multocida* strains isolated from bovine, swine, rabbit, buffalo, alpaca, ovine, human, and dog. Comparative analysis of *P. multocida* strains identified accessory genomes, genomic islands, and positive selection, which could be involved in the phenotypic diversity between strains (*Wilson & Ho, 2013*; *Boyce et al., 2012*; *Moustafa et al., 2015*; *Lefébure & Stanhope, 2007*).

Currently, few studies tried to decipher the genetic diversity among *P. multocida* strains, and none of them strived to correlate their pathogenic phenotypes with their gene repertoires. In this work, we analyzed 23 genomes of *P. multocida* to determine their phylogeny, intraspecific diversity, and to evaluate the correlation between gene content, pathogenicity islands and positive selection as adaptation mechanisms in disease/host specialization.

2. Material and methods

2.1. Obtaining data

The genome sequences (draft and complete) of 23 *P. multocida* strains were downloaded from the National Center for Biotechnology Information Database (NCBI) (www.ncbi.nlm.nih.gov) (Table 1). The software Prokka was used to homogenize and improve the genome annotation (Seemann, 2014).

2.2. Orthologous clustering

All coding sequences (CDSs) from 23 *P. multocida* genomes were subjected to orthologous prediction using OrthoMCL (Fischer et al., 2012). All against all BLASTp comparison was performed to predict orthologous groups using the Markov Clustering Algorithm (MCL). The parameters for the prediction of orthologous genes were defined as 50% sequence identity and 50% coverage with each other. The paralogs were removed from all gene families. The distribution of core, accessory and unique genes, in different strains, were analyzed through the construction of a phylogenetic tree associated with a heatmap using R language (R Development Core Team, 2016) with a script available in GitHub (https://github.com/katholt/plotTree). Additionally, we used an in-house script to identify exclusive genes by a specific group, by considering only genes that were shared by at least two members of a specific group and absented in the other groups (disease) analyzed. The established groups were chosen according to the monophyletic group in the phylogenomic tree (Fig. 1). Furthermore, we performed a functional annotation of exclusive genes by a specific group to better characterize through Gene Ontology and the integration of several databases using GO FEAT (Araujo et al., 2018). A second pan-genome analysis and visualization was performed by Anvi'o v.4 (Eren et al., 2015), which proteins were clustered with MCL algorithm and protein similarity calculation by DIAMOND (Buchfink et al., 2015).

2.3. Pan and core genome analysis

The pan-genome and core genome size and extrapolations were obtained from the models and regression algorithms given by Tettelin (2005). The curve fitting of the pan-genome was performed using a power-law regression based on Heaps' law [$y = A_{pan}x^{Bpan} + C_{pan}$]. As previously described (Tettelin et al., 2005), y is the pan-genome size, x is the genome number and A_{pan}, B_{pan}, and C_{pan} fit parameters. The core, accessory, and pan-genomes were calculated with the software OrthoMCL. The PanGP program (Zhao et al., 2014) was used for pan-genome analysis of 23 *P. multocida* strains, generating the cumulative curve, where B_{pan} is equivalent to the parameter γ, and $\alpha = 1 - \gamma$. When $0 < \gamma < 1$ and $\alpha \leq 1$ the size of the pan-genome increases unboundedly with the sequential addition of new genomes and can be considered open, whereas a $\gamma > 1$ and $\alpha > 1$ is representative of a closed pan-genome, where the addition of news genomes will not increase the pan-genome substantially. Curve fitting of the core-genome was performed using an exponential regression model [$y = A_{core}e^{(Bcore\ x)} + C_{core}$], where, y is the gene number, x is the genome number, e is the Euler number and A_{core}, B_{core} and C_{core} fit parameters.

2.4. Functional classification of the core and accessory genome

The core, accessory and unique genes of 23 *P. multocida* strains were assigned to functional categories. Protein sequences were compared against the Clusters of Orthologous Groups database (Galperin et al., 2015) using rpsblast+ (Altschul et al., 1990) with an e-value of 0.0001. COG categories were assigned for the best hits. Furthermore, complementing the functional annotation was realized through the comparison of protein sequences against KEGG database using BLASTKOALA (Kanehisa et al., 2016).

2.5. Phylogenetic analysis of the P. multocida strains

The phylogenetic relationships of *P. multocida* strains, isolated from

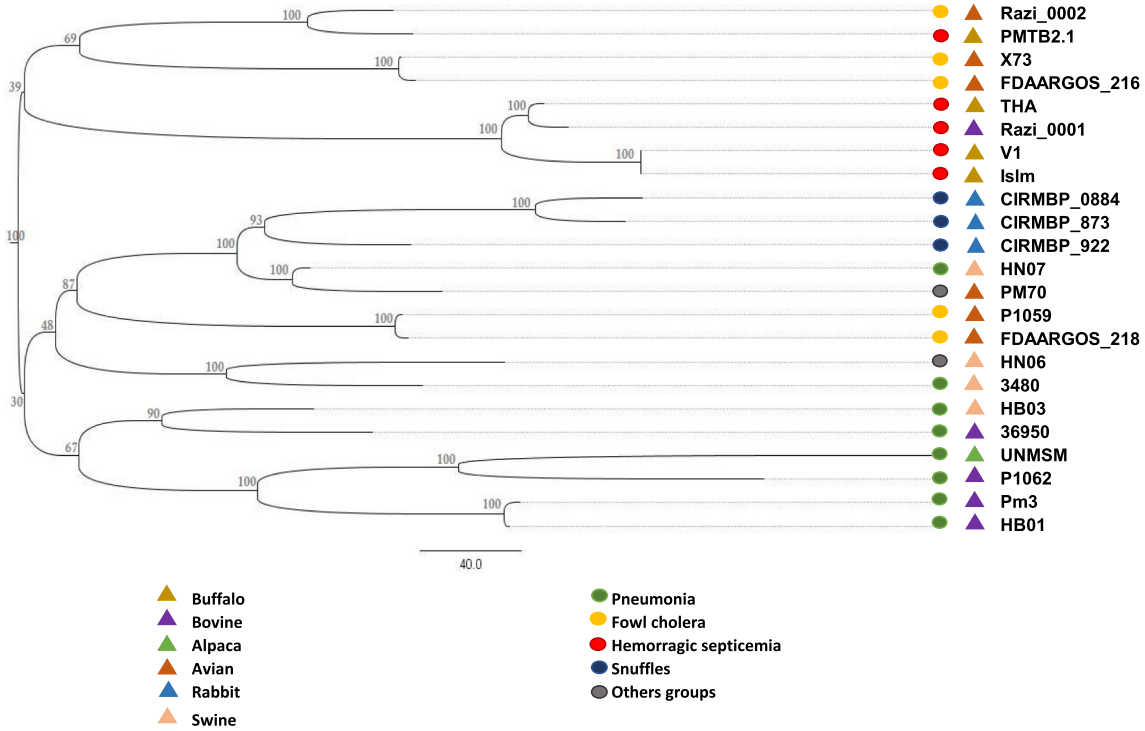


Fig. 1. Phylogenomic analysis constructed using a presence/absence gene matrix of accessory genome from 23 *Pasteurella multocida* strains. The distance was calculated using binary methods and the hierarchical cluster was constructed with a neighbor-joining method based on the Manhattan distance with 1000 bootstrap replicates.

different hosts (bovine, avian, swine, rabbits, buffalo and alpaca) and associated to specific diseases (pneumonia, fowl cholera, hemorrhagic septicemia, atrophic rhinitis and snuffles), were analyzed based in two phylogenomic approaches: core genome and gene content (presence/absence of genes). In the phylogenomic analysis of genes present in the core genome, all the genes containing paralogous sequences were excluded. The amino acid sequences of the core genes were extracted with an *in-house* script, aligned with MUSCLE (Edgar, 2004), and the alignment was trimmed using trimAL (Capella-Gutiérrez et al., 2009). The substitution model was predicted with Prottest (Abascal et al., 2005), and these results were concatenated with FASCONCAT (Kück & Longo, 2014). The tree was constructed with the maximum likelihood method using RaxML (Stamatakis, 2014) with 1000 iterations using the PROTGAMMA substitution model. The resulting tree was visualized using Figtree v 1.4 (Rambaut, 2012). Pan-genome phylogeny was constructed based on the pan-matrix (binary gene presence/absence matrix) using PAST3 (Hammer et al., 2001). The hierarchical cluster was constructed with a neighbor-joining method based on the Manhattan distance with 100 bootstrap replicates.

2.6. Genome plasticity analysis of P. multocida

To improve understanding of role horizontal gene transfer in pathogen adaptation process at specific host and diseases, we performed the prediction of genomic islands and prophages in the 23 *P. multocida* strains. GIPSy (Genomic Island Prediction Software) was used to predict genomic islands (GEIs) (Soares et al., 2016). GEI predictions are based on commonly shared features that are marks of incorporation event: the regions acquired by horizontal transfer were determined by GC Content and codon usage deviation, the presence of transposase, tyrosine recombinase and serine recombinase genes, flanking direct repeats and flanking tRNA genes (Bellanger et al., 2014). Additionally, these islands can be classified into four categories due to its genomic repertoires, such as pathogenicity islands (PAIs), metabolic islands (MIs), resistance islands (RIs) and symbiotic islands (SIs). The presence of prophages regions was predicted using PHASTER (PHAge Search Tool Enhanced Release) (Arndt et al., 2016). For the prediction of genomic islands in the four draft genomes, for each scaffold genome, the contigs were manually assembled. Once the GEIs were predicted, the islands localized in regions linking contigs were eliminated to avoid the prediction of false positives. To find shared GEIs, we performed a local BLAST between their nucleotide sequences using BLASTn. To define homologs GEIs, we considered an identity higher than 60% and used two strategies. First, we performed a comparison between the GEIs of each strain, among members of each group (diseases); and secondly, we compared GEIs between the groups. Finally, GEIs were mapped using BLAST Ring Image Generator (BRIG) (Alikhan et al., 2011). Overall, islands shared by at least two members of a group and absent in the other pathogenic-group analyzed were considered to be specific to the former group.

2.7. Prediction of the accessory genome through Genomic Islands and phages analysis

All the CDSs were extracted from the genomic islands of each genome. Each CDS belonging to the Genomic Islands of each genome was subjected to orthologous prediction using OrthoMCL. A binary gene presence/absence matrix was retrieved from the OrthoMCL resulting data. The distribution of GEI genes in different strains was analyzed through the construction of a phylogenetic tree associated with a heatmap using R language. Furthermore, we performed a functional annotation of genomic islands to better characterize the hypothetical proteins through Gene Ontology and the integration of several databases using GO FEAT (Araujo et al., 2018).

2.8. Virulence factors in core and accessory genome

The presence of virulence genes in core and accessory genome was identified using BLASTp (Altschul et al., 1990) against the Virulence Factor Database (VFDB) (Chen et al., 2004). The parameters considered were an *E*-value of $1e^{-5}$, a minimum identity percentage of 30% and minimum coverage of 70%, between the query and subject sequences. The functional annotations were obtained from the categories provided by the Virulence Factor Database.

2.9. Positive selection analysis

Besides the presence and absence of genes, we searched for sequence variations in *P. multocida* that could explain host adaptation and disease manifestations. Positive selection analysis using branch-site models can identify adaptive mutations that occurred in specific branches of phylogeny in specific amino acid positions (sites) of a protein, i.e., directional selection. The branch to be tested for positive selection is the foreground, while the remaining branches are the background. The test identifies the protein sites in which $\omega > 1$ (positive selection) for the foreground and $\omega < 1$ or $\omega = 1$ (negative selection or neutral evolution) for the background (Yang & Nielsen, 2002; Zhang, 2005). This information identifies the different selective pressures for each branch and can be used to generate hypothesis for functional studies (Yang & Dos Reis, 2011).

The 23 genomes of *P. multocida* were reannotated using RASTtk (Brettin et al., 2015) in PATRIC bacterial Bioinformatics Resource Center (Wattam et al., 2017). The PosiGene pipeline (Sahm et al., 2017) was used to perform a genome-scale positive selection analysis. Strain CIRMBP_0884 (snuffles) was used as the reference and anchor genome. Protein homology was identified by BLASTp bidirectional hits, using the identification of the reference genome to name the orthologous groups. Only protein sequences present in the anchor genome and with at least 50% of identity were analyzed. A species tree required to identify the branches to be tested generated by filtering each alignment using GBLOCKS (Castresana, 2000), phylogeny reconstruction by the maximum parsimony method and jackknifing using DNAPARS from PHYLIP package (Felsenstein, 2005) and calculating a consensus tree using CONSENSE (Felsenstein, 2005). Each ortholog group was realigned at codon level by PRANK (Loytynoja & Goldman, 2008) using the species tree. The last step was the analysis of positive selection using branch site-models present in codeml (Yang, 2007). Only significant results for False Discovery Rate ($p < 0.05$) were considered. The Phipkack was used to perform recombination analysis with the methods Pairwise Homoplasy Index (PHI) (Bruen et al., 2006), Neighbor Similarity Score (NSS) (Jakobsen & Easteal, 1996) and Maximum Chi-Square (Smith, 1992). We combined the results of at least two methods (Posada et al., 2002) and considered the sequences as recombinant when $p < 0.05$ for PHI and at least another test (Hongo et al., 2015).

The foregrounds, required to be monophyletic groups, were chosen according to the species tree (Fig. 2). The snuffles foreground (blue) has strains that caused snuffles in the rabbit. The pneumonia foreground (green) has strains that caused pneumonia in bovines, alpaca and swine. The serotype B subgroup (red) has strains that caused hemorrhagic septicemia in buffalo and bovine and belong to serotype B, what excludes strain PMTB2.1. The serotype F foreground (orange) has the strain that caused swine pneumonia (HN07) and the avirulent strain Pm70, isolated from avian, both from serotype F.

The functions of the positively selected genes were identified using RASTtk annotation, PATRIC's Pathway Summary and BLASTp against InterProScan Database (https://www.ebi.ac.uk/interpro/search/sequence-search).

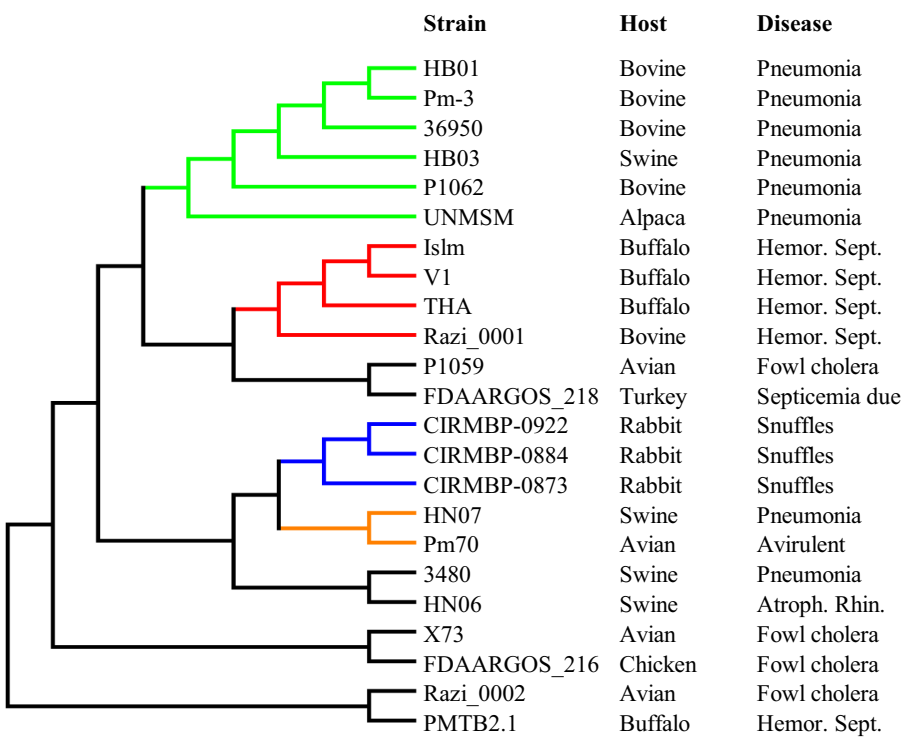


Fig. 2. The phylogenetic tree was reconstructed based on SNPs of core genome of 23 *Pasteurella multocida* strains. The tree was inferred using the maximum parsimony method and jackknifing using DNAPARS. Phylogenetic groups were designated as snuffles foreground (blue), pneumonia foreground (green) and the serotype F foreground (orange). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Results and discussions

3.1. Genome features

In the present work, we performed the pan-genome analysis of 23 *P. multocida* strains isolated from distinct diseases and hosts (Table 1). The sizes of the analyzed genomes are in a range between 2.13 and 2.69 Mb, and the number of coding sequences of the strains varies between the ranges of 2031–2412 (Table 1). All strains were assigned to a group, where each group is related to the disease and host from where they were isolated. The ×73, P1059, Razi_0002, FDAARGOS_216 and FDAARGOS_218 strains are fowl cholera-associated group; the Pm-3, 3480, 36,950, P1062, UNMSM, HB01 and HB03 strains are pneumonia-associated group; the CIRMBP_0873, CIRMBP_0884 and CIRMBP_0992 strains are snuffles-associated group and Razi_0001, V1, Lslm, THA, and PMTB2.1 are a septicemia-associated group. We analyzed the strains in groups for the identification of accessory genes and genomic islands and seek to correlate their presence with their adaptive potential in host specificities and disease manifestations.

3.2. Pan-genome and core-genome profile

The genes distributed in all genomes were denominated as part of the core genome, whereas the ones present in only one strain were denominated as unique genes and the genes present in two or more genomes, but not in all, were denominated accessory genes. In this analysis, we identified a pan-genome of 3585 genes, 1688 (47.08%) of which were shared by all strains, 1200 (33.47%) were part of the accessory genome, and 697 (19.44%) were unique genes. The result shows 52.91% of variable genome content, showing the high genomic diversity of *P. multocida*, as previously reported (Moustafa et al., 2015; Cao et al., 2017). The fitting curve was generated for all 23 genomes (Fig. 3) supported by the formula $\alpha = 1 - \gamma$, were inferred an α of 0.73, results suggest that the *P. multocida* pan-genome is open, whence by each newly added genome, the number of genes will increase significantly. Open pan-genomes are frequently found in bacterial species that are more prone to high horizontal gene transfer (HGT).

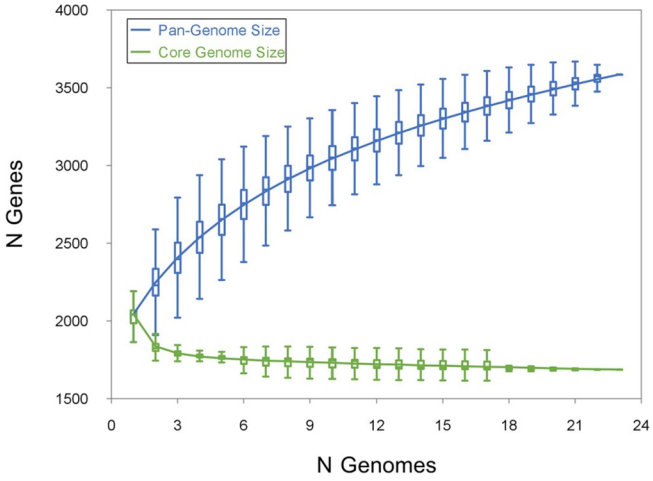


Fig. 3. Core and pan-genome development for 23 *Pasteurella multocida* genomes. The number of genes in the pan (blue) and core (green) are plotted in function of the number of genomes sequentially added. The curve for the pan-genome is the least-squares fit of the function y = AxB + C with the best fit obtained with a correlation r2 = 0.999 for A = 1121.01 ± 3.26, B = 0.27, C = 899.43 ± 7.24. The curve for the core genome is the least-squares fit of the function y = AeBx + C with the best fit obtained with a correlation r2 = 0.935 for A = 1713.94 ± 52.07, B = −0.66, C = 1713.94 ± 1.09. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Functional annotation of genes in the pangenome

The functional annotation of the pan-genome was carried out using the Clusters of Orthologous Groups of proteins (COG) database. The result showed the relative abundance and distribution of the functional categories across the core, accessories and unique genomes of *P. multocida*. Briefly, the unique genes are composed by genes related to metabolism (13.11%), replication, recombination and repair (2.97%), transcription (3.51%), cell wall/membrane/envelope biogenesis (5%) and unknown function (46.35%). The accessory genes are related to

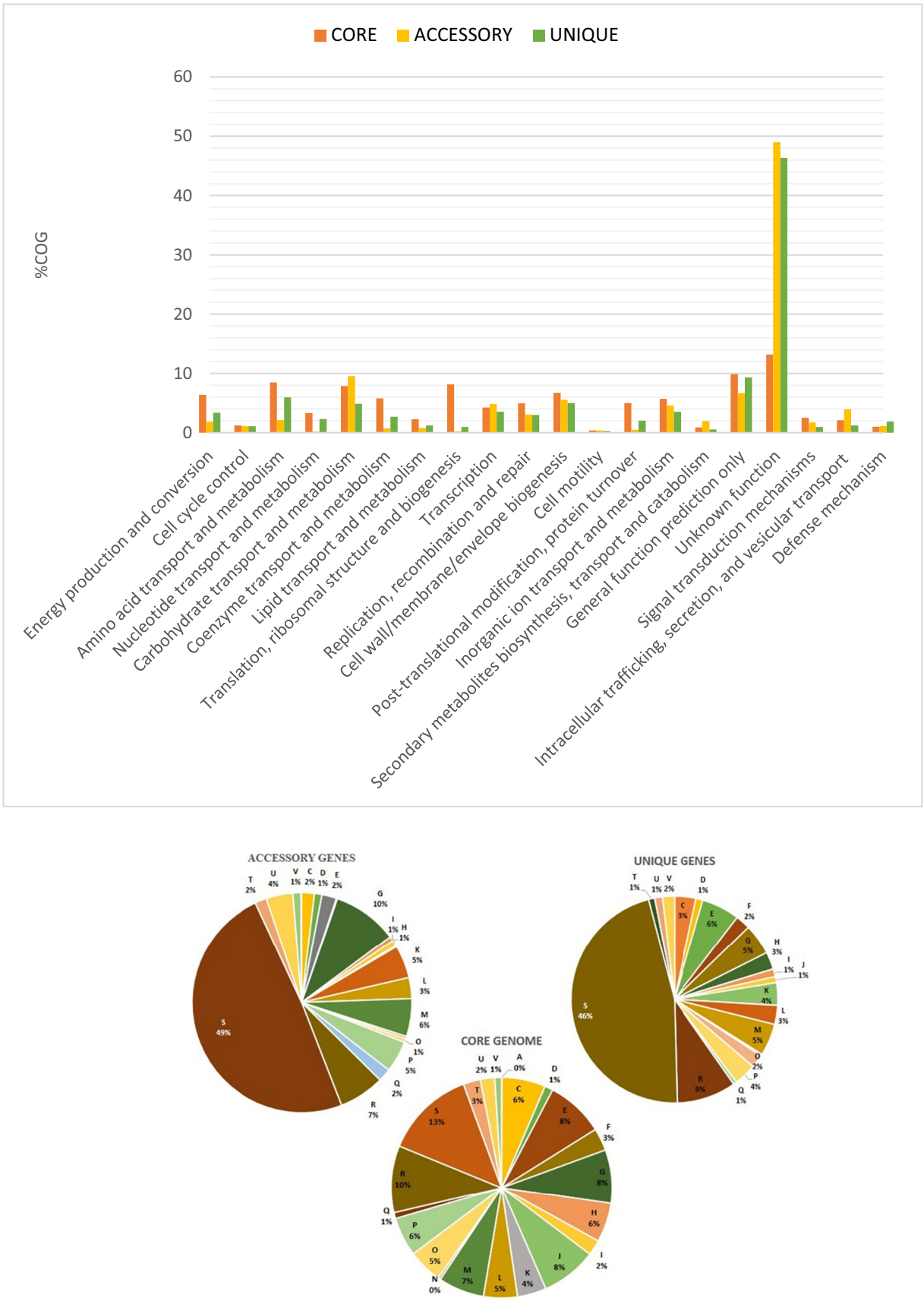


Fig. 4. Relative abundance and distribution of COG categories between the core genome, accessory genome and unique genes of *P. multocida* strains.

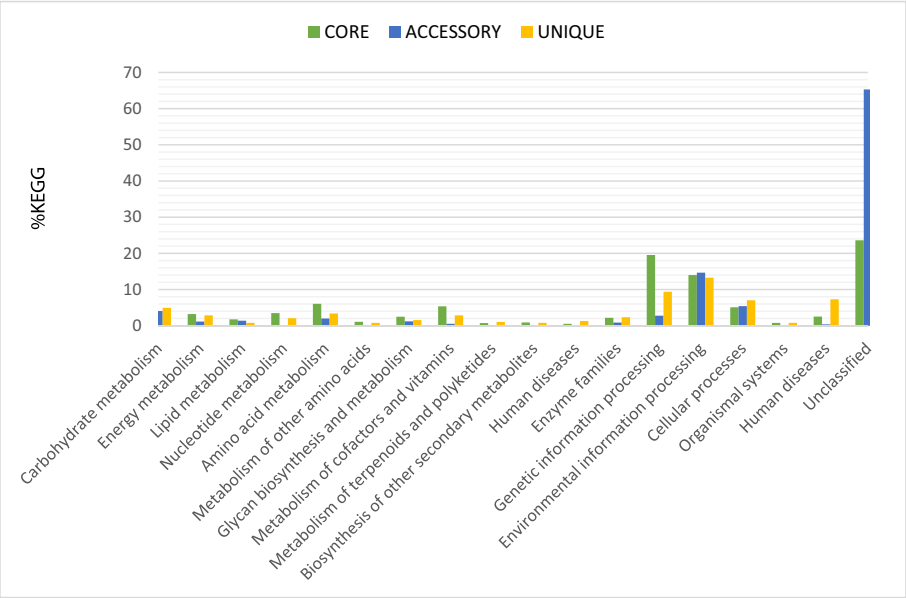


Fig. 5. Distribution of KEGG categories of the *P. multocida* strains in the core-genome, accessory genome, and unique genes.

carbohydrate transport and metabolism (9.54%), cell wall/membrane/envelope biogenesis (5.55%), transcription (4.85%), replication, recombination and repair (3.08%), inorganic ion transport and metabolism (4.6%), and unknown function (49%) (Fig. 4). These genes related to metabolism and replication, and transcription systems, could be related to mechanism adaptations that allow their high invasive spectrum. The core genome present, in its majority, genes assigned to metabolic functions (33.4%). In the core genome, 86% of the genes were assigned to COG function category, whereas 46.35% and 49% of the genes in the unique and accessory gene sets were assigned to the unknown function, suggesting that there is a large number of non-characterized proteins that may be involved in genetic diversification. Futhermore, the KEGG pathway analysis of the core-genomes showed similar result to GO analysis presenting 30.83% of genes related to metabolic function (Fig. 5). Gene categories (Fig. 5) have been described to be involve receptors for sensing environmental factors, transcriptional and translation process to control mechanisms that confer the adaptation to environmental niches (Brooks et al., 2011; Simon et al., 2009).

3.4. Analysis of exclusive genes from established pathogenic group

The exclusiveness of accessory genes in specific host-expresses phenotypes would provide signs of selective advantage of adaptation on the host (Levin & Bergstrom, 2000). Briefly, we identified 127, 79, 18 and 155 CDSs exclusively shared by the groups hemorrhagic septicemia, fowl cholera, snuffles and pneumonia, respectively (S1 Tables S1, S2, S3, and S4). For hemorrhagic septicemia group, we identified 110 hypothetical proteins and virulence-related genes such as TonB-dependent heme receptor A, XRE family transcriptional regulator, ArsR family transcriptional regulator (S1 Table 1). For the avian group, we identified 44 hypothetical proteins and a set of CDSs involved in fructose, galactitol, citrate transport and metabolism, which are described in genomic islands (S1 Table 2). Additionally, we identified proteins associated at virulence as succinyl ornithine transaminase, a protein that encodes a hemoglobin receptor of TonB-dependent family (Prasannavadhana et al., 2014) and two UDP-glucose 6-dehydrogenase, an enzyme involved in the synthesis of different surface structures (Jiang et al., 2010). In the snuffles group, we identified CDSs such as regulator transcriptional glucitol operon repressor and putative HTH-type transcriptional. (S1 Table 3). In the pneumonia group, we identified 126 hypothetical proteins and a putative trehalose utilization locus consisting of a transcriptional regulator, *treR*; a trehalose phosphoenol pyruvate transferase system (PTS) transporter, *treB*; and a trehalose-6-phosphate hydrolase, *treC* (S1 Table 4). Reports indicate that the possession of trehalose biosynthesis pathways has potential to contribute to *Mycobacterium tuberculosis* virulence (Tournu et al., 2013). Also, we identified virulence-related genes as transcriptional regulators, like RCS-specific HTH-type transcriptional activator RclR, HTH-type transcriptional regulator Prtr, Deoxyribose operon repressor, and LexA repressor. We identified an auto-transporter adhesion EhaG, a Trimeric auto-transporter protein, considered an important virulence factor that mediates specific adherence to intestinal epithelial cells, reported in enterohemorrhagic *Escherichia coli* (Totsika et al., 2012). Moreover, also, we identified a mannosylglucosyl-3-phosphoglycerate phosphatase, TonB-dependent receptor, and UDP-glucose 4-epimerase. (S1 Table 4). However still will be elucidating the relation of the presence of them as a mechanism of advantage adaptive in a specific group.

3.5. Virulence factors distributed in the core genome

We identified 172 virulence factors (VFs) in the core genome detailed in the S2 Table 1. Among these genes, some are involved in the cell wall composition. The lipopolysaccharides (LPS) play a prominent role in the pathogenesis of the disease. The structure of LPS for *P. multocida* is formed by a hydrophobic lipid A domain, a non-repeating inner and outer core oligosaccharide (Raetz & Whitfield, 2002). Twenty genes encoding enzymes for LPS biosynthesis were present in all strains.

The conformation of the inner core is composed of residues of 3-deoxy-D-manno-oct-2-ulosonic acid (KDO) and L-*glycerol*-D-*manno*heptose for assembly of the nascent LPS (Raetz & Whitfield, 2002). We identified a cluster of genes coding for enzymes that catalyze the biosynthesis of the inner core, which includes *kdtA*, *kdkA*, *opsX*, *rfaF and waaC* (Raetz & Whitfield, 2002; Johnson et al., 2013). The *kdtA* encodes 3-deoxy-D-manno-octulosonic-acid transferase required for addition of the Kdo molecules to lipid A. The phosphorylation of the first Kdo residue is performed by a kinase, *KdkA*. We observed the presence of a conserved biosynthesis of lipid A, where the first reaction is the acylation of the sugar nucleotide UDP-GlcNAc by LpxA. We also identified the other acyltransferases, like *kdsA*, *lpxB*, *kdsB*, *lpxH*, *lpxK*, *lpxD*, *kpsF*, *lpxL* and *msbB* (Raetz & Whitfield, 2002). The *kdsA* and the *kdsB* serve as the Kdo donors. The transferases and L,

D-heptose precursors are required for the biosynthesis and the transfer of L, D-heptose; among these, there is an isomerase *gmhA*, *rfaD*, and rfaE, phosphatase. The external core structure is different among strains.

Others VFs involved in iron-dependent regulation and acquisition were identified, such as the ferric iron transport-related genes *fecCD*. Also, we identified heme iron utilization and exporter proteins, such as *ccmABCEF* (Harper et al., 2006; Paustian et al., 2002).

The outer membrane proteins enhance *P. multocida* ability to colonize the host and their coding genes have been identified in the core genome, such as *P6* and *ompP5*; lipoprotein, *lolD*, *IcsA*, *metQ* an immunogenic lipoprotein A, two secretory proteins as secretory immunoglobulin A-binding protein EsiB and a protein E of Type II secretion system.

3.6. Phylogenetic analysis of the P. multocida strains

P. multocida strains are adapted to a broad range of hosts with diverse disease manifestations. However, the genetic diversity that there is behind these adaptive mechanisms is not completely known, and the elucidation of their phylogenetic relationship with various strategies could shed light on its pathogenic specialization. The genome of 23 *P. multocida* strains isolated from different hosts and disease were used to estimate their phylogenetic relationship.

The core phylogenomic tree was constructed using 1688 core genes that represent the orthologous genes harbored in a single copy by all the genomes. The strains from the hemorrhagic septicemia group were clustered together and were very distantly related to any other group, except for PMTB2.1 strain, which is closely related to the fowl cholera group (Fig. 6). The avian strains were also clustered in two groups separated except the avirulent strain, Pm70. Pneumonia group formed a closely related group with UNMSM strain more evolutionarily distant and with the strains 3480 and HN07, isolated of swine host, outside the group. The snuffles strains were clustered together. This analysis allows to show evolutionary divergences among strains isolated from different spectrum hosts or diseases but is no sufficient to support the correlation of them, as previous analysis (Boyce et al., 2012).

An alternative strategy of phylogenomic analysis was based on the presence/absence gene matrix of the accessory genome (Fig. 1). A total of 33.47% of genes were identified as accessory genome and could have played a key role in the evolution of *P. multocida*. The phylogeny shows a close relation with phylogeny analysis based on core genes but shows a major variability in the level of genes content on the different groups. The pneumonia cluster has allowed the clustering of strains isolated from pneumonia cases in diverse hosts showing a great variability at the gene level. The avian cholera group is more segregated, but it still tends to cluster in two groups. As mentioned above, hemorrhagic septicemia members are strictly related but profoundly diverge from PMTB2.1 and the other groups. Snuffles group showed a significant variation in genetic content, with major divergence in CIRMBP_0992. The phylogeny based on the accessory genome shows a high diversity at level genes on *P. multocida* strains but no a stronger correlation between gene content, and specificity host and disease manifestation. The clusterization of strains that present a common niche point for an evolutionary convergence, due to the presence of accessory genes acquired by horizontal gene transfer (Lawrence, 1997; Marri et al., 2006). The distribution of accessory genes and pan-genome supports the phylogenetic

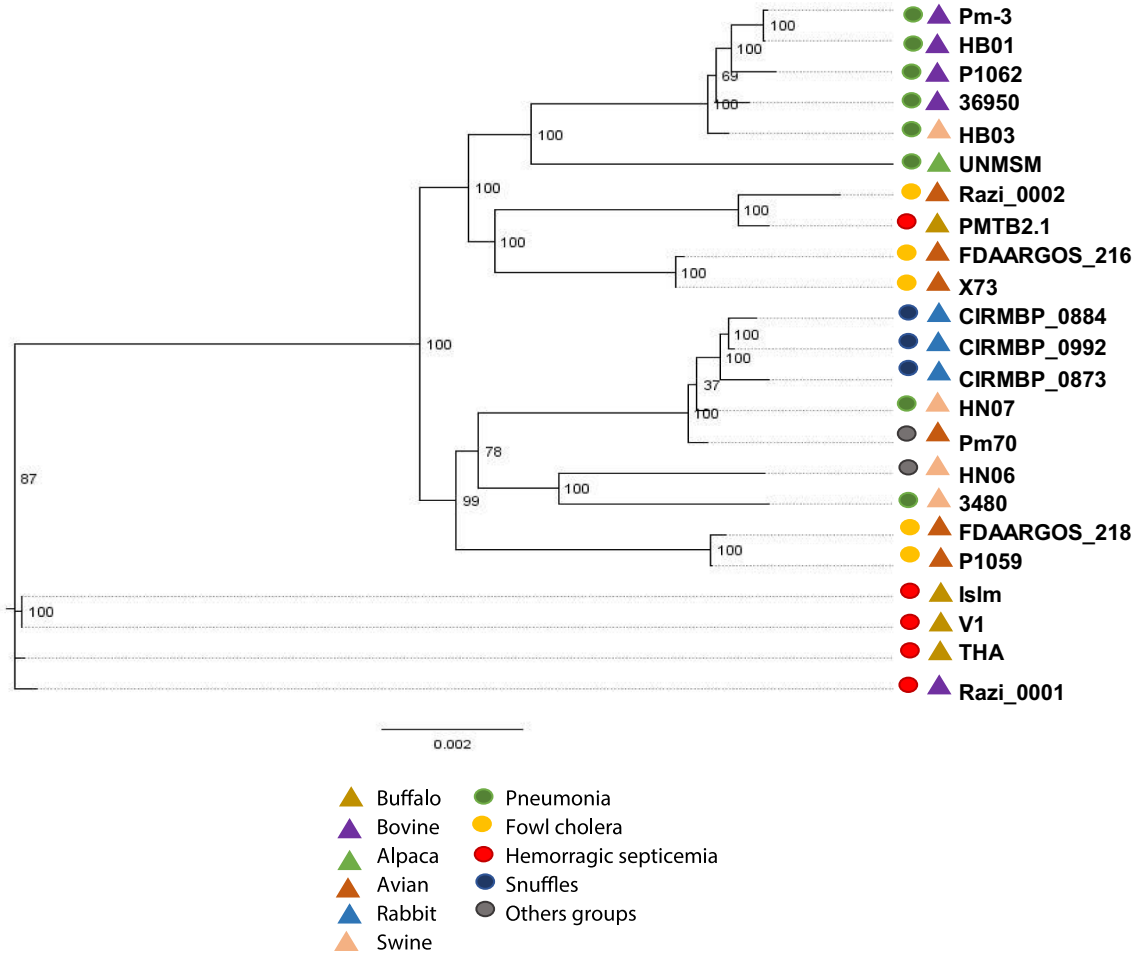


Fig. 6. Phylogenomic tree analysis based on a total of 1688 protein sequences from the core genome of 23 *Pasteurella multocida* strains. The tree was inferred from core genes using the maximum likelihood with a model of protein substitution using RaxML. The percentages for the bootstrap test were 1000 iterations.

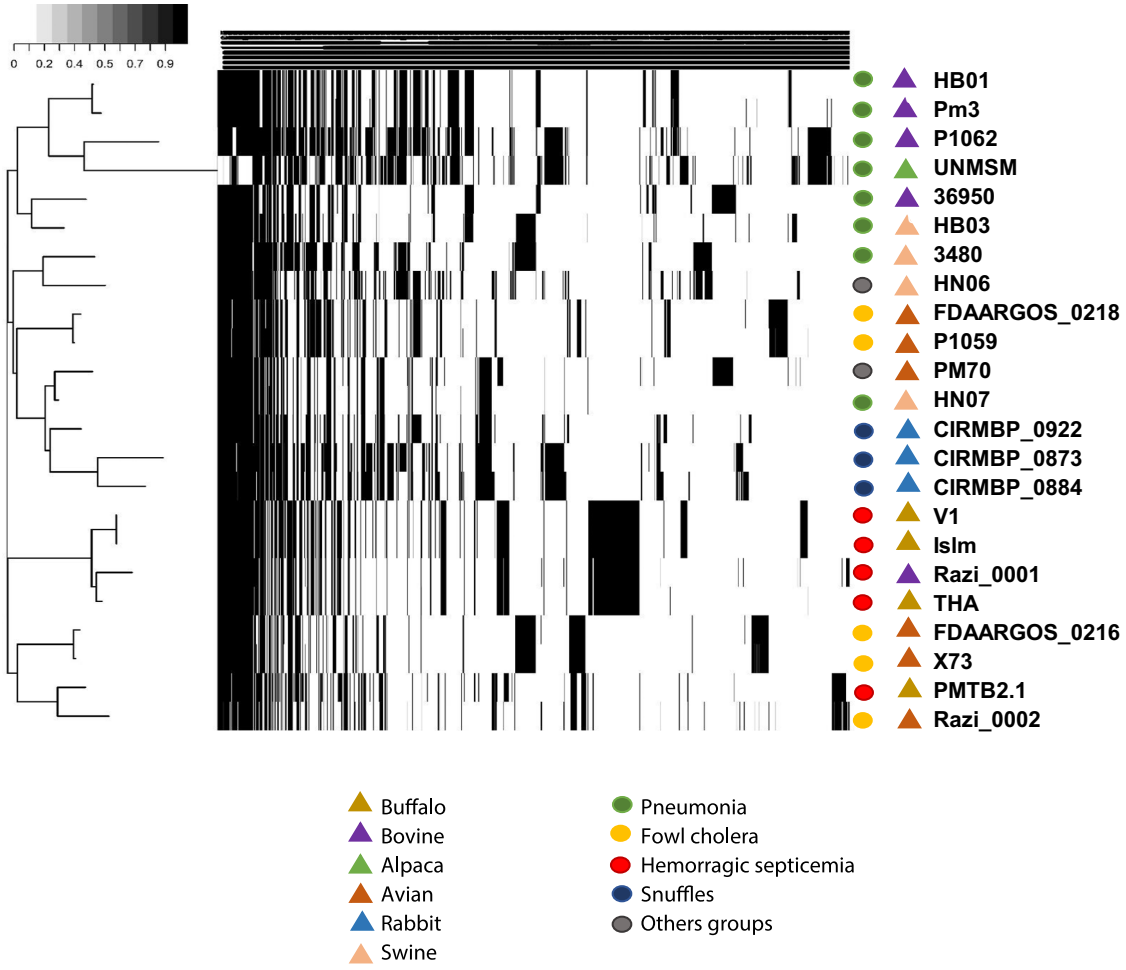


Fig. 7. Distribution of accessory genes in *P. multocida* strains. The heatmap indicating the presence (black) or absence (white) of accessory genes. Each row in the matrix represents one genome and each column represent the accessory genes. The dendrogram was generated by applying a hierarchical clustering algorithm. The phylogeny based on the accessory genome no show a stronger correlation between gene content and specificity host and disease manifestation.

clusterization among groups (Figs. 7 and 8). This finding could support the presence of gene sets related to adaptations of specific groups. Additionally, the pan-genome analysis performed by Anvi'o tool, show the clusterization from genomes based on the presence/absence pattern of genes clusters (Fig. 9). The Anvi'o program uses another heuristic for the determination of orthology and, for this reason, small differences that do not interfere in the final result were visualized.

3.7. Assessment of the pathogenic potential of P. multocida strains based on the Genomic Island content

The 23 genomes from *P. multocida* were analyzed for the prediction of genomic islands (GEIs) and prophages (S3 Table 1). This analysis showed a considerable variation in the number of genomic islands, highlighting an essential intraspecific diversity in *P. multocida*. GEIs were predicted for each genome, being grouped, by association with the disease (hemorrhagic septicemia, fowl cholera, pneumonia and snuffles). The comparative analysis between the GEIs was performed within the same group and between groups. A GEI was considered the genomic island of a group if it were at least in two members of the group.

3.7.1. Hemorrhagic septicemia group

The comparative analysis among GEIs from hemorrhagic septicemia-associated strains (Razi_0001, V1, Islm and THA with exception PMTB2.1) identified seven specific GEIs (S3 Table 2) (Fig. S1). For the description of islands, Razi_0001 and PMTB2.1 were used as reference strains (S3 Tables 3, 4). The genes of each exclusive GEI are detailed on S4. R1_PGI10 is an incomplete phage of approximately 46.6 kb harboring 44 hypothetical proteins, an IgA FC receptor with function antiphagocytic (Nordenfelt et al., 2012) and a protein that encodes cell filamentation protein Fic, component of the toxin-antitoxin module (Garcia-Pino et al., 2008). Also, the prophage R1_ PGI11 has approximately 15.3 kb (S3 Table 2), which include a capsule biosynthesis locus present in all strains belonging to capsular serogroup B, as previously reported (Moustafa et al., 2015). Additionally, a small island R1_PGI2 of 5.7 kb encodes four transferases of glycosyl groups. For PMTB2.1 strain, six PGI were predicted, which are commonly shared with another group. The presence of different islands explains the segregation of this strain from the hemorrhagic septicemia group (Fig.7), and its close relationship with the avian group, whose it shares the islands PMTB2_PGI1, 4, 6 (S3 Table 2).

3.7.2. Fowl cholera group

The comparative analysis of GEIs from fowl cholera-associated strains (×73, P1059, Razzi_0002, F216 and F218) identified six specific regions (S3 Table 5, Figs. S2, S3). For the description of islands ×73 and P1059, strains were used as a reference (S3 Tables 6, 7). The genes of each exclusive GEI are detailed on S5. We found two islands shared between ×73 and F216 strains. ×73_GEI1 with a size of 12.78 kb harbored 12 CDSs, six involved in galactitol transport and metabolism and described in previous studies (Johnson et al., 2013) and three glycosyltransferases. ×73_GEI4 (10.95 kb) contains 10 CDSs, five

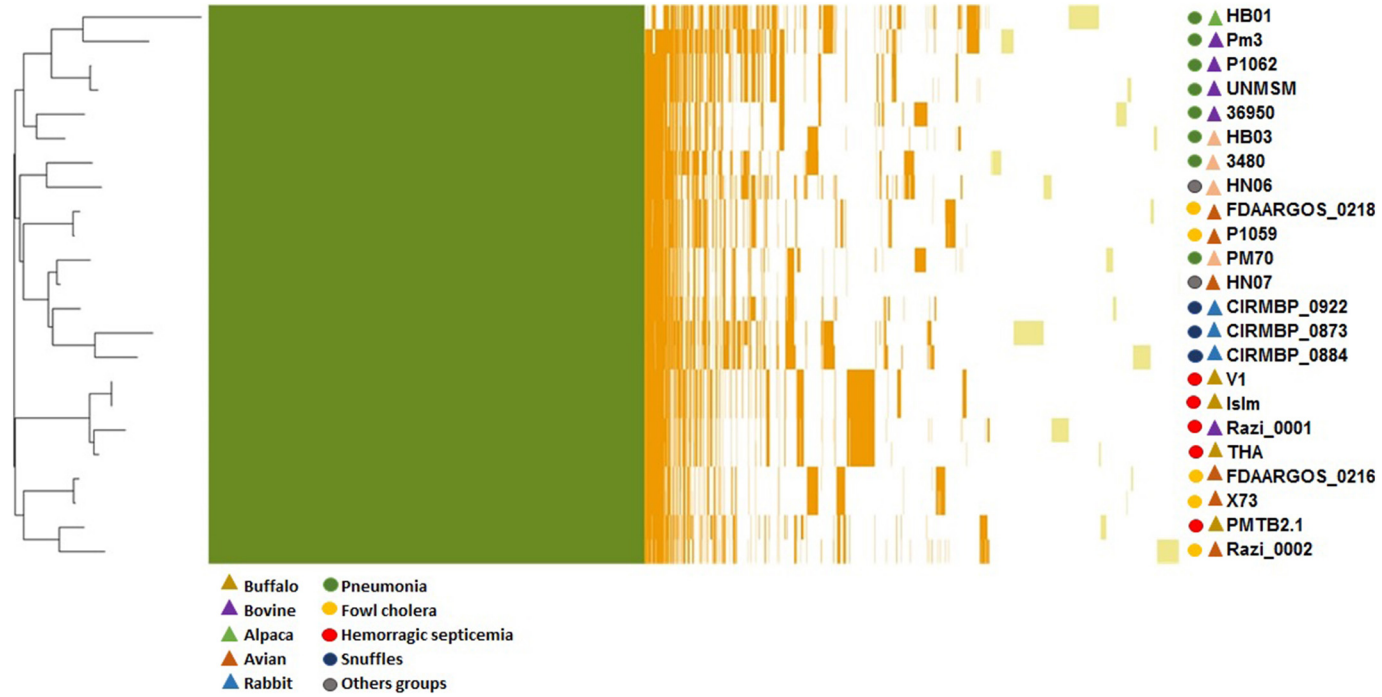


Fig. 8. Distribution of pan-genome in *P. multocida* strains. The heatmap indicating the presence or absence of all genes in each individual genome. Each row in the matrix represents one genome and each column represent the core (green), accessory (orange) and unique (mustard) genes. The dendrogram was generated by applying a hierarchical clustering algorithm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

involved in the transport and utilization of L-fructose and 3 CDSs involve in Ribose ABC transport system. The capacity for using different sugars could be important to host invasion as other reported have been described before (Johnson et al., 2013). Between F216, R2, PMTB2.1, and ×73, we identified two islands ×73_GEI3 (7.28 Kb) with 5 CDSs involved in the allose transport and metabolism. The X73_GEI6 (9.13 Kb) with 7 CDSs, involved in the transport and metabolism of arabinose. We identified two specific GEIs in P1059 and F218. The P1059_GEI1 (8.5 kb) has 4 CDSs involved in the transport and modification of citrate, described in previous studies (Fig. S3) (Johnson et al., 2013).

3.7.3. Snuffles group

There were no specific islands in the rabbit-associated strains (C873, C884, and C992). For the description of islands, the reference strains CIRMBP_0873 and CIRMBP_0992 were used (S3 Tables 9, 10). All GEIs were shared between C873, C884, C0992 and with strains of other groups (Figs. S4, S5). Furthermore, there are several genes uniquely present in snuffles group, scattered throughout the genome (S1 Table 3).

3.7.4. Pneumonia group

The comparative analysis of GEIs from pneumonia-associated strains (Pm-3, 3480, 36,950, P1062, UNMSM, HB01 and HB03) identified 10 specific GEIs (S3 Table 11). For the description of the islands, strains Pm-3, P1062, UNMSM and 36,950 were used as a reference (S3 Tables 12, 13, 14, 15). The genes in each exclusive GEI are detailed on S6. The island Pm-3_GEI2 (5.35 kb) was identified as a putative prophage present in 36,950, UNMSM and Pm-3 strains (Fig. S6) with eight CDSs, including a rhodanese-like domain-containing protein and *AraC* family transcriptional regulator.

A GEI was shared between Pm-3 and HB01. Pm-3_GEI11 (46.1 kb) is an intact phage with 63 CDSs, harboring the cI − cro set of genes, which are regulators lysogenic and lytic states maintenance, respectively. We also identified a CII for stimulation of CI transcription (Oppenheim et al., 2005). Also, we identified a host specificity tail protein J, which binds to the host receptor, protein holing, is a transmembrane protein that conforms a pore for exporting endolysins from the bacterial cell wall (Reddy & Saier, 2013). Also, operon hicAB was identified, it encodes a toxin-antitoxin (TA) system, that has a role regulating bacterial programmed cell death, biofilm formation, cope with nutritional stress, establish persistent subpopulations and offer protection from phage attack (Ghafourian et al., 2014). A P1062_GEI5 (Fig. S7), which encodes three trehalose metabolism genes, Chromate reductase and Quercetin 2,3-dioxygenase which are shared with 36,950, HB01 and HB03.

Although the number of islands gives us a sight of high diversity among strains, an analysis of the gene content is necessary. The distribution of GEI genes allows the corroboration of phylogenetic clustering among some lineages of the set groups (Fig. 10). The set of members of septicemia hemorrhagic group (THA, Razi_0001, Islm and V1), pneumonia group (HB01, Pm-3, P1062 and UNMSM), snuffles group (C0884, C0873) and the two avian groups (F218, P1059 and F216, ×73) were possible clustering at the level of an island gene analysis. This finding could support the presence of gene set at specific groups from adaptation.

3.8. Positive selection and its association with pathogenic adaptation

A previous positive selection analysis using 33 genomes of *Pasteurella multocida* identified 35 genes related to virulence processes, such as adhesion, host immune system evasion, LPS biosynthesis, but also related to gene regulation and DNA and protein metabolism (Cao et al., 2017). The analysis used site models, suitable to identify diversifying selection in specific codons (Yang et al., 2000). Here, we used branch-site models, suitable to identify adaptive mutations by testing for positive selection only in specific codons of pre-determined lineages (Yang & Nielsen, 2002; Zhang, 2005). From twelve genes were initially identified as positive selected, only two were true positive results. The only gene detected in Rabbit foreground was discarded due to

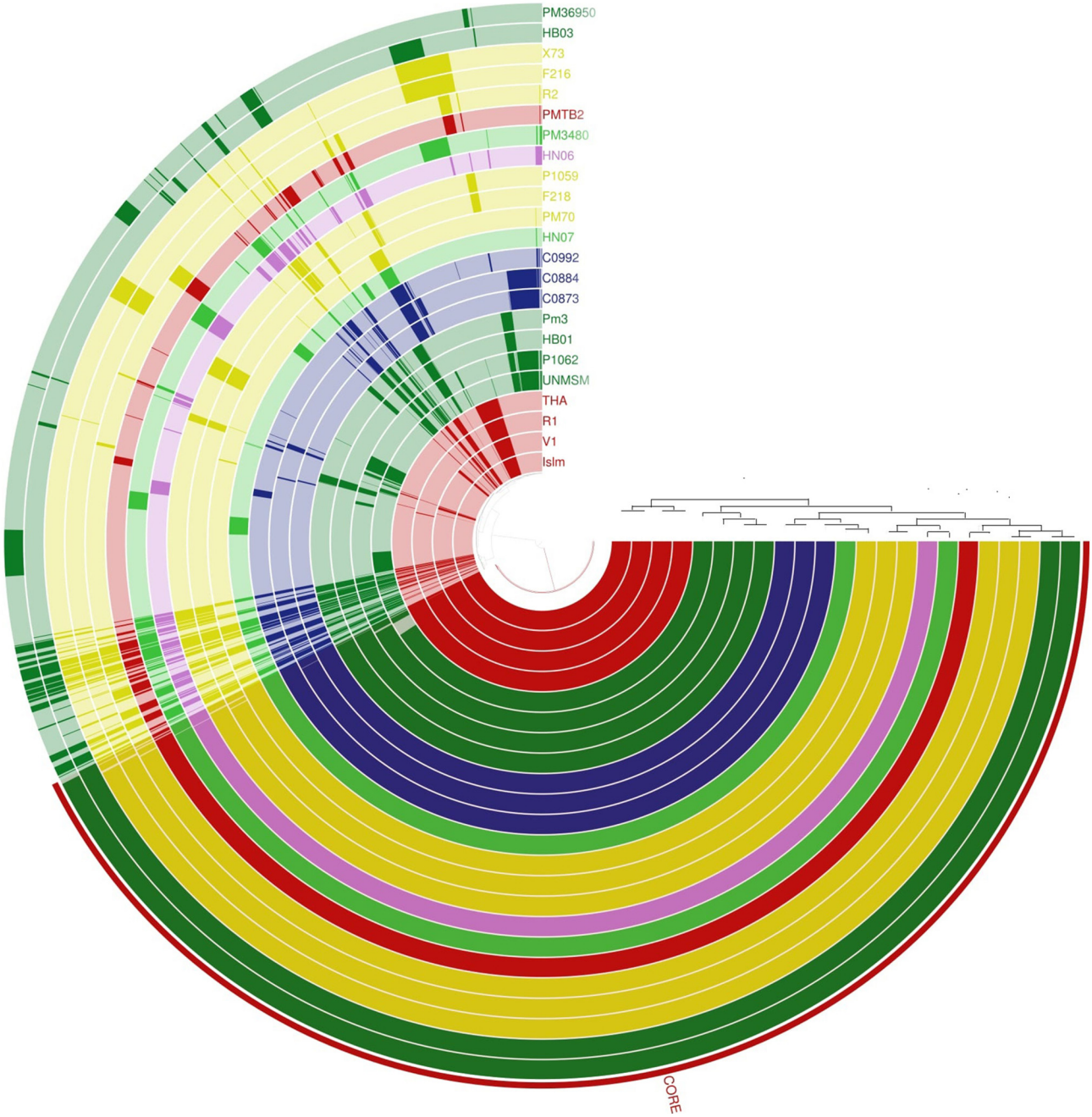


Fig. 9. Pan-genome visualization of 23 *P. multocida* strains was obtained using Anvi'o tool. Each ring indicates a genome, color variations as dark or light show presence or absence of genes by genome. The outer-most track in red show the core genome, while the different members of established groups are indicated by colors: members of hemorrhagic septicemia group (red color), pneumonia group (green gradients color), snuffles group (blue color), fowl cholera group (yellow color) and HN06 strain (purple color). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

alignment issues, and nine other genes had evidence of recombination. Recombination causes false positive results in positive selection analysis due to the alignment of non-homologous codons (Garcia-Pino et al., 2008; Oppenheim et al., 2005). The complete results for each group and the initial 12 detected genes are in S7 Table 6.

Positive selection was identified only for two genes from serotype B foreground. The TonB-dependent hemin ferrichrome receptor (*cir*A, CIRMBP0884_107) is an outer membrane protein that transports iron-containing proteins and colicin Ia (Jakes & Finkelstein, 2010; Zárate-Bonilla et al., 2014). It had 4 PS sites within its functional domains. The alpha/beta hydrolase (CIRMBP0884_803) has seven sites under positive selection, and its function is unknown.

The results of positive selection analysis suggest directional selection only for an alpha/beta hydrolase, an iron and colicin uptake systems from serotype B strains. Most of the adaptations in *P. multocida* appear to be related to the variation of gene content.

4. Conclusion

In summary, the comparative analysis of 23 *P. multocida* strains shows an open pan-genome. Accessory and unique genes present (46.35% and 49%) a high content of proteins of unknown functions, which reveal a high intraspecific diversity and a lack of knowledge of functional mechanism. Additionally, we performed a phylogenetic analysis using accessory genome content that allows showing at great diversification and this could reflect a clusterization among the specific

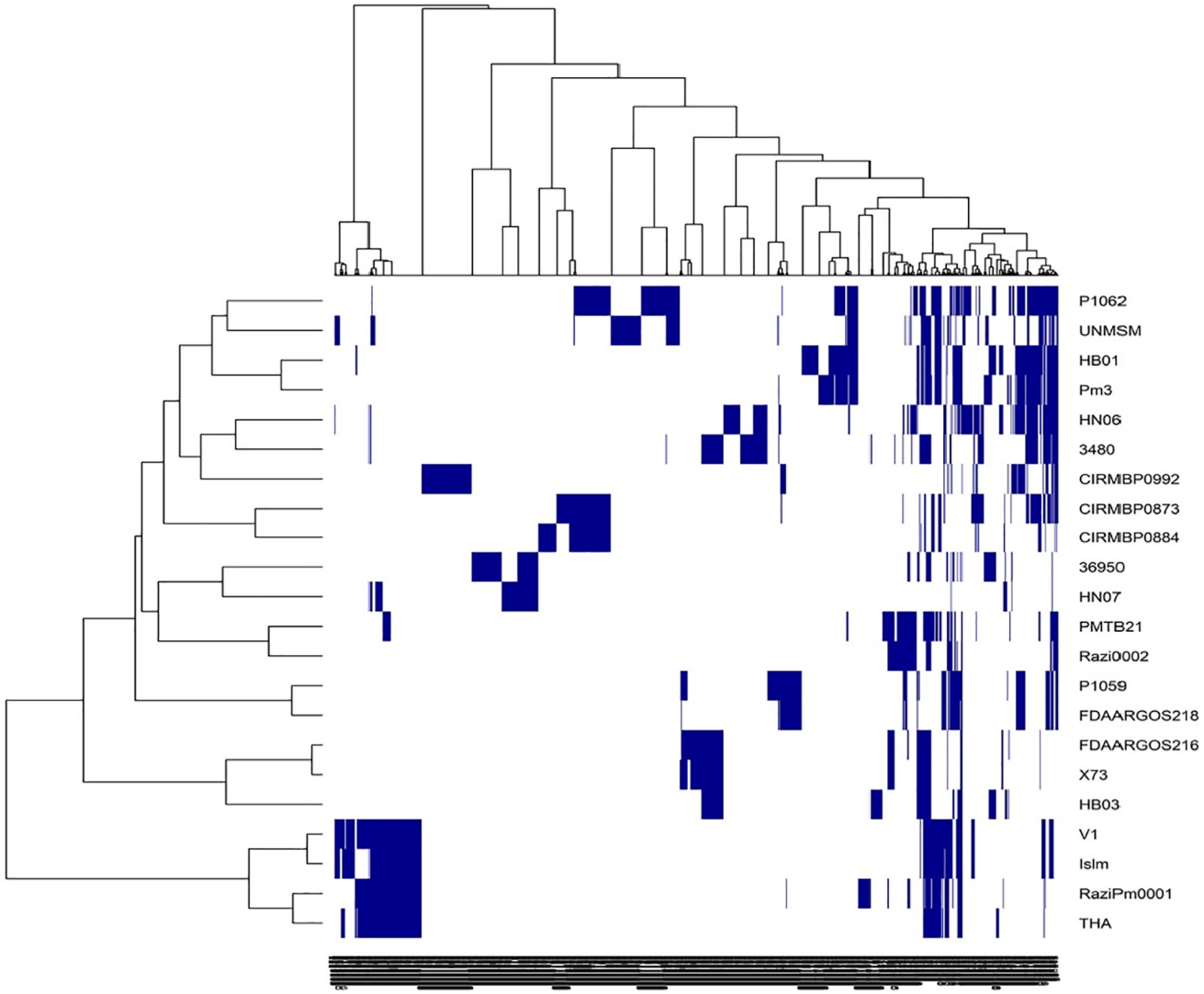


Fig. 10. Distribution of genomic island (GEI) genes in *P. multocida* strains. The heatmap indicating the presence (blue) or absence (white) of GEI genes. Each row in the matrix represents one genome and each column represent the GEI genes. The dendrogram was generated by applying a hierarchical clustering algorithm. The genome no was grouped about gene content. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

disease (hemorrhagic septicemia, fowl cholera, pneumonia and snuffles), but not host. On the base of core genome and their association with virulence genes, we identify a known set of genes encoding enzymes for LPS biosynthesis, set encoding involved in iron-dependent regulation and acquisition and outer membrane proteins, with prominent roles in the pathogenesis.

Analysis based on the content of GEIs identified 7, 6 and 10 GEIs exclusive on set groups (hemorrhagic septicemia, fowl cholera and pneumonia). On both analysis of exclusivity of genomic islands and genes, were corroborated with the presence of a majority of phage-related genes, hypothetical proteins and virulence-associated genes. The positive selection analysis suggests adaptive mutations on an alpha/beta hydrolase, an iron and colicin uptake system, uniquely for the serotype B sub-group, the remaining genes were not considered by having evidence for recombination.

Finally, our finding report an analysis comprehensive of content genomic as strategies of adaptation and diversification of *P. multocida* on pathogenic specialization.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2018.05.084.

Competing interests

The authors of this manuscript have no competing interests.

Authors' contributions

RH, DC, MC, FFA worked on comparative genomics analyzes. FFA and RH wrote the manuscript. MC, FFA, VA, and SS revised the manuscript. FFA, VA, and LM supervised this work.

Acknowledgments

The research was supported by Coordination for the Improvement of Higher Education Personnel (CAPES) (CAPES Bio Computacional 051/2013) and Brazilian National Research Council (CNPq) (CNPq 150203/2016-0, 168804/2017-4). This work was supported by the Programa Nacional de Innovación para la Competitividad y Productividad (Innóvate Perú), under the contract 133-FINCyT-IB-2013, as part of the project “Reverse Vaccinology: Developing a new generation vaccine for the control and/or prevention pasteurellosis pneumonia in alpacas”.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.

Alikhan, N.F., Petty, N.K., Ben Zakour, N.L., Beatson, S.A., 2011. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Araujo, F.A., Barh, D., Silva, A., Guimarães, L., Ramos, R.T.J., 2018. GO FEAT: a rapid web-based functional annotation tool for genomic and transcriptomic data. Sci. Rep. 8, 1794. Internet. Available from. http://www.nature.com/articles/s41598-018-20211-9.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44, W16–W21.

Bellanger, X., Payot, S., Leblond-Bourget, N., Guédon, G., 2014. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiol. Rev. 720–760.

Boyce, J.D., Seemann, T., Adler, B., Harper, M., 2012. Pathogenomics of *Pasteurella multocida*. Curr. Top. Microbiol. Immunol. 361, 23–38.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., et al., 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci. Rep. 5, 8365.

Brooks, A.N., Turkarslan, S., Beer, K.D., Yin Lo, F., Baliga, N.S., 2011. Adaptation of cells to new environments. Wiley Interdiscip. Rev. Syst. Biol. Med. 544–561.

Bruen, T.C., Philippe, H.H., Bryant, D., 2006. A simple and robust statistical test for detecting the presence of recombination. Genetics 172, 2665–2681.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60. Internet. Available from: http://www.nature.com/doifinder/10.1038/nmeth.3176%5Cnhttp://dx.doi.org/10.1038/nmeth.3176%5Cnhttp://www.nature.com/doifinder/10.1038/nmeth.3176%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/25402007.

Cao, P., Guo, D., Liu, J., Jiang, Q., Xu, Z., Qu, L., 2017. Genome-wide analyses reveal genes subject to positive selection in *Pasteurella multocida*. Front. Microbiol. 8.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Carter, G.R., 1955. Studies on *Pasteurella multocida*. I. A hemagglutination test for the identification of serological types. Am. J. Vet. Res. 16, 481–484.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., et al., 2004. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 33, D325–D328. Internet. cited 2018 Jan 7. Available from. http://www.ncbi.nlm.nih.gov/pubmed/15608208.

Christensen, H., Kuhnert, P., Olsen, J.E., Bisgaard, M., 2004. Comparative phylogenies of the housekeeping genes atpD, infB and rpoB and the 16S rRNA gene within the Pasteurellaceae. Int. J. Syst. Evol. Microbiol. 54, 1601–1609.

Edgar, R.C., 2004. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5, 1–19.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., et al., 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data'. PeerJ 3, e1319. Internet. Available from. https://peerj.com/articles/1319.

Ewers, C., Lübke-Becker, A., Bethe, A., Kießling, S., Filter, M., Wieler, L.H., 2006. Virulence genotype of *Pasteurella multocida* strains isolated from different hosts with various disease status. Vet. Microbiol. 114, 304–317.

Felsenstein, J., 2005. PHYLIP (Phylogeny Inference Package) Version 3.6. Department of Genome Sciences, University of Washington, Seattle.

Fischer, S., Brunk, B.P., Chen, F., Gao, X., Harb, O.S., John, B., et al., 2012. NIH Public Access. pp. 1–23.

Galperin, M.Y., Makarova, K.S., Wolf, Y.I., Koonin, E.V., 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res. 43, D261–D269.

Garcia-Pino, A., Christensen-Dalsgaard, M., Wyns, L., Yarmolinsky, M., Magnuson, R.D., Gerdes, K., et al., 2008. Doc of prophage P1 is inhibited by its antitoxin partner Phd through fold complementation. J. Biol. Chem. 283, 30821–30827.

Ghafourian, S., Raftari, M., Sadeghifard, N., Sekawi, Z., 2014. Toxin-antitoxin systems: classification, biological function and application in biotechnology. Curr. Issues Mol. Biol. 16, 9–14.

Hammer, Ø., DAT, Harper, Ryan, P.D., 2001. Past: paleontological statistics software package for education and data analysis. Palaeontol. Electron. 4 (1), 1–9. Internet. Available from. http://palaeo-electronica.org/2001_1/past/issue1_01.htm.

Harper, M., Boyce, J.D., Adler, B., 2006. *Pasteurella multocida* pathogenesis: 125 years after Pasteur. FEMS Microbiol. Lett. 1–10.

Harper, M., John, M., Turni, C., Edmunds, M., St. Michael, F., Adler, B., et al., 2015. Development of a rapid multiplex PCR assay to genotype *pasteurella multocida* strains by use of the lipopolysaccharide outer core biosynthesis locus. J. Clin. Microbiol. 53, 477–485.

Hongo, J.A., de Castro, G.M., Cintra, L.C., Zerlotini, A., Lobo, F.P., 2015. Potion: an end-to-end pipeline for positive Darwinian selection detection in genome-scale data through phylogenetic comparison of protein-coding genes. BMC Genomics 16, 567.

Hurtado, R.E., Aburjaile, F., Mariano, D., Canário, M.V., Benevides, L., Fernandez, D.A., et al., 2017. Draft genome sequence of a virulent strain of *Pasteurella Multocida* isolated from alpaca. J. Genomics 5, 68–70. Internet. Available from. http://www.jgenomics.com/v05p0068.htm.

Jakes, K.S., Finkelstein, A., 2010. The colicin Ia receptor, Cir, is also the translocator for colicin Ia. Mol. Microbiol. 75, 567–578.

Jakobsen, I.B., Easteal, S., 1996. A program for calculating and displaying compatibility matrices as an aid in determining reticulate evolution in molecular sequences. Comput. Appl. Biosci. 12, 291–295.

Jiang, S.S., Lin, T.Y., Wang, W.B., Liu, M.C., Hsueh, P.R., Liaw, S.J., 2010. Characterization of UDP-glucose dehydrogenase and UDP-glucose pyrophosphorylase mutants of proteus mirabilis: defectiveness in polymyxin B resistance, swarming, and virulence. Antimicrob. Agents Chemother. 54, 2000–2009.

Johnson, T.J., Abrahante, J.E., Hunter, S.S., Hauglund, M., Tatum, F.M., Maheswaran, S.K., et al., 2013. Comparative genome analysis of an avirulent and two virulent strains of avian *Pasteurella multocida* reveals candidate genes involved in fitness and pathogenicity. BMC Microbiol. 13 (1) Internet. (Available from: BMC Microbiology).

Kanehisa, M., Sato, Y., Morishima, K., 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J. Mol. Biol. 726–731.

Kück, P., Longo, G.C., 2014. FASconCAT-G: Extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool. 11, 1–8.

Lawrence, J.G., 1997. Selfish operons and speciation by gene transfer. Trends Microbiol. (9), 355.

Lefébure, T., Stanhope, M.J., 2007. Evolution of the core and pan-genome of streptococcus: positive selection, recombination, and genome composition. Genome Biol. 8.

Levin, B.R., Bergstrom, C.T., 2000. Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. Proc. Natl. Acad. Sci. 97, 6981–6985. Internet. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.97.13.6981.

Loytynoja, A., Goldman, N., 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science 320 (80), 1632–1635.

Marri, P.R., Hao, W., Golding, G.B., 2006. Gene gain and gene loss in streptococcus: is it driven by habitat? Mol. Biol. Evol. 23, 2379–2391.

May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whittam, T.S., Kapur, V., 2001. Complete genomic sequence of *Pasteurella multocida*, Pm70. Proc. Natl. Acad. Sci. U. S. A. 98, 3460–3465. Internet. Available from. http://www.ncbi.nlm.nih.gov/pubmed/11248100.

Michael, G.B., Kadlec, K., Sweeney, M.T., Brzuszkiewicz, E., Liesegang, H., Daniel, R., et al., 2012. ICEPmu1, an integrative conjugative element (ICE) of *pasteurella multocida*: structure and transfer. J. Antimicrob. Chemother. 67, 91–100.

Moustafa, A.M., Seemann, T., Gladman, S., Adler, B., Harper, M., Boyce, J.D., et al., 2015. Comparative genomic analysis of Asian haemorrhagic septicaemia-associated strains of *Pasteurella multocida* identifies more than 90 haemorrhagic septicaemia-specific genes. PLoS One 10, 1–16. Internet. Available from: https://doi.org/10.1371/journal.pone.0130296.

Nordenfelt, P., Waldemarson, S., Linder, A., Mörgelin, M., Karlsson, C., Malmström, J., et al., 2012. Antibody orientation at bacterial surfaces is related to invasive infection. J. Exp. Med. 209, 2367–2381. Internet. Available from. http://www.jem.org/lookup/doi/10.1084/jem.20120325.

Okay, S., Kurt Kizildoğan, A., 2015. Comparative genome analysis of five *Pasteurella multocida* strains to decipher the diversification in pathogenicity and host specialization. Gene 567, 58–72.

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., Adhya, S., 2005. Switches in bacteriophage lambda development. Annu. Rev. Genet. 39, 409–429. Internet. Available from: http://www.annualreviews.org/doi/10.1146/annurev.genet.39.073003.113656.

Paustian, M.L., May, B.J., Cao, D., Boley, D., 2002. Transcriptional Response of *Pasteurella multocida* to Defined Iron Sources. 184. pp. 6714–6720.

Posada, D., Crandall, K.A., Holmes, E.C., 2002. Recombination in evolutionary genomics. Annu. Rev. Genet. 36, 75–97.

Prasannavadhana, A., Kumar, S., Thomas, P., Sarangi, L.N., Gupta, S.K., Priyadarshini, A., et al., 2014. Outer membrane proteome analysis of indian strain of *pasteurella multocida* serotype B:2 by MALDI-TOF/MS analysis. Sci. World J. 2014.

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. Vienna Austria. Internet. (0:{ISBN} 3-900051-07-0. Available from). http://www.r-project.org/.

Raetz, C.R.H., Whitfield, C., 2002. Lipopolysaccharide Endotoxins. Annu. Rev. Biochem. 71, 635–700. Internet. Available from: http://www.annualreviews.org/doi/10.1146/annurev.biochem.71.110601.135414.

Rambaut, A., 2012. FigTree v1. 4.0. A Graphical Viewer of Phylogenetic Trees. Inst. Evol. Biol. Univ. Edinburgh Internet. Available from: http://scholar.google.pl/scholar?q=Rambaut,+A.+FigTree+v1.4.0.&btnG=&hl=en&as_sdt=0,5#0.

Reddy, B.L., Saier, M.H., 2013. Topological and phylogenetic analyses of bacterial holin families and superfamilies. Biochim. Biophys. Acta Biomembr. 1828, 2654–2671.

Sahm, A., Bens, M., Platzer, M., Szafranski, K., 2017. PosiGene: Automated and easy-to-use pipeline for genome-wide detection of positively selected genes. Nucleic Acids Res. 45, 1–11.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069.

Simon, C., Wiezer, A., Strittmatter, A.W., Daniel, R., 2009. Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. Appl. Environ. Microbiol. 75, 7519–7526.

Smith, J.M., 1992. Analyzing the mosaic structure of genes. J. Mol. Evol. 34, 126–129.

Soares, S.C., Geyik, H., Ramos, R.T.J., de Sá, P.H.C.G., Barbosa, E.G.V., Baumbach, J., et al., 2016. GIPSy: genomic island prediction software. J. Biotechnol. 232, 2–11. [Internet]. Elsevier B.V.. Available from. https://doi.org/10.1016/j.jbiotec.2015.09.008.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., et al., 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-Genome". Proc. Natl. Acad. Sci. 102, 13950–13955. Internet. Available from. http://www.pnas.org/cgi/doi/10.1073/pnas.0506758102.

Totsika, M., Wells, T.J., Beloin, C., Valle, J., Allsopp, L.P., King, N.P., et al., 2012. Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic Escherichia coli. Appl. Environ. Microbiol. 78, 2179–2189.

Tournu, H., Fiori, A., Van Dijck, P., 2013. Relevance of Trehalose in pathogenicity: some general rules, yet many exceptions. PLoS Pathog. 9.

Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., et al., 2017. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res. 45, D535–D542.

Wilkie, I.W., Harper, M., Boyce, J.D., Adler, B., 2012. Pasteurella multocida: diseases and pathogenesis. Curr. Top. Microbiol. Immunol. 361, 1–22.

Wilson, B.A., Ho, M., 2013. *Pasteurella multocida*: From Zoonosis to cellular microbiology. Clin. Microbiol. Rev. 26, 631–655.

Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.

Yang, Z., Dos Reis, M., 2011. Statistical properties of the branch-site test of positive selection. Mol. Biol. Evol. 28, 1217–1228.

Yang, Z., Nielsen, R., 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol. Biol. Evol. 19, 908–917.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.

Zárate-Bonilla, L.J., del Portillo, P., Sáenz-Suárez, H., Gonzáles-Santos, J., Barreto-Sampaio, G.E., Poutou-Piñales, R.A., et al., 2014. Computational modeling and preliminary *iro*N, *fep*A, and *cir*A gene expression in *Salmonella* Enteritidis under iron-deficiency-induced conditions. Poult. Sci. 93, 221–230.

Zhang, J., 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. 22, 2472–2479.

Zhao, Y., Jia, X., Yang, J., Ling, Y., Zhang, Z., Yu, J., et al., 2014. PanGP: a tool for quickly analyzing bacterial pan-genome profile. Bioinformatics 30, 1297–1299.

Resultados adicionais

Neste trabalho, como resultado adicional aos artigos publicados, visou-se fazer uma análise de *MLST* (*Multilocus Sequence Typing*) para complementar as análises filogenômicas, e avaliar a relação filogenética entre os genótipos e a capacidade de infectar um hospedeiro e/ou provocar uma doença. Para isso, 23 linhagens de *P. multocida* foram submetidas ao servidor *MLST* (cge.cbs.dtu.dk//services/MLST) (Larsen *et al*., 2012). Essa análise determina a sequência tipo (*ST*) ou perfil alélico de acordo com os alelos de cada um dos 7 locos. Esses resultados foram adicionados à ferramenta *eBURST* (Spratt *et al*., 2004) para a predição de grupos de sequências tipo (*STs)* e complexo clonal (*CC*). Isso permite o agrupamento das *STs* que compartilham um ancestral comum recente. Os resultados foram conferidos na base de dados de *STs* de *P. multocida* (https://pubmlst.org/) e descritos na Tabela 1. Os dados de *CCs* foram adicionados à arvore filogenética baseado no *SNPs* do genoma central das 23 linhagens (Figura 1). A árvore foi usada para entender a relação entre os genótipos e os fenótipos das linhagens do estudo. Do estudo a relação filogenética de algumas linhagens relacionada a doença, localização geográfica e *ST*. Como é mostra na Figura 1, as linhagens aviárias de *P. multocida* (X73, P1059, F216, F218) estão divididas em dois grupos de acordo com as *STs*. Sendo identificadas com os complexos clonais ST158 e ST8, reportadas em estudos prévios em amostras isoladas de aves (Subaaharan *et al*., 2010; Li *et al*., 2018; Peng *et al*., 2018). Porém, a linhagem Pm70 e R2 (linhagens aviárias) pertencem a ST9 e ST129 respectivamente, e não se encontram agrupadas. As *STs* preditas em isolados de aves com cólera aviária foram um resultado esperado, uma vez que as *STs* encontradas em nosso estudo (neste grupo), foram reportadas previamente em aves (*Submission in process*: Hurtado, 2019). Entretanto, a maioria dos isolados, com exceção da linhagem R2, proveem dos Estados Unidos, sugerindo que possivelmente existe uma variedade de *STs* nesse país. Ainda assim, estas *STs* também já foram reportadas na China, Austrália e Canadá (*Submission in process*: Hurtado, 2019), evidenciando a propagação destas *STs* entre países e continentes. Além disso, as *STs* analisadas foram preditas em diferentes hospedeiros (*Submission in process*: Hurtado, 2019), sendo não exclusivas de aves com exceção da ST158 (Harms, 2015).

As linhagens provenientes do grupo septicemia hemorrágica de *P. multocida* (V1, THA, Ilsm e R1) foram agrupadas correspondendo com o mesmo complexo clonal ST122 e pertenecendo ao mesmo continente asiático. A ST122 é o genótipo frequentemente reportado em casos de septicemia hemorrágica e não restrito a um único hospedeiro (Hotchkiss *et al*., 2011; Moustafa *et al*., 2013; Petersen *et al*., 2014). A ST122 foi reportado também em países da África e da Ásia (Moustafa *et al*., 2013), com isto é possível deduzir que essa *ST*, pode estar

correlacionada nessas localidades geográficas. Porém, a linhagem PMTB2 (isolada de septicemia hemorrágica) foi agrupada com a linhagem R2, pertencendo ambas ao mesmo complexo clonal ST129.

As linhagens provenientes do grupo pasteurelose (rinite e pneumonia) (CIRMBP-0922, CIRMBP-0884, CIRMBP-0873) foram agrupadas, correspondendo também ao complexo clonal ST9, e tipo de doença. Sendo que o complexo clonal ST9 é frequentemente reportado em casos de pneumonia e presentes em diferentes hospedeiros, mas frequentemente em coelhos (*Submission in process*: Hurtado, 2019). Esse genótipo estaria relacionado à doença e ao hospedeiro, mas poderia estar relacionado a localização geográfica, já que essa *ST* foi reportada nos países da Europa (*Submission in process*: Hurtado, 2019).

As linhagens provenientes do grupo pneumonia (Pm3, P1062, HB03, HB01, UNMSM) foram agrupados, correspondendo também aos complexos clonais ST13 e ST203, e mesmo tipo de doença. Sendo que o complexo clonal ST13 está presente na China, Alemanha e Estados Unidos, reportado previamente (*Submission in process*: Hurtado, 2019), considerando a propagação dessas *STs* em um amplo espectro de países e continentes. Porém, as linhagens HN07 e 3480 pertencem ao ST9 e ST74 respectivamente, e não pertencem ao grupo pneumonia. Os complexos clonais ST13, ST50 e ST74 são reportados em maior prevalência nos casos de pneumonias, e associadas ao sistema respiratório em diferentes hospedeiros (*Submission in process*: Hurtado, 2019). Nesse contexto, é possível inferir que existam algumas *STs* associadas aos casos de pneumonias, das quais algumas são reportadas unicamente com a doença (*Submission in process*: Hurtado, 2019). Contudo, o número de relatos e estudos que confirmem essa hipótese é escasso (Hotchkiss *et al*., 2011; García-Alvarez *et al*., 2017). Além disso, a linhagem UNMSM isolada de pneumonia em alpaca no Peru, pertence ao grupo clonal ST203. Sendo a ST203, reportado em ovinos com pneumonia na Espanha (García-Alvarez *et al*., 2017). O fato da ST203 ser reportada em uma linhagem do Peru e da Espanha condiz com a propagação deste CC nestes países, e que poderia estar associado ao tipo de doença, mas faltariam estudos para sustentar está afirmação. Além disso, a linhagem HN06 (rinite atrófica no coelho) pertence a ST50, sendo ele o primeiro registro de *STs* em casos de rinite atrófica.

Dessa forma, os resultados adicionais nos permitem inferir que conhecendo grupos genotípicos associados a um tipo de doença, hospedeiro, distribuição geográfica e relação com os surtos, é possível planejar estratégias de controle e prevenção para combater essas doenças.

Tabela 1. Sequências tipo e complexos clonais de 23 linhagens de *P. multocida* utilizadas no estudo "*Pan-genomic approach shows insight of genetic divergence and pathogenic adaptation of Pasteurella multocida"*(Hurtado *et al*., 2018).

Linhagens ST CC Hospedeiro Doenças Origem

Linhagens	Origem	Hospedeiro	Doenças	MLST	Alelos							CC
					adk	*Est*	*gdh*	*mdh*	*pgi*	*pmi*	*zwf*	
X73	EUA/América do Norte	Aves	Cólera Aviária	60	4	24	2	11	21	25	25	ST158
FDAARGOS216	EUA/América do Norte	Aves	Cólera Aviária	60	4	24	2	11	21	25	25	ST158
FDAARGOS 218	EUA/América do Norte	Aves	Cólera Aviária	8	4	5	3	5	5	2	5	ST8
P1059	EUA/América do Norte	Aves	Cólera Aviária	8	4	5	3	5	5	2	5	ST8
HN07	China/Ásia	Porco	Pneumonia	9	2	7	5	4	7	5	6	ST9
Pm70	EUA/América do norte	Aves	Cólera Aviária	9	2	7	5	4	7	5	6	ST9
CIRMPB-0884	França/Europa	Coelhos	Pasteurelose (Pulmão)	9	2	7	5	4	7	5	6	ST9
CIRMPB-0873	França/Europa	Coelhos	Pasteurelose (Pulmão)	204	2	11	5	4	27*	5	6	ST9
CIRMPB-0992	França/Europa	Coelhos	Pasteurelose (Cavidade Nasal)	9	2	7	5	4	7	5	6	ST9
Razi0002	Irã/Ásia	Aves	Cólera Aviária	129?	21	33	20	17	20	26?	2	ST129
PMTB2.1	Malásia/Ásia	Búfalo	Septicemia Hemorrágica	201	21	33	20	17	35	26	2	ST129
V1	Tailândia/Ásia	Búfalo	Septicemia Hemorrágica	122	23	37	2	4	17	21	17	ST122
THA	Tailândia /Ásia	Búfalo	Septicemia Hemorrágica	122	23	37	2	4	17	21	17	ST122
Ilsm	Tailândia /Ásia	Búfalo	Septicemia Hemorrágica	122	23	37	2	4	17	21	17	ST122
Razi0001	Irã/Ásia	Bovino	Septicemia Hemorrágica	122	23	37	2	4	17	21	17	ST122
Pm3	China/Ásia	Bovino	Pneumonia	79	26	11	7	4	8	9	10	ST13
P1062	EUA/América do Norte	Bovino	Pneumonia	80	7	11	7	4	8	40	10	ST13
HB03	China/Ásia	Bovino	Pneumonia	13	7	11	7	4	8	9	10	ST13
HB01	China/Ásia	Bovino	Pneumonia	79	26	11	7	4	8	9	10	ST13
36950	Alemanha/Europa	Bovino	Pneumonia	79	26	11	7	4	8	9	10	ST13
3480	China/Ásia	Porco	Pneumonia	74	22	13	3	8	31	8	29	ST74
UNMSM	Peru/América do Sul	Alpaca	Pneumonia	321	23	44	53	15	43	39	36	ST203
HN06	China/Ásia	Porco	Rinite Atrófica	50	14	10	3	8	20	20	19	ST50

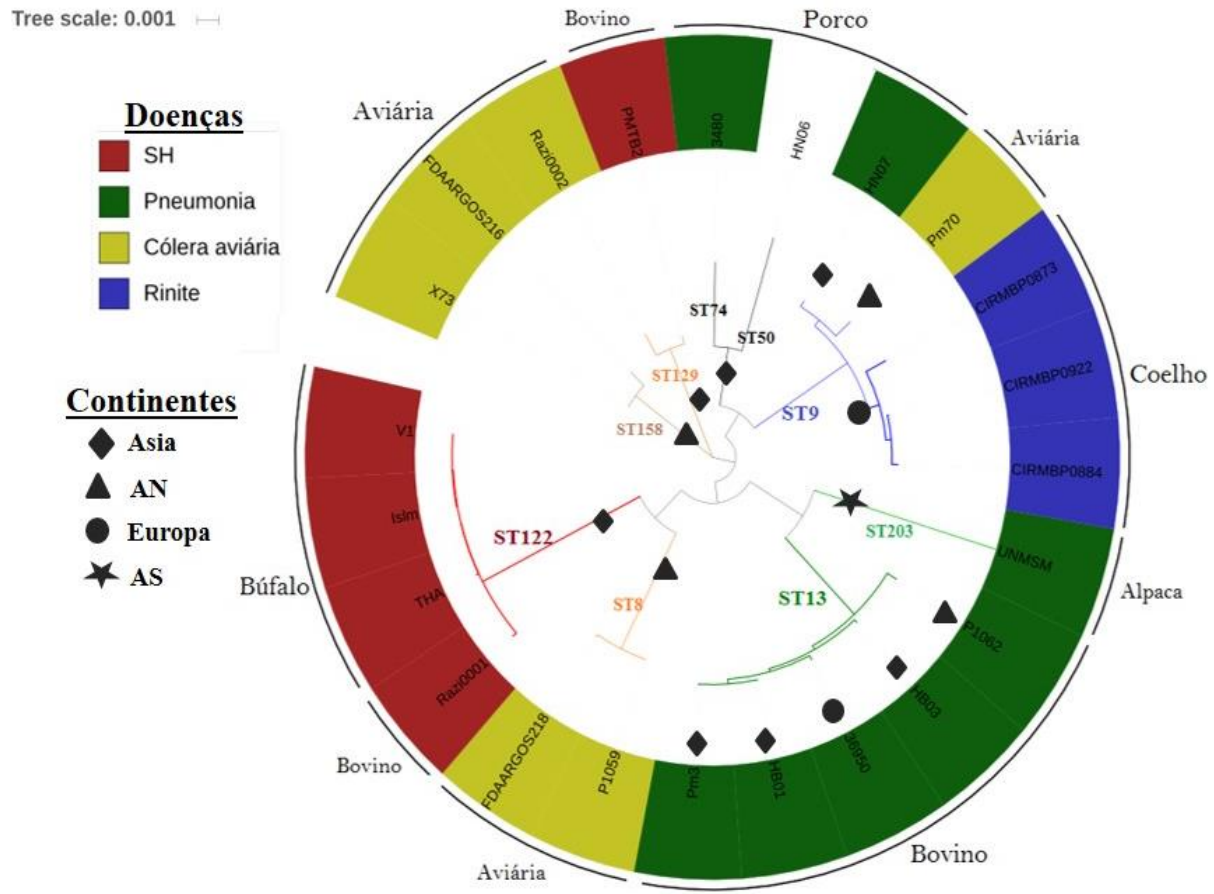


Figura 1. Árvore filogenética baseada no *SNPs* do genoma central de 23 linhagens de *P. multocida* utilizadas no estudo "***Pan-genomic approach shows insight of genetic divergence and pathogenic adaptation of Pasteurella multocida*" (Hurtado *et al.*, 2018)**. Complexos clonais, hospedeiros, doenças, continentes e sequências tipo foram adicionadas na árvore para melhor interpretação da relação filogenética.

CAPÍTULO 3: Discussão geral, conclusão e perspectivas

Discussão geral

- A linhagem UNMSM de *P. multocida* isolada de um pulmão de alpaca com quadro de pneumonia na província de Cuzco, é considerada uma linhagem patogênica, devido ao quadro clínico de isolamento e presença de genes de virulência, que foram identificados em trabalhos prévios (Rímac *et al*., 2017), cumprindo as características definidas para um organismo patogênico (Ewers *et al*., 2006; Harper *et al*., 2006; Wilson and Ho, 2013). Foram identificadas 14 ilhas genômicas, as quais 6 eram regiões de profagos, sendo essa uma das linhagens que apresenta um maior número de elementos móveis. Nas análises filogenéticas, a linhagem UNMSM apresenta-se mais divergente dentro do grupo de isolados de pneumonia (Hurtado et al., 2018), isso poderia ser consequência do isolamento geográfico, sendo a única linhagem isolada e sequenciada da América do Sul.
- Dentre os resultados adicionais, foram analisadas as STs e os CC na linhagem UNMSM, a qual pertence ao ST321 e ao complexo clonal ST203, sendo este CC já reportado em casos de pneumonia em ovino (García-Alvarez *et al.*, 2017). Isso reafirma que certos grupos de ST poderiam estar relacionados a sua procedência fenotípica. Adicionalmente, essa linhagem foi isolada de um agrupamento de linhagens clonais, procedente de análises de diversidade convencionais por PCR (Rocío *et al*., 2017)(Rímac et al., 2017; Rocío et al., 2017)(Rímac et al., 2017; Rocío et al., 2017). No entanto, as análises genômicas seriam importantes para conhecimento da diversidade de linhagens de camelídeos em regiões endêmicas na América do Sul, nos diferentes anos e locais.
- As análises de diversidade baseadas na filogenia de substituição de nucleotídeos, aminoácidos e presença/ausência de genes mostram a concordância no agrupamento de linhagens provenientes de grupos fenotípicos estabelecidos (doença) (Hurtado et al., 2018). Isso corrobora com estudos prévios baseados em análises de MLST, genes associados a virulência e sorotipos, como tentativa de correlacionar grupos genéticos a uma mesma procedência fenotípica (Davies *et al*., 2004; Ewers *et al*., 2006; Dabo *et al*., 2007; Christensen and Bisgaard, 2010; Pors e*t al.*, 2011; Hotchkiss *et al*., 2011; Cardoso-Toset *et al*., 2013; Moustafa *et al*., 2013; Petersen *et al*., 2014; García-Alvarez *et al*., 2015, 2017).
- As análises de população mostram a relação filogenética com as STs, genogrupos e sorogrupos, e relaciona algumas linhagens com o tipo de hospedeiro (Peng *et al*., 2018). Em um estudo recente se sugere que a filogenia tem relação com a procedência do hospedeiro e

localização geográfica das linhagens (Zhu *et al*., 2019). Nesses dois últimos artigos publicados, o estudo foi abordado desde filogenia ao genoma central (core), entretanto é essencial considerar os elementos móveis no processo de evolução. A filogenia baseada no genoma acessório complementa essas análises, as quais, obteve-se a filogenia do genoma central (core) e presença e ausência de genes, os quais mostram concordância no agrupamento de grupos estabelecidos (Hurtado *et al*., 2018). O agrupamento de linhagens poderia ser explicado pela presença de elementos móveis, que foram herdados de um ancestral comum, ou que foram adquiridas horizontalmente de microrganismos pertencentes a um mesmo nicho (teoria de convergência evolutiva). Isso explicaria a porcentagem variada de homologia entre algumas ilhas genômicas dentro de agrupamentos estabelecidos (Brüssow *and* Kutter, 2004; Hurtado *et al*., 2018). Contudo, sobre as análises de filogenias é importante mencionar que existem linhagens que não foram agrupadas de acordo com o esperado, e isso sugere que algumas linhagens seriam capazes de infectar, devido a condições de susceptibilidade do hospedeiro, sendo essa uma característica própria de um patógeno oportunista.

- As análises de diversidade por filogenia baseada na presença e ausência de genes mostra uma maior variabilidade entre linhagens dos grupos estabelecidos (Hurtado *et al*., 2018) comparada a variação nucleotídica/aminoacídica. As análises de pangenoma apresentaram um genoma acessório de 52.91%, sendo que esses genes estavam, em sua maioria, em ilhas genômicas. Esses resultados corroboram com um estudo de 114 linhagens de *P. multocida*, o qual 42% dos genes funcionais apresentam-se dentro das ilhas genômicas, e sua variação estaria determinada pela localização geográfica ou ecológica (Zhu *et al*., 2019). O pangenoma dessa espécie encontra-se aberto, em um aumento regular e constante de novos genes, e isso é esperado, devido à natureza de estilo de vida de *P. multocida*, sendo não restrita a hospedeiro e patógeno facultativo.
- O conjuntos de genes que compõem o pangenoma podem expandir-se através da aquisição de elementos móveis (GEIs, CEIs, profagos) e pressão aos diferentes ambientes (Hurtado *et al*., 2018; Zhu *et al*., 2019)(Hurtado et al., 2018; Zhu et al., 2019)(Hurtado et al., 2018; Zhu et al., 2019). Neste trabalho foram identificados 159 GEIs de 23 linhagens, das quais 7, 6, 10 GEIs foram exclusivas dos grupos pertencentes a septicemia hemorrágica, cólera aviária e pneumonia, respectivamente. Baseado nas análises funcionais foram identificados genes relacionados a fagos, proteínas hipotéticas e genes de virulência. Na maioria dos casos as ilhas pertenciam a profagos, sendo considerado um importante elemento móvel para a variabilidade da espécie (Hurtado *et al*., 2018; Zhu *et al*., 2019). O estudo de Zhu e colaboradores, revelou a predição de ilhas genômicas de 114 genomas de *P. multocida*, identificando 280 ilhas, as

quais 123 eram ilhas de patogenicidade (Zhu *et al.*, 2019), considerando a importância dos elementos móveis na patogenicidade e definição da espécie.

- Em nosso trabalho, análises referente ao conteúdo do genoma acessório revelaram em maior porcentagem os seguintes grupos funcionais: (i) metabolismo e transporte de carboidratos e aminoácidos; (ii) biogênese da parede celular e membrana; (iii) reguladores transicionais e (iv) proteínas de função desconhecida (Hurtado *et al*., 2018). Os mesmos grupos funcionais foram identificados em diversos estudos patogenômicos (Ford *et al*., 2013; Johnson *et al*., 2013; Wilson *and* Ho, 2013; Moustafa *et al*., 2015; Du et al., 2016; Yu *et al*., 2016)(Johnson et al., 2013; Wilson and Ho, 2013; Moustafa et al., 2015; Du et al., 2016; Yu et al., 2016)(Johnson et al., 2013; Wilson and Ho, 2013; Moustafa et al., 2015; Du et al., 2016; Yu et al., 2016). A presença desses grupos funcionais, como parte do genoma acessório, sugerem a sua importância no processo de adaptação, inferindo a fixação no genoma por pressão seletiva (McInerney *et al.*, 2017). Esses mesmos grupos funcionais foram identificados em estudos de expressão diferencial quando foram avaliados em casos de doença *in vivo* em coelhos (Guo *et al*., 2012). A presença de genes e maior expressão diferencial, pertencentes a esses grupos funcionais, são essenciais no processo de regulação e adaptação a determinada condição ambiental. Genes envolvidos no transporte e metabolismo de aminoácidos e carboidratos são importantes para uso das maquinarias metabólicas na aquisição de nutrientes necessários para adaptar-se a um novo hospedeiro (Johnson et al., 2013). Também é possível que reguladores transcricionais sejam selecionados e ativados para uma rápida adaptação a diferentes condições ambientais (Cases *et al*., 2003; Guo *et al.*, 2012). Proteínas com função desconhecida estão presentes na espécie *P. multocida,* demostrando a escassez de estudos funcionais bacterianos. Essas proteínas foram identificadas, principalmente, em regiões de profagos, as quais as bases de dados funcionais ainda não estão completas.
- As análises de polimorfismo de nucleotídeo único demostraram que a espécie encontra-se em um constante processo de diversificação e submetidos a seleção positiva (Fitzpatrick *et al*., 2005; Chen *et al*., 2006; Johnson *et al.*, 2013; Okay *and* Kurt Kızıldoğan, 2015; Cao *et al*., 2017). A maioria das categorias funcionais encontradas estão envolvidas na interação entre o patógeno e hospedeiro, porém não estão associadas com uma determinada patogenia. Entretanto, em análises de seleção positiva de grupos estabelecidos não foi possível identificar genes submetidos a seleção por grupo, com exceção para o subgrupo B HS, em que os genes não estão implicados no processo adaptativo da espécie (Hurtado *et al*., 2018).

Conclusão

P. multocida, linhagem UNMSM, é uma bateria patogênica, que por análises filogenéticas se mostra mais divergente em relação a outras linhagens isoladas de pneumonia, possivelmente por isolamento geográfico e especialização ao hospedeiro. Além disso, a linhagem pertence ao ST203 reportado em casos de pneumonias em ovino. As análises de pangenômica e predição de ilhas genômicas demostraram que a espécie está em um constante processo de adaptação submetido a pressão seletiva, considerando a importância dos elementos móveis na evolução e definição da espécie. As análises funcionais do genoma acessório mostraram que os genes envolvidos em processos como o metabolismo e transporte de carboidratos e aminoácidos, biogênese da parede celular e membrana, reguladores transcricionais são fundamentais no processo adaptativo. Os genes submetidos a seleção positiva são relevantes na interação com hospedeiro, mas não estão implicados no processo de adaptação por patogenias. As análises filogenômicas e pangenômicas em *P. multocida* demonstram que a presença de ilhas genômicas e algumas linhagens podem ter a capacidade de causar uma doença e infectar preferentemente um hospedeiro. A presença de ilhas genômicas e *SNPs* desempenham um papel crucial no processo de diversificação e adaptação da espécie *P. multocida*.

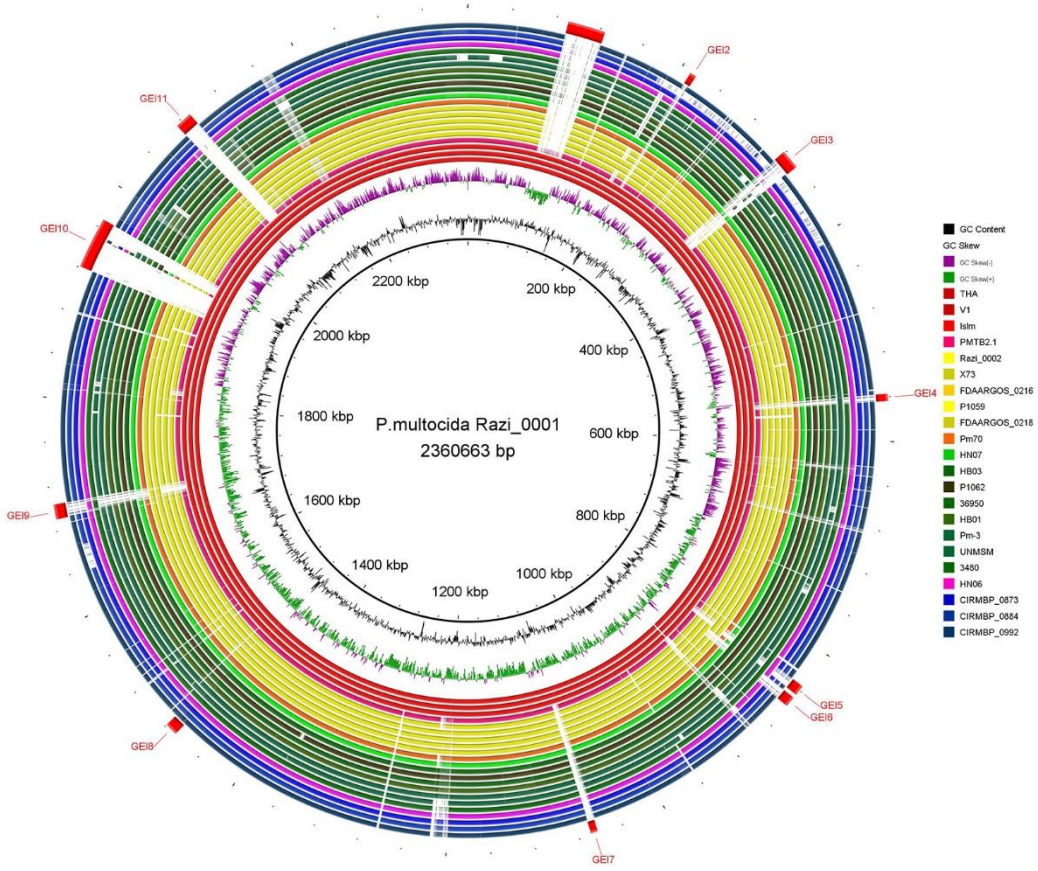
Perspectivas

Este trabalho tem como perspectivas:

- Realizar o sequenciamento das linhagens restantes isoladas do centro experimental da Raya (Puno), proveniente de diferentes regiões de produção de Alpaca no Peru;
- Realizar análises de vacinologia reversa, com enfoque em casos de pneumonia, e identificar alvos para o amplo espectro de doenças e hospedeiros;
- Analisar os reguladores transcricionais, envolvidos em rede regulatórias e redes metabólicas de *P. multocida* em cada perfil de doença associada;
- Analisar a diversidade em nível de polimorfismo de nucleotídeo único que possa estar associada ao nível regulatório do transcriptoma.

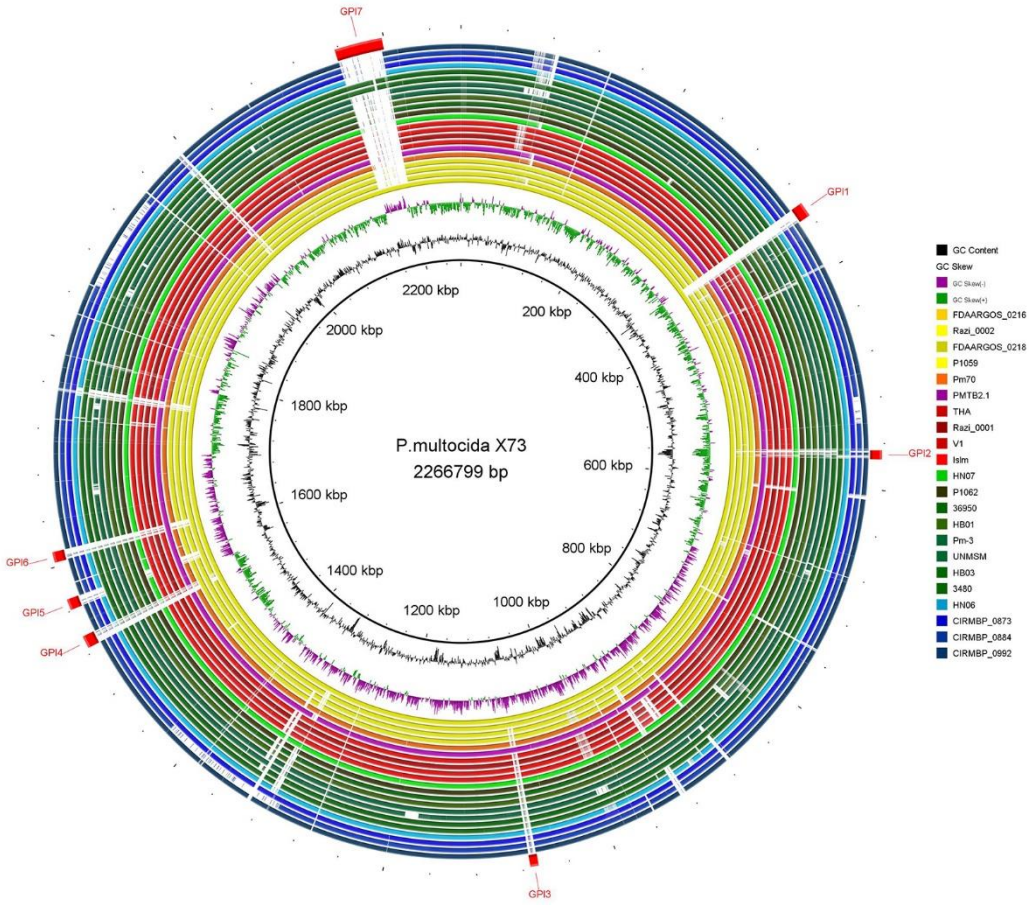
ANEXOS: Material suplementar

Figure. S1. Circular maps comparing genomic islands (GEIs) between *Pasteurella multocida* strain Razi_0001 as reference with others members of hemorrhagic septicemia group (red gradients color), and fowl cholera group (yellow gradients color), pneumonia group (green gradients color), snuffles group (blue gradients color), Pm70 strain (orange color), HN06 strain (purple color).



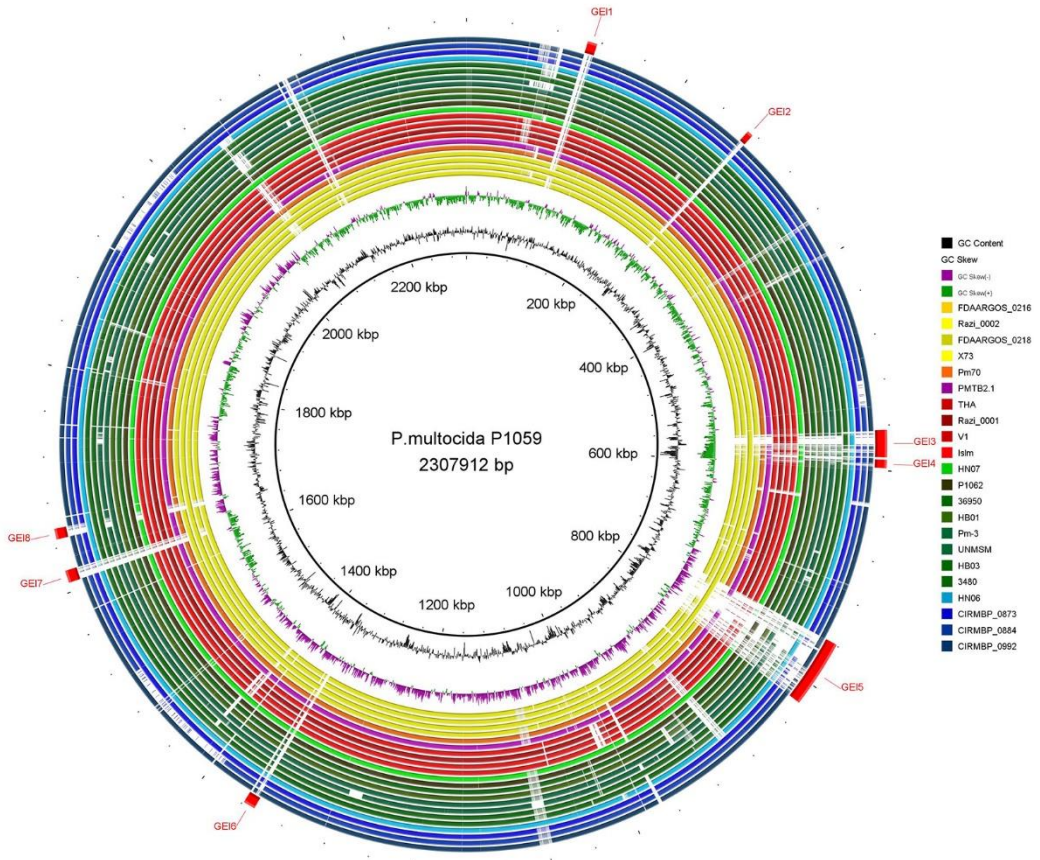
Fonte: (Hurtado *et al*., 2018)

Figure. S2. Circular genomic maps comparing genomic islands (GEIs) of *Pasteurella multocida* strain X73 as reference with others members of fowl cholera group (yellow gradients color), and hemorrhagic septicemia group (red gradients color), pneumonia group (green gradients color), snuffles group (blue gradients color), Pm70 strain (orange color) and HN06 strain (purple color).



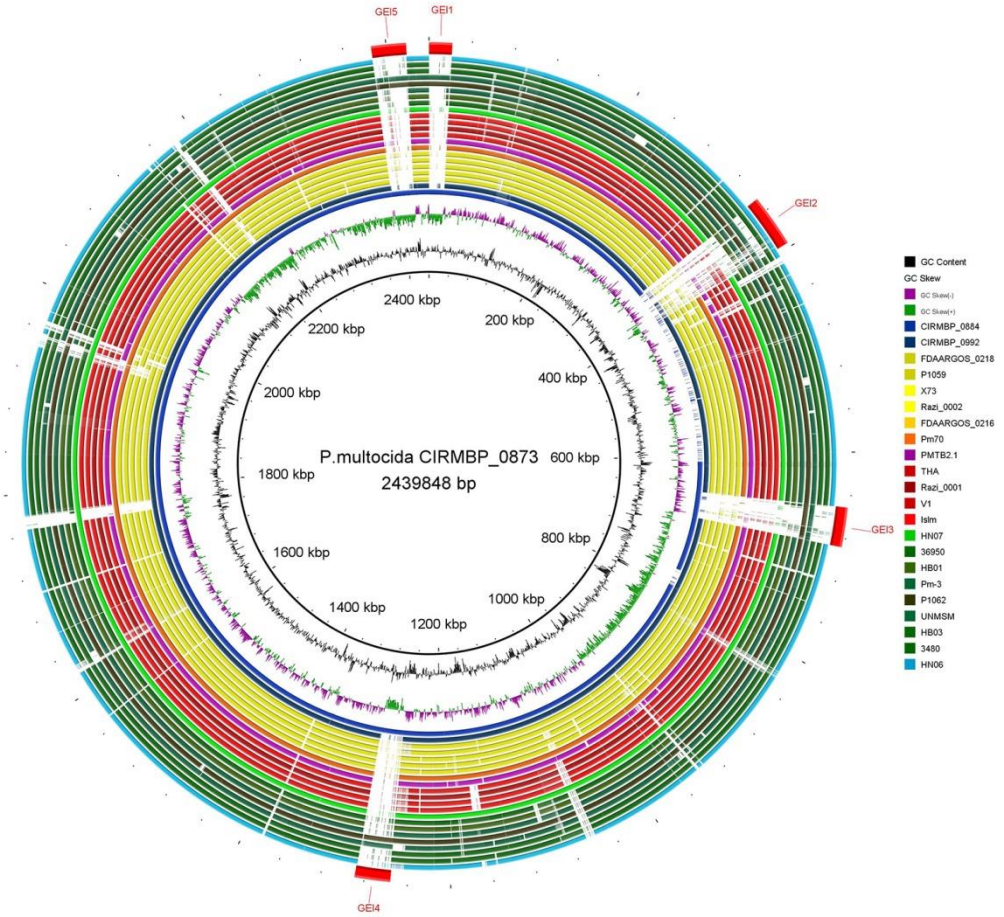
Fonte: (Hurtado *et al*., 2018)

Fig. S3. Circular maps comparing genomic islands (GEIs) between *P. multocida* P1059 strain as reference with others members of fowl cholera group (yellow gradients color), and hemorrhagic septicemia group (red gradients color), pneumonia group (green gradients color), snuffles group (blue gradients color), Pm70 strain (orange color) and HN06 strain (purple color).



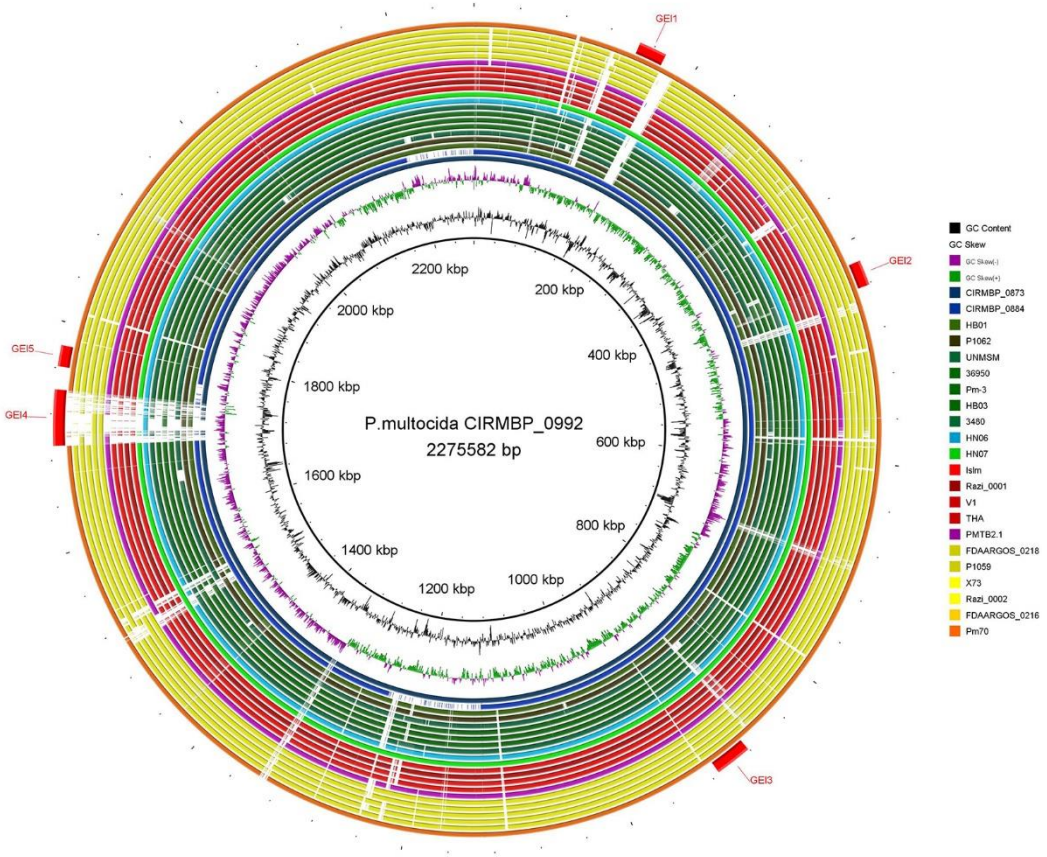
Fonte: (Hurtado *et al*., 2018)

Figure. S4. Circular genomic maps comparing genomic islands (GEIs) between *P. multocida* CIRMBP_0873 strain as reference with others members snuffles group (blue gradients color), and fowl cholera group (yellow gradients color), hemorrhagic septicemia group (red gradients color), pneumonia group (green gradients color), Pm70 strain (orange color) and HN06 strain (purple color).



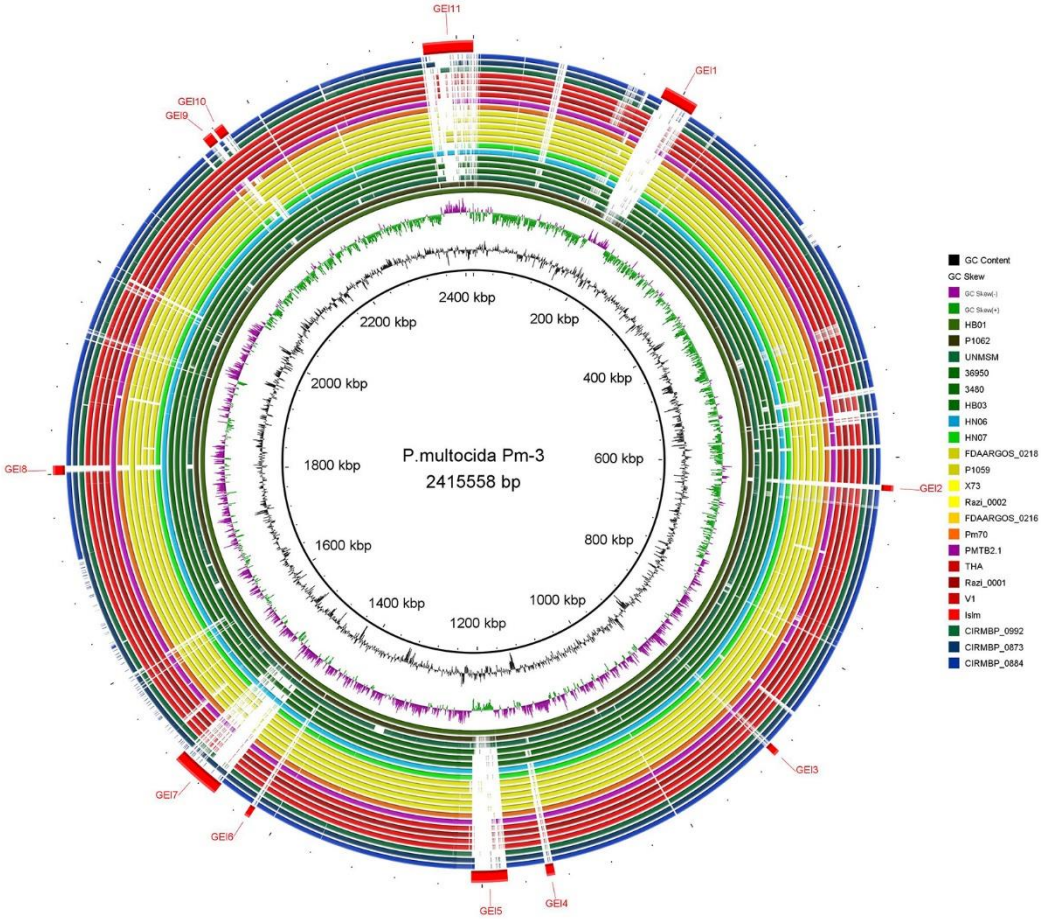
Fonte: (Hurtado *et al*., 2018)

Figure. S5. Circular genomic maps comparing genomic islands (GEIs) between *P. multocida* CIRMBP_0992 strain as reference with others members of snuffles group (blue gradients color), and fowl cholera group (yellow gradients color), septicemia hemorrhagic (red gradients color), pneumonia (green gradients color), Pm70 strain (orange color) and HN06 strain (purple color).



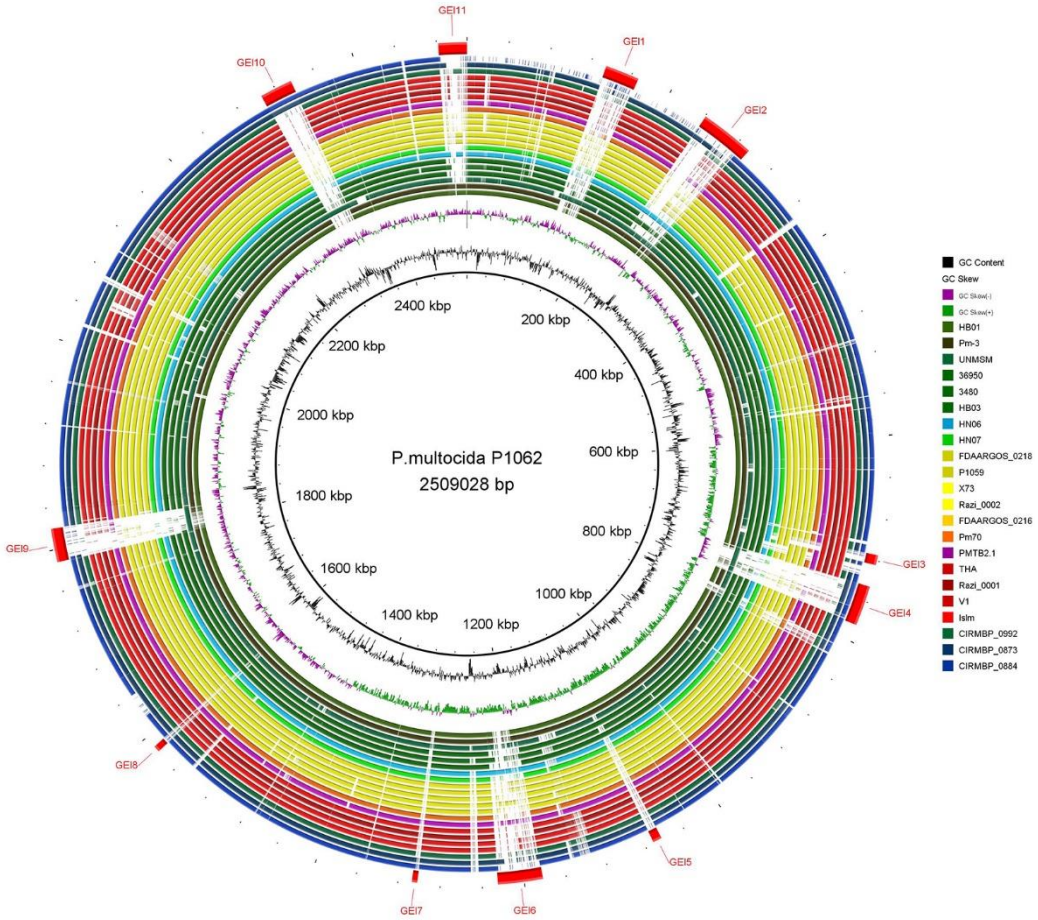
Fonte: (Hurtado *et al*., 2018)

Figure. S6. Circular genomic maps comparing genomic islands (GEIs) between *P. multocida* Pm-3 strain as reference with others members of pneumonia group (green gradients color), and fowl cholera group (yellow gradients color), hemorrhagic septicemia group (red gradients color), snuffles group (blue gradients color), rhinitis atrophic (HN06), Pm70 strain (orange color) and HN06 strain (purple color).



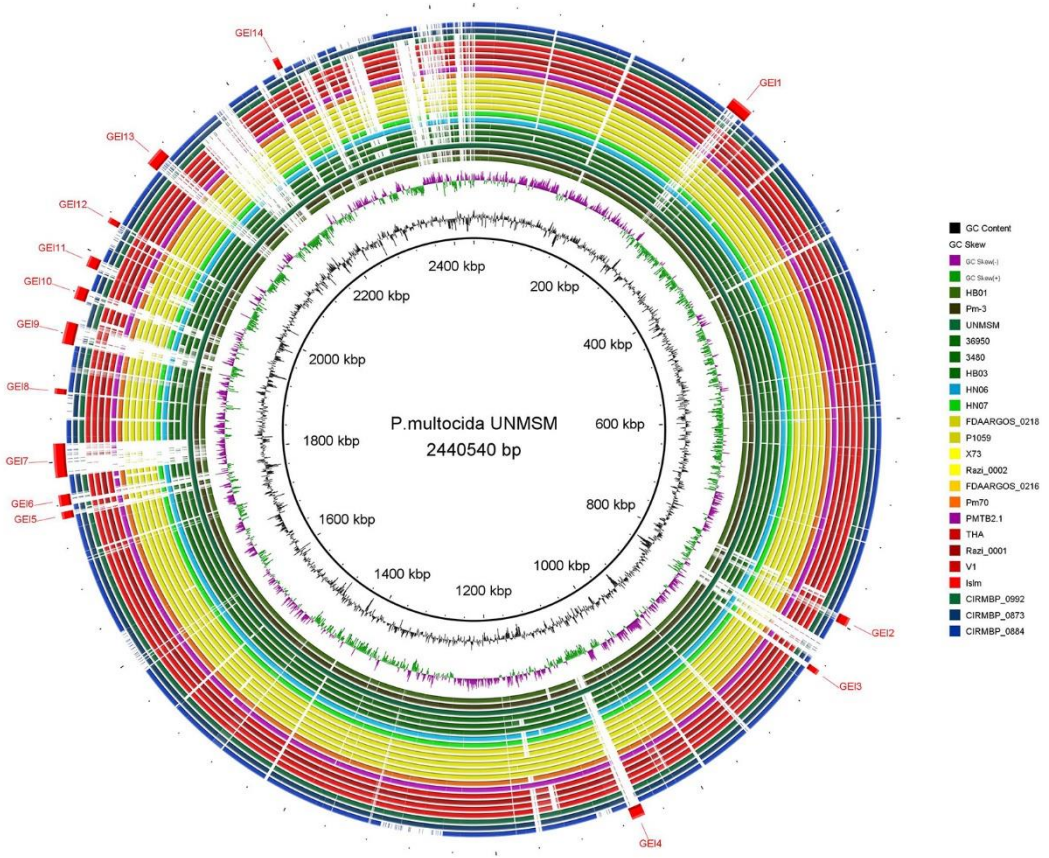
Fonte: (Hurtado *et al*., 2018)

Figure. S7. Circular genomic maps comparing genomic islands (GEIs) of *Pasteurella multocida* strain P1062 as reference with others members of pneumonia group (green gradients color), and fowl cholera group (yellow gradients color), hemorrhagic septicemia group (red gradients color), snuffles group (blue gradients color), Pm70 strain (orange color) and HN06 strain (purple color).



Fonte: (Hurtado *et al*., 2018)

Figure. S8. Circular genomic maps comparing genomic islands (GEIs) of *Pasteurella multocida* UNMSM strain as reference with other members of pneumonia group (green gradients color), and fowl cholera group (yellow gradients color), hemorrhagic septicemia group (red gradients color), snuffles group (blue gradients color), rhinitis atrophic (HN06), Pm70 strain (orange color) and HN06 strain (purple color).



Fonte: (Hurtado *et al*., 2018)

Supplementary Material 2

S1 Table 1. Description of the exclusive genes for hemorrhagic septicemia group. **Fonte:** (Hurtado *et al*., 2018)
Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S1 Table 2. Description of the exclusive genes for Fowl cholera group. **Fonte:** (Hurtado *et al*., 2018)
Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S1 Table 3. Description of the exclusive genes for snuffles group. **Fonte:** (Hurtado *et al*., 2018)
Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

Table 4. Description of the exclusive genes by pneumonia group. **Fonte:** (Hurtado *et al*., 2018)
Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

Supplementary Material 2

S2 Table 1. Virulence factors distributed in the core genome. **Fonte:** (Hurtado *et al*., 2018)
Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

Supplementary Material 3

S3 Table 1. General information of the number of genomic islands predicted in the 23 genomes of *P. multocida***.**

STRAINS	GEIs	PHAGE	GEI TOTAL
3480	8	2	8
36950	6	2	7
HB01	8	3	8
HB03	7	1	7
HN06	7	4	7
P1059	8	1	8
P1062	9	8	11
UNMSM	14	6	14
X73	7	1	7
Pm-3	10	4	11
Razi_0001	11	3	11
HN07	2	1	2
CIRMBP-0884	2	4	4
FDAARGOS_218	4	3	4
FDAARGOS_216	8	1	8
PMTB2.1	6	2	6
THA	10	3	10
Islm	11	4	11
V1	11	4	11
Razi_0002	5	1	5
CIRMBP-0873	3	5	5
CIRMBP-0922	2	10	10

Fonte: (Hurtado *et al*., 2018)

S3 Table 2. Description and comparison of predicted Genomic Island of each GEI of hemorrhagic septicemia group.

Reference GEIs	GC (%)	CDs	Size (kb)	V1 (pident)	Islm (pident)	THA (pident)	PMTB2.1 (pident)	Others members Groups (pident)
R1_GEI1	42.31	49	34.03	V1_GEI6 (100%)	Lslm_GEI6 (100%)	THA_GEI5 (100%)	-	-
R1_GEI2	31.29	6	5.71	V1_GEI7 (100%)	Lslm_GEI7 (100%)	THA_GEI6 (100%)	-	-
R1_GEI3	37.56	29	17.54	V1_GEI11 (100%)	Lslm_GEI11 (100%)	THA_GEI10 (100%)	-	-
R1_GEI7	34.69	5	6.339	V1_GEI1 (100%)	Lslm_GEI1 (100%)	THA_GEI1 (100%)	-	-
R1_GEI9	38.09	10	13.394	V1_GEI3 (100%)	-	THA_GEI2 (100%)	-	P1059_GEI3 (57.6%), F218_GEI4 (46.3%), HN06_GEI2 (53.3%)
R1_GEI10	43.22	63	46.683	V1_GEI4 (63.84%)	Lslm_GEI3 (30.73)	THA_GEI3 (63.8%)	-	-
R1_GEI11	34.16	10	15.371	V1_GEI5 (100%)	Lslm_GEI5 (100%)	THA_GEI4 (100%)	-	-
Ilslm_GEI2	37.78%	58	51.42	V1_GEI2 (100%)	-	-	-	-
PMTB_GEI4	32.87%	6	6.76	-	-	-	-	R0002_GEI5 (100%),
PMTB_GEI6	33.3%	7	9.13	-	-	-	-	X73_GEI6 (100%), F216_GEI4 (100%),

- The GEIs of red color are unique to the group. The GEIs of black color are islands partially shared in the group or that share similarity partially with other groups.
- **pident**: Percent identity

Fonte: (Hurtado *et al*., 2018)

S3 Table 3. Predicted islands of the genome Razi_0001 (R1).

LABEL	#START	#STOP
R1_GEI1	90568	124605
R1_GEI2	209145	214857
R1_GEI3	318557	336101
R1_GEI4	557193	563706
R1_GEI5	835489	844941
R1_GEI6	848368	857926
R1_GEI7	1062663	1069001
R1_GEI8	1468538	1479907
R1_GEI9	1690515	1703908
R1_GEI10	1920930	1967612
R1_GEI11	2074198	2089568

Fonte: (Hurtado *et al*., 2018)

S3 Table 4. Predicted islands of the genome PMTB2.1 (PMTB).

LABEL	#START	#STOP
PMTB_GEI1	268718	286747
PMTB_GEI2	566074	600012
PMTB_GEI3	700662	739845
PMTB_GEI4	899897	906659
PMTB_GEI5	1701213	1710766
PMTB_GEI6	1744068	1753201

Fonte: (Hurtado *et al*., 2018)

S3 Table 5. Description and comparison of predicted Genomic Island of each GEI of fowl cholera group.

Reference GEIs	GC (%)	CDs	Size (kb)	FDAARGOS_216 (pident)	FDAARGOS_218 (pident)	X73 (pident)	Razi_0002 (pident)	Others members Groups (pident)
X73_GEI1	37.59	15	12.8	F216_ GEI8 (100%)	-	-	-	-
X73_GEI2	30.41	13	7.88	F216_ GEI1 (100%)	-	-	-	Pm-3_GEI3 (28.4%), 36950_GEI2 (28.5%), HB01_GEI2 (30%)
X73_GEI3	32.17	7	7.29	F216_ GEI2 (93%)	-	-	R0002_GPI3 (93%)	PMTB2_GEI1 (100%)
X73_GEI6	33.31	7	9.13	F216_ GEI4 (100%)	-	-	R0002_ GEI2 (100%)	PMTB2_ GEI6 (100%),
P1059_GEI1	41.05	8	8.53	-	F218_ GEI 2(100%)	-	-	-
P1059_GEI3	38.38	24	24	-	F218_ GEI4 (81.9%)	-	-	R1_GEI9(32.18%) , HB06_GEI3 (36.5%)
P1059_GEI4	36.91	7	10.6	F216_ GEI1 (63.3%)	-	X73_ GEI 2(63.4%)	-	-
P1059_GEI5	37.12	83	69.27	-	F218_ GEI 3(100%)	-	R0002_ GEI 4(10.4%)	PMTB_GEI3(10%); P1062_GEI1(14%), HN06_GEI6_ (21%), C0992_GEI9 (17%)
P1059_GEI7	41.27	9	10.95	F216_ GEI3 (100%)	-	X73_ GEI 4(100%)	-	-
R2_GEI5	32.89	6	6.76	-	-	-	-	PMTB_ GEI4 (100%)

- The GEIs of red color are unique to the group. The GEIs of black color are islands partially shared in the group or that share similarity partially with other groups.
- **pident**: Percent identity

Fonte: (Hurtado *et al.*, 2018)

S3 Table 6. Predicted islands of the genome X73.

Reference GEI X73	#START	#STOP
X73_GEI1	339169	351953
X73_GEI2	565638	573518
X73_GEI3	1066332	1073616
X73_GEI4	1525058	1536015
X73_GEI5	1560853	1570411
X73_GEI6	1603712	1612845
X73_GEI7	2156689	2198379

Fonte: (Hurtado *et al*., 2018)

S3Table 7. Predicted islands of the genome P1059.

Reference GEI P1059	#START	#STOP
P1059_GEI1	107623	116147
P1059_GEI2	269157	274606
P1059_GEI3	564124	588130
P1059_GEI4	590325	597759
P1059_GEI5	758990	819268
P1059_GEI6	1348162	1358728
P1059_GEI7	1608101	1619057
P1059_GEI8	1647399	1656852

Fonte: (Hurtado *et al*., 2018)

S3 Table 8. Predicted islands of the genome Razi_0002.

Reference GEI R2	#START	#STOP
R2_GEI1	457849	467401
R2_GEI2	500702	509834
R2_GEI3	1340328	1358346
R2_GEI4	1739520	1774910
R2_GEI5	1934945	1941708

Fonte: (Hurtado *et al*., 2018)

S3 Table 9. Predicted islands of the genome CIRMBP_0873 (C873).

LABEL	#START	#STOP
C873_GEI1	3	21667
C873_GEI2	346621	395990
C873_GEI3	651247	688242
C873_GEI4	1255813	1289149
C873_GEI5	2385993	2418620

Fonte: (Hurtado *et al*., 2018)

S3 Table 10. Predicted islands of the genome CIRMBP_0992 (C992).

LABEL	#START	#STOP
C992_GEI1	134514	189473
C992_GEI3	419429	441584
C992_GEI5	880921	913161
C992_GEI9	1692470	1741407
C992_GEI10	1761436	1779844

Fonte: (Hurtado *et al*., 2018)

S3 Table 11. Description and comparison of predicted Genomic Island of each Pneumonia–group.

Reference	GC (%)	CDs	Size (kb)	3480 (pident)	P1062 (pident)	UNMSM (pident)	36950 (pident)	HB01 (pident)	HB03 (pident)	Others members groups (pident)
Pm3_GEI1	41.07	45	31.4	-	P1062_GEI9(10.27 %)	UNMSM_GEI 8 (20%)	-	HB01_GEI8 (100%)	-	-
Pm3_GEI2	37.29	8	5.35	-	-	UNMSM_GEI 3 (92.9%)	36950_GE I3 (100%)	-	-	-
Pm3_GEI3	30.97	10	6.0	-	P1062_GEI8 (100%)	-	36950_GE I2 (100%)	HB01_GEI2 (100%)	HB03_GE I2 (100%)	-
Pm3_GEI4	30.89	6	7.74	-	-	-	36950_GE I5 (100%)	HB01_GEI3 (100%)	HB03_GE I3 (100%)	-
Pm3_GEI5	41.07	45	31.4	-	P1062_GEI9(10%)	UNMSM_GEI 8 (19.6%)	-	HB01_GEI8 (100%)	-	-
Pm3_GEI6	32.86	5	5.22	-	-	UNMSM _GEI6 (31.75%)	-	HB01_GEI5 (34%)	-	-
Pm3_GEI7	38.42	63	46.1	3480_GE I1 (20.2%)	P1062_GEI2 (48.12%)	UNMSM_GEI 8 (16.5%)	-	HB01_GEI1(64%)	-	HN06_GE I1 (19%), C0873_GE I2 (25.97%)
Pm3_GEI11	39.02	70	46.7	3480_GE I1 (9.2%)	P1062_GEI11 (29.15%),P1062_G EI6 (13.46%)	-	-	HB01_GEI5 (90.6%), HB01_GEI1(36.7 8%)	-	-
P1062_GEI1	39.95	42	31.7	-	-	UNMSM_GEI 13 (48.4%)	-	-	-	-
P1062_GEI5	36.73	7	8.57	-	-	-	36950_GE I6 (100%)	HB01_GEI4 (100%)	HB03_GE I4 (100%)	-
P1062_GEI9	40.79	29	32.11	-	-	UNMSM_GEI (100%)	-	-	-	-
P1062_GEI11	40.08	25	31.45	-	-	-	-	HB01_GEI5 (49.68%)	-	HN06_GE I6 (30.66%), C0992_GE I9 (26.7%), F218_GEI 3 (30.8%)
UNMSM_G EI7	39.36	43	31.71	-	P1062_GEI9 (98.4%)	-	-	-	-	-

- The GEIs of red color are unique to the group. The GEIs of black color are islands partially shared in the group or that share similarity partially with other groups.
- **pident**: Percent identity

Fonte: (Hurtado *et al*., 2018)

S3 Table 12. Predicted islands of the genome Pm-3.

REFERENCE	#START	#STOP
PM3_GEI1	182864	215842
PM3_GEI2	625869	631222
PM3_GEI3	895141	901141
PM3_GEI4	1132516	1140257
PM3_GEI5	1176078	1209903
PM3_GEI6	1424112	1429336
PM3_GEI7	1463038	1509613
PM3_GEI8	1799159	1807795
PM3_GEI9	2146789	2156242
PM3_GEI10	2159668	2169225
PM3_GEI11	2368606	2415312

Fonte: (Hurtado *et al*., 2018)

S3 Table 13. Predicted islands of the genome P1062.

REFERENCE	**#START**	**#STOP**
P1062_GEI1	134656	167436
P1062_GEI2	240664	292873
P1062_GEI3	714550	724003
P1062_GEI4	745258	783672
P1062_GEI5	1062784	1071361
P1062_GEI6	1181943	1225145
P1062_GEI7	1301837	1307043
P1062_GEI8	1582477	1588471
P1062_GEI9	1787839	1819957
P1062_GEI10	2305216	2336673
P1062_GEI11	2481611	2509016

Fonte: (Hurtado *et al*., 2018)

S3 Table 14. Predicted islands of the genome UNMSM.

REFERENCE	#START	#STOP
UNMSM_GEI1	258148	279516
UNMSM_GEI2	790931	799448
UNMSM_GEI3	846777	851950
UNMSM_GEI4	1057677	1070365
UNMSM_GEI5	1746728	1753572
UNMSM_GEI6	1758786	1769302
UNMSM_GEI7	1785541	1817250
UNMSM_GEI8	1863525	1868561
UNMSM_GEI9	1910585	1931354
UNMSM_GEI10	1953329	1965613
UNMSM_GEI11	1985391	1995841
UNMSM_GEI12	2030356	2035643
UNMSM_GEI13	2096073	2114029
UNMSM_GEI14	2246523	2251955

Fonte: (Hurtado *et al*., 2018)

S3 Table 15. Predicted islands of the genome 36950.

REFERENCE	#START	#STOP
36950_GEI1	266444	359969
36950_GEI2	569425	575426
36950_GEI3	839320	844679
36950_GEI4	1140019	1149376
36950_GEI5	1284272	1292014
36950_GEI6	1420951	1429528
36950_GEI7	1723676	1733231

Fonte: (Hurtado *et al*., 2018)

Supplementary Material 4

S4 Table 1. Description of GEI1 genes on R1 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S4 Table 2. Description of GEI2 genes on R1 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	OOHFKHOD_00214	50S ribosomal protein L31	210	J6CN71 50S ribosomal protein L31 OS=Pasteurella multocida subsp. multocida str. Anand1_buffalo GN=rpmE PE=3 SV=1	98.59% [70/71]	GO:0003735 - structural constituent of ribosome|GO:0005840 - ribosome|GO:0006412 - translation|GO:0019843 - rRNA binding|GO:0046872 - metal ion binding
2	OOHFKHOD_00213	Lipooligosaccharide biosynthesis protein lex-1	840	Beta-1,4 galactosyltransferase	98.59% [280/284]	GO:0016757 - transferase activity, transferring glycosyl groups
3	OOHFKHOD_00215	hypothetical protein	840	Beta-1,2 N-acetylglucosaminyltransferase	98.59% [280/284]	GO:0016757 - transferase activity, transferring glycosyl groups
4	OOHFKHOD_00217	hypothetical protein	882	D0EAD8 Alpha 1,6-LD-heptosyltransferase OS=Pasteurella multocida GN=hptE PE=4 SV=1	98.33% [294/299]	GO:0008152 - metabolic process|GO:0016757 - transferase activity, transferring glycosyl groups
5	OOHFKHOD_00218	Formamidopyrimidine-DNA glycosylase	810	A0A1D2NWD2 Formamidopyrimidine-DNA glycosylase OS=Pasteurella multocida GN=mutM PE=3 SV=1	98.54% [270/274]	GO:0003684 - damaged DNA binding|GO:0006284 - base-excision repair|GO:0006289 - nucleotide-excision repair|GO:0008270 - zinc ion binding|GO:0008534 - oxidized purine nucleobase lesion DNA N-glycosylase activity|GO:0140078 - class I DNA-(apurinic or apyrimidinic site) lyase activity|GO:0140080 - class III/IV DNA-(apurinic or apyrimidinic site) lyase activity
6	OOHFKHOD_00216	hypothetical protein	1056	Alpha 1,6-DD-heptosyltransferase	98.60% [352/357]	GO:0008152 - metabolic process|GO:0016757 - transferase activity, transferring glycosyl groups

Fonte: (Hurtado *et al*., 2018)

S4 Table 3. Description of GEI3 genes on R1 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	OOHFKHOD_00332		Blast error			
2	OOHFKHOD_00336	hypothetical protein	216	Uncharacterized protein	98.63% [72/73]	
3	OOHFKHOD_00337	hypothetical protein	213	Uncharacterized protein	98.61% [71/72]	
4	OOHFKHOD_00338	hypothetical protein	216	Uncharacterized protein	98.63% [72/73]	
5	OOHFKHOD_00339	hypothetical protein	147	Uncharacterized protein	84.48% [49/58]	GO:0006355 - regulation of transcription, DNA-templated
6	OOHFKHOD_00340	hypothetical protein	201	AlpA family transcriptional regulator	98.53% [67/68]	
7	OOHFKHOD_00335	hypothetical protein	423	Uncharacterized protein	98.60% [141/143]	
8	OOHFKHOD_00341	hypothetical protein	591	Uncharacterized protein	98.50% [197/200]	
9	OOHFKHOD_00342	hypothetical protein	462	Uncharacterized protein	98.72% [154/156]	
10	OOHFKHOD_00344	hypothetical protein	366	Terminase	98.39% [122/124]	
11	OOHFKHOD_00346	hypothetical protein	354	Uncharacterized protein	99.16% [118/119]	GO:0003676 - nucleic acid binding|GO:0004519 - endonuclease activity
12	OOHFKHOD_00347	hypothetical protein	324	Uncharacterized protein	99.08% [108/109]	
13	OOHFKHOD_00348	hypothetical protein	354	Phage head-tail adapter protein	99.16% [118/119]	
14	OOHFKHOD_00350	hypothetical protein	552	Peptidase	98.40% [184/187]	
15	OOHFKHOD_00352	hypothetical protein	405	Phage tail protein	98.54% [135/137]	
16	OOHFKHOD_00353	hypothetical protein	261	Transglycosylase	98.86% [87/88]	GO:0016021 - integral component of membrane
17	OOHFKHOD_00333	Prophage integrase IntA	1197	Preprotein translocase	98.52% [399/405]	GO:0003677 - DNA binding|GO:0006310 - DNA recombination|GO:0015074 - DNA integration
18	OOHFKHOD_00334	hypothetical protein	1602	DNA primase	98.52% [534/542]	
19	OOHFKHOD_00343	hypothetical protein	1665	Terminase	98.40% [555/564]	GO:0003779 - actin binding
20	OOHFKHOD_00345	hypothetical protein	2727	Uncharacterized protein	98.38% [909/924]	
21	OOHFKHOD_00349	hypothetical protein	1209	Phage portal protein	98.53% [403/409]	
22	OOHFKHOD_00351	hypothetical protein	1197	Capsid protein	98.52% [399/405]	

Fonte: (Hurtado *et al*., 2018)

S4 Table 4. Description of GEI7 genes on R1 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	OOHFKHOD_01072	hypothetical protein	738	Uncharacterized protein	98.40% [246/250]	
2	OOHFKHOD_01074	putative metal-dependent hydrolase YjjV	729	Hydrolase TatD	98.38% [243/247]	GO:0016888 - endodeoxyribonuclease activity, producing 5'-phosphomonoesters
3	OOHFKHOD_01071	hypothetical protein	1764	KAP family P-loop domain protein	98.49% [588/597]	
4	OOHFKHOD_01073	7-cyano-7-deazaguanine synthase	1311	DNA-binding protein	97.11% [437/450]	GO:0003677 - DNA binding|GO:0005524 - ATP binding|GO:0016874 - ligase activity

Fonte: (Hurtado *et al*., 2018)

S4 Table 5. Description of GEI10 genes on R1 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S4 Table 6. Description of GEI11 genes on R1 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	OOHFKHOD_02015	hypothetical protein	708	Q9L9L0 BcbE OS=Pasteurella multocida GN=bcbE PE=4 SV=1	98.74% [236/239]	
2	OOHFKHOD_02016	hypothetical protein	372	Q9L9L1 BcbF OS=Pasteurella multocida GN=bcbF PE=4 SV=1	98.41% [124/126]	
3	OOHFKHOD_02019	Spore protein YkvP	978	BcbI	98.49% [326/331]	
4	OOHFKHOD_02011	UDP-N-acetylglucosamine 2-epimerase	1056	Q9L9K6 BcbA OS=Pasteurella multocida GN=bcbA PE=3 SV=1	98.05% [352/359]	GO:0008761 - UDP-N-acetylglucosamine 2-epimerase activity
5	OOHFKHOD_02012	UDP-N-acetyl-D-mannosamine dehydrogenase	1266	Q9L9K7 BcbB OS=Pasteurella multocida GN=bcbB PE=3 SV=1	98.37% [422/429]	GO:0000271 - polysaccharide biosynthetic process|GO:0016616 - oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor|GO:0016628 - oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor|GO:0051287 - NAD binding
6	OOHFKHOD_02013	putative glycosyltransferase EpsJ	3099	Uncharacterized protein	98.38% [1033/1050]	
7	OOHFKHOD_02014	hypothetical protein	1425	BcbD	98.55% [475/482]	
8	OOHFKHOD_02017	hypothetical protein	1563	BcbG	98.49% [521/529]	
9	OOHFKHOD_02018	hypothetical protein	2163	BcbH	98.36% [721/733]	
10	OOHFKHOD_02020	hypothetical protein	2046	LipA	98.41% [682/693]	GO:0000271 - polysaccharide biosynthetic process|GO:0015774 - polysaccharide transport

Fonte: (Hurtado *et al*., 2018)

S4 Table 7. Description of GEI4 genes on PMTB2.1 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	IAKPAGJH_00859	hypothetical protein	270	A0A292DLE1 Putative addiction module antidote protein OS=Pasteurella multocida GN=CRN72_01610 PE=4 SV=1	91.84% [90/98]	
2	IAKPAGJH_00860		Blast error			
3	IAKPAGJH_00861	hypothetical protein	225	ATP-dependent RNA helicase HrpA	81.52% [75/92]	GO:0004386 - helicase activity
4	IAKPAGJH_00862	hypothetical protein	333	Uncharacterized protein	73.51% [111/151]	
5	IAKPAGJH_00863	ATPase RavA	1203	ATPase family protein	74.67% [401/537]	GO:0005524 - ATP binding|GO:0016887 - ATPase activity
6	IAKPAGJH_00864	Protein ViaA	1260	Uncharacterized protein	86.42% [420/486]	

Fonte: (Hurtado *et al*., 2018)

S4 Table 8. Description of GEI6 genes on PMTB2.1 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	IAKPAGJH_01633	Arabinose operon regulatory protein	855	Arabinose operon transcriptional regulator AraC	98.62% [285/289]	
2	IAKPAGJH_01634	Autoinducer 2-binding protein LsrB	972	Autoinducer 2 ABC transporter substrate-binding protein	98.48% [324/329]	
3	IAKPAGJH_01636	Ribose import permease protein RbsC	957	ABC transporter permease	98.46% [319/324]	
4	IAKPAGJH_01637	Autoinducer 2 import system permease protein LsrD	945	ABC transporter permease	98.44% [315/320]	
5	IAKPAGJH_01635	Arabinose import ATP-binding protein AraG	1485	Sugar ABC transporter ATP-binding protein	98.41% [495/503]	GO:0005524 - ATP binding
6	IAKPAGJH_01638	Ribulokinase	1593	ATPase	98.52% [531/539]	
7	IAKPAGJH_01639	L-arabinose isomerase	1485	A0A290LTP6 L-arabinose isomerase OS=Pasteurella multocida GN=CLD33_01210 PE=4 SV=1	98.41% [495/503]	GO:0016853 - isomerase activity

Fonte: (Hurtado *et al*., 2018)

S4 Table 9. Description of GEI6 genes on Islm strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

Supplementary Material 5

S5 Table 1. Description of GEI1 genes on X73 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
2	FBFOMBLG_00315	hypothetical protein	99	Glycosyl transferase family 1	47.83% [33/69]	GO:0016740 - transferase activity
3	FBFOMBLG_00318	UDP-glucose 6-dehydrogenase	141	Nucleotide sugar dehydrogenase	68.12% [47/69]	GO:0000271 - polysaccharide biosynthetic process|GO:0003979 - UDP-glucose 6-dehydrogenase activity|GO:0051287 - NAD binding
4	FBFOMBLG_00319		75	Uncharacterized protein	89.29% [25/28]	
5	FBFOMBLG_00304	Tagatose 1,6-diphosphate aldolase	717	Tagatose 1,6-diphosphate aldolase	74.22% [239/322]	GO:0009024 - tagatose-6-phosphate kinase activity|GO:0016829 - lyase activity|GO:0019512 - lactose catabolic process via tagatose-6-phosphate
6	FBFOMBLG_00306	PTS system galactitol-specific EIIA component	330	PTS system galactitol-specific transporter subunit IIA	71.90% [110/153]	GO:0016301 - kinase activity
7	FBFOMBLG_00307	PTS system galactitol-specific EIIB component	231	PTS system galactitol-specific transporter subunit IIB	81.05% [77/95]	GO:0008982 - protein-N(PI)-phosphohistidine-sugar phosphotransferase activity|GO:0009401 - phosphoenolpyruvate-dependent sugar phosphotransferase system
8	FBFOMBLG_00310	Glucitol operon repressor	579	DeoR family transcriptional regulator	75.39% [193/256]	GO:0003677 - DNA binding|GO:0003700 - DNA binding transcription factor activity|GO:0005622 - intracellular|GO:0006351 - transcription, DNA-templated
9	FBFOMBLG_00311	hypothetical protein	297	Uncharacterized protein	75.57% [99/131]	GO:0016758 - transferase activity, transferring hexosyl groups
10	FBFOMBLG_00312	hypothetical protein	213	Uncharacterized protein	82.56% [71/86]	GO:0016758 - transferase activity, transferring hexosyl groups
11	FBFOMBLG_00313	hypothetical protein	651	Glycosyltransferase, group 1 family protein	78.91% [217/275]	GO:0016757 - transferase activity, transferring glycosyl groups
12	FBFOMBLG_00314	N-acetyl-alpha-D-glucosaminyl L-malate synthase	285	Glycosyltransferase, group 1 family protein	75.40% [95/126]	GO:0016757 - transferase activity, transferring glycosyl groups
13	FBFOMBLG_00316	hypothetical protein	357	Uncharacterized protein	57.49% [119/207]	
14	FBFOMBLG_00317	UDP-glucose 6-dehydrogenase	639	Nucleotide sugar dehydrogenase	78.60% [213/271]	GO:0000271 - polysaccharide biosynthetic process|GO:0003979 - UDP-glucose 6-dehydrogenase activity|GO:0051287 - NAD binding
15	FBFOMBLG_00305	D-tagatose-1,6-bisphosphate aldolase subunit GatZ	1005	D-tagatose-1,6-bisphosphate aldolase subunit GatZ	76.66% [335/437]	GO:0003824 - catalytic activity|GO:0019402 - galactitol metabolic process
16	FBFOMBLG_00308	PTS system galactitol-specific EIIC component	1164	PTS system, galactitol-specific IIC component	84.72% [388/458]	GO:0009401 - phosphoenolpyruvate-dependent sugar phosphotransferase system|GO:0015577 - galactitol transmembrane transporter activity|GO:0016021 - integral component of membrane

17	FBFOMBLG_00309	Galactitol-1-phosphate 5-dehydrogenase	780	Galactitol-1-phosphate dehydrogenase	73.86% [260/352]	GO:0008270 - zinc ion binding|GO:0016491 - oxidoreductase activity

Fonte: (Hurtado *et al*., 2018)

S5 Table 2. Description of GEI3 genes on X73 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	FBFOMBLG_00952	Ribose-5-phosphate isomerase B	441	Ribose 5-phosphate isomerase B	98.66% [147/149]	GO:0016853 - isomerase activity
2	FBFOMBLG_00953	D-allose kinase	891	Allose kinase	98.02% [297/303]	GO:0016301 - kinase activity
3	FBFOMBLG_00954	D-allulose-6-phosphate 3-epimerase	666	Allulose-6-phosphate 3-epimerase	98.23% [222/226]	
4	FBFOMBLG_00955	D-allose transport system permease protein AlsC	942	Allose ABC transporter	98.13% [314/320]	
5	FBFOMBLG_00957	D-allose-binding periplasmic protein	930	Allose ABC transporter	97.79% [310/317]	
6	FBFOMBLG_00958	HTH-type transcriptional regulator RpiR	849	MurR/RpiR family transcriptional regulator	98.26% [283/288]	
7	FBFOMBLG_00956	Galactose/methyl galactoside import ATP-binding protein MglA	1518	Allose ABC transporter ATP-binding protein	98.25% [506/515]	GO:0005524 - ATP binding

Fonte: (Hurtado *et al*., 2018)

S5 Table 3. Description of GEI1 genes on P1059 strain.

0	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	EHKLGOCA_00082	Citrate lyase acyl carrier protein	216	A0A179CY03 Citrate lyase acyl carrier protein OS=Bibersteinia trehalosi Y31 GN=citD PE=3 SV=1	74.23% [72/97]	GO:0005737 - cytoplasm|GO:0016829 - lyase activity
2	EHKLGOCA_00083	Citrate lyase subunit beta	756	Citrate lyase subunit beta	85.42% [252/295]	GO:0006084 - acetyl-CoA metabolic process|GO:0008816 - citryl-CoA lyase activity|GO:0009346 - citrate lyase complex|GO:0046872 - metal ion binding
3	EHKLGOCA_00087	hypothetical protein	357	Uncharacterized protein	75.80% [119/157]	GO:0016021 - integral component of membrane
4	EHKLGOCA_00088	hypothetical protein	402	Uncharacterized protein	86.45% [134/155]	GO:0016021 - integral component of membrane
5	EHKLGOCA_00081	[Citrate [pro-3S]-lyase] ligase	1086	A0A1E3XHX0 [Citrate [pro-3S]-lyase] ligase OS=Pasteurella multocida GN=BGK37_10525 PE=4 SV=1	98.10% [362/369]	GO:0005524 - ATP binding|GO:0008080 - N-acetyltransferase activity|GO:0008771 - |GO:0009058 - biosynthetic process|GO:0016829 - lyase activity
6	EHKLGOCA_00084	Citrate lyase alpha chain	1290	A0A1V3K0A8 Citrate lyase alpha chain OS=Rodentibacter heylii GN=BKG89_10325 PE=4 SV=1	84.15% [430/511]	GO:0006084 - acetyl-CoA metabolic process|GO:0008814 - citrate CoA-transferase activity|GO:0008815 - citrate (pro-3S)-lyase activity|GO:0009346 - citrate lyase complex
7	EHKLGOCA_00085	2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase	768	A6VNL5 Probable 2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase	52.24% [256/490]	GO:0005524 - ATP binding|GO:0016310 - phosphorylation|GO:0016757 - transferase activity, transferring glycosyl groups|GO:0016779 - nucleotidyltransferase

				OS=Actinobacillus succinogenes (strain ATCC 55618 / 130Z) GN=citG PE=3 SV=1		activity|GO:0046917 - triphosphoribosyl-dephospho-CoA synthase activity|GO:0051191 - prosthetic group biosynthetic process
8	EHKLGOCA_00086	Inner membrane protein YbhI	1227	Uncharacterized protein	84.16% [409/486]	GO:0005215 - transporter activity|GO:0006814 - sodium ion transport|GO:0016021 - integral component of membrane|GO:0055085 - transmembrane transport

Fonte: (Hurtado *et al*., 2018)

S5 Table 4. Description of GEI2 genes on P1059 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	EHKLGOCA_00249	hypothetical protein	0		0.00% [0/45]	
2	EHKLGOCA_00248	hypothetical protein	414	Uncharacterized protein	89.03% [138/155]	GO:0016021 - integral component of membrane
3	EHKLGOCA_00250	hypothetical protein	573	Uncharacterized protein	90.52% [191/211]	GO:0016021 - integral component of membrane
4	EHKLGOCA_00251	hypothetical protein	558	Type I secretion target GGXGXDXXX repeat-containing domain protein	35.29% [186/527]	
5	EHKLGOCA_00252	Bifunctional hemolysin/adenylate cyclase	810	RTX toxin protein	46.55% [270/580]	

Fonte: (Hurtado *et al*., 2018)

S5 Table 5. Description of GEI7 genes on P1059 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	EHKLGOCA_01504	Glycerol-3-phosphate regulon repressor	738	DeoR family transcriptional regulator	97.23% [246/253]	GO:0003677 - DNA binding|GO:0003700 - DNA binding transcription factor activity|GO:0005622 - intracellular|GO:0006351 - transcription, DNA-templated
2	EHKLGOCA_01507	L-fucose mutarotase	399	A0A1W9FTA2 L-fucose mutarotase OS=Pasteurella multocida GN=fucU PE=3 SV=1	91.10% [133/146]	GO:0005737 - cytoplasm|GO:0016854 - racemase and epimerase activity|GO:0042354 - L-fucose metabolic process|GO:0042806 - fucose binding
3	EHKLGOCA_01508	L-fuculose phosphate aldolase	636	A0A1W9FT98 L-fuculose phosphate aldolase OS=Pasteurella multocida GN=fucA PE=3 SV=1	97.25% [212/218]	GO:0008270 - zinc ion binding|GO:0008738 - L-fuculose-phosphate aldolase activity|GO:0042355 - L-fucose catabolic process
4	EHKLGOCA_01511	Ribose import binding protein RbsB	954	A0A292DN35 LacI family transcriptional regulator OS=Pasteurella multocida GN=CRN72_07240 PE=4 SV=1	98.45% [318/323]	

5	EHKLGOCA_01505	L-fucose isomerase	1755	A0A292DP21 L-fucose isomerase OS=Pasteurella multocida GN=fucI PE=4 SV=1	98.32% [585/595]	
6	EHKLGOCA_01506	L-fuculokinase	1398	A0A1W9FT64 L-fuculokinase OS=Pasteurella multocida GN=fucK PE=3 SV=1	92.64% [466/503]	GO:0005524 - ATP binding|GO:0008737 - L-fuculokinase activity|GO:0042355 - L-fucose catabolic process
7	EHKLGOCA_01509	Galactose/methyl galactoside import ATP-binding protein MglA	1500	A0A1W9FSW4 Galactose/methyl galactoside import ATP-binding protein MglA OS=Pasteurella multocida GN=mglA PE=3 SV=1	96.90% [500/516]	GO:0005354 - galactose transmembrane transporter activity|GO:0005524 - ATP binding|GO:0005886 - plasma membrane|GO:0015407 - monosaccharide-transporting ATPase activity
8	EHKLGOCA_01510	Ribose import permease protein RbsC	1026	Ribose ABC transporter permease	98.56% [342/347]	GO:0005215 - transporter activity|GO:0005886 - plasma membrane|GO:0006810 - transport|GO:0016021 - integral component of membrane
9	EHKLGOCA_01512	Lactaldehyde dehydrogenase	1425	Aldehyde dehydrogenase	97.54% [475/487]	GO:0016491 - oxidoreductase activity

Fonte: (Hurtado *et al*., 2018)

Supplementary Material 6

S6 Table 1. Description of GEI1 genes on Pm-3 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S6 Table 2. Description of GEI2 genes on Pm-3 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	BKFAJCHN_00576	hypothetical protein	288	Uncharacterized protein	98.97% [96/97]	
2	BKFAJCHN_00577	hypothetical protein	378	Uncharacterized protein	98.44% [126/128]	
3	BKFAJCHN_00578	Inner membrane protein RclC	636	Membrane protein	98.60% [212/215]	GO:0016021 - integral component of membrane
4	BKFAJCHN_00579	Thiosulfate sulfurtransferase PspE	351	Rhodanese-like domain-containing protein	99.15% [117/118]	
5	BKFAJCHN_00580	hypothetical protein	339	A0A1D7R1Q7 Alkyl hydroperoxide reductase AhpD OS=Pasteurella multocida GN=AZI96_02845 PE=3 SV=1	99.12% [113/114]	GO:0004601 - peroxidase activity|GO:0051920 - peroxiredoxin activity
6	BKFAJCHN_00581	RCS-specific HTH-type transcriptional activator RclR	867	AraC family transcriptional regulator	98.63% [289/293]	GO:0003700 - DNA binding transcription factor activity|GO:0006351 - transcription, DNA-templated|GO:0043565 - sequence-specific DNA binding

7	BKFAJCHN_00582	hypothetical protein	651	Uncharacterized protein	74.32% [217/292]	
8	BKFAJCHN_00575	hypothetical protein	180	Uncharacterized protein	43.48% [60/138]	

Fonte: (Hurtado *et al*., 2018)

S6 Table 3. Description of GEI3 genes on Pm-3 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	BKFAJCHN_00812	hypothetical protein	0		0.00% [0/39]	
2	BKFAJCHN_00813	hypothetical protein	204	Uncharacterized protein	98.55% [68/69]	GO:0016021 - integral component of membrane
3	BKFAJCHN_00808	hypothetical protein	402	Uncharacterized protein	98.53% [134/136]	GO:0016021 - integral component of membrane
4	BKFAJCHN_00809	hypothetical protein	441	Uncharacterized protein	98.66% [147/149]	
5	BKFAJCHN_00810	hypothetical protein	384	Uncharacterized protein	98.46% [128/130]	
6	BKFAJCHN_00811	hypothetical protein	513	Uncharacterized protein	98.84% [171/173]	
7	BKFAJCHN_00814	hypothetical protein	369	Uncharacterized protein	98.40% [123/125]	
8	BKFAJCHN_00815	hypothetical protein	351	Uncharacterized protein	99.15% [117/118]	
9	BKFAJCHN_00816	hypothetical protein	507	Uncharacterized protein	98.83% [169/171]	
10	BKFAJCHN_00817	hypothetical protein	351	Uncharacterized protein	99.15% [117/118]	

Fonte: (Hurtado *et al*., 2018)

S6 Table 4. Description of GEI4 genes on Pm-3 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	BKFAJCHN_01036	hypothetical protein	441	Uncharacterized protein	98.66% [147/149]	
2	BKFAJCHN_01038	hypothetical protein	558	Uncharacterized protein	98.41% [186/189]	
3	BKFAJCHN_01035	hypothetical protein	1095	Uncharacterized protein	97.33% [365/375]	
4	BKFAJCHN_01037	hypothetical protein	2004	Uncharacterized protein	98.38% [668/679]	
5	BKFAJCHN_01039	hypothetical protein	1269	Uncharacterized protein	98.37% [423/430]	

6	BKFAJCHN_01040	hypothetical protein	1740	Uncharacterized protein	98.47% [580/589]	
7	BKFAJCHN_01034	hypothetical protein	0		0.00% [0/77]	

Fonte: (Hurtado *et al*., 2018)

S6 Table 5. Description of GEI5 genes on Pm-3 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S6 Table 6. Description of GEI6 genes on Pm-3 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	BKFAJCHN_01331	hypothetical protein	207	Uncharacterized protein	98.57% [69/70]	
2	BKFAJCHN_01330	hypothetical protein	414	Transposase	98.57% [138/140]	GO:0003677 - DNA binding|GO:0004803 - transposase activity|GO:0006313 - transposition, DNA-mediated
3	BKFAJCHN_01333	hypothetical protein	570	Uncharacterized protein	98.45% [190/193]	
4	BKFAJCHN_01329	hypothetical protein	1134	Transposase	98.44% [378/384]	
5	BKFAJCHN_01332	hypothetical protein	1236	Uncharacterized protein	98.56% [412/418]	

Fonte: (Hurtado *et al*., 2018)

S6 Table 7. Description of GEI7 genes on Pm-3 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S6 Table 8. Description of GEI11 genes on Pm-3 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S6 Table 9. Description of GEI7 genes on 3480 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	FDIJDFMK_01749	hypothetical protein	720	Virulence protein	70.59% [240/340]	

2	FDIJDFMK_01747	hypothetical protein	603	Anticodon nuclease	43.51% [201/462]	
3	FDIJDFMK_01748	hypothetical protein	654	Type I restriction enzyme EcoR124II specificity protein	57.98% [218/376]	GO:0003677 - DNA binding|GO:0006304 - DNA modification
4	FDIJDFMK_01750	putative type I restriction enzymeP M protein	1254	Type I restriction-modification system subunit M	79.47% [418/526]	GO:0003677 - DNA binding|GO:0008170 - N-methyltransferase activity|GO:0009007 - site-specific DNA-methyltransferase (adenine-specific) activity|GO:0009307 - DNA restriction-modification system
5	FDIJDFMK_01746	Type I restriction enzyme EcoR124II R protein	2430	A0A1A7NTN1 Type I restriction enzyme R Protein OS=Gallibacterium genomosp. 3 GN=QV01_03605 PE=3 SV=1	77.59% [810/1044]	GO:0003677 - DNA binding|GO:0005524 - ATP binding|GO:0009035 - Type I site-specific deoxyribonuclease activity|GO:0009307 - DNA restriction-modification system

Fonte: (Hurtado *et al*., 2018)

S6 Table 10. Description of GEI8 genes on 3480 strain.

#	Locus tag	Prokka annotation	Length	Product	Completeness	Gene onthology
1	FDIJDFMK_02028	hypothetical protein	243	Transposase	80.20% [81/101]	
2	FDIJDFMK_02029	hypothetical protein	243	Transposase	72.32% [81/112]	
3	FDIJDFMK_02030	hypothetical protein	189	Uncharacterized protein	38.18% [63/165]	
4	FDIJDFMK_02032	hypothetical protein	249	Transcriptional regulator	72.17% [83/115]	GO:0043565 - sequence-specific DNA binding
5	FDIJDFMK_02033	hypothetical protein	402	DDE transposase	62.62% [134/214]	
6	FDIJDFMK_02034	hypothetical protein	285	ISSst2, family IS1595	76.00% [95/125]	
7	FDIJDFMK_02035	hypothetical protein	204	DNA repair protein	45.95% [68/148]	
8	FDIJDFMK_02036	hypothetical protein	0		0.00% [0/139]	
9	FDIJDFMK_02031	hypothetical protein	618	ATP-binding protein	46.40% [206/444]	GO:0005524 - ATP binding

Fonte: (Hurtado *et al*., 2018)

S6 Table 11. Description of GEI1 genes on P1062 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S6 Table 12. Description of GEI5 genes on P1062 strain.

#	Locus tag	Prokka	Length	Product	Completeness	Gene onthology
1	FAKKMGJE_01094	Hypothetical protein	354	Uncharacterized protein	99.16% [118/119]	
2	FAKKMGJE_01096	Quercetin 2,3-dioxygenase	894	Quercetin 2,3-dioxygenase	98.68% [298/302]	GO:0051213 - dioxygenase activity
3	FAKKMGJE_01097	Chromate reductase	528	NADPH-dependent FMN reductase	98.88% [176/178]	GO:0016491 - oxidoreductase activity
4	FAKKMGJE_01093	PTS system trehalose-specific EIIBC component	1416	PTS maltose transporter subunit IIBC	98.54% [472/479]	GO:0005886 - plasma membrane|GO:0008982 - protein-N(PI)-phosphohistidine-sugar phosphotransferase activity|GO:0009401 - phosphoenolpyruvate-dependent sugar phosphotransferase system|GO:0015574 - trehalose transmembrane transporter activity|GO:0016021 - integral component of membrane|GO:0016301 - kinase activity
5	FAKKMGJE_01095	Trehalose-6-phosphate hydrolase	1638	Alpha,alpha-phosphotrehalase	98.38% [546/555]	GO:0005737 - cytoplasm|GO:0005993 - trehalose catabolic process|GO:0008788 - alpha,alpha-phosphotrehalase activity
6	FAKKMGJE_01098	Glutathione-regulated potassium-efflux system protein KefC	1800	Potassium transporter	98.52% [600/609]	GO:0006813 - potassium ion transport|GO:0015299 - solute:proton antiporter activity|GO:0016021 - integral component of membrane
7	FAKKMGJE_01092	HTH-type transcriptional regulator TreR	942	Trehalose repressor	98.43% [314/319]	GO:0003677 - DNA binding|GO:0005991 - trehalose metabolic process|GO:0006351 - transcription, DNA-templated|GO:0045892 - negative regulation of transcription, DNA-templated

Fonte: (Hurtado *et al*., 2018)

S6 Table 13. Description of GEI9 genes on P1062 strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S6 TABLE 14. Description of GEI7 genes on UNMSM strain. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

Supplementary Material 7

S7 Table 1. Complete results for positive selection analysis using the target group BovinePneumonia. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S7 Table 2. Complete results for positive selection analysis using the target group Snuffles. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S7 Table 3. Complete results for positive selection analysis using the target group SerotypeB. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S7 Table 4. Complete results for positive selection analysis using the target group SerotypeF. **Fonte:** (Hurtado *et al*., 2018)

Disponível em https://www.dropbox.com/sh/ybzwijizlck8mh1/AABt_jRnX9MFGhogxz1jm_cLa?dl=0

S7 Table 5. Recombination analysis for positively selected genes in *Pasteurella multocida*.

Locus	**NSS**	**Chi2**	**Phipack**
CIRMBP0884_626	0	0	0
CIRMBP0884_1043	0.002	0.091	0.0108
CIRMBP0884_803	0.066	0.009	0.162
CIRMBP0884_1057	0	0	0.398
CIRMBP0884_300	0.002	0.002	0.825
CIRMBP0884_1860	0	0	3.08E-23
CIRMBP0884_78	0	0	3.06E-13
CIRMBP0884_1954	0.001	0	0.00134
CIRMBP0884_360	0	0	0.00000231
CIRMBP0884_107	0.054	0.207	0.0888
CIRMBP0884_1277	0	0	1.49E-15
CIRMBP0884_1278	1	0.001	1

Fonte: (Hurtado *et al*., 2018)

S7 Table 6. Genes and number of sites under positive selection in different lineages of *Pasteurella multocida*.

Gene	Product	FDR	ω	PS sites	Localization	Foreground	RASTtk ID	GenBankID
purL	*Phosphoribosylformylglycinamidine synthase	2.40E-14	12.91	17	NA	BovinePneumonia	PmCIRMBP0884_626	BTV67_03100
talA	*Transaldolase	5.70E-03	5.51	2	CP	BovinePneumonia	PmCIRMBP0884_1043	BTV67_05150
-	Hypothetical protein (Alpha/beta hydrolase domain)	3.70E-10	24.29	7	NA	SerotypeB	PmCIRMBP0884_803	BTV67_03975
murB	*UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.3.1.98)	1.00E-05	13.57	5	NA	SerotypeB	PmCIRMBP0884_1057	BTV67_05215
lolE	*Lipoprotein releasing system transmembrane protein LolE	1.30E-04	14.18	7	CM	SerotypeB	PmCIRMBP0884_300	BTV67_01460
-	*Hypothetical protein (Membrane anchored)	3.90E-04	30.37	12	NA	SerotypeB	PmCIRMBP0884_1860	BTV67_09135
hrpA	*ATP-dependent RNA helicase	7.60E-04	8.83	9	CP	SerotypeB	PmCIRMBP0884_78	BTV67_00380
ilvB	*Acetolactate synthase large subunit (EC 2.2.1.6)	2.10E-03	7.2	3	CP	SerotypeB	PmCIRMBP0884_1954	BTV67_09595
-	*Hypothetical protein	3.30E-03	2.35	14	NA	SerotypeB	PmCIRMBP0884_360	BTV67_01755
cirA	TonB-dependent hemin, ferrichrome receptor	3.70E-02	4.13	4	OM	SerotypeB	PmCIRMBP0884_107	BTV67_00535
cI	*Bacteriophage P4, Ash protein	0.0257	9.2	1	NA	SerotypeF	PmCIRMBP0884_1277	BTV67_06300

* - recombinant, PS - positively selected sites, CP - cytoplasm, CM - cytoplasmic membrane, OM - outer membrane, NA – unknown.

Fonte: (Hurtado *et al*., 2018)

REFERÊNCIAS

Brüssow, H., Kutter, E., 2004. Phage Ecology, in: Bacteriophages. Biology and Applications. doi:10.1071/EG04251

Cao, P., Guo, D., Liu, J., Jiang, Q., Xu, Z., Qu, L., 2017. Genome-wide analyses reveal genes subject to positive selection in Pasteurella multocida. Front. Microbiol. 8. doi:10.3389/fmicb.2017.00961

Cardoso-Toset, F., Gómez-Laguna, J., Callejo, M., Vela, A.I., Carrasco, L., Fernández-Garayzábal, J.F., Maldonado, A., Luque, I., 2013. Septicaemic pasteurellosis in free-range pigs associated with an unusual biovar 13 of Pasteurella multocida. Vet. Microbiol. doi:10.1016/j.vetmic.2013.08.005

Cases, I., De Lorenzo, V., Ouzounis, C.A., 2003. Transcription regulation and environmental adaptation in bacteria. Trends Microbiol. doi:10.1016/S0966-842X(03)00103-3

Chen, S.L., Hung, C.-S., Xu, J., Reigstad, C.S., Magrini, V., Sabo, A., Blasiar, D., Bieri, T., Meyer, R.R., Ozersky, P., Armstrong, J.R., Fulton, R.S., Latreille, J.P., Spieth, J., Hooton, T.M., Mardis, E.R., Hultgren, S.J., Gordon, J.I., 2006. Identification of genes subject to positive selection in uropathogenic strains of Escherichia coli: A comparative genomics approach. Proc. Natl. Acad. Sci. doi:10.1073/pnas.0600938103

Christensen, H., Bisgaard, M., 2010. Molecular classification and its impact on diagnostics and understanding the phylogeny and epidemiology of selected members of Pasteurellaceoe of veterinary importance. Berl. Munch. Tierarztl. Wochenschr. doi:10.2376/0005-9366-123-20

Dabo, S.M., Taylor, J.D., Confer, a W., 2007. Pasteurella multocida and bovine respiratory disease. Anim. Health Res. Rev. doi:10.1017/S1466252307001399

Davies, R.L., MacCorquodale, R., Reilly, S., 2004. Characterisation of bovine strains of Pasteurella multocida and comparison with isolates of avian, ovine and porcine origin. Vet. Microbiol. doi:10.1016/j.vetmic.2003.11.013

Du, H., Fang, R., Pan, T., Li, T., Li, N., He, Q., Wu, R., Peng, Y., Zhou, Z., 2016. Comparative Genomics Analysis of Two Different Virulent Bovine *Pasteurella multocida* Isolates. Int. J. Genomics 2016, 1–14. doi:10.1155/2016/4512493

Ewers, C., Lübke-Becker, A., Bethe, A., Kießling, S., Filter, M., Wieler, L.H., 2006. Virulence genotype of Pasteurella multocida strains isolated from different hosts with various disease status. Vet. Microbiol. 114, 304–317. doi:10.1016/j.vetmic.2005.12.012

Fitzpatrick, D.A., Creevey, C.J., McInerney, J.O., 2005. Evidence of positive Darwinian selection in putative meningococcal vaccine antigens. J. Mol. Evol. doi:10.1007/s00239-004-0290-6

Ford, C.B., Shah, R.R., Maeda, M.K., Gagneux, S., Murray, M.B., Cohen, T., Johnston, J.C., Gardy, J., Lipsitch, M., Fortune, S.M., 2013. Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. Nat. Genet. doi:10.1038/ng.2656

García-Alvarez, A., Chaves, F., Fernández, A., Sanz, C., Borobia, M., Cid, D., 2015. An ST11 clone of Pasteurella multocida, widely spread among farmed rabbits in the Iberian Peninsula, demonstrates respiratory niche association. Infect. Genet. Evol. doi:10.1016/j.meegid.2015.07.018

García-Alvarez, A., Vela, A.I., San Martín, E., Chaves, F., Fernández-Garayzábal, J.F., Lucas, D., Cid, D., 2017. Characterization of Pasteurella multocida associated with ovine pneumonia using multi-locus sequence typing (MLST) and virulence-associated gene profile analysis and comparison with porcine isolates. Vet. Microbiol. doi:10.1016/j.vetmic.2017.04.015

Guo, D., Lu, Y., Zhang, A., Liu, J., Yuan, D., Jiang, Q., Lin, H., Si, C., Qu, L., 2012. Identification of genes transcribed by Pasteurella multocida in rabbit livers through the selective capture of transcribed sequences. FEMS Microbiol. Lett. doi:10.1111/j.1574-6968.2012.02559.x

Harms, N.E., 2015. DYNAMICS OF DISEASE: ORIGINS AND ECOLOGY OF AVIAN CHOLERA IN THE EASTERN CANADIAN ARCTIC. University of Saskatchewan.

Harper, M., Boyce, J.D., Adler, B., 2006. Pasteurella multocida pathogenesis: 125 Years after

Pasteur. FEMS Microbiol. Lett. doi:10.1111/j.1574-6968.2006.00442.x
Hotchkiss, E.J., Hodgson, J.C., Lainson, F.A., Zadoks, R.N., 2011. Multilocus sequence typing of a global collection of Pasteurella multocida isolates from cattle and other host species demonstrates niche association. BMC Microbiol. doi:10.1186/1471-2180-11-115
Hurtado, R., Carhuaricra, D., Soares, S., Viana, M.V.C., Azevedo, V., Maturrano, L., Aburjaile, F., 2018. Pan-genomic approach shows insight of genetic divergence and pathogenic-adaptation of Pasteurella multocida. Gene. doi:10.1016/j.gene.2018.05.084
Johnson, T.J., Abrahante, J.E., Hunter, S.S., Hauglund, M., Tatum, F.M., Maheswaran, S.K., Briggs, R.E., 2013. Comparative genome analysis of an avirulent and two virulent strains of avian Pasteurella multocida reveals candidate genes involved in fitness and pathogenicity. BMC Microbiol. 13, 1. doi:10.1186/1471-2180-13-106
Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., Jelsbak, L., Sicheritz-Pontén, T., Ussery, D.W., Aarestrup, F.M., Lund, O., 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J. Clin. Microbiol. doi:10.1128/JCM.06094-11
Li, Z., Cheng, F., Lan, S., Guo, J., Liu, W., Li, X., Luo, Z., Zhang, M., Wu, J., Shi, Y., 2018. Investigation of genetic diversity and epidemiological characteristics of Pasteurella multocida isolates from poultry in southwest China by population structure, multi-locus sequence typing and virulence-associated gene profile analysis. J. Vet. Med. Sci. doi:10.1292/jvms.18-0049
McInerney, J.O., McNally, A., O'Connell, M.J., 2017. Why prokaryotes have pangenomes. Nat. Microbiol. doi:10.1038/nmicrobiol.2017.40
Moustafa, A.M., Bennett, M.D., Edwards, J., Azim, K., Mesaik, M.A., Choudhary, M.I., Pathanasophon, P., Worarach, A., Ali, Q., Abubakar, M., Anjum, R., 2013. Molecular typing of haemorrhagic septicaemia-associated Pasteurella multocida isolates from Pakistan and Thailand using multilocus sequence typing and pulsed-field gel electrophoresis. Res. Vet. Sci. doi:10.1016/j.rvsc.2013.07.003
Moustafa, A.M., Seemann, T., Gladman, S., Adler, B., Harper, M., Boyce, J.D., Bennett, M.D., 2015. Comparative genomic analysis of Asian haemorrhagic septicaemia-associated strains of Pasteurella multocida identifies more than 90 haemorrhagic septicaemia-specific genes. PLoS One 10, 1–16. doi:10.1371/journal.pone.0130296
Okay, S., Kurt Kızıldoğan, A., 2015. Comparative genome analysis of five Pasteurella multocida strains to decipher the diversification in pathogenicity and host specialization. Gene 567, 58–72. doi:10.1016/j.gene.2015.04.063
Peng, Z., Liang, W., Wang, F., Xu, Z., Xie, Z., Lian, Z., Hua, L., Zhou, R., Chen, H., Wu, B., 2018. Genetic and Phylogenetic Characteristics of Pasteurella multocida Isolates From Different Host Species. Front. Microbiol. 9, 1–13. doi:10.3389/fmicb.2018.01408
Petersen, A., Bisgaard, M., Townsend, K., Christensen, H., 2014. MLST typing of Pasteurella multocida associated with haemorrhagic septicaemia and development of a real-time PCR specific for haemorrhagic septicaemia associated isolates. Vet. Microbiol. doi:10.1016/j.vetmic.2014.02.022
Pors, S.E., Hansen, M.S., Bisgaard, M., Jensen, H.E., 2011. Occurrence and associated lesions of Pasteurella multocida in porcine bronchopneumonia. Vet. Microbiol. doi:10.1016/j.vetmic.2011.01.005
Rímac, R., Luna, L., Hurtado, R., Rosadio, R., Maturrano, L., 2017. Detection and genetic characterization of Pasteurella multocida from alpaca (Vicugna pacos) pneumonia cases. Trop. Anim. Health Prod. doi:10.1007/s11250-017-1309-5
Rocío, R.B., Raquel, H.C., Luis, L.E., Raúl, R.A., Lenin, M.H., 2017. Análisis de diversidad genética de cepas de pasteurella multocida aisladas de alpacas con signos de neumonía. Rev. Investig. Vet. del Peru 28, 723–729. doi:10.15381/rivep.v28i3.13358
Rosadio, R., Cirilo, E., Manchego, A., Rivera, H., 2011. Respiratory syncytial and parainfluenza type 3 viruses coexisting with Pasteurella multocida and Mannheimia hemolytica in acute pneumonias of neonatal alpacas. Small Rumin. Res. doi:10.1016/j.smallrumres.2011.02.001
Spratt, B.G., Hanage, W.P., Li, B., Aanensen, D.M., Feil, E.J., 2004. Displaying the relatedness

among isolates of bacterial species - The eBURST approach. FEMS Microbiol. Lett. doi:10.1016/j.femsle.2004.11.015
Subaaharan, S., Blackall, L.L., Blackall, P.J., 2010. Development of a multi-locus sequence typing scheme for avian isolates of Pasteurella multocida. Vet. Microbiol. doi:10.1016/j.vetmic.2010.01.017
Wilson, B.A., Ho, M., 2013. Pasteurella multocida: From Zoonosis to cellular microbiology. Clin. Microbiol. Rev. 26, 631–655. doi:10.1128/CMR.00024-13
Yu, C., Sizhu, S., Luo, Q., Xu, X., Fu, L., Zhang, A., 2016. Genome sequencing of a virulent avian Pasteurella multocida strain GX-Pm reveals the candidate genes involved in the pathogenesis. Res. Vet. Sci. 105, 23–27. doi:10.1016/j.rvsc.2016.01.013
Zhu, D., He, J., Yang, Z., Wang, M., Jia, R., Chen, S., Liu, M., Zhao, X., Yang, Q., Wu, Y., Zhang, S., Liu, Y., Zhang, L., Yu, Y., You, Y., Chen, X., Cheng, A., 2019. Comparative analysis reveals the Genomic Islands in Pasteurella multocida population genetics: on Symbiosis and adaptability. BMC Genomics 20, 63. doi:10.1186/s12864-018-5366-6

Cao, P., Guo, D., Liu, J., Jiang, Q., Xu, Z., Qu, L., 2017. Genome-wide analyses reveal genes subject to positive selection in Pasteurella multocida. Front. Microbiol. 8. doi:10.3389/fmicb.2017.00961
Cardoso-Toset, F., Gómez-Laguna, J., Callejo, M., Vela, A.I., Carrasco, L., Fernández-Garayzábal, J.F., Maldonado, A., Luque, I., 2013. Septicaemic pasteurellosis in free-range pigs associated with an unusual biovar 13 of Pasteurella multocida. Vet. Microbiol. doi:10.1016/j.vetmic.2013.08.005
Chen, S.L., Hung, C.-S., Xu, J., Reigstad, C.S., Magrini, V., Sabo, A., Blasiar, D., Bieri, T., Meyer, R.R., Ozersky, P., Armstrong, J.R., Fulton, R.S., Latreille, J.P., Spieth, J., Hooton, T.M., Mardis, E.R., Hultgren, S.J., Gordon, J.I., 2006. Identification of genes subject to positive selection in uropathogenic strains of Escherichia coli: A comparative genomics approach. Proc. Natl. Acad. Sci. doi:10.1073/pnas.0600938103
Christensen, H., Bisgaard, M., 2010. Molecular classification and its impact on diagnostics and understanding the phylogeny and epidemiology of selected members of Pasteurellaceoe of veterinary importance. Berl. Munch. Tierarztl. Wochenschr. doi:10.2376/0005-9366-123-20
Dabo, S.M., Taylor, J.D., Confer, a W., 2007. Pasteurella multocida and bovine respiratory disease. Anim. Health Res. Rev. doi:10.1017/S1466252307001399
Davies, R.L., MacCorquodale, R., Reilly, S., 2004. Characterisation of bovine strains of Pasteurella multocida and comparison with isolates of avian, ovine and porcine origin. Vet. Microbiol. doi:10.1016/j.vetmic.2003.11.013
Du, H., Fang, R., Pan, T., Li, T., Li, N., He, Q., Wu, R., Peng, Y., Zhou, Z., 2016. Comparative Genomics Analysis of Two Different Virulent Bovine *Pasteurella multocida* Isolates. Int. J. Genomics 2016, 1–14. doi:10.1155/2016/4512493
Ewers, C., Lübke-Becker, A., Bethe, A., Kießling, S., Filter, M., Wieler, L.H., 2006. Virulence genotype of Pasteurella multocida strains isolated from different hosts with various disease status. Vet. Microbiol. 114, 304–317. doi:10.1016/j.vetmic.2005.12.012
Fitzpatrick, D.A., Creevey, C.J., McInerney, J.O., 2005. Evidence of positive Darwinian selection in putative meningococcal vaccine antigens. J. Mol. Evol. doi:10.1007/s00239-004-0290-6
García-Alvarez, A., Chaves, F., Fernández, A., Sanz, C., Borobia, M., Cid, D., 2015. An ST11 clone of Pasteurella multocida, widely spread among farmed rabbits in the Iberian Peninsula, demonstrates respiratory niche association. Infect. Genet. Evol. doi:10.1016/j.meegid.2015.07.018
García-Alvarez, A., Vela, A.I., San Martín, E., Chaves, F., Fernández-Garayzábal, J.F., Lucas, D., Cid, D., 2017. Characterization of Pasteurella multocida associated with ovine pneumonia using multi-locus sequence typing (MLST) and virulence-associated gene profile analysis and comparison with porcine isolates. Vet. Microbiol. doi:10.1016/j.vetmic.2017.04.015
Guo, D., Lu, Y., Zhang, A., Liu, J., Yuan, D., Jiang, Q., Lin, H., Si, C., Qu, L., 2012. Identification

of genes transcribed by Pasteurella multocida in rabbit livers through the selective capture of transcribed sequences. FEMS Microbiol. Lett. doi:10.1111/j.1574-6968.2012.02559.x

Harms, N.E., 2015. DYNAMICS OF DISEASE: ORIGINS AND ECOLOGY OF AVIAN CHOLERA IN THE EASTERN CANADIAN ARCTIC. University of Saskatchewan.

Harper, M., Boyce, J.D., Adler, B., 2006. Pasteurella multocida pathogenesis: 125 Years after Pasteur. FEMS Microbiol. Lett. doi:10.1111/j.1574-6968.2006.00442.x

Hotchkiss, E.J., Hodgson, J.C., Lainson, F.A., Zadoks, R.N., 2011. Multilocus sequence typing of a global collection of Pasteurella multocida isolates from cattle and other host species demonstrates niche association. BMC Microbiol. doi:10.1186/1471-2180-11-115

Hurtado, R., Carhuaricra, D., Soares, S., Viana, M.V.C., Azevedo, V., Maturrano, L., Aburjaile, F., 2018. Pan-genomic approach shows insight of genetic divergence and pathogenic-adaptation of Pasteurella multocida. Gene. doi:10.1016/j.gene.2018.05.084

Johnson, T.J., Abrahante, J.E., Hunter, S.S., Hauglund, M., Tatum, F.M., Maheswaran, S.K., Briggs, R.E., 2013. Comparative genome analysis of an avirulent and two virulent strains of avian Pasteurella multocida reveals candidate genes involved in fitness and pathogenicity. BMC Microbiol. 13, 1. doi:10.1186/1471-2180-13-106

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., Jelsbak, L., Sicheritz-Pontén, T., Ussery, D.W., Aarestrup, F.M., Lund, O., 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J. Clin. Microbiol. doi:10.1128/JCM.06094-11

Li, Z., Cheng, F., Lan, S., Guo, J., Liu, W., Li, X., Luo, Z., Zhang, M., Wu, J., Shi, Y., 2018. Investigation of genetic diversity and epidemiological characteristics of Pasteurella multocida isolates from poultry in southwest China by population structure, multi-locus sequence typing and virulence-associated gene profile analysis. J. Vet. Med. Sci. doi:10.1292/jvms.18-0049

Massacci, F.R., Magistrali, C.F., Cucco, L., Curcio, L., Bano, L., Mangili, P.M., Scoccia, E., Bisgaard, M., Aalbæk, B., Christensen, H., 2018. Characterization of Pasteurella multocida involved in rabbit infections. Vet. Microbiol. doi:10.1016/j.vetmic.2017.11.023

McInerney, J.O., McNally, A., O'Connell, M.J., 2017. Why prokaryotes have pangenomes. Nat. Microbiol. doi:10.1038/nmicrobiol.2017.40

Moustafa, A.M., Bennett, M.D., Edwards, J., Azim, K., Mesaik, M.A., Choudhary, M.I., Pathanasophon, P., Worarach, A., Ali, Q., Abubakar, M., Anjum, R., 2013. Molecular typing of haemorrhagic septicaemia-associated Pasteurella multocida isolates from Pakistan and Thailand using multilocus sequence typing and pulsed-field gel electrophoresis. Res. Vet. Sci. doi:10.1016/j.rvsc.2013.07.003

Moustafa, A.M., Seemann, T., Gladman, S., Adler, B., Harper, M., Boyce, J.D., Bennett, M.D., 2015. Comparative genomic analysis of Asian haemorrhagic septicaemia-associated strains of Pasteurella multocida identifies more than 90 haemorrhagic septicaemia-specific genes. PLoS One 10, 1–16. doi:10.1371/journal.pone.0130296

Okay, S., Kurt Kızıldoğan, A., 2015. Comparative genome analysis of five Pasteurella multocida strains to decipher the diversification in pathogenicity and host specialization. Gene 567, 58–72. doi:10.1016/j.gene.2015.04.063

Peng, Z., Liang, W., Wang, F., Xu, Z., Xie, Z., Lian, Z., Hua, L., Zhou, R., Chen, H., Wu, B., 2018a. Genetic and Phylogenetic Characteristics of Pasteurella multocida Isolates From Different Host Species. Front. Microbiol. 9, 1408. doi:10.3389/fmicb.2018.01408

Peng, Z., Wang, H., Liang, W., Chen, Y., Tang, X., Chen, H., Wu, B., 2018b. A capsule/lipopolysaccharide/MLST genotype D/L6/ST11 of Pasteurella multocida is likely to be strongly associated with swine respiratory disease in China. Arch. Microbiol. doi:10.1007/s00203-017-1421-y

Petersen, A., Bisgaard, M., Townsend, K., Christensen, H., 2014. MLST typing of Pasteurella multocida associated with haemorrhagic septicaemia and development of a real-time PCR specific for haemorrhagic septicaemia associated isolates. Vet. Microbiol. doi:10.1016/j.vetmic.2014.02.022

Pors, S.E., Hansen, M.S., Bisgaard, M., Jensen, H.E., 2011. Occurrence and associated lesions of

Pasteurella multocida in porcine bronchopneumonia. Vet. Microbiol. doi:10.1016/j.vetmic.2011.01.005

Rímac, R., Luna, L., Hurtado, R., Rosadio, R., Maturrano, L., 2017. Detection and genetic characterization of Pasteurella multocida from alpaca (Vicugna pacos) pneumonia cases. Trop. Anim. Health Prod. doi:10.1007/s11250-017-1309-5

Rocío, R.B., Raquel, H.C., Luis, L.E., Raúl, R.A., Lenin, M.H., 2017. Análisis de diversidad genética de cepas de pasteurella multocida aisladas de alpacas con signos de neumonía. Rev. Investig. Vet. del Peru 28, 723–729. doi:10.15381/rivep.v28i3.13358

Rosadio, R., Cirilo, E., Manchego, A., Rivera, H., 2011. Respiratory syncytial and parainfluenza type 3 viruses coexisting with Pasteurella multocida and Mannheimia hemolytica in acute pneumonias of neonatal alpacas. Small Rumin. Res. doi:10.1016/j.smallrumres.2011.02.001

Spratt, B.G., Hanage, W.P., Li, B., Aanensen, D.M., Feil, E.J., 2004. Displaying the relatedness among isolates of bacterial species - The eBURST approach. FEMS Microbiol. Lett. doi:10.1016/j.femsle.2004.11.015

Subaaharan, S., Blackall, L.L., Blackall, P.J., 2010. Development of a multi-locus sequence typing scheme for avian isolates of Pasteurella multocida. Vet. Microbiol. doi:10.1016/j.vetmic.2010.01.017

Wilson, B.A., Ho, M., 2013. Pasteurella multocida: From Zoonosis to cellular microbiology. Clin. Microbiol. Rev. 26, 631–655. doi:10.1128/CMR.00024-13

Yu, C., Sizhu, S., Luo, Q., Xu, X., Fu, L., Zhang, A., 2016. Genome sequencing of a virulent avian Pasteurella multocida strain GX-Pm reveals the candidate genes involved in the pathogenesis. Res. Vet. Sci. 105, 23–27. doi:10.1016/j.rvsc.2016.01.013

Zhu, D., He, J., Yang, Z., Wang, M., Jia, R., Chen, S., Liu, M., Zhao, X., Yang, Q., Wu, Y., Zhang, S., Liu, Y., Zhang, L., Yu, Y., You, Y., Chen, X., Cheng, A., 2019. Comparative analysis reveals the Genomic Islands in Pasteurella multocida population genetics: on Symbiosis and adaptability. BMC Genomics 20, 63. doi:10.1186/s12864-018-5366-6

Hurtado, R., V., Maturrano, M.V.C., Azevedo, L., Aburjaile, F., 2018. “Insights of *Pasteurella multocida*: genomic diversity, epidemiology and evolutionary adaptation”. 2019. Submission in process. International Journal of Medical Microbiology.