

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Departamento de Biologia Geral
Programa de Pós-Graduação em Bioinformática

Leandro Gomes Alves

Pan-genômica do patógeno zoonótico negligenciado *Bartonella henselae*

BELO HORIZONTE

2021

Leandro Gomes Alves

Pan-genômica do patógeno zoonótico negligenciado *Bartonella henselae*

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

ORIENTADOR: Prof. Dr. Siomar de Castro Soares

CO-ORIENTADORA: Dr^a. Letícia de Castro Oliveira

BELO HORIZONTE

2021

043

Alves, Leandro Gomes.

Pan-genômica do patógeno zoonótico negligenciado *Bartonella henselae* [manuscrito] / Leandro Gomes Alves. – 2021.

62 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Siomar de Castro Soares. Co-orientadora: Dr^a. Letícia de Castro Oliveira.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. *Bartonella henselae*. 2. Doença da Arranhadura de Gato. 3. Genoma. I. Soares, Siomar de Castro. II. Oliveira, Letícia de Castro. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 573:004



UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa Interunidades de Pós-Graduação em Bioinformática da UFMG

PARECER Nº 10/2021
PROCESSO Nº 23072.217504/2021-51

FOLHA DE APROVAÇÃO

"Pan-genômica do patógeno zoonótico negligenciado Bartonella henselae"

Leandro Gomes Alves

Dissertação aprovada pela banca examinadora constituída pelos Professores:

Prof. Siomar de Castro Soares - Orientador
UFTM

Profa. Leticia de Castro Oliveira - Coorientadora
UFTM

Prof. Luis Carlos Guimarães
UFPA

Prof. Vasco Ariston de Carvalho Azevedo
UFMG

Belo Horizonte, 05 de abril de 2021.



Documento assinado eletronicamente por **Vasco Ariston de Carvalho Azevedo, Professor do Magistério Superior**, em 05/04/2021, às 14:50, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Siomar de Castro Soares, Usuário Externo**, em 05/04/2021, às 16:44, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Letícia de Castro Oliveira, Usuário Externo**, em 06/04/2021, às 13:49, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Luis Carlos Guimarães, Usuário Externo**, em 06/04/2021, às 15:32, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0656738** e o código CRC **C3EE3BA6**.

À minha querida mãe, que sempre me incentivou, me apoiou, e me ensinou a trilhar progressivamente pelo caminho da formação acadêmica, não deixando que nenhuma situação me deixasse desanimar.

AGRADECIMENTOS

Agradeço à Deus, o Criador de tudo, por tudo que tem me propiciado ao longo da minha caminhada;

Aos meus amados pais, José Alves (*in memoriam*) e Maria do Socorro, que, em toda a minha vida, me apoiaram nas minhas escolhas e me deram uma vida cheia de amor incondicional. À minha mãe, por me dar todo o suporte que eu precisava nos momentos mais difíceis. Ao meu pai, que infelizmente não está aqui para presenciar esse marco, mas é a quem devo todo o meu respeito e agradecimento por me fazer tornar a pessoa que sou hoje;

Ao meu irmão, Denis, e minha cunhada, Ivete, pela amizade e incentivo, e por sempre vibrarem pelas minhas vitórias;

À todos os meus familiares, que, cada um à sua maneira, sempre me acompanharam e incentivaram para que eu chegasse até aqui;

À minha noiva, Liliane, por todo o companheirismo e amizade, por não me deixar abalar e desistir dos meus sonhos hora alguma, por estar sempre presente nos momentos mais importantes da minha vida e, principalmente, pelo incentivo e paciência. Amo você! Agradeço também à sua família, que me acolheu e se tornou minha segunda morada;

Ao professor e amigo, Siomar, pela orientação durante toda a minha carreira científica. Sempre me incentivando e abrindo horizontes à pesquisa. Seu encorajamento só me fez seguir em frente, me fazendo acreditar no que realmente sou capaz;

À minha co-orientadora, Letícia, me ajudando e norteando para que eu fosse capaz de realizar este trabalho;

Aos amigos que a UFMG me apresentou, Wyll, Arun, Sandeep, companheiros essenciais durante minha jornada em Belo Horizonte e que levarei para a vida. Obrigado por todos os momentos de trabalho e descontração;

À Pós-Graduação em Bioinformática e a todos os professores, pela carga didática adquirida durante toda a minha passagem pelo programa;

Ao grupo de pesquisa em Bioinformática da UFTM, por serem colegas especiais que me acolheram e só acrescentaram à minha vida;

À todos vocês, a minha imensa gratidão!

“Não são nossas habilidades que mostram quem realmente somos, são as nossas escolhas.”

J. K. Rowling

“A verdadeira viagem de descobrimento não consiste em procurar novas paisagens, mas em ter novos olhos.”

Marcel Proust

RESUMO

O gênero *Bartonella* é composto por bactérias Gram-negativas e reemergentes, como a *Bartonella henselae*, que infecta humanos e sobrevive dentro de eritrócitos. Essa espécie é transmitida por arranhões ou mordidas de gatos domésticos e usualmente causa infecções sintomáticas em humanos, conhecidas como Doença da Arranhadura do Gato. A doença causa múltiplos sinais clínicos em humanos, como doenças dermatológicas, cardiovasculares, linfáticas, hepáticas e do sistema nervoso, principalmente em indivíduos imunossuprimidos. Embora a bactéria seja altamente relevante por sua importância zoonótica em todo o mundo, poucos estudos objetivaram caracterizar os genomas dessa espécie e ainda não há estudos pan-genômicos disponíveis. Aqui, foram realizadas análises de filogenômica, pan-genômica e plasticidade genômica para determinar aspectos epidemiológicos, o tamanho do pan-genoma e sua variabilidade nas ilhas de patogenicidade identificadas. Os 24 genomas de *B. henselae* analisados apresentaram similaridade variando de 96 a 100%. Foram encontrados dois subconjuntos de genomas nas análises filogenômicas, um composto exclusivamente de linhagens isoladas de gatos e outro de hospedeiros felinos e humanos. Identificou-se um pan-genoma de 1655 genes não redundantes, onde 70% destes genes integram seu core genoma, composto por genes essenciais ao microrganismo. Além disso, o desenvolvimento do pan-genoma do conjunto completo de dados identificou um valor de $\alpha = 0,94$. A análise de plasticidade genômica revelou a presença de 6 ilhas de patogenicidade e 1 região de fago no genoma de referência *B. henselae* Houston-1. No geral, os resultados mostraram que os genomas são altamente similares, com um pan-genoma tendendo a tornar-se fechado. Ainda, as regiões de plasticidade genômica analisadas são compostas por genes relacionados à virulência da bactéria. Para a validação dos padrões de distribuição genômica, aponta-se a necessidade de sequenciamento de mais genomas em todo o mundo para melhor caracterizar variações no pan-genoma e compreender os padrões de adaptação dessa espécie. A alta conservação dos genomas dessa espécie é muito importante para o desenvolvimento de novas vacinas e análises de alvos de drogas contra esse patógeno. Esses dados poderão ser utilizados em trabalhos futuros, de grande relevância para a contenção da doença em todo o mundo.

Palavras-chave: *Bartonella henselae*. Doença da arranhadura do gato. Pan-genoma. Plasticidade genômica. Filogenômica.

ABSTRACT

The genus *Bartonella* is comprised of Gram-negative re-emerging bacteria like *Bartonella henselae*, which infects humans and survives inside erythrocytes. This species is transmitted by scratches and bites from domestic cats and usually causes a symptomatic infection in humans, known as Cat-Scratch Disease (CSD). The disease causes multiple clinical signs in humans, such as dermatologic, cardiovascular, lymphatic, hepatic and nervous system diseases mainly in immunosuppressed individuals. Although the bacteria are highly relevant for its zoonotic importance worldwide, few studies aimed at characterizing these species genomes and there is still no pan-genome study available. Here, phylogenomic, pan-genome and genome plasticity analyses were performed to determine the epidemiological aspects, the size of the pan-genome and its variability in the identified pathogenicity islands. The 24 genomes of *B. henselae* analyzed showed similarity ranging from 96 to 100%. Two subsets of genomes were found in phylogenomic analyzes, one composed exclusively of strains isolated from cats and the other from feline and human hosts. A pan-genomic analysis identified 1655 non-redundant genes, where 70% of these genes integrate their core genome, composed of genes essential to the microorganism. In addition, the pan-genome development of the complete data set identified a α value = 0.94. The analysis of genomic plasticity revealed the presence of 6 pathogenicity islands and 1 phage region in the reference genome *B. henselae* Houston-1. Overall, the results showed that the genomes are highly similar, with a pan-genome tending to become closed. Furthermore, the regions of genomic plasticity analyzed are composed of genes related to the virulence of the bacteria. For the validation of genomic distribution patterns, the need of sequencing more genomes worldwide is pointed out in order to better characterize the variations in the pan-genome and understand the adaptation patterns of this species. The high conservation of genomes of this species is very important for the development of new vaccines and analyzes of drug targets against this pathogen. These data can be used in future works, of great relevance for the containment of the disease worldwide.

Keywords: *Bartonella henselae*. Cat scratch disease. Pan-genome. Genomic plasticity. Phylogenomics.

LISTA DE FIGURAS

Figure 1: Phylogenomic tree correlating the evolutionary distance among the strains of <i>Bartonella henselae</i>	37
Figure 2: Dendrogram created with wgMLST profile of the 24 genomes of <i>Bartonella henselae</i>	37
Figure 3: Gene synteny analysis of the complete and draft genomes of <i>Bartonella henselae</i> .	38
Figure 4: Gene synteny analysis of the completed genomes of <i>Bartonella henselae</i>	38
Figure 5: Pan-genome, core genome and singletons development of <i>Bartonella henselae</i>	39
Figure 6: Subset classification of the pan-genome according to the functional categories of the Cluster of Orthologous Groups	39
Figure 7: Circular representation of the genomic comparison among the strains of the <i>Bartonella henselae</i> species	40
Figure 8: Gene content of the phage region found in <i>Bartonella henselae</i> Houston-1	40
Figure 9: Circular representation of the genomic comparison between all strains of <i>Bartonella henselae</i> using the strain BM1374165 as reference	41
Figure 10: Circular representation of the genomic comparison between all strains of <i>Bartonella henselae</i> using the strain Houston-I as reference	41

LISTA DE TABELAS

Table 1: Genomic features of the strains used in this study	36
Table 2: Locations and sizes of pathogenicity islands predicted in the reference strain <i>Bartonella henselae</i> Houston-1	39
Table 3: List of all genes located in the pathogenicity islands of the reference strain <i>Bartonella henselae</i> Houston-1	42

LISTA DE ABREVIATURAS, SIGLAS E TERMOS EM INGLÊS

AB – Angiomatose Bacilar

ACT – *Artemis Comparison Tool*

AIDS – Síndrome da Imunodeficiência Adquirida

BLAST – *Basic Local Alignment Search Tool*

BRIG – *BLAST Ring Image Generator*

CDSs – *Coding DNA Sequences*

COG – *Clusters of Orthologous Groups*

CSD – *Cat-Scratch Disease*

DAG – Doença da Arranhadura do Gato

DNA – Ácido desoxirribonucleico

ELISA – Ensaio de imunoabsorção enzimática

GC – Conteúdo de guanina-citosina

GI – Ilha genômica

GIPSy – *Genomic Island Prediction Software*

HGT – Transferência horizontal de genes (do inglês, “*Horizontal Gene Transfer*”)

LCBs – Blocos colineares locais (do inglês, “*Locally Collinear Blocks*”)

Mb – Milhões de bases

MCL – *Markov Clustering*

MI – Ilha genômica metabólica

MLST – Tipagem por sequenciamento de múltiplos locus gênicos (do inglês, “*Multilocus Sequence Typing*”)

MLVA – *Multiple locus variable number of tandem repeats analysis*

NCBI – *National Center for Biotechnology Information*

PAI – Ilha genômica de patogenicidade

PAMPs – Padrões moleculares relacionados a patógenos

Pb – Pares de bases

PCR – Reação em cadeia da polimerase

RI – Ilha genômica de resistência

RNA – Ácido ribonucleico

rRNA – Ácido ribonucleico ribossômico

SI – Ilha genômica simbiótica

ssDNA – DNA fita simples (do inglês, “*single-stranded DNA*”)

T4SS – Sistema de Secreção Tipo IV

TLR – Receptores do tipo Toll

tRNA – Ácido ribonucleico transportador

USA – *United States of America*

VEGF – Fatores de crescimento endotelial vascular

VFDB – Virulence Factor Database

wgMLST – *Whole genome multilocus sequence typing*

SUMÁRIO

1. INTRODUÇÃO	15
1.1. Gênero Bartonella	15
1.2. Bartonella henselae.....	17
1.2.1. Transmissão	17
1.2.2. Epidemiologia	18
1.2.3. Fisiopatologia e fatores de virulência.....	19
1.2.4. Sintomatologia	21
1.2.5. Diagnóstico.....	22
1.2.6. Tratamento e profilaxia	23
1.3. BIOINFORMÁTICA	24
1.3.1. Genômica comparativa.....	24
1.3.2. Filogenômica	26
1.3.3. Pan-genômica.....	27
1.3.4. Plasticidade genômica.....	28
1.4. Abordagens genômicas em Bartonella henselae	29
2. JUSTIFICATIVA.....	32
3. OBJETIVOS	33
3.1. OBJETIVO GERAL	33
3.2. OBJETIVOS ESPECÍFICOS	33
4. The pan-genome of the zoonotic neglected pathogen <i>Bartonella henselae</i> reveals two groups with different patterns of adaptation to hosts.....	34
5. CONCLUSÕES	49
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	50
7. PRODUÇÃO CIENTÍFICA.....	61

1. INTRODUÇÃO

1.1. Gênero *Bartonella*

Atualmente, as bactérias do gênero *Bartonella* são consideradas patógenos zoonóticos reemergentes e negligenciados. Apesar de sua ampla distribuição mundial, afetam principalmente a população com condições precárias de saneamento, e que estão em contato direto com animais domésticos e artrópodes transmissores desses microrganismos.

O gênero foi assim nomeado em homenagem ao microbiologista peruano Dr. Alberto Leonardo Barton Thompson. Em 1905, Thompson investigou pacientes que apresentavam febre e anemia grave após uma viagem a La Oroya, no Peru. Tal estudo levou à identificação da bactéria *Bartonella bacilliformis*, agente etiológico da Febre de Oroya (fase aguda, caracterizada por bacteremia que cursa com anemia hemolítica fatal) e verruga peruana (fase crônica, com a presença de tumores vasculares) (GOMES; RUIZ, 2017). Por décadas, o gênero foi composto exclusivamente pela espécie supracitada.

Durante a Primeira Guerra Mundial, emergiu uma doença que acometia diversos soldados por meses. Caracterizada por febre de início súbito com cinco dias de duração e recorrências durante o ano, a doença foi denominada “Febre de Trincheiras”. A causa e a forma de transmissão dessa patologia permaneceram incertas, uma vez que o organismo causador não pôde ser cultivado. Em 1918, a transmissão por inoculação sanguínea foi associada ao piolho, quando foram encontradas bactérias denominadas, até então, como “corpos de *Rickettsias*” no intestino e fezes dos piolhos que estavam em contato com pacientes doentes. Após inúmeras discussões e tentativas de isolamento do agente etiológico, houve em 1920, um consenso geral de que a gênese da Febre de Trincheiras advinha da bactéria *Rickettsia quintana* (ANSTEAD, 2016). Outros estudos levaram à transferência de *R. quintana* para o gênero *Rochalimaea*, até a reclassificação completa deste para o gênero *Bartonella* em 1993.

O primeiro relato de uma manifestação clínica da Doença da Arranhadura do Gato (DAG) se deu por Parinaud *et al.* (1889) em pacientes com conjuntivite e

linfadenopatia regional (VALOR; HUBER, 2018). Em 1931, o professor de pediatria Robert Debré começou uma investigação de um menino com adenite supurativa associado com múltiplas arranhaduras de gato. Debré *et al.* (1950) realizaram a divulgação formal e escrita da síndrome clínica da DAG, cuja nomenclatura se fixou devido à relação de gatos domésticos com a patogênese da doença. O agente etiológico da DAG foi observado em linfonodos de pacientes estudados por Wear *et al.* (1983), quando bacilos pleomórficos Gram-negativos estavam presentes nas regiões próximas aos linfonodos acometidos, principalmente nas paredes capilares. *Rochalimaea henselae* foi isolado e assim denominado em homenagem à Diane Hensel, uma microbiologista que contribuiu consideravelmente para o isolamento inicial da bactéria causadora da DAG (ANDERSON; NEUMAN, 1997).

Até 1993, quatro espécies compunham o gênero *Rochalimaea*, sendo elas: *Rochalimaea quintana*, *Rochalimaea vinsonii*, *Rochalimaea elizabethae* e *R. henselae*. Nesse ano, os gêneros *Rochalimaea* e *Bartonella* foram reclassificados filogeneticamente considerando a estreita relação entre conteúdo de guanina-citosina (GC) e homologia do RNA ribossômico 16S. Sendo assim, o gênero *Rochalimaea* foi fundido ao gênero *Bartonella* (HÖRDT *et al.*, 2020).

Patologias associadas ao atual gênero *Bartonella* permaneceram negligenciadas até meados dos anos 90, quando infecções oportunistas acometiam pessoas vulneráveis, principalmente pacientes com a Síndrome da Imunodeficiência Adquirida (AIDS) e pessoas sem teto em áreas urbanas (KOEHLER, 1995). Desde então, o número de espécies aumentou consideravelmente, assim como a associação de diferentes espécies com doenças humanas e uma variedade de manifestações clínicas associadas. A evolução dos métodos diagnósticos contribuiu consideravelmente para a identificação e documentação de agentes de infecções em animais, doentes e saudáveis, e seres humanos, imunocompetentes e imunocomprometidos.

De acordo com a classificação taxonômica atual, o gênero de bactérias *Bartonella* está incluso no filo das Proteobacteria, classe das Alphaproteobacteria, ordem Rhizobiales e família Bartonellaceae. Os membros dessa família possuem a habilidade de infectar eritrócitos e causar bacteremias intra-eritrocíticas prolongadas em seus hospedeiros reservatórios. As bartonelas são caracterizadas por serem

bactérias Gram-negativas, de formato cocobacilar, pleomórficas, aeróbicas obrigatórias e de crescimento fastidioso (DEHIO, 2004; MADA *et al.*, 2020).

1.2. *Bartonella henselae*

B. henselae é um organismo catalase, urease e indol negativos, não fermentador de carboidratos, positivo para fosfatase alcalina e arginina, e negativo também para acetoina e redução de nitrato a nitrito. Além disso, essa espécie possui fímbrias, que estão associadas com a citoaderência da bactéria (DIDDI *et al.*, 2013). A espécie é uma das mais conhecida dentre as Bartoneloses patogênicas causadoras de infecções em humanos, sendo a DAG e a Angiomatose Bacilar (AB) suas manifestações mais comuns.

1.2.1. Transmissão

Os gatos domésticos (*Felis catus*), principalmente jovens, são os principais reservatórios naturais de *B. henselae*, assim como são para outras espécies de *Bartonella*. Raramente desenvolvem manifestações clínicas, permanecendo infectados com altos níveis de bacteremia intravascular assintomática (ÁLVAREZ-FERNÁNDEZ *et al.*, 2018; CHOMEL *et al.*, 2006). Cães também são relatados na literatura como reservatórios e potenciais transmissores da espécie para humanos (BREITSCHWERDT, EDWARD B., 2017; IANNINO *et al.*, 2018). A associação desse grupo de microrganismos com reservatórios felinos e humanos já está estabelecida há muito tempo, mas, devido ao crescente interesse e importância do gênero no ramo médico e veterinário, novos reservatórios têm sido relatados e investigados, como: mamíferos marinhos, tais como golfinhos (MAGGI *et al.*, 2005) e baleias brancas (MAGGI *et al.*, 2008); tartarugas marinhas (VALENTINE *et al.*, 2007); bovinos (CHERRY *et al.*, 2009a); e aves (MASCARELLI *et al.*, 2014).

Muitos estudos sugerem a transferência desse patógeno para artrópodes hematófagos, flebotomíneos, piolhos, pulgas e carrapatos durante o repasto sanguíneo (ANGELAKIS *et al.*, 2010; BILLETER *et al.*, 2008; KAISER *et al.*, 2011). O principal vetor de transmissão da bactéria entre gatos é a pulga *Ctenocephalides felis*. A bactéria pode ser expelida na saliva e excretada junto às fezes da pulga, como demonstrado por Bouhsira *et al.* (2013), durante um experimento de alimentação

artificial. Consequentemente, a contaminação acontece através do contato com as fezes desse artrópode infectado por *B. henselae*.

A transmissão para humanos pode ocorrer através de arranhões, lambeduras ou mordidas de animais contaminados pelas fezes de vetores. O processo efetivo de transmissão ocorre, então, por meio da inoculação da bactéria no tecido humano.

Os gatos, domésticos ou selvagens, além de reservatórios naturais, são os principais transmissores da doença ao humano. Além disso, existem relatos de transmissão por transfusão sanguínea e por agulhas contaminadas entre veterinários (DINIZ *et al.*, 2016; LIN *et al.*, 2011; OLIVEIRA *et al.*, 2010). Um estudo de ampliações de DNA por reação em cadeia da polimerase (PCR) de regiões do DNAs ribossomais 16S e 23S e de análises do gene *rpoB* e de imunofluorescência aponta a possibilidade de transmissão da bactéria da mãe para o filho durante o parto ou *in utero* (BREITSCHWERDT, EDWARD B *et al.*, 2010).

1.2.2. Epidemiologia

A espécie *B. henselae*, assim como as demais espécies do gênero *Bartonella*, possui distribuição mundial e é descrita na maioria dos países do mundo (BLANCO; RAOULT, 2005). De acordo com Bass *et al.* (1997), a incidência da DAG é diretamente proporcional à quantidade de gatos presentes na população observada (principais reservatórios da bactéria), grau de infestação por pulgas (transmissores inter e intraespécies) e regiões de clima quente e úmido (ambiente propício para a propagação de artrópodes transmissores da bactéria).

Nos Estados Unidos, estima-se que a incidência anual de DAG seja de 4,7 diagnósticos para cada 100.000 habitantes com idade inferior à 65 anos, sendo que, cerca de 12.500 pacientes são diagnosticados ambulatorialmente e quase 500 pacientes são internados pelo agravamento da doença por ano no país (NELSON *et al.*, 2016). Essa estimativa é considerada baixa, uma vez que muitos pacientes acometidos pela bactéria desenvolvem casos autolimitados e não procuram atendimento médico, levando à uma subnotificação da doença e seu negligenciamento.

No Brasil, Costa *et al.* (2005) encontraram uma soropositividade para *B. henselae* em 60 (13,7%) das 457 amostras de soro de indivíduos saudáveis analisadas em um estudo na cidade de Piau, no estado de Minas Gerais

Um estudo retrospectivo em um banco de sangue em Campinas, São Paulo, demonstrou a bacteremia prolongada em pacientes assintomáticos doadores de sangue, onde 15 dos 500 doadores analisados apresentaram PCR positiva para *B. henselae* e 16% dos doadores apresentavam anticorpos contra a referida bactéria. Estes dados demonstram a importância de se analisar o sangue de doadores para a possibilidade de infecção assintomática por *B. henselae* (PITASSI *et al.*, 2015). Um estudo de Diniz *et al.* (2016) visou a procura por fatores de risco associados a infecções subclínicas em doadores de sangue também no sudeste brasileiro. Dentre os doadores, a soropositividade para espécies de *Bartonella* foi maior dentre aqueles que tiveram contato com gatos.

Visando estudar a prevalência da bactéria em gatos domésticos, Souza *et al.* (2010) realizaram um estudo no município de Vassouras, no estado do Rio de Janeiro (RJ), onde se evidenciou a presença de DNA de *Bartonella spp.* em 36 das 37 (97,3%) amostras sanguíneas de gatos por meio de análises de PCR para o gene *htrA*. Essa alta prevalência foi justificada pelo fato de todos os gatos pertencerem a um abrigo, sem acesso ao exterior, facilitando a transmissão direta (por meio de arranhaduras ou mordeduras) ou indireta (através de vetores) entre eles. Outro estudo brasileiro, do mesmo grupo de pesquisa, realizado na região metropolitana da cidade do Rio de Janeiro – RJ, visou a análise de amostras sanguíneas de 163 gatos domésticos de propriedades privadas colhidas no último trimestre de 2009. Realizando ensaios de PCR, foram encontrados DNA de *Bartonella henselae* em quase 75% dos gatos analisados (SOUZA *et al.*, 2017).

1.2.3. Fisiopatologia e fatores de virulência

As bactérias da espécie *B. henselae* colonizam o intestino médio dos artrópodes sugadores de sangue e são expelidas em suas fezes. Durante o repasto sanguíneo, ocasionalmente as fezes contaminadas penetram na pele dos mamíferos reservatórios, como gatos, principalmente (CHOMEL *et al.*, 2009). É necessário que elas colonizem um nicho primário onde são transportadas para células do endotélio vascular antes da invasão da corrente sanguínea. Uma vez dentro do sistema vascular

do hospedeiro, as bactérias apresentam tropismo por células eritrocíticas, envolvendo um sistema complexo de adesão e invasão celular, gerando um novo nicho intra-eritrocítico competente para ser transmitido para um novo vetor artrópode durante a sucção sanguínea deste. A localização intra-eritrocítica não hemolítica e intra-endotelial garante a preservação da infecção bacteriana. A resposta imune desses hospedeiros contra o patógeno é imunomodulada pelo mesmo, se tornando ineficaz em eliminá-lo do organismo, ocasionando bacteremia intra-eritrocítica assintomática de longa duração (CHOMEL *et al.*, 2009; SEUBERT; SCHULEIN; DEHIO, 2001).

Após a transmissão da bactéria para humanos, a mesma utiliza estratégias de fuga do sistema imune inato do hospedeiro. No caso de *B. henselae*, os receptores celulares especializados em reconhecimento de potenciais bactérias e vírus invasores, como os receptores *Toll-like* (TLR) não conseguem reconhecer os padrões moleculares relacionados a patógenos (PAMPs, do inglês) localizados superficialmente à membrana externa da bactéria, que evade do sistema de defesa inato do organismo (ZÄHRINGER *et al.*, 2004).

Assim como nos hospedeiros reservatórios, o processo de invasão endotelial se repete em humanos, uma vez que os microrganismos invadem células endoteliais por mecanismos dependentes do Sistema de Secreção Tipo IV (T4SS), principalmente VirB / VirD4 (SIAMER; DEHIO, 2015). T4SS representa um conjunto de fatores de virulência composto por moléculas responsáveis pela transferência eficaz de complexos de nucleoproteínas, como ssDNA relaxases, para células hospedeiras, criando um mecanismo benéfico anti-apoptótico das células infectadas (HARMS *et al.*, 2017; WAGNER; DEHIO, 2019).

Adesinas do tipo A de *Bartonellas* (BadA) também são fatores de virulência importantes presentes na superfície de *B. henselae* associadas com a autoaglutinação bacteriana seguida de adesão da bactéria à glóbulos vermelhos, além de inibir apoptose e induzir o processo angiogênico (FRANZ; KEMPF, 2011; RIESS *et al.*, 2007). A estimulação da proliferação vascular é característica da infecção por esta bactéria, decorrente do processo de indução de angiogênese e inibição da apoptose de células endoteliais que ocasiona na produção de fatores de crescimento endotelial vascular (VEGF) (BERRICH *et al.*, 2011; DEHIO, 2003). O sistema de secreção Trw também se mostra importante para a invasão de eritrócitos pela espécie *B. henselae*,

uma vez que sua localização na superfície da bactéria favorece a interação com os glóbulos vermelhos e sua posterior invasão (VAYSSIER-TAUSSAT *et al.*, 2010).

1.2.4. Sintomatologia

A adaptação específica de *B. henselae* ao hospedeiro reservatório não causa sintomas ao mesmo, ou apresentam-se apenas na forma leve da doença, levando a uma infecção subclínica. Enquanto isso, a infecção de hospedeiros acidentais, como no caso do ser humano, está associada a um amplo espectro de doenças e múltiplos sinais clínicos (FLORIN; ZAOUTIS; ZAOUTIS, 2008).

As manifestações clínicas da infecção por *B. henselae* dependem do estado imunológico da pessoa acometida. A DAG e a AB são as principais patogenias envolvendo o patógeno (CHANG *et al.*, 2002).

Entre três e dez dias após a exposição à bactéria por meio da mordida ou arranhão de um gato infectado, a maioria dos pacientes com DAG desenvolvem, no local do trauma, lesões caracterizadas por pápulas eritematosas com crostas. A linfadenopatia regional é a principal manifestação clínica primária decorrente da DAG, devido à drenagem linfática do local de inoculação da bactéria para o linfonodo mais próximo da lesão, podendo cursar com sinais inflamatórios (sensibilidade, dor, edema, eritema e hiperemia) cerca de 1 a 3 semanas após a exposição. Microscopicamente, esses linfonodos apresentam agregados bacterianos, granulomas com ou sem necrose e microabscessos.

Aproximadamente metade dos pacientes cursam com febre de origem desconhecida, dor abdominal, fadiga, mal-estar, náuseas e vômitos, e uma minoria exibe outros sinais, como dor de cabeça, dor de garganta, emagrecimento, hepatoesplenomegalia, anemia hemolítica e erupção cutânea. Em indivíduos imunocompetentes, esse quadro geralmente é autolimitado, sendo resolvido sem antibioticoterapia em cerca de 2 a 5 meses (GARCÍA *et al.*, 2014).

Em indivíduos imunocomprometidos, *B. henselae* pode causar uma ampla variedade de complicações sistêmicas, como a própria AB potencialmente fatal, endocardite (FLORIN; ZAOUTIS; ZAOUTIS, 2008), osteomielite (DONÀ *et al.*, 2018), doenças oculares (BIANCARDI; CURI, 2014; CELIKER *et al.*, 2018; KSIAA *et al.*,

2019; ORAY *et al.*, 2017; VALOR; HUBER, 2018), neurológicas (CANNETI *et al.*, 2019; GANESAN; MIZEN, 2005; GERBER *et al.*, 2002; HMAIMESS *et al.*, 2004; ZAKHOUR *et al.*, 2018), hematológicas (VIEIRA-DAMIANI; ERICSON, 2016), cutâneas (LINS *et al.*, 2019), entre outras.

A AB é caracterizada por lesões cutâneas derivadas da proliferação do epitélio vascular em virtude da infecção por *B. henselae* (e também por *B. quintana*) principalmente em indivíduos imunocomprometidos, como portadores de AIDS ou transplantados em terapia de imunossupressão (DINIZ, *et al.*, 2016). As lesões são altamente vasculares graças à angiogênese anormal estimulada por essas espécies de *Bartonella* (VELHO *et al.*, 2003). Além disso, essas lesões podem adentrar pelo tecido cutâneo e levar ao envolvimento visceral, como o fígado, baço ou medula óssea. Nesses órgãos, o quadro de múltiplas lesões vasculares, parenquimatosas e hemorrágicas passa a ser denominado de Peliose Bacilar (LINS *et al.*, 2019).

1.2.5. Diagnóstico

O diagnóstico imediato da infecção por *B. henselae* é bastante desafiador, uma vez que a maioria dos sinais e sintomas são inespecíficos e podem ser observados em outras infecções. A identificação da DAG clássica se baseia, principalmente, em uma análise conjunta de sinais e sintomas do paciente, história clínica e exame físico. A exposição a gatos é um critério chave para o curso da investigação, seguido do relato de mordidas e arranhões pelos mesmos e presença de lesão no local de inoculação e linfonodos aumentados (MAGGI *et al.*, 2011).

A análise laboratorial para confirmação diagnóstica de infecções por *B. henselae* não se dá por um método padrão ouro, mas sim, através de um conjunto de métodos para evitar falsos-negativos. As metodologias mais utilizadas são: testes sorológicos e moleculares, cultura de tecidos e isolamento bacteriano, microscopia eletrônica, e ensaios imuno-histoquímicos (FLORIN; ZAOUTIS; ZAOUTIS, 2008).

Exames sorológicos para detecção de anticorpos anti *B. henselae*, como técnicas de imunofluorescência indireta e ELISA (ensaio de imunoabsorção enzimática), são os melhores testes para o diagnóstico da DAG, pela rapidez e

facilidade de execução em comparação com as demais metodologias (VERMEULEN *et al.*, 2007).

Testes moleculares são utilizados para elucidar casos em que o resultado da sorologia é negativa e ainda há suspeita clínica que precisa de confirmação. A PCR é altamente sensível e específica para detecção de fragmentos de DNA de *B. henselae* em amostras provindas de linfonodos infectados. De modo geral, a região intergênica 16S-23S do DNA ribossômico é um dos principais alvos para determinação e diferenciação de espécies de *Bartonella* durante os testes de PCR, todavia a utilização de duas ou mais regiões alvo aumentam consideravelmente a sensibilidade do teste (CHERRY *et al.*, 2009b). A principal desvantagem dessa metodologia é o uso de amostras obtidas de forma invasiva, como aspirados e biópsias de órgãos e tecidos acometidos, principalmente linfonodos. Também se faz necessário o uso de equipamentos e profissionais treinados para execução das técnicas moleculares (CHONDROGIANNIS *et al.*, 2012).

O isolamento de *B. henselae* em cultura a partir de amostras sanguíneas aumenta consideravelmente a sensibilidade e especificidade do diagnóstico da infecção (BREITSCHWERDT, E. B., 2014; KWON *et al.*, 2018). Porém, essa metodologia é difícil, exigindo período de incubação de aproximadamente 6 a 8 semanas em placas de ágar sangue à 35°C, com 5% de CO₂ e ambiente saturado de água (OTEO *et al.*, 2017). Além disso, em rotinas laboratoriais, o sistema de hemocultura automatizada não permite o crescimento da bactéria, já que as mesmas dependem, dentre outros fatores, de um longo período de incubação, e, em geral, os frascos de hemoculturas são desprezados após 7 dias de ausência de crescimento (BLANCO *et al.*, 2008).

1.2.6. Tratamento e profilaxia

As manifestações clínicas de infecções por *B. henselae* determinam o processo terapêutico apropriado. Na DAG clássica, em pacientes imunocompetentes, a doença tende a ser autolimitada e o uso de antimicrobianos não é indicado, sendo o tratamento basicamente sintomático, como, por exemplo, administração de analgésicos. Em contraste, é aconselhável o uso de antimicrobianos em pacientes imunocomprometidos, principalmente para redução

do risco de disseminação sistêmica da bactéria e acometimento dos órgãos (BREITSCHWERDT, EDWARD B., 2008). Os esquemas de antibióticos recomendados para estes pacientes envolvem a administração de azitromicina, doxiciclina e eritromicina, em dosagens e durações específicas para evitar bacteremias (SHORBATLI; KORANYI; NAHATA, 2018). Em casos de AB e peliose hepática, a eritromicina é o medicamento de primeira escolha para o tratamento (ROLAIN *et al.*, 2004).

A infecção bacteriana por *B. henselae* resulta em imunidade permanente, uma vez que o processo infeccioso ativa o sistema imune adaptativo com a consequente formação de anticorpos contra a referida bactéria. Porém, recidivas de bacteremias já foram relatadas em pacientes com AIDS, mesmo após administração de antimicrobianos (ROLAIN *et al.*, 2004).

Dado que a bactéria é transmitida de gatos para o ser humano causando a DAG, uma das formas de prevenção da infecção é evitar contato direto com estes, especialmente gatos de rua e de origem desconhecida, casos em que se desconhece o histórico de ectoparasitose. A higienização após contato com os animais de estimação também é um fator de cautela para evitar transmissão da bactéria. Para minimizar a transmissão entre felinos, a erradicação de pulgas é extremamente eficaz, além de manter os animais em casa, evitando a exposição e contato com gatos de rua que podem ser possíveis vetores. (BRUNT *et al.*, 2006; CHOMEL; BOULOUIS; BREITSCHWERDT, 2004).

1.3. BIOINFORMÁTICA

1.3.1. Genômica comparativa

A genômica comparativa consiste na comparação do conteúdo genético de um organismo contra outro e seu principal propósito é obter um melhor entendimento biológico da espécie em estudo (GENEREUX *et al.*, 2020; SIVASHANKARI; SHANMUGHAVEL, 2007). Esta metodologia ajuda a determinar as sequências de regiões codificantes do genoma e suas funções, assim como caracterizar a frequência de eventos evolucionários, plasticidade genômica e estabelecer as relações filogenéticas entre linhagens e/ou espécies (OGIER *et al.*, 2010). A maior parte das

análises comparativas tem como objetivo identificar similaridades e diferenças entre os organismos (ABBY; DAUBIN, 2007).

A comparação genômica é usada com frequência em muitos aspectos diferentes da ciência, como, por exemplo, na comparação dos genes de *Drosophila melanogaster* (mosca das frutas - organismo modelo) com os genes humanos. Foi relatado que 548 genes do genoma da mosca apresentam homologia a genes humanos que estão ligados a doenças humanas de diferentes naturezas (doenças cardiovasculares, visuais, auditivas, endócrinas e ósseas) (REITER *et al.*, 2001). A descoberta de genes homólogos comumente compartilhados entre humanos e outros organismos abre a possibilidade de testar novas terapias em organismos modelo (SIVASHANKARI; SHANMUGHAVEL, 2007).

Da mesma forma, essa abordagem pode ser usada em organismos procariotos como, por exemplo, na comparação de *Bacillus licheniformis* (bactéria gram-positiva de interesse farmacêutico e biotecnológico, usada para expressão de proteínas e produção de antibióticos) com duas espécies correlatas: *Bacillus subtilis* e *Bacillus halodurans*. A comparação entre estas três bactérias não só contribuiu para a montagem do genoma de *B. licheniformis* como também ajudou nos estudos evolucionários e na identificação de eventos de transferência horizontal de genes (HGT, do inglês) entre elas (REY *et al.*, 2004).

Estendendo o conceito, a genômica comparativa pode ser usada com um grande número de bactérias com estilos de vida distintos. Um estudo usando 317 genomas, de diferentes espécies, gêneros e filos de bactérias, foi realizado visando estabelecer padrões entre seus estilos de vida, repertórios gênicos e tamanhos dos genomas. Como resultado, os autores observaram que patógenos intracelulares são mais propensos à perda gênica ou evolução genômica reducionista, contribuindo consideravelmente para os mecanismos de especiação dos mesmos (MERHEJ *et al.*, 2009).

Finalmente, a disponibilidade de genomas bacterianos em bancos de dados e o uso de abordagens de genômica comparativa permitem a realização de estudos mais aprofundados utilizando novas abordagens, tais como: filogenômica, pan-genômica e plasticidade genômica (LAPIERRE; GOGARTEN, 2009; MUZZI; MASIGNANI; RAPPUOLI, 2007; TETTELIN, H. *et al.*, 2005).

1.3.2. Filogenômica

A filogenética é um dos ramos de estudo da filogenia que se refere à determinação da distância evolutiva de espécies baseando-se em suas relações genéticas. No passado, as reconstruções evolucionárias da árvore da vida eram realizadas baseando-se, principalmente, na identificação do ponto de divergência entre as espécies, considerando-se unicamente características homólogas compartilhadas. Contudo, tal metodologia poderia tornar-se muito complicada devido às evoluções convergente e divergente (PATANÉ; MARTINS; SETUBAL, 2018).

Com o advento das técnicas moleculares, a filogenética foi aperfeiçoada enormemente pelo uso de diferenças entre nucleotídeos e marcadores universais de referência, criando a área da filogenômica (DELSUC *et al.*, 2005; PÉREZ-LOSADA; ARENAS; CASTRO-NALLAR, 2018). Na era pós-genômica, uma segunda onda de mudanças trouxe novas abordagens à filogenômica, que agora infere a divergência evolucionária tirando vantagem de dados de genomas inteiros, como: conteúdo e ordem gênica; ortologia; e estrutura da fita de DNA ou assinatura do DNA (KUMAR *et al.*, 2012). Neste sentido, a filogenômica pode ser definida como a junção da filogenética e genômica para reconstruir árvores confiáveis de espécies, analisar a distribuição e espalhamento bacteriano e predizer genes ortólogos e parálogos (CHAN; RAGAN, 2013; OCAÑA; DÁVILA, 2011).

A metodologia de tipagem genotípica por sequenciamento *multilocus* (*Multilocus Sequence Typing* – MLST) é classicamente utilizada para análises epidemiológicas por possuir alto poder de discriminação entre cepas bacterianas. Inicialmente descrita por Maiden *et al.* (1998), o método se baseia na análise de, normalmente, sete genes constitucionais, chamados também de genes *housekeeping*, compostos por cerca de 450 pares de base (pb) cada. O perfil criado pelo conjunto dos alelos de cada gene define o chamado perfil alélico do isolado, ou *sequence typing* (ST) (MAIDEN, MARTIN C. J. *et al.*, 2013; PÉREZ-LOSADA; ARENAS; CASTRO-NALLAR, 2018).

Atualmente, uma ferramenta nova vem ganhando espaço na área de filogenômica pela dinâmica de análise de quantidades pré-determinadas de genes e, inclusive, todo o genoma. O chamado *whole genome MLST* (wgMLST), que trabalha

como um MLST estendido, possui poder discriminatório notavelmente superior se comparado aos usuais sete genes da metodologia clássica, sendo capaz de distinguir clones entre isolados de uma mesma espécie (JOLLEY; MAIDEN, 2014).

1.3.3. Pan-genômica

Uma metodologia criada para prover uma visão mais ampla do genoma de uma espécie ou gênero é a metodologia de pan-genômica. A ideia de pan-genoma foi inicialmente introduzida por Tettelin *et al.* (2005) utilizando as sequências genômicas de oito linhagens de *Streptococcus agalactiae*. Após esta pesquisa, outros estudos foram realizados utilizando análises pan-genômicas de diferentes microrganismos, incluindo *Bacillus cereus* (RASKO *et al.*, 2005), *Escherichia coli* (RASKO *et al.*, 2008), *Sulfolobus islandicus* (RENO *et al.*, 2009), *Streptococcus pneumoniae* (DONATI *et al.*, 2010), *Methanobrevibacter smithii* (HANSEN *et al.*, 2011), *Corynebacterium diphtheriae* (TROST *et al.*, 2012), *Corynebacterium pseudotuberculosis* (SOARES *et al.*, 2013), *Pantoea ananatis* (DE MAAYER *et al.*, 2014), entre outros.

Ultimamente, a disponibilidade de um grande número de genomas de diferentes isolados do mesmo patógeno tem aberto a possibilidade de investigar várias características genômicas intrínsecas à espécie (TETTELIN, H. *et al.*, 2005). Neste contexto, o objetivo principal dessa metodologia é a comparação genômica de diferentes linhagens de uma dada espécie ou até mesmo gênero, visando identificar similaridades e diferenças entre elas (ALCARAZ *et al.*, 2010; SNIPEN; ALMØY; USSERY, 2009).

O pan-genoma é definido como o repertório completo e não-redundante de genes de uma dada espécie ou gênero e é composto de três subgrupos:

- I. *Core genome* (ou genoma central), que corresponde a todos os genes compartilhados pelo conjunto de dados estudado. Usualmente, são genes altamente conservados e estão envolvidos em processos celulares essenciais relacionados à garantia do funcionamento do microrganismo (MUZZI; MASIGNANI; RAPPUOLI, 2007);
- II. *Shared genome* (também chamado de *core* estendido), que consiste de genes que são compartilhados por duas ou mais linhagens, mas não estão presentes

em todas. Geralmente são genes envolvidos em funções adaptativas e virulentas de determinados nichos (LAPIERRE; GOGARTEN, 2009; MIRA *et al.*, 2010); e,

- III. *Singletons* (ou genoma exclusivo), que são genes espécie/linhagem específicos, presente apenas em determinada espécie/linhagem e ausente nas demais. Normalmente, os genes que estão presentes neste subgrupo dos *singletons* estão envolvidos na adaptação da bactéria ao ambiente ou hospedeiro. São pouco conservados com relação a outros grupos, e são relacionados, também, com HGT, além de estarem intimamente envolvidos em aquisição e perda de material genético de virulência e patogenicidade (JORDAN, 2001; MEDINI *et al.*, 2005; TETTELIN, H. *et al.*, 2005; TETTELIN, HERVÉ *et al.*, 2008).

Além da classificação em subgrupos, o pan-genoma também pode ser classificado em aberto ou fechado, dependendo da quantidade de novos genes que são adicionados ao pan-genoma a cada novo genoma adicionado na análise (MEDINI *et al.*, 2005). Ou seja:

- I. O pan-genoma aberto representa aquele cujo número de novos genes adicionados ao pan-genoma aumenta a cada novo genoma incluído na pesquisa, tendendo ao platô;
- II. O pan-genoma fechado é aquele cujo número de novos genes tende a zero, conservando o número de genes já presentes no pan-genoma antes da adição de genomas.

1.3.4. Plasticidade genômica

Procariotos são organismos muito promíscuos, comparados a eucariotos, no sentido de que os mesmos podem adquirir novas adaptações ao ambiente através da incorporação de DNA recebido de diferentes organismos via HGT (BOTO, 2010).

Vários mecanismos diferentes podem estar envolvidos em eventos de HGT, exercendo um papel importante na evolução através da incorporação de plasmídeos, bacteriófagos, transposons, elementos de inserção e ilhas genômicas. Devido a características específicas da fonte e a mecanismos utilizados na incorporação

genômica, regiões adquiridas horizontalmente possuem em comum: (i) um desvio na assinatura genômica (conteúdo G+C e uso de códon), que reflete a assinatura genômica do organismo doador; (ii) a presença de sequências de inserção e/ou tRNAs flanqueadores, que, por sua vez, podem apresentar uma sequência de inserção específica em suas regiões 3'-terminais; e, (iii) a presença de transposases (AZEVEDO *et al.*, 2011).

Ilhas genômicas (GIs, do inglês: *Genomic Islands*), grandes regiões adquiridas por eventos de HGT, podem estar ausentes em organismos do mesmo gênero ou espécie e ainda albergarem altas concentrações de genes com funções correlatas e específicas, que as classificam em: ilhas de resistência (RIs), com altas concentrações de genes de resistência a antibióticos (KRIZOVA; NEMEC, 2010); ilhas simbióticas (SIs), que podem estar correlacionadas com a associação de bactérias à plantas hospedeiras da família Leguminosae, por exemplo (BARCELLOS *et al.*, 2007); ilhas metabólicas (MIs), com vários genes associados com a biossíntese de metabólitos secundários (TUMAPA *et al.*, 2008); e, ilhas de patogenicidade (PAIs), que possuem altas concentrações de fatores de virulência, estão presentes em bactérias patogênicas e ausentes em espécies não-patogênicas do mesmo gênero e/ou espécie correlata, além de estarem envolvidas na re-emergência de vários organismos patogênicos devido à inserção de novos determinantes de virulência (DOBRINDT *et al.*, 2000).

Em vista de sua alta instabilidade e da incorporação de vários genes em bloco, GIs são muito interessantes para análises de genômica comparativa, principalmente para pan-genômica, onde a informação para várias linhagens está disponível. As análises pan-genômicas associadas com a predição de GIs podem ajudar a compreender a plasticidade genômica do gênero ou espécie como um todo e ajudar a correlacionar a presença/ausência de diferentes regiões com a adaptabilidade ao hospedeiro ou ambiente.

1.4. Abordagens genômicas em *Bartonella henselae*

As linhagens de *B. henselae* são divididas em dois genótipos diferentes de acordo com o rRNA 16S: (I) tipo 1 ou Houston-1; e (II) tipo 2, ou Marseille (LA SCOLA *et al.*, 2002). Ambos estão distribuídos mundialmente, com prevalências relativas em

determinadas regiões, sendo o genótipo Houston-1 encontrado comumente na Ásia e o genótipo Marseille em gatos europeus (BOULLOUIS *et al.*, 2005).

Embora as bactérias do gênero *Bartonella* apresentem alta importância médica, veterinária e zoonótica, atualmente existem poucos estudos sobre o perfil genômico das espécies.

Alguns dos trabalhos relatados na literatura visam, por exemplo, a caracterização de genes específicos. Fiskus *et al.* (2003) encontraram, através de alinhamentos múltiplos de sequências, genes presentes em *B. henselae* e *B. bacilliformis* que codificam proteínas com considerável homologia com espécies da ordem *Rizhobiales*, como os gêneros *Brucella* e *Agrobacterium*. As proteínas (DdIB, FtsQ e FtsA) são relacionadas com a síntese da parede celular e no processo de divisão celular das bactérias. Riess *et al.* (2007) contribuíram para o entendimento da variabilidade do gene que expressa a Adesina A, um importante fator de virulência em linhagens de *B. henselae* (BadA). O trabalho revelou que eventos de plasticidade genômica, como inserções e deleções, são capazes de alterar a expressão de BadA em linhagens da bactéria, coexistindo linhagens BadA-positivas e BadA-negativas.

Lindroos *et al.* (2006) avaliaram o conteúdo genômico e estrutura de 38 genomas de *B. henselae* isolados de gatos e humanos. O estudo demonstrou o tamanho homogêneo do genoma da espécie, variando de 1,9 a 2,0 Mb, além da alta frequência de rearranjos em regiões de GIs. De maneira geral, o trabalho sugere que as diferenças derivadas da plasticidade genômica entre as linhagens determina diferentes padrões de expressão proteica que podem favorecer ou dificultar a evasão ao sistema imune e alterar consideravelmente a patogenicidade do microrganismo no hospedeiro durante sua permanência.

Com o propósito de avaliar as taxas de mutações e recombinações em paralelo com a análises do pan-genoma de *B. henselae*, Guy *et al.* (2012) investigaram a variação de genes desse microrganismo. O estudo foi promissor para a dedução da elevada taxa de recombinação gênica intra-espécie, onde as regiões com maior predominância desses eventos foram identificadas em segmentos de genes dos sistemas T4SS. Relativamente, pelo fato dos genomas apresentarem variabilidade nessas regiões específicas relacionadas com virulência, os dados são sugestivos de recombinação dos genes de virulência decorrentes de processos de adaptação a

hospedeiros. Além disso, o pan-genoma foi classificado como fechado, com um conjunto delimitado de genes, dada a baixa variação gênica nos genomas da espécie como um todo.

Estudos anteriores revelaram a possibilidade de HGT, de adaptabilidade e patogenicidade envolvendo espécies do gênero *Bartonella* (BERGLUND *et al.*, 2009). Uma revisão bibliográfica de 2019 reuniu estudos acerca de eventos relacionados à transferência de genes e suas possíveis origens e impactos nas bactérias do gênero *Bartonella*. A base para as mudanças que ocorrem nos genomas é multifatorial, com origens em HGT e em recombinações gênicas dentro de uma espécie com pan-genoma fechado, como no caso das bartonelas. Conseqüentemente, levando em consideração o estilo de vida dessas bactérias, principalmente pelo desenvolvimento intracelular facultativo, a restrição a ambientes específicos favorece a conferência de adaptabilidade a hospedeiros (QUÉBATTE; DEHIO, 2019).

2. JUSTIFICATIVA

De forma geral, as bartoneloses apresentam-se como doenças negligenciadas, além de emergentes, com um amplo espectro de manifestações clínicas importantes para a saúde de pessoas e animais ao redor do mundo, principalmente por causar condições crônicas e debilitantes para pessoas de baixos recursos.

Com o advento da era genômica, novas portas se abrem para pesquisa no gênero de bactérias patogênicas. A busca por conhecimentos genômicos da espécie causadora da DAG, utilizando novas estratégias não antes descritas na literatura, poderiam ajudar a identificar: (1) os padrões de distribuição dos complexos clonais da espécie utilizando técnicas atuais de epidemiologia molecular de genoma inteiro; (2) potenciais fatores de virulência que possam estar relacionados com o processo patológico e com os diferentes sintomas causados pelas linhagens; e (3), métodos de aquisição de fatores de virulência, via identificação dos mecanismos de HGT, por meio de análises de plasticidade genômica voltadas para a predição de ilhas genômicas, principalmente ilhas de patogenicidade.

3. OBJETIVOS

3.1. OBJETIVO GERAL

Realizar análises de sintenia gênica, filogenômica e pan-genômica de 24 linhagens de *Bartonella henselae* disponíveis publicamente, incluindo a identificação e investigação de genes de fatores de virulência presentes em regiões adquiridas horizontalmente.

3.2. OBJETIVOS ESPECÍFICOS

- Realizar análises de filogenia por fragmentação-alinhamento-filogenia, gene a gene e wgMLST;
- Analisar os rearranjos genômicos intraespécie via sintenia gênica;
- Caracterizar genes ortólogos;
- Classificar os grupos de genes ortólogos nos subconjuntos do pan-genoma;
- Desenvolver o pan-genoma intraespécie, incluindo seus subconjuntos
- Classificar os subconjuntos de genes do pan-genoma nas categorias funcionais do *clusters of orthologous groups* (COG);
- Identificar regiões adquiridas horizontalmente, localizar ilhas genômicas de patogenicidade e de integração de fagos e analisar o conteúdo das mesmas.

4. The pan-genome of the zoonotic neglected pathogen *Bartonella henselae* reveals two groups with different patterns of adaptation to hosts

As análises de genômica comparativa decorrentes do trabalho foram suficientes para preparação do artigo intitulado “The pan-genome of the zoonotic neglected pathogen *Bartonella henselae* reveals two groups with different patterns of adaptation to hosts”, publicado na revista Journal of Bacteriology and Mycology, do grupo Austin Publishing. Em geral, foram utilizadas análises de pangenômica voltadas para a identificação do conteúdo gênico conservado (*core* genoma) e variável (*shared* genoma e *singletons*), avaliando as possíveis regiões advindas da plasticidade genômica do organismo via análises de ilhas genômicas.

Research Article

The Pan-Genome of the Zoonotic Neglected Pathogen *Bartonella henselae* Reveals Two Groups with Different Patterns of Adaptation to Hosts

Alves LG¹, Oliveira LC², Benevides LJ³, Zen FL², Tiwari S¹, Jaiswal A¹, Ghosh P⁴, Barh D⁵, Azevedo V¹ and Soares SC^{2*}

¹Department of Microbiology, Immunology and Parasitology, Institute of Biological Sciences, Federal University of Minas Gerais, Brazil

²Department of General Biology, Institute of Biological Sciences and Natural Sciences, Federal University of Triângulo Mineiro, Brazil

³Bioinformatics Laboratory, National Laboratory for Scientific Computing, Brazil

⁴Department of Computer Science, Virginia Commonwealth University, USA

⁵Center for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), India

*Corresponding author: Soares SC, Department of General Biology, Institute of Biological Sciences and Natural Sciences, Federal University of Triângulo Mineiro, 30 Frei Paulino St, Uberaba, MG, 38025180, Brazil

Received: May 15, 2020; Accepted: June 04, 2020;

Published: June 11, 2020

Abstract

The genus *Bartonella* is comprised of Gram-negative re-emerging bacteria like *Bartonella henselae*, which mainly infects humans and survives inside erythrocytes. This species is transmitted by scratches and bites from domestic cats and usually causes a symptomatic infection in humans, known as Cat-Scratch Disease (CSD). The disease causes multiple clinical signs in humans, such as dermatologic, cardiovascular, lymphatic, hepatic and nervous system diseases in immunosuppressed individuals. Although the bacteria are highly relevant for its zoonotic importance worldwide, few studies aimed at characterizing these species genomes and there is still no pan-genome study available. Here, we performed phylogenomic, pan-genome and genome plasticity analyses to determine the epidemiological aspects, the size of the pan-genome and its variability in the identified pathogenicity and resistance islands. Altogether, our results showed that the genomes are highly similar, with an almost closed pan-genome. Also, we found two subsets of genomes, composed of 7 and 17 genomes of bacteria. Our results point to the need of sequencing more genomes worldwide to better characterize these variations in the pan-genome and understand the patterns of adaptation of this species. The highly conserved genomes from this species are very important for the development of new vaccines and analyses of drug targets against this pathogen. Furthermore, these data may then be used in future works, which will be highly relevant for containing the disease worldwide.

Keywords: *Bartonella henselae*; Cat-Scratch Disease; Pan-Genome; Genome Plasticity; Phylogenomics

Abbreviations

ACT (Artemis Comparison Tool); *B. apis* (*Bartonella apis*); *B. henselae* (*Bartonella henselae*); BRIG (BLAST Ring Image Generator); CDSs (Coding DNA Sequences); COG (Cluster of Orthologous Groups); CSD (Cat-Scratch Disease); DNA (deoxyribonucleic acid); GC-content (guanine-cytosine content); GIPSY (Genomic Island Prediction Software); LCBs (Locally Collinear Blocks); Mb (mega base pairs); kb (kilo base pairs); MCL (Markov Clustering); NCBI (National Center for Biotechnology Information); PAIs (Pathogenicity Islands); RIs (Resistance Islands); T4SS (type IV secretion system); USA (United States of America); VFDB (Virulence Factor Database); wgMLST (whole genome Multilocus Sequence Typing).

Introduction

The genus *Bartonella* is comprised of Gram-negative, fastidious, intracellular and reemerging bacteria [1]. Some species from this genus mainly infect cats, but is also an opportunistic pathogen for humans and are frequently acquired through hematophagous arthropod vectors. Those species are able to infect and survive inside erythrocytes, through a long intra-erythrocytic and intra-endothelial infection, which results in recidivist bacteremia both in humans and other mammals [2]. The genus presents at least 13 human pathogenic

species [3], among which the 3 most relevant ones being: *Bartonella bacilliformis*, which causes the Oroya fever [4]; *Bartonella quintana*, causing trench fever [5]; and, *Bartonella henselae*, the causative agent of the Cat-Scratch Disease (CSD).

Domestic cats are the main reservoirs of *B. henselae* and both immunocompetent and immunosuppressed cats are frequently asymptomatic to the infection, though they can present subclinical infections and suffer from recidivist bacteremia through their lives [6]. *B. henselae* transmission between cats occurs through arthropod vectors, mainly fleas [7] and dogs are also potential reservoirs of *B. henselae* [8]. The transmission route between humans and cats occurs mainly through the dermic inoculation of the bacteria through scratches or bites of contaminated cats [9]. However, other studies show *B. henselae* infections of immunocompetent humans through the possible inoculation by ticks and spider bites [10,11]. In the bloodstream, the bacteria invade erythrocytes, where they persist intracellularly, causing erythrocytic and vascular endothelial alterations [12].

B. henselae infection in humans causes multiple clinical signs, such as dermatologic, cardiovascular, lymphatic, hepatic and nervous system diseases [13]. CSD is characterized by circumscribed regional lymphadenopathy in the inoculation site and previous studies revealed an immune-dependent pattern of clinical manifestations of

CSD, mainly in immunosuppressed humans [14].

Although the bacteria present a high veterinary, medical, and zoonotic importance, there are currently few studies about the genomic profile of the species and the few existing ones aim at characterizing specific genes and analyzing the recombination and mutation rates [15-18]. Previous studies revealed the possibility of a horizontal gene transfer of adaptability and virulence genes in the species from the genus *Bartonella* [19]. Altogether, the fact that *B. henselae* is distributed worldwide and considered as a neglected zoonotic pathogen, opens doors for new genome plasticity studies of the species. In this work, we performed phylogenomics, pan-genomics, and genome plasticity analyses to find the possible epidemiological relationships between the strains, the conserved and variable subsets of genes and the pathogenicity and antibiotic resistance islands that may be involved in the pathogenesis process.

Material and Methods

Genomes

We retrieved 24 genomes of different strains of *B. henselae* publicly available at the GenBank on NCBI (<https://www.ncbi.nlm.nih.gov>). The genomes were retrieved through FTP in .gbff, .fna and .faa file formats. The genome of *Bartonella apis* was used as a non-pathogenic reference where applicable.

Phylogenomic analyses

The software Gegenees was used to compare the percentage of similarity between the 24 genomes [20]. The .fna files containing the complete and draft genomes of *B. henselae* were imported to Gegenees and analyzed using the default parameters. Briefly, Gegenees fragments all genomes in pre-defined sizes; performs similarity analyses using BLASTn among the genomes to identify the commonly shared regions; and, finally, the software creates a heatmap with the percentage of similarity among the genomes. Here, we used a fragmentation size of 200 bp and a similarity threshold of 40%.

The similarity matrix generated by Gegenees was then exported in .nexus format and further used in the software SplitsTree [21] to create a neighbour-joining phylogenetic tree based on the similarity among the strains using the option Equal Angle.

For a greater resolution on the phylogenomic tree analysis, a whole genome Multilocus sequence typing (wgMLST) analysis [22] was performed via the online software PGADB-builder [23]. The 24 genomes of *B. henselae*, in the .fasta format, were compared with the PGADB profile through BLASTn, using the module Build_PGADB. Next, the PGADB profile of the genomes, in the .scheme format, was used to build a wgMLST tree using the module Build_wgMLSTtree, with a 90% coverage and 90% identity filter. After the BLASTn analyses, the output file .newick was exported and used as input file in the software MEGA7 [24] for the reconstruction of the phylogenomic tree.

Gene Synteny

The gene synteny analysis was performed using the software Mauve [25]. Briefly, Mauve fragments the genomes in pre-defined sizes, creates Locally Collinear Blocks (LCBs) using the sequence alignments and exports the results as a figure where the rearrangement events are represented. Here, we imported the .fna files from the 24

strains and used the Houston-1 strain as reference along with the “progressiveMauve” algorithm.

Pan-genomic analyses

The prediction of orthologous genes was performed with the software OrthoFinder [26] and further classified with the use of an in-house script in: core genome, containing only genes that are commonly shared by all strains; shared genome, containing genes that are shared by two or more strains, but not all; and singletons, with strain-specific genes. Briefly, OrthoFinder used the .faa files with the amino acid sequences of all Coding DNA Sequences (CDSs) in each genome to perform an *all-vs-all* BLASTp analysis. The sequences were then grouped using the Markov Clustering (MCL) algorithm to determine the orthologous genes [27].

Pan-genome and subsets development

The pan-genome, core genome and singletons development were calculated based on the mean values of the permutations of all genomes using the method described in Soares et al. [28]. The final curves were then fitted using an *in-house* script to estimate the fixed parameters for Heap’s Law (pan-genome analyses) and least-squares fit of the exponential regression decay (core-genome and singletons). The extrapolations of the pan-genomes from the different datasets were calculated by curve fitting based on Heap’s Law with the formula $n = \kappa * N^{-\alpha}$, where n is the expected number of genes for a given number of genomes, N is the number of genomes, and the other terms are constants defined to fit the curve. The extrapolations of the core genomes and singletons for all datasets were calculated by curve fitting based on least-squares fit of the exponential regression decay with the formula $n = \kappa * \exp[-x/\tau] + tg(\theta)$, where n is the expected subset of genes for a given number of genomes, x is the number of genomes, \exp is Euler’s number, and the other terms are constants defined to fit the curve.

We used 7 subsets of genes for all the analyses: the complete genome dataset; 2 subsets with genomes isolated from different hosts (cats and humans); two subsets of genomes from different locations (France and USA); and, two subsets containing the genomes of the two groups identified on Gegenees and PGADB (genomes 2-8 and 9-25 on Gegenees). The group 2-8 includes the strains: A242, A244, A121, A112, A233, U4 and BM1374165, while the group 9-25 consists of the strains: A235, JK41, JK42, Zeus, JK53, A76, A74, BM1374163, MVT02, A20, A71, F1, FDAARGOS175, Houston-I, Houston-1, JK51, and JK50.

Classification of CDSs of the pan-genome subsets into the cluster of orthologous groups

The subsets of the pan-genome (core, shared and singletons) were classified according to the functional categories of the cluster of orthologous groups (COG) into 1 – Information storage and processing; 2 – signaling and cellular processes; 3 – metabolism; and, 4 – poorly characterized. To perform this classification, the CDSs of the subsets were BLAST aligned against the myva database of COG, using an e-value of $1e-6$, and the result was crosschecked with the WHOG information from COG [29].

Horizontally acquired regions

The software GIPSY (Genomic Island Prediction Software) [30,31] was used to perform the prediction of Pathogenicity Islands

Table 1: Genomic features of the strains used in this study.

Species	Strain	Assembly	Genome size (pb)	%GC	No. of contigs	No. of scaffolds	No. of genes	No. of proteins	Isolation country	Host	Collection date
<i>Bartonella henselae</i>	Houston-1	GCA_000046705	1.931.050	38.2	1	1	1.743	1.52	USA	Human	
	BM1374165	GCA_000612765	1.975.500	38.1	1	1	1.767	1.538	France	Cat	
	BM1374163	GCA_000612965	1.905.380	38.2	1	1	1.705	1.505	France	Cat	
	MVT02	GCA_001291465	1.905.380	38.2	1	1	1.706	1.505	France	Human	
	Houston-I	GCA_001525625	2.014.710	38.3	1	1	1.757	1.52	USA	Human	Jan-95
	JK 51	GCA_000516675	1.926.860	38.1	26	7	1.731	1.511	USA	Cat	Oct-94
	JK 50	GCA_000516695	1.921.450	38.2	28	5	1.732	1.505	USA	Human	Sep-94
	JK 42	GCA_000516715	1.921.970	38.3	47	13	1.691	1.469	USA	Cat	Jul-93
	JK 41	GCA_000516735	1.875.760	38	33	9	1.675	1.466	USA	Human	Jul-93
	Zeus	GCA_000708485	1.940.920	38.2	60	19	1.727	1.465	USA	Cat	1990/1995
	JK 53	GCA_000708545	1.940.190	38.3	53	12	1.711	1.462	USA	Human	1995/2005
	FDAARGOS_175	GCA_001525625	2.014.750	38.3	2	2	1.803	1.547	USA	Human	1990
	A20	GCA_001932075	1.840.740	38	66	66	1.596	1.454	France	Cat	Apr-96
	A74	GCA_001932085	1.855.280	38	71	71	1.61	1.471	France	Cat	Mar-96
	A76	GCA_001932095	1.862.680	37.9	58	58	1.613	1.466	France	Cat	Mar-96
	A71	GCA_001932135	1.839.690	38	74	74	1.607	1.463	France	Cat	Apr-96
	A112	GCA_001932145	1.838.410	37.9	64	64	1.63	1.467	France	Cat	May-96
	A233	GCA_001932165	1.835.600	37.9	46	46	1.618	1.461	Denmark	Cat	Oct-98
	A121	GCA_001932175	1.837.810	37.9	53	53	1.615	1.458	France	Cat	May-96
	A242	GCA_001932215	1.860.020	38	46	46	1.638	1.486	Denmark	Cat	Mar-98
A235	GCA_001932225	1.821.810	38	42	42	1.584	1.448	Denmark	Cat	Mar-98	
A244	GCA_001932235	1.857.850	38	43	43	1.628	1.47	Denmark	Cat	Mar-98	
F1	GCA_001932245	1.864.200	38	67	67	1.621	1.479	USA	Cat	1996	
U4	GCA_001932295	1.812.180	38	56	56	1.576	1.42	USA	Cat	1996	
<i>Bartonella apis</i>	BBC0122	GCA_002007565	2.907.210	45.7	1	1	2.546	2.445	Switzerland	<i>Apis mellifera</i>	Dec-14

(PAIs) and Resistance Islands (RIs) in the genomes of *B. henselae*. Briefly, GIPSY is a multi-pronged tool that identifies the most common features of the pathogenicity islands, such as genomic signature deviation, tRNA flanking genes, transposases and a high concentration of virulence factors. We used .gbk files and default parameters in our analyses. The genome of *Bartonella apis* [32] was used as the non-pathogenic reference.

The phage regions, i.e. regions harboring phage sequences, were predicted and annotated by the software PHASTER [33] using the .fasta file of the Houston-1 strain. PHASTER performs gene prediction in bacterial genome followed by BLAST analyses against a customized phage database on NCBI and prophage database developed by Srividhya et al. [34,35], which contains phage genes. PAI and phage regions were plotted in comparative circular genome representations on the software BRIG (BLAST Ring Image Generator) [36].

We used the software BRIG to create a circular genome mapping using different genomes as references. The genomic sequences in .fasta format were used on BRIG, which performed the comparative analysis using BLAST to compare the strains of *B. henselae* with the reference Houston-1 strain. Furthermore, the species *B. apis* was also added to the analysis. The results were plotted on a circular comparative map.

The image generated is composed of rings, where each genome is represented by a ring of a given color. The intensity of such color is related to the degree of similarity with the reference genome. Deleted regions of the genomes, compared with the reference, are represented by empty regions (blanks). The coordinates of the islands predicted by GIPSY and the phages predicted by PHASTER were added to BRIG's circular map for visualization of the genomic plasticity events. Two additional plasticity analyses were performed using, as references, strains isolated from cats and humans (BM1374165 and Houston-I, respectively). Both were selected as references because they are strains with the largest complete genomes among the 24 available genomes for each host group.

The coordinates of the pathogenicity islands predicted by GIPSY were analyzed at the ACT (Artemis Comparison Tool) software [37] and exported to a list of the genes present in each region. The tool PATRIC [38] was used to find information on the virulence genes found in the analysis. The tool uses BLASTp to compare the genes found on the islands against a database that integrates the Virulence Factor Database (VFDB) [39], MvirDB [40] and manually annotated virulence genes, mainly of the genus *Mycobacterium*, *Shigella*, *Salmonella*, *Escherichia*, *Listeria* and *Bartonella* [38].

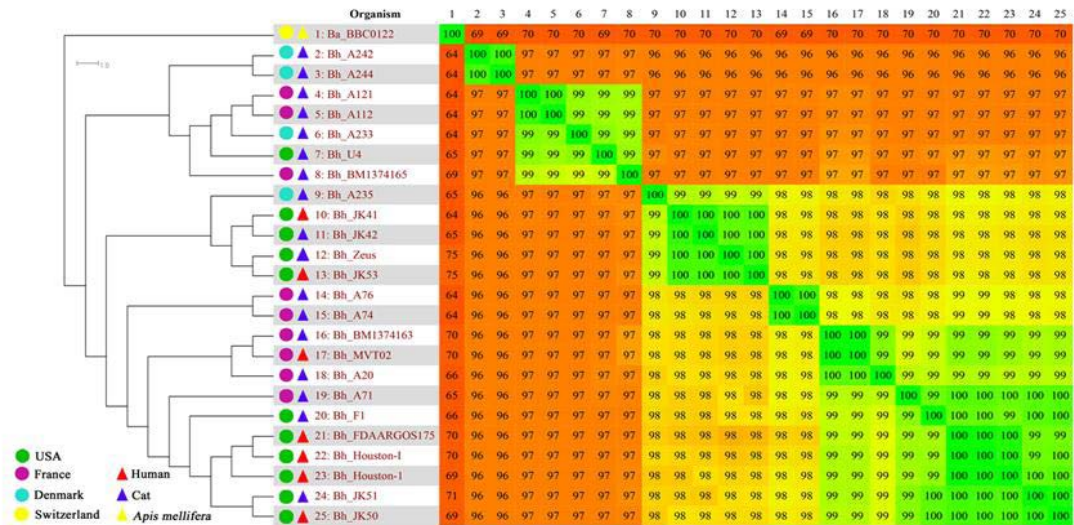


Figure 1: Phylogenomic tree correlating the evolutionary distance among the strains of *Bartonella henselae*. The genomes of the bacteria are enumerated and named in the 25 lines, whereas the genomes are found in the same order in the columns at the upper region of the chart. The heatmap values indicate the percentage of similarity among the genomes. The colors indicate the degree of similarity, varying from red (low similarity) to green (high similarity). The phylogenomic reconstruction was performed using the SplitsTree software based upon the distance matrix generated by the Gegenees software. The circle next to the name of the strain indicates the country of isolation of the same, according to the respective color of the legend. The triangle represents the host from which the bacterium was isolated for sequencing, also following the color of the legend. Bh: *Bartonella henselae*. Ba: *Bartonella apis*. N/A: Not applicable. USA: United States of America.

Results

General Features

The genomic features of all the genomes used for the analyses are presented in Table 1. Regarding the species *B. henselae*, the genomes have a relatively homogenous size, varying between 1,81 and 2,02 Mb, with a gene content in the range from 1576 to 1803 genes and a GC-content that varies from 37,9% to 38,3%. The amount of contigs on the draft genomes varies from 2 to 74 (Table 1).

Phylogenomic analyses

The similarity analysis was made using Gegenees software to determine the degree of genomic variability among the strains of *B. henselae*. The generated heatmap (Figure 1) shows a high similarity among the strains of *B. henselae*, varying from 96 to 100%. The strains were sub-grouped on clusters, generating 2 clusters on the heatmap. The first cluster comprises of the strains 2 to 8, with the similarity varying from 99% to 100% among them, where all the strains were isolated from cats. The second cluster is formed by the other strains (9-25), with similarities varying from 98% to 100%, where the strains were isolated both from cats or humans. A phylogenomic tree based on Gegenees analysis was also plotted on Figure 1. Noteworthy, the strains A244 and A242, isolated from cats, are more distantly related to the others, whereas they cluster together with other 5 strains (A121, A112, A233, U4 and BM1374165), also isolated from cats.

A phylogenomic tree based on wgMLST analysis is shown in Figure 2, explaining the evolutionary correlation among the strains, where the short evolutionary distance among the various strains is noticeable. As observed at the SplitsTree reconstruction, the correlation of ancestry among the strains corresponds to the same pattern previously found on Gegenees, with a high bootstrap value, while the strains A112, A121, A233, BM1374165, U4, A242, and

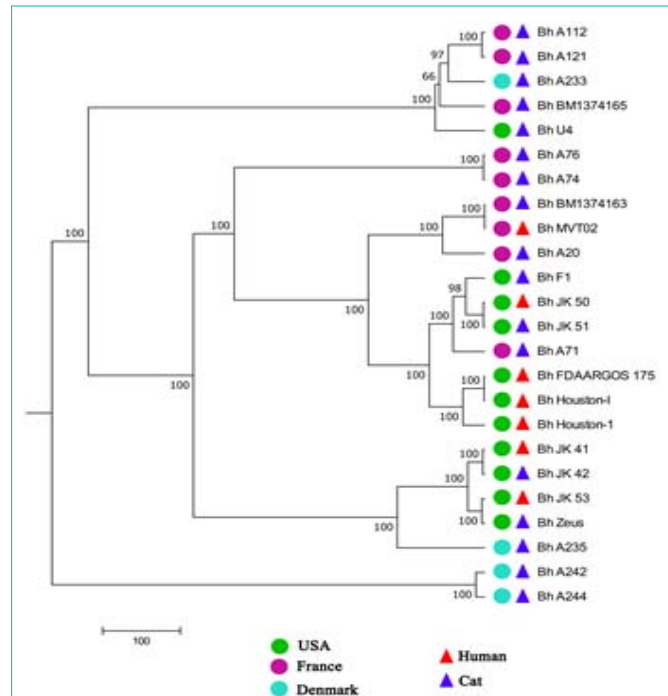


Figure 2: Dendrogram created with the wgMLST profile of the 24 genomes of *Bartonella henselae*. The PGADB profile of the genomes of *Bartonella henselae* was used to construct a wgMLST tree using the Build_wgMLSTree module. The bootstrap values are shown at the nodes. The phylogenomic tree was built via the online software PGADB-builder and the file ".newick" was exported and used in MEGA7 for the tree visualization. The circle next to the name of the strain indicates the country of isolation of the same, according to the respective color of the legend. The triangle represents the host from which the bacterium was isolated for sequencing, also following the color of the legend. Bh: *Bartonella henselae*. N/A: Not applicable. USA: United States of America.

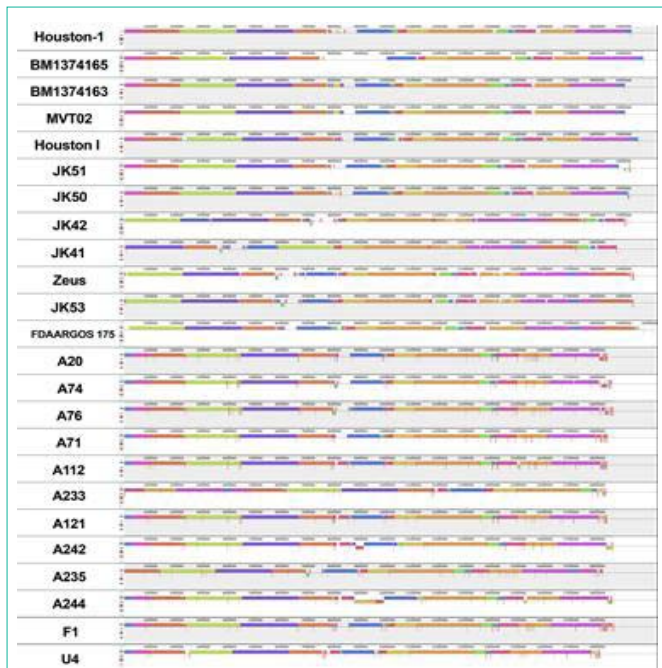


Figure 3: Gene synteny analysis of the complete and draft genomes of *Bartonella henselae*. The limits among the contigs are represented by red lines. The different Locally Collinear Blocks (LCBs) that are conserved among the strains are represented by different colors. The deleted regions of the genome are represented by blank spaces between the blocks, and the low similarity regions among the genomes are represented by blank spaces inside the blocks. The enumerated ruler above the LCBs represents the genomic position.

A244 are more distantly related from the others, including the strains isolated from humans. According to these results, other analyses were carried out using subsets based on the strains present in clusters 2-8 and 9-25.

Gene Synteny Analysis

At Figure 3, the prediction of LCBs on all the strains has shown large and numerous regions of homology, mainly among the complete genomes. All the draft genomes present small and large inverted regions represented by a change in the orientation of the LCBs from one DNA strand to another. The strain BM1374165 has a large new region compared to Houston-1.

The use of fragmented genomes hinders the visualization of the synteny among the strains. A more accurate analysis was made using only the completed genomes, as represented in Figure 4. There is a predominance of 3 LCBs: the first one (represented by the red color) with approximately 1.46 Mb, a second one (blue) with around 0.43 Mb and a small one (green) between the other two, with just 4.85 kb. Only the strain BM1374165 presents small deletions on a pathogenicity island (highlighted by the black box). Besides that, it is interesting to notice that MVT02 and BM1374163, although presenting 100% similarity on Gegenees, are still very similar on the analysis run by Mauve, as opposed to Houston 1 and Houston I, that present regions of deletion on different regions of the genome.

Pan-genomic analysis

The pan-genomic analysis through Ortho Finder software identified a pan-genome composed of 1655 non-redundant genes.

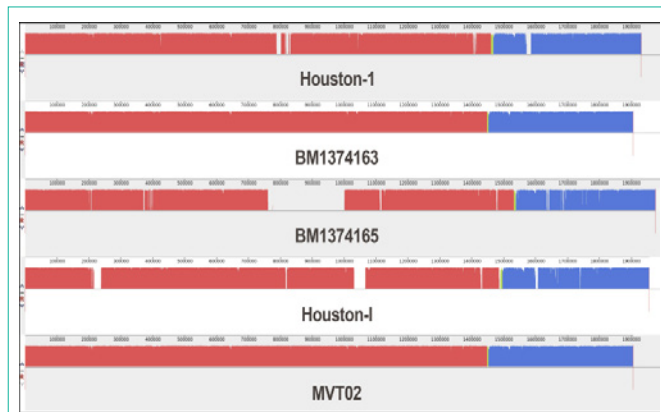


Figure 4: Gene synteny analysis of the completed genomes of *Bartonella henselae*. The genomes of the strains of *B. henselae* were aligned using the Mauve software. The limits among the contigs are represented by red lines. The different Locally Collinear Blocks (LCBs), that are conserved among the strains are represented by different colors. The deleted regions of the genome are represented by blank spaces between the blocks. The low similarity regions between the genomes are represented by blank spaces inside the block. The enumerated ruler above the LCBs represents the genomic position.

The core genome, composed of genes commonly shared by all species, contains 1160 genes, or 70.1% of the total number of genes in the pan-genome. The shared genome, with genes shared by two or more strains, but not all, has 388 genes (23.44% of the total pan-genome). Lastly, 107 genes (6.46%) are strain specific, being denominated singletons.

Pan-genome, core genome and singletons development

Considering $\alpha = 1 - \gamma$, where the γ value for the extrapolation of the pan-genome of the complete dataset was 0.06, the α value of the complete dataset is 0.94, whereas the subsets USA, France, cats, humans, 2-8 and 9-25 have the following α values, respectively: 0.94, 0.93, 0.93, 0.92, 0.90 and 0.95. According to Heap's Law, an α lower than 1 is representative of an open pan-genome, where each added genome will contribute some genes to the pan-genome, whereas an α higher than 1 is representative of a closed pan-genome, where newly added genomes will not contribute significantly to the pan-genome. Considering this rule, all pan-genomes are still open, although varying at different rates, where the subset of 2-8 presents the fastest changing pan-genomes, opposing to the subset 9-25, which present the slowest growing pan-genome.

The core genomes of the complete dataset and the subsets USA, France, humans, 2-8 and 9-25 have the following $tg(\theta)$ values, respectively: ~1136, ~1023, ~1141, ~1165, ~1230, ~1137 and ~1211. Concerning the singletons of the complete dataset and the subsets USA, France, Cats, humans, 2-8 and 9-25, they have the following $tg(\theta)$ values, respectively: ~1, ~6, ~3, ~3, ~12, ~16 and ~4.

According to the least-squares fit of the exponential regression decay, the $tg(\theta)$ represents the point where the curve will stabilize, which may be correlated to the number of genes in the core genome after stabilization and the number of singletons that will be added to the pan-genome for each newly sequenced genome. Considering this rule, the core genome of the complete dataset will have 1136 genes after stabilization, whereas the subset from the USA will be the

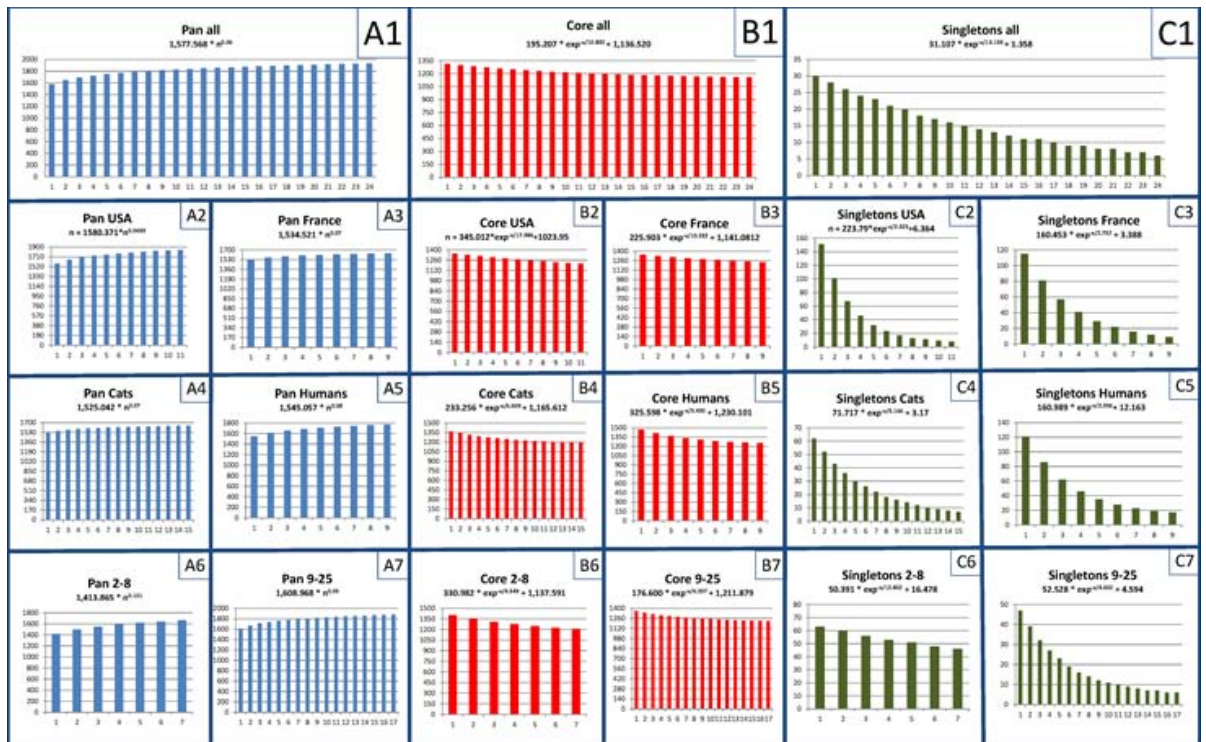


Figure 5: Pan-genome, core genome and singletons development of *Bartonella henselae*. A1/B1/C1, respectively, the pan-genome, core genome and singletons development using all 24 strains of *B. henselae*; A2/B2/C2, pan-genome, core genome and singletons development of the strains isolated from USA; A3/B3/C3, the pan-genome, core genome and singletons development of the strains isolated from France; A4/B4/C4, the pan-genome, core genome and singletons development of the strains isolated of cats; A5/B5/C5, the pan-genome, core genome and singletons development using of the strains isolated of humans; A6/B6/C6, the pan-genome, core genome and singletons development using of the clusters 2-8 generated from the software Gegenees; A7/B7/C7, the pan-genome, core genome and singletons development using of the clusters 9-25 generated from Gegenees.

Table 2: Locations and sizes of pathogenicity islands predicted in the reference strain *Bartonella henselae* Houston-1.

ISLAND	LOCATION	SIZE (pb)	
PAI 1	126.888	136.627	9.739
PAI 2	356.6	428.14	71.54
PAI 3	1.277.004	1.300.302	23.298
PAI 4	1.437.361	1.505.591	68.23
PAI 5	1.524.451	1.542.024	17.573
PAI 6	1.799.481	1.819.889	20.408

smallest one, with 1023 genes and the core genome of humans will be the largest one, with 1230 genes. On the singletons subset, on the other hand, only one gene will be added to the complete dataset for each newly sequenced genome, whereas the subsets from France and Cats will have 3 newly added genes and the subsets humans and 2-8 will have 12 and 16 newly added genes, respectively (Figure 5).

Classification of the CDSs in the subsets of the pan-genome on the cluster of orthologous groups

The classification of subsets of the pan-genome on the functional categories of the COG is represented in Figure 6. Most of the genes at the subsets shared and singletons are composed of “poorly characterized” genes. On the other hand, the core genome is mostly composed of genes that integrate the cellular “metabolism”, in addition to genes responsible for “cellular processes and signaling”

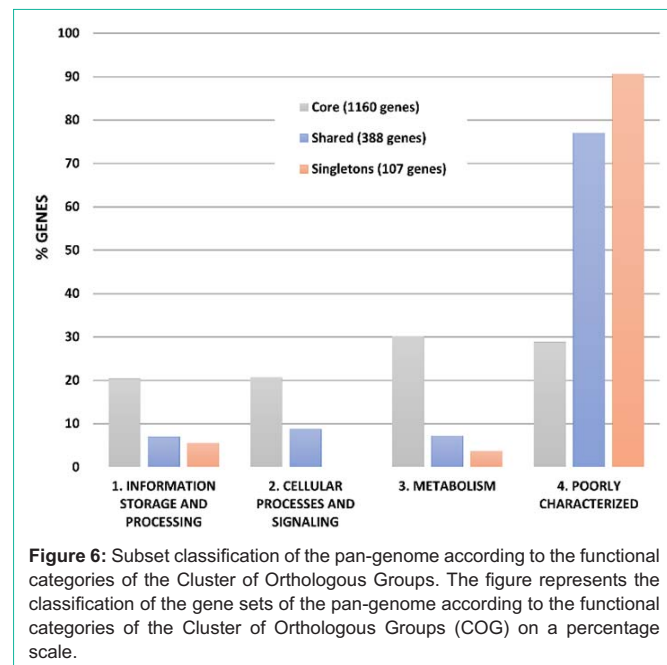


Figure 6: Subset classification of the pan-genome according to the functional categories of the Cluster of Orthologous Groups. The figure represents the classification of the gene sets of the pan-genome according to the functional categories of the Cluster of Orthologous Groups (COG) on a percentage scale.

and “information storage and processing”.

Genomic Plasticity

The analysis of the circular comparison of the genomes of the

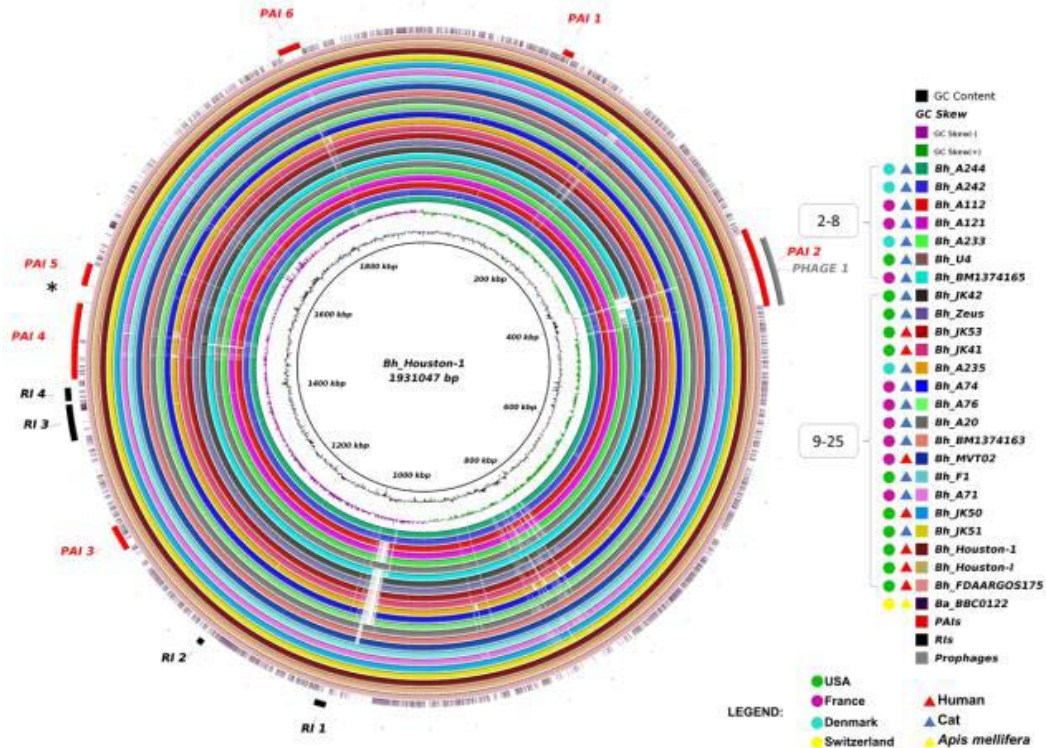
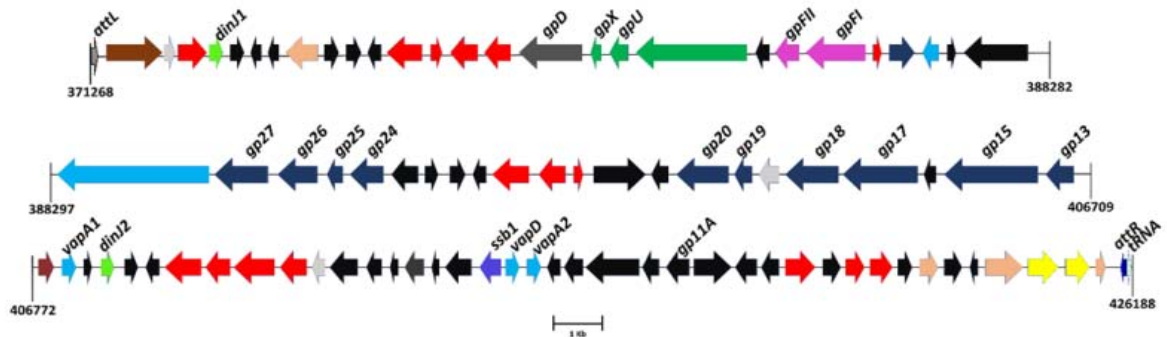


Figure 7: Circular representation of the genomic comparison among the strains of the *Bartonella henselae* species. The circular genomic comparisons were created with the BRIG software. The image represents the genomic plasticity among the genomes of *B. henselae* using the strain Houston-1 as a reference. Each ring in the image corresponds to one genome of one strain of *B. henselae*, with the respective color corresponding to the legend at the right. The deleted regions are represented by blank spaces inside the circles, while the shared genes regions are filled with color. The PAIs, RIs, and phage regions were predicted by the GIPSy and PHASTER software, respectively. The circle next to the name of the strain indicates the country of isolation of the same, according to the respective color of the legend. The triangle represents the host from which the bacterium was isolated for sequencing, also following the color of the legend. The numbers next to the keys indicate the numbering of the cluster generated by Gegenees and PGAdB analysis. The asterisk (*) represents the location of the cluster of virulence genes *virB*. Bh: *Bartonella henselae*. Ba: *Bartonella apis*. PAI: pathogenicity island. RI: resistance island N/A: Not applicable. USA: United States of America.



Legend:

Hypothetical prophage integrase	Late control gene D protein	Killer protein	<i>attL</i>
Hypothetical protein	Phage tail protein	DNA-damage-inducible protein	<i>tRNA</i>
Anti-repressor protein	Major tail tube protein	DNA-binding protein	<i>attR</i>
Hypothetical prophage protein	Phage protein	Single-strand binding protein (<i>ssb</i>)	
Phage related protein	Virulence-associated protein	Exonuclease	

Figure 8: Gene content of the phage region found in *Bartonella henselae* Houston-1. The phage region determined by the PHASTER software, with approximately 54.9 Kb, possesses 89 genes that are represented by colored arrows. The segment was cut in 3 parts to facilitate the representation. Arrows pointing to the right and left sides represent the forward and reverse strands of DNA. Gene names are represented right above their correspondent arrows. To facilitate the representation, the intergenic regions were not dimensioned. The final region of the segment, composed of *attR* and *tRNA* are overlapping on the genome.

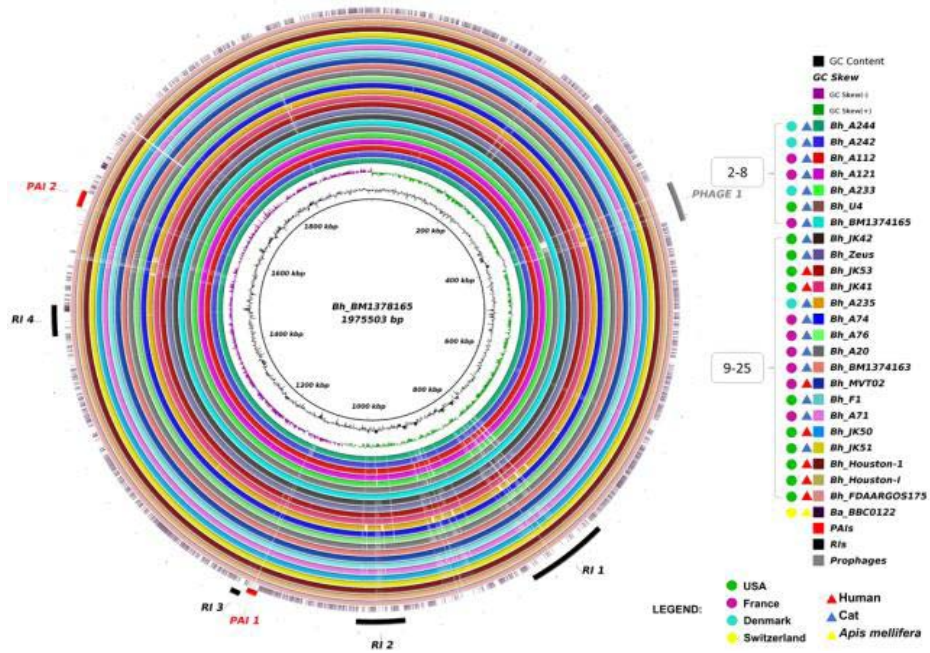


Figure 9: Circular representation of the genomic comparison between all strains of *Bartonella henselae* using the strain BM1374165 as reference. The image represents the genomic plasticity among the genomes of *B. henselae* using the strain BM1374165 as a reference. Each ring in the image corresponds to one genome of one strain of *B. henselae*, with the respective color corresponding to the legend at the right. The deleted regions are represented by blank spaces inside the circles, while the shared genes regions are filled with color. The circle next to the name of the strain indicates the country of isolation of the same, according to the respective color of the legend. The triangle represents the host from which the bacterium was isolated for sequencing, also following the color of the legend. The numbers next to the keys indicate the numbering of the clusters generated by Gegenees and PGAdB analysis. Bh: *Bartonella henselae*. Ba: *Bartonella apis*. PAI: pathogenicity island. RI: resistance island N/A: Not applicable. USA: United States of America.

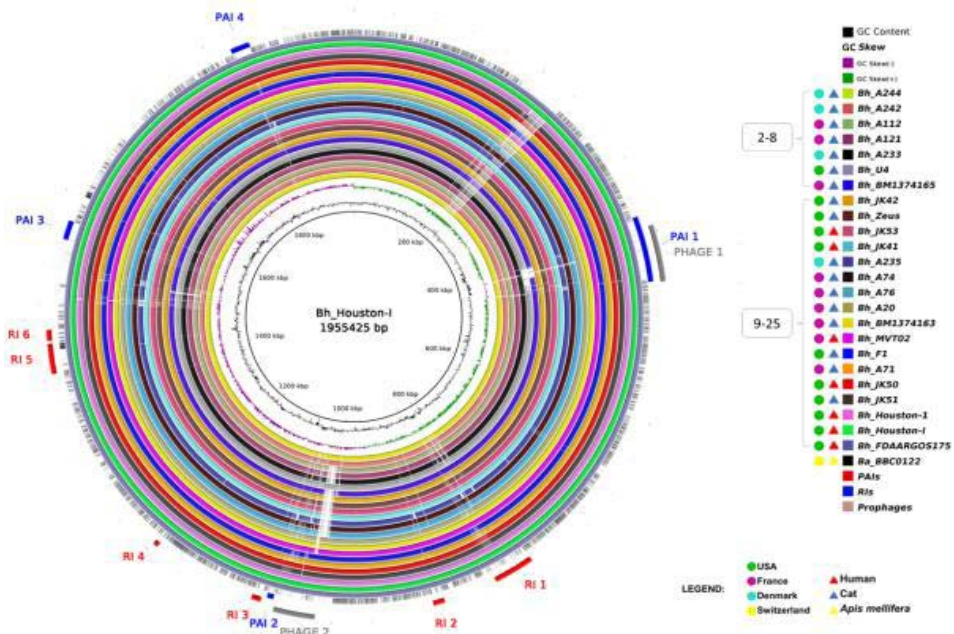


Figure 10: Circular representation of the genomic comparison between all strains of the *Bartonella henselae* using the strain Houston-I as reference. The image represents the genomic plasticity among the genomes of *B. henselae* using the strain Houston-I as a reference. Each ring in the image corresponds to one genome of one strain of *B. henselae*, with the respective color corresponding to the legend at the right. The deleted regions are represented by blank spaces inside the circles, while the shared genes regions are filled with color. The circle next to the name of the strain indicates the country of isolation of the same, according to the respective color of the legend. The triangle represents the host from which the bacterium was isolated for sequencing, also following the color of the legend. The numbers next to the keys indicate the numbering of the cluster generated by Gegenees and PGAdB analysis. Bh: *Bartonella henselae*. Ba: *Bartonella apis*. PAI: pathogenicity island. RI: resistance island N/A: Not applicable. USA: United States of America.

Table 3: List of all genes located in the pathogenicity islands of the reference strain *Bartonella henselae* Houston-1.

ISLAND	GENE	LOCATION	PRODUCT	
PAI 1	<i>rpsA</i>	127083	128783	30S ribosomal protein S1
	<i>cmk</i>	128916	129563	Cytidylate kinase
	<i>aroA</i>	129560	130888	3-phosphoshikimate 1-carboxyvinyltransferase
	BH00960	131075	131455	Hypothetical protein
	<i>ilvD</i>	132510	134348	Dihydroxy-acid dehydratase
	BH00990	134491	134639	Prophage integrase
PAI 2	<i>lldD</i>	357436	358587	L-lactate dehydrogenase
	BH02730	360289	360414	Hypothetical protein
	BH02740	361075	361395	Hypothetical protein
	<i>pheP</i>	361576	363003	Phenylalanine-specific permease
	BH02760	363045	363725	ABC transporter ATP-binding protein
	BH02770	363722	364360	Hypothetical protein
	BH02780	364374	364928	Biotin synthase
	BH02790	366252	366632	exonuclease
	BH02800	367274	367522	[no /product]
	BH02810	368313	368600	[no /product]
	BH02820	368624	368975	Phage related
	BH02830	369436	370097	Phage related
	BH02840	370293	370475	Hypothetical protein
	BH02850	370522	370674	Hypothetical protein
	BH02860	370808	371101	Hypothetical protein
	BH02870	371023	371178	Hypothetical protein
	BH02880	371485	372642	Hypothetical phage integrase
	BH02890	372656	372886	Hypothetical protein
	BH02900	372914	373504	Anti-repressor protein
	<i>dinJ1</i>	373624	373899	DNA-damage-inducible protein
	BH02920	373912	374202	Hypothetical prophage protein
	BH02930	374325	374555	Hypothetical prophage protein
	BH02940	374666	374893	Hypothetical prophage protein
	BH02950	374890	375552	Phage related lysozyme
	BH02960	375712	376014	Hypothetical prophage protein
	BH02970	376001	376321	Hypothetical prophage protein
	BH02980	376365	376643	Hypothetical prophage protein
	BH02990	376640	377371	Anti-repressor protein
	BH03000	377391	377618	Anti-repressor
	BH03010	377623	378195	Anti-repressor protein
	BH03020	378223	378771	Anti-repressor protein
	<i>gpD</i>	379083	380390	Late control gene D protein
	<i>gpX</i>	380387	380611	Phage tail protein
	<i>gpU</i>	380608	380988	Phage tail protein
	BH03060	380994	383294	Phage tail protein
	BH03070	383404	383688	Hypothetical prophage protein
	<i>gpFII</i>	383690	384196	Major tail tube protein FII
	<i>gpFI</i>	384196	385428	Major tail sheath protein FI
	BH03100	385668	385844	Anti-repressor
	BH03110	385898	386419	Phage protein
	BH03120	386465	386794	Virulence-associated protein
	BH03130	386882	387061	Hypothetical prophage protein
	BH03140	387357	388682	Hypothetical prophage protein
	BH03150	388697	391840	Virulence-associated protein
	<i>gp27</i>	391842	392948	Phage protein gp27
	<i>gp26</i>	392948	393775	Phage protein gp26
	<i>gp25</i>	393772	394101	Phage protein gp25

<i>gp24</i>	394110	394799	Phage-related baseplate assembly protein
BH03200	394780	395322	Hypothetical prophage protein
BH03210	395374	395646	Hypothetical prophage protein
BH03220	395636	395953	Hypothetical prophage protein
BH03230	396194	396472	Hypothetical prophage protein
BH03240	396469	397227	Anti-repressor protein
BH03250	397246	397800	Anti-repressor protein
BH03260	397821	398012	Anti-repressor
BH03270	398185	399258	Hypothetical prophage protein
BH03280	399294	399647	Hypothetical prophage protein
<i>gp20</i>	399649	400725	Phage protein gp20
<i>gp19</i>	400738	401106	Phage protein gp19
BH03310	401103	401516	Hypothetical protein
<i>gp18</i>	401366	402475	Phage protein gp18
<i>gp17</i>	402399	403955	Phage protein gp17
BH03340	403955	404203	Hypothetical prophage protein
<i>gp15</i>	404207	406135	Phage terminase large subunit (gp15)
<i>gp13</i>	406128	406709	Phage protein gp13
BH03370	406772	407086	Killer protein
<i>vapA1</i>	407103	407399	Virulence-associated protein a
BH03390	407511	407684	Hypothetical prophage protein
<i>dinJ2</i>	407809	408072	DNA-damage-inducible protein
BH03410	408059	408340	Hypothetical prophage protein
BH03420	408383	408661	Hypothetical prophage protein
BH03430	408658	409416	Anti-repressor protein
BH03440	409476	409991	Anti-repressor protein
BH03450	410008	410847	Anti-repressor protein
BH03460	410879	411421	Anti-repressor protein
BH03470	411438	411704	Hypothetical protein
BH03480	412018	412596	Hypothetical prophage protein
BH03490	412600	412920	Hypothetical prophage protein
BH03500	413179	413355	Hypothetical prophage protein
BH03510	413393	413788	DNA-binding protein
BH03520	413785	413943	Hypothetical prophage protein
BH03530	414079	414618	Hypothetical prophage protein
<i>ssb1</i>	414599	415045	Single-strand binding protein SSB
BH03550	415237	415517	<i>vapD</i>
<i>vapA2</i>	415540	415830	Virulence-associated protein A
BH03570	415927	416199	Hypothetical prophage protein
BH03580	416364	416762	Hypothetical prophage protein
BH03590	416776	417888	Hypothetical prophage protein
BH03600	417878	418225	Hypothetical prophage protein
<i>gp11A</i>	418218	418694	Hypothetical prophage protein
BH03620	418822	419598	Hypothetical prophage protein
BH03630	419607	420059	Hypothetical prophage protein
BH03640	420072	420437	Hypothetical prophage protein
BH03650	420515	421132	Phage repressor protein
BH03660	421374	421748	Hypothetical prophage protein
BH03670	421813	422205	Anti-repressor protein
BH03680	422215	422682	Anti-repressor protein
BH03690	422672	422974	Hypothetical prophage protein
BH03700	422987	423370	Phage related protein
BH03710	423374	423739	Hypothetical prophage protein
BH03720	423736	423903	Hypothetical prophage protein
BH03730	423812	424591	Phage related protein
BH03740	424593	425216	Exonuclease
BH03750	425245	425736	Exonuclease

	BH03760	425809	426012	Phage related protein
	<i>serC</i>	426588	427751	Phosphoserine aminotransferase
PAI 3	BH11550	1277140	1278357	Hypothetical protein
	BH11560	1278427	1280100	Hypothetical protein
	BH11570	1280717	1280968	Hypothetical protein
	<i>ubiC</i>	1282863	1283378	Chorismate-pyruvate lyase
	<i>carA</i>	1284573	1285772	Carbamoyl-phosphate synthase small chain
	BH11600	1285837	1286849	putative transport protein
	BH11610	1287680	1288519	Hypothetical protein
	<i>carB</i>	1289195	1292680	Carbamoyl-phosphate synthase large chain
	BH11630	1293067	1293276	Cold shock protein
	<i>aatA</i>	1293574	1294776	Aspartate aminotransferase a
	BH11650	1295556	1296461	LysR family transcriptional regulator
	<i>trxB</i>	1296518	1297471	
	<i>trkH</i>	1297784	1299178	trk system potassium uptake protein trkH
	BH11680	1299870	1299953	[no /product]
	PAI 4	BH12870	1439270	1439626
BH12880		1440117	1440584	Prophage integrase
<i>asd</i>		1440905	1441921	Aspartate-semialdehyde dehydrogenase
BH12900		1442300	1442926	Hypothetical protein
<i>yaeC</i>		1443453	1444271	Lipoprotein
BH12920		1444299	1445333	ABC transporter, ATP-binding protein
BH12930		1445326	1445985	ABC transporter, permease protein
BH12940		1446396	1447533	Prophage integrase
<i>nodN</i>		1447916	1448395	Nodulation protein n
BH12960		1448563	1449900	Hypothetical protein
BH12970		1454740	1455480	Hypothetical protein
BH12980		1455496	1456524	Hypothetical protein
BH12990		1456535	1457506	Hypothetical protein
BH13000		1458309	1458431	Prophage integrase
BH13010		1459285	1459866	Probable surface protein
BH13020		1459863	1462346	Hypothetical protein
BH13030		1463064	1469090	Probable surface protein
BH13040		1470331	1471284	<i>fatD</i>
BH13050		1471232	1472220	<i>fatC</i>
BH13060		1472178	1472962	<i>ceuE</i>
BH13080		1473184	1473395	[no /product]
BH13090		1474501	1475505	Hypothetical protein
BH13100		1476693	1477628	[no /product]
BH13110		1477431	1479672	[no /product]
BH13120		1480666	1484991	Hypothetical protein
BH13130	1485515	1485661	Prophage integrase	
BH13140	1486276	1488897	Probable surface protein	
BH13150	1489421	1489567	Prophage integrase	
BH13160	1490182	1492524	Hypothetical protein	
BH13170	1493412	1493714	Hypothetical protein	
BH13180	1494963	1497581	Hypothetical protein	
BH13190	1499323	1499766	Hypothetical protein	
<i>parA1</i>	1503848	1504477	<i>parA</i> protein	
PAI 5	BH13360	1521519	1521725	Hypothetical protein
	BH13370	1521736	1523370	Putative cell filamentation protein
	<i>traG</i>	1523533	1525452	Conjugial transfer protein traG
	BH13390	1525994	1527622	Hypothetical protein
	BH13400	1528692	1530290	Putative cell filamentation protein
	BH13410	1530764	1532368	Hypothetical protein
BH13420	1532694	1534088	Hypothetical protein	
BH13430	1534424	1536928	Hypothetical protein	

	BH13440	1537113	1540244	Hypothetical protein
	BH13450	1541044	1541217	Prophage integrase
PAI 6	BH15520	1799985	1801265	Oxidoreductase
	<i>korB</i>	1802021	1802290	<i>korB</i> protein
	BH15540	1802337	1803065	Hypothetical protein
	<i>trwN</i>	1803062	1803730	<i>trwN</i> protein
	<i>korA</i>	1803874	1804167	<i>korA</i> protein
	<i>trwL1</i>	1804130	1804447	<i>trwL1</i> protein
	<i>trwL2</i>	1804494	1804805	<i>trwL2</i> protein
	<i>trwL3</i>	1804829	1805152	<i>trwL3</i> protein
	<i>trwL4</i>	1805190	1805507	<i>trwL4</i> protein
	<i>trwL5</i>	1805661	1805969	<i>trwL5</i> protein
	<i>trwL6</i>	1806088	1806429	<i>trwL6</i> protein
	<i>trwL7</i>	1806567	1806881	<i>trwL7</i> protein
	<i>trwL8</i>	1807042	1807353	<i>trwL8</i> protein
	<i>trwM</i>	1807366	1807674	<i>trwM</i> protein
	<i>trwK</i>	1807677	1810148	<i>trwK</i> protein
	<i>trwJ1</i>	1810145	1810933	<i>trwJ1</i> protein
	<i>trwI1</i>	1811167	1812051	<i>trwI1</i> protein
	<i>trwH1</i>	1812216	1812362	Hypothetical protein
	<i>trwJ2</i>	1812359	1812985	<i>trwJ2</i> protein
	<i>trwI2</i>	1813241	1814125	<i>trwI2</i> protein
<i>trwH2</i>	1814293	1814439	Hypothetical protein	
<i>trwG</i>	1814436	1815134	<i>trwG</i> protein	
<i>trwF</i>	1815148	1815954	<i>trwF</i> protein	
<i>trwE</i>	1815954	1817141	<i>trwE</i> protein	
<i>trwD</i>	1817116	1818177	<i>trwD</i> protein	

strains of *B. henselae* (Figure 7) shows various regions shared by all the strains, as well as few deleted regions on them. We identified 6 pathogenicity islands (PAIs) through predictions made by the GIPSY software. Their lengths vary from ~9.7 kb (PAI 1) to ~68.2 kb (PAI 4). The location and size of the PAIs are shown in Table 2 and the list of all the genes located at PAIs is presented at Table 3. We also identified one phage region considered intact by the PHASTER software, whose location corresponds, approximately, to the same PAI 2 region (Figure 8). Among the 191 genes predicted into all the islands, 73 (38%) are “hypothetical proteins”, and, among these, 36 are “hypothetical prophage protein”. The strain *B. henselae* U4 presents a deletion at the region predicted as PAI 2 on the reference genome (*B. henselae* Houston-1), that corresponds, also, to the region predicted as Phage 1.

According to the analysis of the genome BM1374165 (Figure 9) and Houston-I (Figure 10), isolated from cat and human, respectively, we predicted 4 RIs, 2 PAIs and one phage region on the former and 6 RIs, 4 PAIs and two phage regions on the latter. Interestingly, the region that is present on the strain BM1374165 and absent on Houston-1, observed on Mauve (Figure 3-4), mostly corresponded to RI 1 predicted by GIPSY on the former, as observed in Figure 9. This is a duplication of the region 840001-903302 of the strain Houston-1, which is equivalent to the regions 810628-873868 and 911748-976210 of the strain BM1374165.

Briefly, the phage region harbours the genes: phage integrase (*int*), which is required for the integration of the element, head genes, terminase subunits, prohead protease, scaffold protein, major capsid subunit, and the tail genes, including the major and minor

tail subunit proteins, overlapping open reading frames, and the tape measure protein [41].

Discussion

Our analyses suggest that the *B. henselae* species present a high similarity among its strains, indicating a certain stability of the genome and low degree of diversity. This fact corroborates previous studies that demonstrated the low rate of recombination of the intraspecies gene content [17,18].

Bouchouicha et al. [42] demonstrate the use of MLVA for classification of *B. henselae* genomes into two genotypes. In the case of our study, where the strains are completely sequenced and deposited in public databases, it is more viable to use wgMLST approaches, which use the whole genome, instead of 16s. This methodology is more precise than MLVA and is able to identify clones and microclones from the samples as shown by Maiden [22]. Although the strains of *B. henselae* have a high similarity at its genetic content, they can still have some small variations at the nucleotide level that cannot be traced by the Gegenees software. In order to identify the differences at such level, we performed a phylogenomic analysis based on the genome polymorphism of the species. As shown by the tree created by Splits Tree, the analysis of wgMLST via PGAdB-builder also presents a high proximity among the species, demonstrating the gene conservation throughout time. The strains A242 and A244 became further apart compared to the others. Both strains were isolated from cats in Denmark on the same period (March 1998) [43].

The proximity of the A74 and A76 strains could be associated to the fact that both were isolated and sequenced at the same period from feline hosts on France (March 1996). The host's data, as well as in the case described above, could not justify the phylogenomic analysis. But, an interesting fact is evidenced by analyzing the U4 and F1 strains. Both were isolated in 1996 from feline hosts in the USA. However, both in the SplitsTree and the PGAdB-builder analyses, the strains presented themselves as phylogenomically apart and present in clusters with strains from 2 different countries (France and the USA). From these results, it might be possible to hypothesize that the phylogenomic proximity among different strains is independent of the fact that they were isolated from the same country or year.

Although highly similar concerning their gene content, different strains of the same species may exhibit genetic rearrangements that allow them to develop different phenotypes. To verify the possible existence of rearrangement at the *B. henselae* genomes, a gene synteny analysis was performed by Mauve. This software identifies LCBs during the alignment process, defined as regions without local rearrangements of probable homologous sequences shared by two or more genomes. From the analysis considering only the completed genomes, we could observe that the species maintain its structure conserved among its strains since there are no inversions and translocations. Although the completed genomes present some deletion regions, none of them are inside pathogenicity islands.

Altogether, the data on the extrapolations show a very homogeneous variation on the pan-genome, core genome and singletons of the complete dataset, with an α value of 0.94, representing a slow growing open pan-genome and $t_g(\theta)$ of 1136 and 1 for the core genome and singletons, respectively, which shows a tendency

for a closed pan-genome. However, the singletons from the genome isolated from the USA and the core genome and singletons from the strains isolated from humans show fast changing values, showing a possible fast changing pan-genome, which is in agreement with the α value of 0.92 from the pan-genome of human. Also, the α value of the strains 2-8 was the smallest one, 0.90, whereas the core genome and singletons present only slight variations on the curve, which is in agreement with the grouping of those strains in the phylogenomic tree.

As expected, the analysis of the subsets of the pan-genome by the COG database demonstrated that the main components of the core genome are genes responsible for essential processes of the bacteria, as metabolism and cellular signaling. In contrast, genes that compose the shared and singletons are poorly characterized genes. This is in agreement with the high importance of the core genes subset, which is mostly composed of housekeeping genes of the species.

Genome plasticity analyses allow the search for regions on the genome called pathogenicity island (PAIs), which are very important for being acquired by horizontal gene transfer and for containing virulence genes, as initially observed by Hacker, in virulence studies with *Escherichia coli* [44]. For PAIs prediction, we included the *B. apis* bacteria genome at GIPSY as a non-pathogenic reference. The choice was based on the fact that this species, until this moment, was not associated to a pathogenic response on a human host or on mammals in general. *B. apis* can be found in the posterior intestine of fully grown bees, where they live in a symbiotic relationship with its host [32]. Besides, the species is intimately related with the *B. henselae* species.

Six PAIs were identified from the analyses of all the *B. henselae* genome through GIPSY (Table 2). The gene list with the respective location can be found in Table 3. We analyzed gene by gene and we focused on the genes that code proteins involved in known virulence mechanisms and described in the literature.

PAI 1: The gene *aroA*, coding for the product 3-phosphoshikimate 1-carboxyvinyltransferase, is an integral part of the Shikimate pathway, responsible for the formation of aromatic amino acids in bacteria [45]. The mutation in this gene is used in *Salmonella* species for the generation of recombinant mutants used as adjuncts in attenuated vaccines [46]. Besides that, other studies using the gene *aroA* as a potential vaccine target were published in literature and used on a variety of microorganisms, such as *Pasteurella multocida* [47], *Bordetella pertussis* [48] and *Yersinia pestis* [49]. However, in the last study, although the mutation in the gene was able to elicit an immune response in guinea pigs, it presented virulence in mice. A recent study revealed that the elimination of the *aroA* gene of *Salmonella typhimurium* led to an increased virulence and consequent elevation of immunogenicity of the strain, such a fact being advantageous to the optimization of the vaccine agents and bacterial immunotherapy [50].

The gene *ilvD* coding for the product dihydroxy-acid dehydratase was described in *Mycobacterium tuberculosis* [51] and *Escherichia coli* [52] as being an enzyme essential to bacteria. Its main function is to take part in the biosynthesis of branched-chain amino acid. Previous studies with anti-tuberculosis agents revealed that the inactivation of the *ilvD* gene led to auxotrophy of three amino acids (isoleucine,

leucine, and valine) and reduced the surviving capabilities of the bacteria [53].

PAI 2: The second pathogenicity island corresponds to a phage region predicted by Phaster as phage integration regions. Several phage proteins can be found in this region, justifying the pathogenicity island predicted by GIPSY.

PAI 3: The operon *carAB* was the main component found in this island. In prokaryotes, the genes *carA* and *carB* are responsible for the coding of the minor and major subunits, respectively, of the product Carbamoyl-phosphate synthase large chain [54]. This enzyme is essential for the metabolism of arginine and pyrimidines. Previous studies revealed that mutations on the *carA* and *carB* genes produce a silenced mutant, which may be used as an oral vaccine against avian colibacillosis [55].

PAI 4: The gene *asd*, that code Aspartate-semialdehyde dehydrogenase, takes part in the formation of the cell wall and was described in studies of the immunologic response of *Salmonella enterica serovar Typhimurium* [56].

PAI 5: Six of the ten genes present in this island are genes that code hypothetical proteins, with no description of those in the literature. It was also found as a prophage integration region.

PAI 6: The genes *trwL1*, *trwL2*, *trwJ2*, *trwF*, *trwE*, and *trwD* were found in this island. The genes of the *trw* family are intimately related with the species of the *Bartonella* genus virulence, mainly *Bartonella tribocorum* [57], being an operon that codes the components of type IV secretion system (T4SS). The *trw* locus was acquired *via* horizontal gene transfer on *Bartonella* species and is currently present in approximately 13 species [58]. It is a conjugation system involved in the bacteria pathogenicity and related with the transport of proteins (mostly virulence factors) and DNA. The bacteria component as a whole eases the dispersion of bacterial genetic material, including the pathogenicity islands to other microorganisms, being able to transform non-virulent species into pathogenic ones, which makes it an important virulence factor present in *B. henselae* [59]. Besides, it was already described that the *trw* system mediates the specific erythrocyte infection of the host, attaching itself to the erythrocyte surface, which is the main developmental factor in the pathogeny of *B. henselae* [57].

It is interesting to notice that the main virulence genes of the T4SS, of the *virB* family, are not harbored by any of the predicted PAIs in our study. However, such genes are present among the PAI 4 and PAI 5, as represented by an asterisk in Figure 7, and such a fact is relevant due to the possibility that these genes are integrated into the islands and become transmissible by horizontal transference to non-pathogenic species.

The additional analysis of the genomes isolated from cat (BM1374165) and human (Houston-I), selected from clusters 2-8 and 9-25, respectively, showed that, in cat isolates, the amount of horizontally acquired regions (PAIs, Ris, Prophage) is lower in relation to the human isolate.

Conclusion

All the results presented here highlighted the high similarity among the genomes of *B. henselae*. We have found two clusters of

genomes, one composed of strains isolated from cats (2-8) with more diverse pan-genome and another composed of strains isolated from cats or humans (9-25), with a more homogeneous pan-genome, where the extrapolations of the pan-genome, core genome, and singletons show that group 2-8 are varying faster than the strains from group 9-25. Also, although many PAIs and RIs were predicted, none of them present deletions in agreement with the subsets studied here and may not be correlated with the epidemiological hypothesis raised here. Altogether, there is still a need for sequencing more strains from other countries to better understand the pangenomic features of this organism. Finally, the highly conserved genomes from this species are very important for the development of new vaccines and analyses of drug targets against this pathogen that is of great importance for causing disease not only in domestic cats but also in humans, and these data may then be used in future works, which will be highly relevant for containing the disease worldwide.

References

- Boulouis H-J, Chao-chin C, Henn JB, Kasten RW, Chomel BB. Factors associated with the rapid emergence of zoonotic Bartonella infections. *Vet Res.* 2005; 36: 383-410.
- Eicher SC, Dehio C. Bartonella entry mechanisms into mammalian host cells. *Cell Microbiol.* 2012; 14: 1166-73.
- Okaro U, Addisu A, Casanas B, Anderson B. Bartonella Species, an Emerging Cause of Blood-Culture-Negative Endocarditis. *Clin Microbiol Rev.* 2017; 30: 709-746.
- Minnick MF, Anderson BE, Lima A, Battisti JM, Lawyer PG, Birtles RJ. Oroya Fever and Verruga Peruana: Bartonellosis Unique to South America. *PLoS Negl Trop Dis.* 2014; 8: e2919.
- Foucault C, Brouqui P, Raoult D. Bartonella quintana characteristics and clinical management. *Emerg Infect Dis.* 2006; 12: 217-223.
- Breitschwerdt EB. Feline bartonellosis and cat scratch disease. *Vet Immunol Immunopathol.* 2008; 123: 167-171.
- Gutiérrez R, Nachum-Biala Y, Harrus S. Relationship between the Presence of Bartonella Species and Bacterial Loads in Cats and Cat Fleas (Ctenocephalides felis) under Natural Conditions. *Appl Environ Microbiol.* 2015; 81: 5613-5621.
- Weeden AL, Cherry NA, Breitschwerdt EB, Cheves AG, Wamsley HL. Bartonella henselae in canine cavity effusions: prevalence, identification, and clinical associations. *Vet Clin Pathol.* 2017; 46: 326-30.
- Regier Y, O'Rourke F, Kempf VAJ. Bartonella spp. - a chance to establish One Health concepts in veterinary and human medicine. *Parasit Vectors* 2016; 9: 261.
- Maggi RG, Ericson M, Mascarelli PE, Bradley JM, Breitschwerdt EB. Bartonella henselae bacteremia in a mother and son potentially associated with tick exposure. *Parasit Vectors.* 2013; 6: 101.
- Mascarelli PE, Maggi RG, Hopkins S, Mozayani BR, Trull CL, Bradley JM, et al. Bartonella henselae infection in a family experiencing neurological and neurocognitive abnormalities after woodlouse hunter spider bites. *Parasit Vectors.* 2013; 6: 98.
- Dehio C. Bartonella interactions with endothelial cells and erythrocytes. *Trends Microbiol.* 2001; 9: 279-285.
- Florin TA, Zaoutis TE, Zaoutis LB. Beyond Cat Scratch Disease: Widening Spectrum of Bartonella henselae Infection. *Pediatrics.* 2008; 121: e1413-e1425.
- Chang C, Chomel BB, Kasten RW, Tappero JW, Sanchez MA, Koehler JE. Molecular Epidemiology of Bartonella henselae Infection in Human Immunodeficiency Virus–Infected Patients and Their Cat Contacts, Using Pulsed Field Gel Electrophoresis and Genotyping. *J Infect Dis.* 2002; 186: 1733-1739.

15. Fiskus W, Padmalayam I, Kelly T, Guibao C, Baumstark BR. Identification and Characterization of the DdlB, FtsQ and FtsA Genes Upstream of FtsZ in *Bartonella bacilliformis* and *Bartonella henselae*. *DNA Cell Biol.* 2003; 22: 743-752.
16. Riess T, Raddatz G, Linke D, Schäfer A, Kempf VAJ. Analysis of *Bartonella* adhesin A expression reveals differences between various *B. henselae* strains. *Infect Immun.* 2007; 75: 35-43.
17. Guy L, Nystedt B, Sun Y, Näslund K, Berglund EC, Andersson SGE. A genome-wide study of recombination rate variation in *Bartonella henselae*. *BMC Evol Biol.* 2012; 12: 65.
18. Lindroos H, Vinnere O, Mira A, Repsilber D, Naslund K, Andersson SGE. Genome Rearrangements, Deletions, and Amplifications in the Natural Population of *Bartonella henselae*. *J Bacteriol.* 2006; 188: 7426-7439.
19. Berglund EC, Frank AC, Calteau A, Vinnere Pettersson O, Granberg F, Eriksson A-S, et al. Run-Off Replication of Host-Adaptability Genes Is Associated with Gene Transfer Agents in the Genome of Mouse-Infecting *Bartonella grahamii*. *PLoS Genet.* 2009; 5: e1000546.
20. Ågren J, Sundström A, Håfström T, Segerman B. Gegenees: Fragmented Alignment of Multiple Genomes for Determining Phylogenomic Distances and Genetic Signatures Unique for Specified Target Groups. *PLoS One.* 2012; 7: e39107.
21. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 2006; 23: 254-267.
22. Maiden MCJ, van Rensburg MJJ, Bray JE, Earle SG, Ford SA, Jolley KA, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol.* 2013; 11: 728-736.
23. Liu Y-Y, Chiou CS, Chen CC. PGADB-builder: A web service tool for creating pan-genome allele database for molecular fine typing. *Sci Rep.* 2016; 6: 36213.
24. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 2016; 33: 1870-1874.
25. Darling AE, Mau B, Perna NT. Progressive Mauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS One.* 2010; 5: e11147.
26. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 2015; 16: 157.
27. Enright AJ, Van Dongen S, Ouzounis CA. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* 2002; 30: 1575-1584.
28. Soares SC, Silva A, Trost E, Blom J, Ramos R, Carneiro A, et al. The Pan-Genome of the Animal Pathogen *Corynebacterium pseudotuberculosis* Reveals Differences in Genome Plasticity between the Biovar ovis and equi Strains. *PLoS One.* 2013; 8: e53818.
29. Galperin MY, Kristensen DM, Makarova KS, Wolf YI, Koonin EV. Microbial genome analysis: The COG approach. *Brief Bioinform.* 2018.
30. Soares SC, Geyik H, Ramos RTJ, de Sá PHCG, Barbosa EGV, Baumbach J, et al. GIPSY: Genomic island prediction software. *J Biotechnol.* 2016; 232: 2-11.
31. da Silva Filho AC, Raittz RT, Guizelini D, De Pierri CR, Augusto DW, dos Santos-Weiss ICR, et al. Comparative Analysis of Genomic Island Prediction Tools. *Front Genet.* 2018; 9.
32. Kešnerová L, Moritz R, Engel P. *Bartonella apis* sp. nov., a honey bee gut symbiont of the class Alphaproteobacteria. *Int J Syst Evol Microbiol.* 2016; 66: 414-421.
33. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 2016; 44: W16-W21.
34. Srividhya K V., Rao G V., L R, Mehta P, Prilusky J, Manicka S, et al. Database and Comparative Identification of Prophages. *Intell. Control Autom.*, vol. 344, Springer Berlin Heidelberg. 2006; 863-868.
35. Srividhya K V., Alaguraj V, Poornima G, Kumar D, Singh GP, Raghavenderan L, et al. Identification of Prophages in Bacterial Genomes by Dinucleotide Relative Abundance Difference. *PLoS One* 2007; 2: e1193.
36. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 2011; 12: 402.
37. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics.* 2012; 28: 464-469.
38. Mao C, Abraham D, Wattam AR, Wilson MJC, Shukla M, Yoo HS, et al. Curation, integration and visualization of bacterial virulence factors in PATRIC. *Bioinformatics.* 2015; 31: 252-258.
39. Chen L, Xiong Z, Sun L, Yang J, Jin Q. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res.* 2012; 40: D641-D645.
40. Zhou CE, Smith J, Lam M, Zemla A, Dyer MD, Slezak T. MvirDB—a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. *Nucleic Acids Res.* 2007; 35: D391-D394.
41. Hatfull GF. Bacteriophage genomics. *Curr Opin Microbiol.* 2008; 11: 447-453.
42. Bouchoucha R, Durand B, Monteil M, Chomel BB, Berrich M, Arvand M, et al. Molecular Epidemiology of Feline and Human *Bartonella henselae* Isolates. *Emerg Infect Dis.* 2009; 15: 813-816.
43. Woudstra C, Fach P, Chomel BB, Haddad N, Boulouis HJ. Draft Genome Sequences of 12 Feline *Bartonella henselae* Isolates. *Genome Announc.* 2017; 5: 12-13.
44. Hacker J, Bender L, Ott M, Wingender J, Lund B, Marre R, et al. Deletions of chromosomal regions coding for fimbriae and hemolysins occur *in vitro* and *in vivo* in various extraintestinal *Escherichia coli* isolates. *Microb Pathog.* 1990; 8: 213-225.
45. Bentley R, Haslam E. The Shikimate Pathway-A Metabolic Tree with Many Branche. *Crit Rev Biochem Mol Biol.* 1990; 25: 307-384.
46. Karasova D, Sebkova A, Vrbas V, Havlickova H, Sisak F, Rychlik I. Comparative analysis of *Salmonella enterica serovar Enteritidis* mutants with a vaccine potential. *Vaccine.* 2009; 27: 5265-5270.
47. Homchampa P, Strugnell RA, Adler B. Molecular analysis of the *aroA* gene of *Pasteurella multocida* and vaccine potential of a constructed *aroA* mutant. *Mol Microbiol.* 1992; 6: 3585-3593.
48. Roberts M, Maskell D, Novotny P, Dougan G. Construction and characterization *in vivo* of *Bordetella pertussis aroA* mutants. *Infect Immun.* 1990; 58: 732-739.
49. Oyston PCF, Russell P, Williamson ED, Titball RW. An *aroA* mutant of *Yersinia pestis* is attenuated in guinea-pigs, but virulent in mice. *Microbiology.* 1996; 142: 1847-1853.
50. Felgner S, Frahm M, Kocijancic D, Rohde M, Eckweiler D, Bielecka A, et al. *aroA* -Deficient *Salmonella enterica Serovar Typhimurium* Is More Than a Metabolically Attenuated Mutant. *MBio.* 2016; 7: e01220-e01216.
51. Singh V, Chandra D, Srivastava BS, Srivastava R. Downregulation of Rv0189c, encoding a dihydroxyacid dehydratase, affects growth of *Mycobacterium tuberculosis* *in vitro* and in mice. *Microbiology.* 2011; 157: 38-46.
52. Flint DH, Emptage MH, Finnegan MG, Fu W, Johnson MK. The role and properties of the iron-sulfur cluster in *Escherichia coli* dihydroxy-acid dehydratase. *J Biol Chem.* 1993; 268: 14732-14742.
53. Grandoni JA, Marta PT, Schloss J V. Inhibitors of branched-chain amino acid biosynthesis as potential antituberculosis agents. *J Antimicrob Chemother.* 1998; 42: 475-482.
54. GUO J, SONG X, ZOU L, ZOU H, CHEN G. The small and large subunits of carbamoyl-phosphate synthase exhibit diverse contributions to pathogenicity in *Xanthomonas citri* subsp. *citri*. *J Integr Agric.* 2015; 14: 1338-1347.
55. Kwaga JKP, Allan BJ, Van den Hurk J V., Seida H, Potter AA. A *carAB* mutant

- of avian pathogenic *Escherichia coli* serogroup O2 is attenuated and effective as a live oral vaccine against colibacillosis in Turkeys. *Infect Immun*. 1994; 62: 3766-3772.
56. Piao HH, Tam VTM, Na HS, Kim HJ, Ryu PY, Kim SY, et al. Immunological responses induced by *asd* and *wzy/asd* mutant strains of *Salmonella enterica* serovar Typhimurium in BALB/c mice. *J Microbiol*. 2010; 48: 486-495.
57. Vayssier-Taussat M, Rhun D Le, Deng HK, Biville F, Cescau S, Danchin A, et al. The Trw type IV secretion system of *Bartonella* mediates host-specific adhesion to erythrocytes. *PLoS Pathog*. 2010; 6: e1000946.
58. Saenz HL, Engel P, Stoeckli MC, Lanz C, Raddatz G, Vayssier-Taussat M, et al. Genomic analysis of *Bartonella* identifies type IV secretion systems as host adaptability factors. *Nat Genet*. 2007; 39: 1469-1476.
59. Saisongkorh W, Robert C, La Scola B, Raoult D, Rolain J-M. Evidence of Transfer by Conjugation of Type IV Secretion System Genes between *Bartonella* Species and *Rhizobium radiobacter* in *Amoeba*. *PLoS One*. 2010; 5: e12666.

5. CONCLUSÕES

De forma geral, o trabalho aponta para a comprovação da alta similaridade intraespécie de *B. henselae*. Seu pan-genoma, apesar de aberto, tende a se tornar fechado. Porém, para a validação dos padrões de distribuição genômica, aponta-se a necessidade de sequenciamento de mais genomas em todo o mundo, de diferentes locais e hospedeiros, como humanos, gatos e cães, para melhor caracterizar variações no pan-genoma e compreender os padrões adaptativos dessa espécie.

A alta conservação dos genomas dessa espécie é muito importante para o desenvolvimento de novas vacinas e análises de alvos de drogas contra esse patógeno. Esses dados poderão ser utilizados em trabalhos futuros, de grande relevância para o tratamento da doença em todo o mundo.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ABBY, Sophie; DAUBIN, Vincent. Comparative genomics and the evolution of prokaryotes. *Trends in Microbiology*, v. 15, n. 3, p. 135–141, mar. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17289390>>. Acesso em: 31 maio 2018.
- ALCARAZ, Luis *et al.* Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. *BMC Genomics*, v. 11, n. 1, p. 332, 26 maio 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20504335>>. Acesso em: 31 maio 2018.
- ÁLVAREZ-FERNÁNDEZ, Alejandra; BREITSCHWERDT, Edward B.; SOLANO-GALLEGO, Laia. Bartonella infections in cats and dogs including zoonotic aspects. *Parasites & Vectors*, v. 11, n. 1, p. 624, 4 dez. 2018. Disponível em: <<https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-018-3152-6>>.
- ANDERSON, Burt E.; NEUMAN, Mark A. Bartonella spp. as emerging human pathogens. *Clinical Microbiology Reviews*, v. 10, n. 2, p. 203–219, abr. 1997. Disponível em: <<http://cmr.asm.org/lookup/doi/10.1128/CMR.10.2.203>>. Acesso em: 19 maio 2020.
- ANGELAKIS, Emmanouil *et al.* Potential for Tick-borne Bartonellosis. *Emerging Infectious Diseases*, v. 16, n. 3, p. 385–391, mar. 2010. Disponível em: <http://wwwnc.cdc.gov/eid/article/16/3/09-1685_article.htm>. Acesso em: 19 maio 2020.
- ANSTEAD, Gregory M. The centenary of the discovery of trench fever, an emerging infectious disease of World War 1. *The Lancet Infectious Diseases*, v. 16, n. 8, p. e164–e172, 1 ago. 2016. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/27375211>>. Acesso em: 8 mar. 2020.
- AZEVEDO, Vasco *et al.* Whole Genome Annotation: In Silico Analysis. *Bioinformatics - Trends and Methodologies*. [S.l.]: InTech, 2011. Disponível em: <<http://www.intechopen.com/books/bioinformatics-trends-and-methodologies/whole-genome-annotation-in-silico-analysis>>. Acesso em: 1 jun. 2018.
- BARCELLOS, F. G. *et al.* Evidence of Horizontal Transfer of Symbiotic Genes from a Bradyrhizobium japonicum Inoculant Strain to Indigenous Diazotrophs Sinorhizobium (Ensifer) fredii and Bradyrhizobium elkanii in a Brazilian Savannah Soil. *Applied and Environmental Microbiology*, v. 73, n. 8, p. 2635–2643, 15 abr. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17308185>>. Acesso em: 1 jun. 2018.
- BASS, James W.; VINCENT, Judy M.; PERSON, Donald A. The expanding spectrum of Bartonella infections: II. Cat-scratch disease. *Pediatric Infectious Disease Journal*, v. 16, n. 2, p. 163–179, fev. 1997.
- BERGLUND, Eva C. *et al.* Run-Off Replication of Host-Adaptability Genes Is Associated with Gene Transfer Agents in the Genome of Mouse-Infecting Bartonella grahamii. *PLoS Genetics*, v. 5, n. 7, p. e1000546, 3 jul. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19578403>>. Acesso em: 16 mar. 2018.
- BERRICH, Moez *et al.* Differential effects of Bartonella henselae on human and feline macro- and micro-vascular endothelial cells. *PloS one*, v. 6, n. 5, p. e20204, 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21637717>>. Acesso em: 20 mar. 2018.

BIANCARDI, Ana Luiza; CURI, Andre Luiz Land. Cat-Scratch Disease. *Ocular Immunology and Inflammation*, v. 22, n. 2, p. 148–154, 9 abr. 2014. Disponível em: <<http://www.tandfonline.com/doi/full/10.3109/09273948.2013.833631>>. Acesso em: 15 maio 2020.

BILLETER, S. A. *et al.* Vector transmission of Bartonella species with emphasis on the potential for tick transmission. *Medical and Veterinary Entomology*, v. 22, n. 1, p. 1–15, mar. 2008. Disponível em: <<http://doi.wiley.com/10.1111/j.1365-2915.2008.00713.x>>. Acesso em: 19 maio 2020.

BLANCO, José Ramón *et al.* Diagnóstico microbiológico de las infecciones por patógenos bacterianos emergentes: Anaplasma, Bartonella, Rickettsia, Tropheryma whipplei. *Enfermedades Infecciosas y Microbiología Clínica*, v. 26, n. 9, p. 573–580, 2008.

BLANCO, José Ramón; RAOULT, Didier. Enfermedades producidas por Bartonella spp. *Enfermedades Infecciosas y Microbiología Clínica*, v. 23, n. 5, p. 313–320, 1 maio 2005.

BOTO, Luis. Horizontal gene transfer in evolution: facts and challenges. *Proceedings. Biological sciences*, v. 277, n. 1683, p. 819–27, 22 mar. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19864285>>. Acesso em: 1 jun. 2018.

BOUHSIRA, Emilie *et al.* Ctenocephalides felis an in vitro potential vector for five Bartonella species. *Comparative Immunology, Microbiology and Infectious Diseases*, v. 36, n. 2, p. 105–111, mar. 2013. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S014795711200118X>>. Acesso em: 16 maio 2020.

BOULOUIS, Henri Jean *et al.* Factors associated with the rapid emergence of zoonotic Bartonella infections. *Veterinary Research*, v. 36, n. 3, p. 383–410, maio 2005.

BREITSCHWERDT, E. B. Bartonellosis: One Health Perspectives for an Emerging Infectious Disease. *ILAR Journal*, v. 55, n. 1, p. 46–58, 1 jan. 2014. Disponível em: <<https://academic.oup.com/ilarjournal/article-lookup/doi/10.1093/ilar/ilu015>>. Acesso em: 31 maio 2020.

BREITSCHWERDT, Edward B. Bartonellosis, One Health and all creatures great and small. *Veterinary Dermatology*, v. 28, n. 1, p. 96–e21, 1 fev. 2017. Disponível em: <<http://doi.wiley.com/10.1111/vde.12413>>. Acesso em: 7 maio 2020.

BREITSCHWERDT, Edward B. *Feline bartonellosis and cat scratch disease. Veterinary Immunology and Immunopathology*. [S.l.: s.n.]. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18295347>>. Acesso em: 19 mar. 2018. , 15 maio 2008

BREITSCHWERDT, Edward B *et al.* Molecular Evidence of Perinatal Transmission of Bartonella vinsonii subsp. berkhoffii and Bartonella henselae to a Child. *Journal of Clinical Microbiology*, v. 48, n. 6, p. 2289–2293, 1 jun. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20392912>>. Acesso em: 6 nov. 2017.

BRUNT, Jane *et al.* American Association of Feline Practitioners 2006 Panel report on diagnosis, treatment, and prevention of Bartonella spp. infections. *Journal of Feline Medicine and Surgery*, 2006.

CANNETI, B. *et al.* Neurological presentations of Bartonella henselae infection. *Neurological Sciences*, v. 40, n. 2, p. 261–268, 1 fev. 2019.

CELIKER, Hande *et al.* Bartonella henselae neuroretinitis in patients without cat scratch. *Japanese Journal of Infectious Diseases*, v. 71, n. 6, p. 397–401, 2018.

CHAN, Cheong Xin; RAGAN, Mark A. Next-generation phylogenomics. *Biology direct*, v. 8, p. 3, 22 jan. 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23339707>>. Acesso em: 31 maio 2018.

CHANG, Chao-Chin *et al.* Molecular Epidemiology of Bartonella henselae Infection in Human Immunodeficiency Virus–Infected Patients and Their Cat Contacts, Using Pulsed-Field Gel Electrophoresis and Genotyping. *The Journal of Infectious Diseases*, v. 186, n. 12, p. 1733–1739, 15 dez. 2002. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12447758>>. Acesso em: 19 mar. 2018.

CHERRY, Natalie A. *et al.* PCR detection of Bartonella bovis and Bartonella henselae in the blood of beef cattle. *Veterinary Microbiology*, v. 135, n. 3–4, p. 308–312, 30 mar. 2009a. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0378113508004409>>. Acesso em: 19 maio 2020.

CHERRY, Natalie A. *et al.* PCR detection of Bartonella bovis and Bartonella henselae in the blood of beef cattle. *Veterinary Microbiology*, v. 135, n. 3–4, p. 308–312, 30 mar. 2009b.

CHOMEL, Bruno B. *et al.* Bartonella Spp. in Pets and Effect on Human Health. *Emerging Infectious Diseases*, v. 12, n. 3, p. 389–394, mar. 2006. Disponível em: <http://wwwnc.cdc.gov/eid/article/12/3/05-0931_article.htm>. Acesso em: 16 maio 2020.

CHOMEL, Bruno B. *et al.* Ecological fitness and strategies of adaptation of Bartonella species to their hosts and vectors. *Veterinary Research*. [S.l.: s.n.], 2009

CHOMEL, Bruno B.; BOULOUIS, Henri Jean; BREITSCHWERDT, Edward B. Cat scratch disease and other zoonotic Bartonella infections. *Journal of the American Veterinary Medical Association*, v. 224, n. 8, p. 1270–1279, 15 abr. 2004. Disponível em: <<http://avmajournals.avma.org/doi/abs/10.2460/javma.2004.224.1270>>. Acesso em: 15 maio 2020.

CHONDROGIANNIS, Konstantinos *et al.* Seronegative cat-scratch disease diagnosed by PCR detection of Bartonella henselae DNA in lymph node samples. *The Brazilian Journal of Infectious Diseases*, v. 16, n. 1, p. 96–99, 1 jan. 2012. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1413867012702842>>. Acesso em: 31 maio 2020.

COSTA, Paulo Sérgio Gonçalves Da; BRIGATTE, Marcos Emilio; GRECO, Dirceu Bartolomeu. Antibodies to Rickettsia rickettsii, Rickettsia typhi, Coxiella burnetii, Bartonella henselae, Bartonella quintana, and Ehrlichia chaffeensis among healthy population in Minas Gerais, Brazil. *Memórias do Instituto Oswaldo Cruz*, v. 100, n. 8, p. 853–859, dez. 2005. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02762005000800006&lng=en&tlng=en>. Acesso em: 15 maio 2020.

DE MAAYER, Pieter *et al.* Analysis of the Pantoea ananatis pan-genome reveals factors underlying its ability to colonize and interact with plant, insect and vertebrate hosts. *BMC Genomics*, v. 15, n. 1, p. 404, 27 maio 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24884520>>. Acesso em: 31 maio 2018.

DEBRÉ, Robert *et al.* La maladie des griffes de chat. *Bull Mem Soc Med Hop Paris*, v. 66, p.

76–79, 1950.

DEHIO, Christoph. Molecular and Cellular Basis of Bartonella Pathogenesis. *Annual Review of Microbiology*, v. 58, n. 1, p. 365–390, out. 2004. Disponível em: <<http://www.annualreviews.org/doi/10.1146/annurev.micro.58.030603.123700>>.

DEHIO, Christoph. *Recent progress in understanding Bartonella-induced vascular proliferation. Current Opinion in Microbiology*. [S.l: s.n.], 2003

DELSUC, Frédéric *et al.* Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics*, v. 6, n. 5, p. 361–375, 1 maio 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15861208>>. Acesso em: 31 maio 2018.

DIDDI, Kavita *et al.* Strategy for identification & characterization of Bartonella henselae with conventional & molecular methods. *The Indian journal of medical research*, v. 137, n. 2, p. 380–7, fev. 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23563383>>. Acesso em: 16 jun. 2020.

DINIZ, Lucia Martins *et al.* Bacillary angiomatosis with bone invasion. *Anais Brasileiros de Dermatologia*, v. 91, n. 6, p. 811–814, 1 dez. 2016. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0365-05962016000600811&lng=en&tlng=en>. Acesso em: 25 maio 2020.

DINIZ, Pedro Paulo Vissotto de Paiva *et al.* Risk Factors for Bartonella species Infection in Blood Donors from Southeast Brazil. *PLOS Neglected Tropical Diseases*, v. 10, n. 3, p. e0004509, 21 mar. 2016. Disponível em: <<https://dx.plos.org/10.1371/journal.pntd.0004509>>. Acesso em: 15 maio 2020.

DOBRINDT, U. *et al.* Toxin genes on pathogenicity islands: impact for microbial evolution. *International Journal of Medical Microbiology*, v. 290, n. 4–5, p. 307–311, out. 2000. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/11111903>>. Acesso em: 1 jun. 2018.

DONÀ, D. *et al.* Osteomyelitis in Cat-Scratch Disease: A Never-Ending Dilemma—A Case Report and Literature Review. *Case Reports in Pediatrics*, v. 2018, p. 1–8, 2018. Disponível em: <<https://www.hindawi.com/journals/cripe/2018/1679306/>>. Acesso em: 23 maio 2020.

DONATI, Claudio *et al.* Structure and dynamics of the pan-genome of Streptococcus pneumoniae and closely related species. *Genome Biology*, v. 11, n. 10, p. R107, 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21034474>>. Acesso em: 31 maio 2018.

FISKUS, Warren *et al.* Identification and Characterization of the *DdlB*, *FtsQ* and *FtsA* Genes Upstream of *FtsZ* in *Bartonella bacilliformis* and *Bartonella henselae*. *DNA and Cell Biology*, v. 22, n. 11, p. 743–752, nov. 2003. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/14659047>>. Acesso em: 22 fev. 2018.

FLORIN, Todd a; ZAOUTIS, T. E.; ZAOUTIS, Lisa B. Beyond Cat Scratch Disease: Widening Spectrum of Bartonella henselae Infection. *PEDIATRICS*, v. 121, n. 5, p. e1413–e1425, 1 maio 2008. Disponível em: <<http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2007-1897>>.

FRANZ, Bettina; KEMPF, Volkhart A.J. *Adhesion and host cell modulation: Critical pathogenicity determinants of Bartonella henselae. Parasites and Vectors*. [S.l: s.n.], 2011

- GANESAN, Kandasamy; MIZEN, Kelvin. Cat scratch disease: An unusual cause of facial palsy and partial ptosis: Case report. *Journal of Oral and Maxillofacial Surgery*, v. 63, n. 6, p. 869–872, 1 jun. 2005.
- GARCÍA, Juan C. *et al.* Hepatosplenic cat scratch disease in immunocompetent adults report of 3 cases and review of the literature. *Medicine (United States)*, v. 93, n. 17, p. 267–279, 1 out. 2014.
- GENEREUX, Diane P. *et al.* A comparative genomics multitool for scientific discovery and conservation. *Nature*, v. 587, n. 7833, p. 240–245, 12 nov. 2020. Disponível em: <<http://www.nature.com/articles/s41586-020-2876-6>>.
- GERBER, John E *et al.* Fatal meningitis and encephalitis due to *Bartonella henselae* bacteria. *Journal of forensic sciences*, v. 47, n. 3, p. 640–4, 1 maio 2002. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12051353>>. Acesso em: 23 maio 2020.
- GOMES, Cláudia; RUIZ, Joaquim. Carrion's Disease: the Sound of Silence. *Clinical Microbiology Reviews*, v. 31, n. 1, 29 nov. 2017. Disponível em: <<https://cmr.asm.org/content/31/1/e00056-17>>.
- GUY, Lionel *et al.* A genome-wide study of recombination rate variation in *Bartonella henselae*. *BMC evolutionary biology*, v. 12, p. 65, 11 maio 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22577862>>. Acesso em: 24 jan. 2018.
- HANSEN, E. E. *et al.* Pan-genome of the dominant human gut-associated archaeon, *Methanobrevibacter smithii*, studied in twins. *Proceedings of the National Academy of Sciences*, v. 108, n. Supplement_1, p. 4599–4606, 15 mar. 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21317366>>. Acesso em: 31 maio 2018.
- HARMS, Alexander *et al.* Evolutionary Dynamics of Pathoadaptation Revealed by Three Independent Acquisitions of the VirB/D4 Type IV Secretion System in *Bartonella*. *Genome Biology and Evolution*, v. 9, n. 3, p. 761–776, mar. 2017. Disponível em: <<https://academic.oup.com/gbe/article-lookup/doi/10.1093/gbe/evx042>>.
- HMAIMESS, G. *et al.* Cat scratch disease presenting as meningomyeloradiculopathy. *Archives of Disease in Childhood*, v. 89, n. 7, p. 691–692, 2004.
- HÖRDT, Anton *et al.* Analysis of 1,000+ Type-Strain Genomes Substantially Improves Taxonomic Classification of Alphaproteobacteria. *Frontiers in Microbiology*, v. 11, 7 abr. 2020. Disponível em: <<https://www.frontiersin.org/article/10.3389/fmicb.2020.00468/full>>.
- IANNINO, Filomena *et al.* *Bartonella* infections in humans dogs and cats. *Veterinaria italiana*, v. 54, n. 1, p. 63–72, 31 mar. 2018. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/29631316/>>. Acesso em: 28 jan. 2021.
- JOLLEY, Keith A.; MAIDEN, Martin C.J. Using MLST to study bacterial variation: prospects in the genomic era. *Future Microbiology*, v. 9, n. 5, p. 623–630, maio 2014. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/24957089/>>. Acesso em: 21 out. 2020.
- JORDAN, I. King. Lineage-Specific Gene Expansions in Bacterial and Archaeal Genomes. *Genome Research*, v. 11, n. 4, p. 555–565, 1 abr. 2001. Disponível em: <<http://www.genome.org/cgi/doi/10.1101/gr.1660R>>.
- KAISER, Patrick O. *et al.* *Bartonella* spp.: Throwing light on uncommon human infections.

International Journal of Medical Microbiology, v. 301, n. 1, p. 7–15, jan. 2011. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1438422110000743>>. Acesso em: 19 maio 2020.

KOEHLER, J E. Bartonella-associated infections in HIV-infected patients. *AIDS clinical care*, v. 7, n. 12, p. 97–102, dez. 1995. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/11362939>>. Acesso em: 21 mar. 2020.

KRIZOVA, Lenka; NEMEC, Alexandr. A 63 kb genomic resistance island found in a multidrug-resistant *Acinetobacter baumannii* isolate of European clone I from 1977. *Journal of Antimicrobial Chemotherapy*, v. 65, n. 9, p. 1915–1918, set. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20558468>>. Acesso em: 1 jun. 2018.

KSIAA, Imen *et al.* Update on Bartonella neuroretinitis. *Journal of Current Ophthalmology*, v. 31, n. 3, p. 254–261, 1 set. 2019.

KUMAR, S. *et al.* Statistics and Truth in Phylogenomics. *Molecular Biology and Evolution*, v. 29, n. 2, p. 457–472, 1 fev. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21873298>>. Acesso em: 31 maio 2018.

KWON, Hea Yoon *et al.* Characterization of Clinical Isolates of Bartonella henselae Strains, South Korea. *Emerging Infectious Diseases*, v. 24, n. 5, p. 912–915, 1 maio 2018. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/35938757/>?report=abstract>. Acesso em: 2 fev. 2021.

LA SCOLA, Bernard *et al.* Genotypic Characteristics of Two Serotypes of Bartonella henselae. *Journal of Clinical Microbiology*, v. 40, n. 6, p. 2002–2008, 1 jun. 2002. Disponível em: <<http://jcm.asm.org/cgi/doi/10.1128/JCM.40.6.2002-2008.2002>>.

LAPIERRE, Pascal; GOGARTEN, J. Peter. Estimating the size of the bacterial pan-genome. *Trends in Genetics*, v. 25, n. 3, p. 107–110, mar. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19168257>>. Acesso em: 31 maio 2018.

LIN, Jen Wei; CHEN, Chih Ming; CHANG, Chao Chin. Unknown fever and back pain caused by bartonella henselae in a veterinarian after a needle puncture: A case report and literature review. *Vector-Borne and Zoonotic Diseases*, v. 11, n. 5, p. 589–591, 1 maio 2011. Disponível em: <<https://www.liebertpub.com/doi/abs/10.1089/vbz.2009.0217>>. Acesso em: 14 abr. 2021.

LINDROOS, Hillevi *et al.* Genome Rearrangements, Deletions, and Amplifications in the Natural Population of Bartonella henselae. *Journal of Bacteriology*, v. 188, n. 21, p. 7426–7439, 1 nov. 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16936024>>. Acesso em: 25 jan. 2018.

LINS, Karina de Almeida; DRUMMOND, Marina Rovani; VELHO, Paulo Eduardo Neves Ferreira. Cutaneous manifestations of bartonellosis. *Anais Brasileiros de Dermatologia*, v. 94, n. 5, p. 594–602, 1 set. 2019. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0365059619300625>>. Acesso em: 16 maio 2020.

MADA, Pradeep Kumar; ZULFIQAR, Hassam; JOEL CHANDRANESAN, Andrew Stevenson. *Bartonellosis*. [S.l.: s.n.], 2020. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28613624>>.

MAGGI, Ricardo G. *et al.* BARTONELLA HENSELAE IN CAPTIVE AND HUNTER-

- HARVESTED BELUGA (DELPHINAPTERUS LEUCAS). *Journal of Wildlife Diseases*, v. 44, n. 4, p. 871–877, out. 2008. Disponível em: <<http://www.jwildlifedis.org/doi/10.7589/0090-3558-44.4.871>>. Acesso em: 19 maio 2020.
- MAGGI, Ricardo G. *et al.* Bartonella henselae in Porpoise Blood. *Emerging Infectious Diseases*, v. 11, n. 12, p. 1894–1898, dez. 2005. Disponível em: <http://wwwnc.cdc.gov/eid/article/11/12/05-0969_article.htm>. Acesso em: 19 maio 2020.
- MAGGI, Ricardo G. *et al.* Bartonella spp. bacteremia in high-risk immunocompetent patients. *Diagnostic Microbiology and Infectious Disease*, v. 71, n. 4, p. 430–437, dez. 2011. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0732889311003555>>. Acesso em: 15 maio 2020.
- MAIDEN, M. C. J. *et al.* Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences*, v. 95, n. 6, p. 3140–3145, 17 mar. 1998. Disponível em: <www.pnas.org>. Acesso em: 21 out. 2020.
- MAIDEN, Martin C. J. *et al.* MLST revisited: the gene-by-gene approach to bacterial genomics. *Nature Reviews Microbiology*, v. 11, n. 10, p. 728–736, 2 out. 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23979428>>. Acesso em: 31 maio 2018.
- MASCARELLI, Patricia E. *et al.* Bartonella henselae and B. koehlerae DNA in Birds. *Emerging Infectious Diseases*, v. 20, n. 3, p. 490–492, mar. 2014. Disponível em: <http://wwwnc.cdc.gov/eid/article/20/3/13-0563_article.htm>. Acesso em: 19 maio 2020.
- MEDINI, Duccio *et al.* The microbial pan-genome. *Current Opinion in Genetics & Development*, v. 15, n. 6, p. 589–594, dez. 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16185861>>. Acesso em: 31 maio 2018.
- MERHEJ, Vicky *et al.* Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. *Biology Direct*, v. 4, n. 1, p. 13, 10 abr. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19361336>>. Acesso em: 31 maio 2018.
- MIRA, Alex *et al.* The bacterial pan-genome: A new paradigm in microbiology. *International Microbiology*, 2010.
- MUZZI, A; MASIGNANI, V; RAPPUOLI, R. The pan-genome: towards a knowledge-based discovery of novel targets for vaccines and antibacterials. *Drug Discovery Today*, v. 12, n. 11–12, p. 429–439, jun. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17532526>>. Acesso em: 31 maio 2018.
- NELSON, Christina A.; SAHA, Shubhayu; MEAD, Paul S. Cat-scratch disease in the United States, 2005–2013. *Emerging Infectious Diseases*, v. 22, n. 10, p. 1741–1746, 1 out. 2016.
- NEVES FERREIRA VELHO, Paulo Eduardo *et al.* Angiomatose bacilar: Revisão da literatura e documentação iconográfica. *Anais Brasileiros de Dermatologia*, v. 78, n. 5, p. 601–609, set. 2003.
- OCAÑA, Kary A C S; DÁVILA, Alberto M R. Phylogenomics-based reconstruction of protozoan species tree. *Evolutionary bioinformatics online*, v. 7, p. 107–21, 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21863127>>. Acesso em: 31 maio 2018.
- OGIER, Jean-Claude *et al.* Units of plasticity in bacterial genomes: new insight from the

comparative genomics of two bacteria interacting with invertebrates, *Photorhabdus* and *Xenorhabdus*. *BMC Genomics*, v. 11, n. 1, p. 568, 15 out. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20950463>>. Acesso em: 31 maio 2018.

OLIVEIRA, A. M. *et al.* Suspected needle stick transmission of *Bartonella vinsonii* subspecies *berkhoffii* to a veterinarian. *Journal of Veterinary Internal Medicine*, v. 24, n. 5, p. 1229–1232, set. 2010.

ORAY, Merih *et al.* Diverse clinical signs of ocular involvement in cat scratch disease. *Turk Oftalmoloji Dergisi*, v. 47, n. 1, p. 9–17, 2017.

OTEO, José A. *et al.* Prevalence of *Bartonella* spp. by culture, PCR and serology, in veterinary personnel from Spain. *Parasites & Vectors*, v. 10, n. 1, p. 553, 7 dez. 2017. Disponível em: <<https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-017-2483-z>>. Acesso em: 31 maio 2020.

PATANÉ, José S. L.; MARTINS, Joaquim; SETUBAL, João C. Phylogenomics. *Methods in Molecular Biology*. [S.l.]: Humana Press Inc., 2018. v. 1704. p. 103–187. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/29277865/>>. Acesso em: 9 jan. 2021.

PÉREZ-LOSADA, Marcos; ARENAS, Miguel; CASTRO-NALLAR, Eduardo. Microbial sequence typing in the genomic era. *Infection, Genetics and Evolution*, v. 63, p. 346–359, 1 set. 2018. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30000000/>>. Acesso em: 9 jan. 2021.

PITASSI, Luiza Helena Urso *et al.* *Bartonella* spp. Bacteremia in Blood Donors from Campinas, Brazil. *PLOS Neglected Tropical Diseases*, v. 9, n. 1, p. e0003467, 15 jan. 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/25590435>>. Acesso em: 6 nov. 2017.

QUÉBATTE, Maxime; DEHIO, Christoph. *Bartonella* gene transfer agent: Evolution, function, and proposed role in host adaptation. *Cellular Microbiology*, v. 21, n. 11, 9 nov. 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/cmi.13068>>.

RASKO, D *et al.* Genomics of the group of organisms. *FEMS Microbiology Reviews*, v. 29, n. 2, p. 303–329, abr. 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15808746>>. Acesso em: 31 maio 2018.

RASKO, David A *et al.* The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *Journal of bacteriology*, v. 190, n. 20, p. 6881–93, 15 out. 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18676672>>. Acesso em: 31 maio 2018.

REITER, L. T. *et al.* A Systematic Analysis of Human Disease-Associated Gene Sequences In *Drosophila melanogaster*. *Genome Research*, v. 11, n. 6, p. 1114–1125, 1 jun. 2001. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/11381037>>. Acesso em: 31 maio 2018.

RENO, M. L. *et al.* Biogeography of the *Sulfolobus islandicus* pan-genome. *Proceedings of the National Academy of Sciences*, v. 106, n. 21, p. 8605–8610, 26 maio 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19435847>>. Acesso em: 31 maio 2018.

REY, Michael W *et al.* Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. *Genome Biology*, v. 5, n. 10, p. r77, 2004. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15461803>>. Acesso em: 31 maio 2018.

em: 31 maio 2018.

RIESS, Tanja *et al.* Analysis of Bartonella adhesin A expression reveals differences between various B. henselae strains. *Infection and immunity*, v. 75, n. 1, p. 35–43, jan. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17060468>>. Acesso em: 22 fev. 2018.

ROLAIN, J. M. *et al.* Recommendations for treatment of human infections caused by Bartonella species. *Antimicrobial Agents and Chemotherapy*. [S.l.: s.n.], 2004

SEUBERT, Anja; SCHULEIN, Ralf; DEHIO, Christoph. Bacterial persistence within erythrocytes: A unique pathogenic strategy of Bartonella spp. *International Journal of Medical Microbiology*, 2001.

SHORBATLI, Loujain A.; KORANYI, Katalin I.; NAHATA, Milap C. Effectiveness of antibiotic therapy in pediatric patients with cat scratch disease. *International Journal of Clinical Pharmacy*, v. 40, n. 6, p. 1458–1461, 1 dez. 2018. Disponível em: <<https://link.springer.com/article/10.1007/s11096-018-0746-1>>. Acesso em: 22 jun. 2020.

SIAMER, Sabrina; DEHIO, Christoph. New insights into the role of Bartonella effector proteins in pathogenesis. *Current Opinion in Microbiology*. [S.l.: s.n.], 2015

SIVASHANKARI, Selvarajan; SHANMUGHAVEL, Piramanayagam. Comparative genomics - a perspective. *Bioinformatics*, v. 1, n. 9, p. 376–8, 27 mar. 2007. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17597925>>. Acesso em: 31 maio 2018.

SNIPEN, Lars; ALMØY, Trygve; USSERY, David W. Microbial comparative pan-genomics using binomial mixture models. *BMC genomics*, v. 10, p. 385, 19 ago. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19691844>>. Acesso em: 31 maio 2018.

SOARES, Siomar C. *et al.* The Pan-Genome of the Animal Pathogen Corynebacterium pseudotuberculosis Reveals Differences in Genome Plasticity between the Biovar ovis and equi Strains. *PLoS ONE*, v. 8, n. 1, p. e53818, 14 jan. 2013. Disponível em: <<http://dx.plos.org/10.1371/journal.pone.0053818>>. Acesso em: 31 maio 2018.

SOUZA, Aline Moreira *et al.* Bartonella spp. and hematological changes in privately owned domestic cats from Rio de Janeiro, Brazil. *Journal of infection in developing countries*, v. 11, n. 8, p. 591–596, 5 set. 2017. Disponível em: <<https://jidc.org/index.php/journal/article/view/8152>>. Acesso em: 7 maio 2020.

SOUZA, Aline Moreira De *et al.* Bartonelose: análise molecular e sorológica em gatos do Rio de Janeiro Brasil. *Revista Brasileira de Ciência Veterinária*, v. 17, n. 1, p. 7–11, 2010. Disponível em: <<http://dx.doi.org/10.4322/rbcv.2014.135>>. Acesso em: 7 maio 2020.

TETTELIN, H. *et al.* Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: Implications for the microbial “pan-genome”. *Proceedings of the National Academy of Sciences*, v. 102, n. 39, p. 13950–13955, 27 set. 2005. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16172379>>. Acesso em: 31 maio 2018.

TETTELIN, Hervé *et al.* Comparative genomics: the bacterial pan-genome. *Current opinion in microbiology*, v. 11, n. 5, p. 472–7, out. 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19086349>>. Acesso em: 31 maio 2018.

TROST, E. *et al.* Pangenomic Study of Corynebacterium diphtheriae That Provides Insights into the Genomic Diversity of Pathogenic Isolates from Cases of Classical Diphtheria,

Endocarditis, and Pneumonia. *Journal of Bacteriology*, v. 194, n. 12, p. 3199–3215, 15 jun. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22505676>>. Acesso em: 31 maio 2018.

TUMAPA, Sarinna *et al.* Burkholderia pseudomallei genome plasticity associated with genomic island variation. *BMC genomics*, v. 9, p. 190, 25 abr. 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18439288>>. Acesso em: 1 jun. 2018.

VALENTINE, K. Hope *et al.* Bartonella DNA in Loggerhead Sea Turtles. *Emerging Infectious Diseases*, v. 13, n. 6, p. 949–950, jun. 2007. Disponível em: <http://wwwnc.cdc.gov/eid/article/13/6/06-1551_article.htm>. Acesso em: 19 maio 2020.

VALOR, Cristhian; HUBER, Katherine. Atypical presentation of cat scratch disease: Parinaud's oculoglandular syndrome with facial nerve paresis. *BMJ Case Reports*, v. 2018, p. bcr-2018-224378, 7 jul. 2018.

VAYSSIER-TAUSSAT, Muriel *et al.* The Trw type IV secretion system of Bartonella mediates host-specific adhesion to erythrocytes. *PLoS Pathogens*, v. 6, n. 6, p. e1000946, 10 jun. 2010. Disponível em: <<http://dx.plos.org/10.1371/journal.ppat.1000946>>. Acesso em: 20 mar. 2018.

VERMEULEN, M.J. *et al.* Serological testing for Bartonella henselae infections in The Netherlands: clinical evaluation of immunofluorescence assay and ELISA. *Clinical Microbiology and Infection*, v. 13, n. 6, p. 627–634, 1 jun. 2007. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1198743X14622308>>. Acesso em: 30 maio 2020.

VIEIRA-DAMIANI, Gislaine; ERICSON, Marna Elise. Bartonella Henselae Initial Infection of Mature Human Erythrocytes Observed in Real Time Using Bacterial Endogenous Fluorescence. *Journal of Tropical Diseases*, v. 04, n. 02, 2016. Disponível em: <<http://www.esciencecentral.org/journals/bartonella-henselae-initial-infection-of-mature-human-erythrocytes-observed-in-real-time-using-bacterial-endogenous-fluorescence-2329-891X-1000207.php?aid=69774>>. Acesso em: 23 maio 2020.

WAGNER, Alexander; DEHIO, Christoph. Role of distinct type-IV-secretion systems and secreted effector sets in host adaptation by pathogenic Bartonella species. *Cellular Microbiology*, v. 21, n. 3, 6 mar. 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1111/cmi.13004>>.

WEAR, Douglas J. *et al.* Cat scratch disease: A bacterial infection. *Science*, v. 221, n. 4618, p. 1403–1405, 30 set. 1983. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/6612349>>. Acesso em: 8 mar. 2020.

ZÄHRINGER, Ulrich *et al.* Structure and biological activity of the short-chain lipopolysaccharide from Bartonella henselae ATCC 49882T. *Journal of Biological Chemistry*, v. 279, n. 20, p. 21046–21054, 14 maio 2004. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/14766898/>>. Acesso em: 28 set. 2020.

ZAKHOUR, Ramia *et al.* Transverse myelitis and guillain-barré syndrome associated with cat-scratch disease, Texas, USA, 2011. *Emerging Infectious Diseases*, v. 24, n. 9, p. 1754–1755, 1 set. 2018.

7. PRODUÇÃO CIENTÍFICA

✓ Artigos publicados:

- ✓ **ALVES, Leandro Gomes**; DE CASTRO OLIVEIRA, Letícia, BENEVIDES, Leandro de Jesus; ZEN, Felipe Lucas; TIWARI, Sandeep; JAISWAL, Arun Kumar; GHOSH, Preetam; Barh D; AZEVEDO, Vasco; SOARES, Siomar de Castro. The pan-genome of the zoonotic neglected pathogen *Bartonella henselae* reveals two groups with different patterns of adaptation to hosts. *Journal of Bacteriology and Mycology*, v. 7, n. 3, p. 1-14, 11 jun. 2020. Disponível em: <<https://austinpublishinggroup.com/bacteriology/download.php?file=fulltext/bacteriology-v7-id1133.pdf>>
- ✓ TREVISAN, Rafael Obata; SANTOS, Malú Mateus; DESIDÉRIO, Chamberttan Souza; **ALVES, Leandro Gomes**; DE JESUS SOUSA, Thiago; DE CASTRO OLIVEIRA, Letícia; JAISWAL, Arun Kumar; TIWARI, Sandeep; BOVI, Wesley Guimarães; DE OLIVEIRA-SILVA, Mariana; COSTA-MADEIRA, Juliana Cristina; CASTELLANO, Lúcio Roberto Caçado; SILVA, Marcos Vinicius; AZEVEDO, Vasco; RODRIGUES JUNIOR, Virmondés; OLIVEIRA, Carlo José Freire; SOARES, Siomar de Castro. In Silico Identification of New Targets for Diagnosis, Vaccine, and Drug Candidates against *Trypanosoma cruzi*. *Disease Markers*, v. 2020, p. 1–15, 10 dez. 2020. Disponível em: <<https://www.hindawi.com/journals/dm/2020/9130719/>>.
- ✓ JAISWAL, Arun Kumar; TIWARI, Sandeep; JAMAL, Syed Babar; DE CASTRO OLIVEIRA, Letícia; **ALVES, Leandro Gomes**; AZEVEDO, Vasco; GHOSH, Preetam; OLIVEIRA, Carlo Jose Freire; SOARES, Siomar de Castro. The pan-genome of *Treponema pallidum* reveals differences in genome plasticity between subspecies related to venereal and non-venereal syphilis. *BMC Genomics*, v. 21, n. 1, p. 33, 10 dez. 2020. Disponível em: <<https://bmcmgenomics.biomedcentral.com/articles/10.1186/s12864-019-6430-6>>.
- ✓ FELICE, Andrei Giacchetto; **ALVES, Leandro Gomes**; FREITAS, Alissa de Sarom Ferreira; RODRIGUES, Thais Cristina Vilela; JAISWAL, Arun Kumar; TIWARI, Sandeep; GOMES, Lucas Gabriel Rodrigues; MIRANDA, Fábio Malcher; RAMOS, Rommel Thiago Jucá; AZEVEDO, Vasco; DE CASTRO OLIVEIRA, Letícia, OLIVEIRA, Carlo José Freire; SOARES, Siomar de Castro; BENEVIDES, Leandro de Jesus. Pan-genomic analyses of 47 complete genomes of the *Rickettsia* genus and prediction of new vaccine targets and virulence factors of the species. *Journal of Biomolecular Structure & Dynamics*, p. 1-15, 15 mar. 2021. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/07391102.2021.1898473>>.

✓ Artigo em revisão:

- ✓ FELICE, Andrei Giacchetto; SANTOS, Leonardo Noronha Queiroz; KOLOSSOWSKI, Ian; ZEN, Felipe Lucas; **ALVES, Leandro Gomes**; RODRIGUES, Thais Cristina Vilela; PRATO, LÍgia C S; JAISWAL, Arun Kumar; TIWARI, Sandeep; MIRANDA, Fabio Malcher; RAMOS, Rommel Thiago Jucá; AZEVEDO, Vasco; OLIVEIRA, Carlo José Freire; BENEVIDES, Leandro de Jesus; SOARES, Siomar de Castro. Comparative genomics of *Bordetella pertussis* and prediction of new vaccines and drug targets. *Journal of Biomolecular Structure & Dynamics*.