



Universidade Federal do Pará
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética e Biologia Molecular

***Corynebacterium*: ABORDAGEM GENÔMICA, RELÓGIO MOLECULAR E
SURGIMENTO DA PATOGENICIDADE**

ALYNE CRISTINA SODRÉ LIMA

Belém - Pará

2019

ALYNE CRISTINA SODRÉ LIMA

***Corynebacterium*: ABORDAGEM GENÔMICA, RELÓGIO MOLECULAR E
SURGIMENTO DA PATOGENICIDADE**

Tese de doutorado apresentado ao Programa de Pós-Graduação em Genética e Biologia Molecular, da Universidade Federal do Pará – UFPA, como pré-requisito para obtenção o título de Doutor em Genética e Biologia Molecular.

Área de concentração: Genética e Biologia Molecular.

Orientador: Prof D.Sc. Artur Luiz da Costa da Silva

BELÉM - PARÁ
2019

ALYNE CRISTINA SODRÉ LIMA

***Corynebacterium*: ABORDAGEM GENÔMICA, RELÓGIO MOLECULAR E
SURGIMENTO DA PATOGENICIDADE**

**Data da apresentação: 30/10/2019.
Belém (PA)**

BANCA EXAMINADORA :

Artur Luiz da Costa da Silva

(UFPA – Orientador)

Adriana Ribeiro Carneiro Folador

(UFPA – Membro interno)

Diego Assis das Graças

(UFPA – Membro interno)

Luís Carlos Guimarães

(UFPA – Membro Interno)

André Guimarães Maciel e Silva

(UFPA – Membro externo)

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

L732c Lima, Alyne Cristina Sodré
Corynebacterium: ABORDAGEM GENÔMICA, RELÓGIO
MOLECULAR E SURGIMENTO DA PATOGENICIDADE /
Alyne Cristina Sodré Lima. — 2019.
88 f. : il. color.

Orientador(a): Prof. Dr. Artur Luiz da Costa da Silva
Tese (Doutorado) - Programa de Pós-Graduação em Genética e
Biologia Molecular, Instituto de Ciências Biológicas, Universidade
Federal do Pará, Belém, 2019.

1. análise temporal, análise filogenética, Corynebacterium
pseudotuberculosis, genômica comparativa. I. Título.

CDD 576

AGRADECIMENTOS

A Deus, por ser.

Aos meus familiares que sempre acreditaram que o estudo é fonte transformadora.

Ao meu orientador, prof. Dr. Artur Luiz da Costa da Silva, pela oportunidade e me auxiliar na construção do meu conhecimento e me integrar seu grupo de pesquisa.

A UFPA e ao Programa de Pós-Graduação em Genética e Biologia Molecular por ter me possibilitado o doutoramento.

A CAPES pela concessão da bolsa.

A Silvanira, Soraya e Vagner por todo auxílio em todos os momentos que os requisitei.

Aos professores Joana Marques, Diego Assis, Adriana Folador, Rommel Silva e em especial ao prof Rafael Baraúna, pela disponibilidade e paciência em me ajudar sempre que precisei.

Aos amigos Carla paixão, Lino, Wana, Patrícia, Cássia, Adriedson e Diego Barreto que conquistei no decorrer desses anos e que tornaram meus dias mais alegres... sem o apoio de vocês tudo seria mais difícil.

Aos que colaboraram imensamente na obtenção do conhecimento de mundo novo, para mim, e que são amigos para toda a vida: Vitória, Larissa, Leonardo e Joriane.

As minhas amigas Marcele, Jéssica, Carla e Aline por me acompanhar desde o início da minha trajetória acadêmica e nunca desistir da minha amizade mesmo eu sumindo, de vez enquanto.

Aos meus colegas de trabalho e alunos, que compreenderam a importância desse momento e da melhor forma possível me apoiaram.

Aos meus amigos do PROSEG – UFRA, em especial ao Hugo e a minha querida professora Luciara, e família, que todo dia me inspira.

Aos meus pais Waldir do Socorro, Maria do Socorro, aos meus irmãos Alain e Alan a minha cunhada Lorena e aos meus sobrinhos Davi e Ícaro, por tudo, hoje e sempre.

E a todos que de forma direta e indireta possibilitaram a minha chegada ate aqui.

A todos, OBRIGADA.

RESUMO

Em um mesmo gênero podemos identificar micro-organismos com capacidade patogênica ou não, com pouca informação relacionada ao período temporal de surgimento da patogenicidade, como é o caso do gênero *Corynebacterium*. A espécie não patogênica com maior relevância *Corynebacterium glutamicum* de grande importância na produção de aminoácidos, e as patogênicas sendo *Corynebacterium diphtheriae* produtora de toxina diftérica, e *Corynebacterium pseudotuberculosis* principal agente causador de linfadenite caseosa. Há uma carência tanto de dados epidemiológicos quanto de estudos publicados no estado do Pará no que diz respeito ao cenário da *C. pseudotuberculosis* na região Norte. Desta forma, objetivou-se com este estudo inicialmente realizar o sequenciamento e análise do genoma da cepa Cp05 de *C. pseudotuberculosis*, além de incluir este e outros exemplares de isolados pertencentes ao gênero *Corynebacterium*, com o objetivo de identificar quais eventos genômicos estão relacionados ao surgimento da patogenicidade de *Corynebacterium pseudotuberculosis* de maneira a explorar a plasticidade genômica dentro deste gênero. Para a divulgação do genoma da *C. pseudotuberculosis* PA05 utilizou-se testes microbiológicos, moleculares e bioquímicos. Foram analisados 64 genomas de diferentes espécies do gênero *Corynebacterium* e mais os genomas da *Mycobacterium tuberculosis* H37Rv, *Nocardia* sp. Y18 e *Rhodococcus jostii* RHA1 como outgroups, para análise filogenética e de relógio molecular. Utilizando os genes *dnaA*, *dnaK*, *gyrB*, *recA*, *rpoB* e o 16S rRNA, Para análise de genômica comparativa utilizou-se três espécies: *C. pseudotuberculosis*, *C. diphtheriae* e *C. glutamicum*. Estes genomas tiveram os genes ortólogos calculados no PGAP v.1.2.1, e os genes espécies específicos, foram analisados no software Blast2GO v.5.0 para determinar compartimentos celulares, funções moleculares e processos biológicos. Para determinar diferenças entre as características genômicas foram comparadas as análises de variância ($p < 0,05$): tamanho do genoma, número de genes, densidade gênica. Na árvore filogenética bayesiana obtida as espécies patogênicas *C. pseudotuberculosis*, *C. diphtheriae* e *C. ulcerans* agruparam-se. Um segundo agrupamento de espécies patogênicas foi formado por *C. kutscheri*, *C. matruchotii* e *C. mustelae*, próximo ao clado de *C. pseudotuberculosis*. *C. glutamicum* formou um clado não patogênico que inclui *C. crudilactis*, *C. allunae*, *C. deserti* e *C. efficiens*. *C. glutamicum* foi uma das espécies não patogênicas mais intimamente relacionadas às espécies patogênicas do clado de *C. pseudotuberculosis*. Divergências entre espécies de *Corynebacterium* foram observadas, a cada 5.000 anos, além das primeiras espécies deste gênero ser bactérias de vida livre e sofreram adaptação aos diferentes hospedeiros. A realização das análises genômicas comparativas nas três espécies de *Corynebacterium* permitiu um entendimento sobre a evolução genômica das corinebactérias patogênicas, diante do compartilhamento de 911 genes de ortólogos. A obtenção da sequência proteica dos genes e a análise dessas sequências apresentaram os principais termos de ontologia gênica, que foram processo de redução da oxidação; Ligação de ATP e DNA; componente integral da membrana. Sequenciamento de novas cepas é necessário, pois com mais exemplares melhor é o entendimento da evolução deste gênero, melhor o entendimento da patogenicidade *C. pseudotuberculosis*, que aqui pode ser confirmada evolutivamente associada a domesticidade animal, redução de tamanho de genoma e priorização de genes essenciais para manutenção de vida da célula.

Palavras-chave: análise temporal, análise filogenética, *Corynebacterium pseudotuberculosis*, genômica comparativa.

ABSTRACT

We can identify microorganisms with pathogenic capacity or not within the same genus, through little information related to the time period of onset of pathogenicity, such as the genus *Corynebacterium*. The most relevant non-pathogenic species are: *Corynebacterium glutamicum* for its great importance in amino acid production, and the pathogenic species *Corynebacterium diphtheriae* producing diphtheria toxin, and *Corynebacterium pseudotuberculosis* leading causative agent of lymphadenitis caseosa. There is a lack of both epidemiological data and studies published in the state of Pará regarding the scenario of *C. pseudotuberculosis* in the northern region. Thus, the objective of this study was initially to disseminate the genome of the strain Cp05 *C. pseudotuberculosis*, as well as to include this and other specimens of isolates belonging to the genus *Corynebacterium*, in order to identify which genomic events are related to the emergence of the pathogenicity in *Corynebacterium pseudotuberculosis*, exploring the genomic plasticity within this genus. For the disclosure of the genome of the Cp05 *C. pseudotuberculosis* strain we used microbiological, molecular and biochemical tests. Using 64 genomes of different species of the genus *Corynebacterium* and *Mycobacterium tuberculosis* H37Rv, *Nocardia* sp. Y18 and *Rhodococcus jostii* RHA1 as outgroups for phylogenetic and molecular clock analysis. Testing the genes *dnaA*, *dnaK*, *gyrB*, *recA*, *rpoB* and 16S rRNA, for comparative genomic analysis from three species: *C. pseudotuberculosis*, *C. diphtheriae* and *C. glutamicum*. These genomes had the ortholog genes calculated in PGAP v.1.2.1, and species specific genes were analyzed in Blast2GO v.5.0 software to determine cell compartments, molecular functions and biological processes. To determine differences between genomic characteristics, variance analysis ($p < 0.05$) was compared: genome size, number of genes and gene density. In the Bayesian phylogenetic tree obtained, the pathogenic species *C. pseudotuberculosis*, *C. diphtheriae* and *C. ulcerans* grouped. A second grouping of pathogenic species was formed by *C. kutscheri*, *C. matruchotii* and *C. mustelae*, close to the *C. pseudotuberculosis* clade. *C. glutamicum* has formed a non-pathogenic clade that includes *C. crudilactis*, *C. allunae*, *C. deserti* and *C. efficiens*. *C. glutamicum* was one of the nonpathogenic species most closely related to the pathogenic species of the *C. pseudotuberculosis* clade. Divergences between *Corynebacterium* species have been observed every 5,000 years, in addition to the first species of this genus being free-living bacteria and undergoing adaptation to different hosts. Comparative genomic analysis of the three *Corynebacterium* species allowed an understanding of the genomic evolution of pathogenic corynebacteria, given the sharing of 911 genes from orthologs. Obtaining the protein sequence of the genes and the analysis of these sequences presented the main terms of gene ontology, which were oxidation reduction process; ATP and DNA binding; and integral component of the membrane. Sequencing of new strains is necessary, because with more specimens the better the understanding of the evolution of this genus, the better the understanding of *C. pseudotuberculosis* pathogenicity, which here can be confirmed evolutionarily associated with animal domesticity, genome size reduction and prioritization of essential genes for maintaining cell life.

Keywords: *Corynebacterium pseudotuberculosis*, temporal analysis, phylogenetic analysis, comparative genomics.

I. ESTRUTURA DO MANUSCRITO

Esta tese está dividida em dois capítulos, cada um baseado em um artigo publicado ou submetido para publicação, como seguem:

- I. O primeiro capítulo trata do relato e apresentação da sequência do genoma preliminar da cepa de *Corynebacterium pseudotuberculosis*, realizado através de cultivo, testes bioquímicos e de bioinformática. Publicado na revista *GENOME ANNOUNCEMENTS*.
- II. O segundo capítulo aborda o período temporal de surgimento da patogenicidade, através de relógio molecular, de *Corynebacterium pseudotuberculosis* de maneira a explorar a plasticidade genômica dentro do gênero *Corynebacterium*. Neste segundo capítulo optamos por submeter o trabalho primeiramente para o preprint BioRxiv, uma vez que o trabalho estará disponível para críticas e melhoramentos antes da submissão final à uma revista.

LISTA DE FIGURAS

Figura 1. Fotomicrografia de *Corynebacterium glutamicum* em microscópio de luz após coloração de gram.

Figura 2. Fotomicrografia de *Corynebacterium diphtheriae* após coloração de Gram. Bactérias Gram-positivas medindo de 0,5 a 1 µm de tamanho e em forma claviforme.

Figura 3. Fotomicrografia de *Corynebacterium pseudotuberculosis* após coloração de gram. As hastes gram-positivas, de formas pleomórficas que variam de cocóides a bastões filamentosos.

Figura 4. Doença edematosa da pele (a) e linfadenite caseosa (b), lesões ocasionadas por *C. pseudotuberculosis*.

CAPÍTULO II

Figure 1. Rooted phylogenetic tree calculated using the six genes *dnaA*, *dnaK*, *gyrB*, *recA*, *rpoB* and 16S rRNA. Bayesian inference was used and the topology was tested using bootstrap analysis. *Nocardia* sp. sp. Y48 (NZ_CP018082), *Rhodococcus* WB9 (NZ_CP041769) e *Mycobacterium tuberculosis* H37Rv (NC_000962) were used as outgroups. Bootstrap values above 0.5 are shown in tree.

Figure 2. Molecular clock of the *Corynebacterium* genus. *Nocardia* sp. Y48 (NZ_CP018082), *Rhodococcus* WB9 (NZ_CP041769) e *Mycobacterium tuberculosis* H37Rv (NC_000962) were used as outgroups. Ruler values represent time at every 5,000 years in the past.

Figure 3. Unrooted phylogenomic tree calculated using the 911 orthologous genes of *C. pseudotuberculosis* (red clade), *C. diphtheriae* (green clade) and *C. glutamicum* (blue clade). Branch weight is shown between each resolved clade. Numbers next to each clade indicate the amount of species-specific genes identified by PGAP. A total of 117 genomes were used, 74 of *C. pseudotuberculosis*, 21 of *C. diphtheriae*, and 22 of *C. glutamicum*. Accession numbers are listed in table S2.

Figure 4. GO terms at level 5 of cellular component category for the specie-specific genes of *C. pseudotuberculosis* 258 (A), *C. diphtheriae* (B) and *C. glutamicum* (C). It is worth noting that for all three species, membrane proteins were the most representative in the dataset.

Figure 5. GO enrichment analysis. Bar graphs show the GO terms that were significantly ($p < 0.05$) enriched in the dataset of genes found only in *C. pseudotuberculosis* (A), *C. diphtheriae* (B), and *C. glutamicum* (C), compared to the set of genes in their genomes. GO terms are ranked starting from the most significantly enriched term.

Figure S1. GO terms of genes found only in the genomes of *C. diphtheriae*.

Figure S2. GO terms of genes found only in the genomes of *C. glutamicum*.

Figure S3. GO terms of genes found only in the genomes of *C. pseudotuberculosis*.

LISTA DE TABELAS

CAPÍTULO II

Table S1. Genome size, number of genes and genome density for the three species analyzed by comparative genomics. Data was compared using ANOVA ($p < 0,05$). Different letter in the same column indicates statistically significant differences.

Table S2. List of genomes of *C. pseudotuberculosis*, *C. diphtheriae*, and *C. glutamicum* available in GenBank at March-2019. These genes were downloaded and used to calculate the unrooted phylogenomic tree of figure 3.

SUMÁRIO

1. INTRODUÇÃO	13
1.1. Bactérias e a evolução.....	13
1.2. Filogenia e o relógio molecular: uso em estudos microbiológicos	14
1.3. Gênero <i>Corynebacterium</i> sp.....	16
1.3.1. <i>Corynebacterium glutamicum</i>	18
1.3.2. <i>Corynebacterium diphtheriae</i>	19
1.3.3. <i>Corynebacterium pseudotuberculosis</i>	21
REFERÊNCIA BIBLIOGRÁFICA	25
2. OBJETIVOS	30
2.1. Objetivo Geral:	30
2.2. Objetivos específicos:.....	30
3. CAPÍTULO I: Draft Genome Sequence of <i>Corynebacterium pseudotuberculosis</i> Strain PA05 Isolated from na Ovine Host in Pará State, Brazil	32
4. CAPÍTULO II: A molecular timescale and comparative genomics of <i>Corynebacterium</i> genus.....	37
4.1. Abstract.....	40
4.2. Introduction.....	41
4.3. Methods.....	42
4.4. Results and Discussion.....	45
Acknowledgments	55
Conflict of interest	55
References.....	55
Supplementary material.....	63
5. DISCUSSÃO GERAL.....	70
6. CONCLUSÃO GERAL	79
7. REFERÊNCIA BIBLIOGRÁFICA.....	80
8. ANEXOS	84

1. INTRODUÇÃO

1.1. Bactérias e a evolução

A história evolutiva é melhor contada por meio da análise de filogenia de grupos tendo como base a ancestralidade desses grupos e mudanças evolutivas ocorridas que refletem a história dessa evolução (Swofford & Olsen, 1996).

O processo evolutivo microbiano ocorre a partir de mutações e recombinações. A mutação, que ocorre naturalmente, é a alteração aleatória (substituição, deleção, inserção e duplicação) na sequência de DNA, fato que prolonga e estende ao longo do tempo e de gerações.

As mutações podem ser neutras quando não altera de forma positiva ou negativa a célula, deletérias quando diminui a habilidade de sobrevivência da célula, e a benéfica que auxilia na melhor adaptação da célula ao ambiente em que vive. Na recombinação há a quebra do DNA religando-se e formando uma nova combinação genética, essa recombinação é essencial para integrar o genoma do DNA adquirido por transferência horizontal (Madigan et al., 2016).

A evolução se dá devido a alterações na frequência de alelos, em uma população, por várias gerações, sendo os principais processos evolutivos associados à seleção natural e deriva genética. A seleção baseia-se na habilidade de adaptação do organismo nas condições as quais estão sujeitos, quando ocorrem mutações benéficas na seleção, há um aumento na capacidade adaptativa da célula (COLAUTTI and LAU, 2015).

Já a deriva genética se dá pela alteração da frequência alélica devido a certo grupo deixar mais descendentes que a média da população. A deriva genética possui grande importância na evolução da patogenicidade de microrganismo, pois novas infecções estão associadas a grupos pequenos de organismos (HAUCK and MAIDEN, 2018).

A elucidação dos mecanismos que coordenam a interação entre cromossomos, replicons acessórios, DNA móvel e mecanismos de transferência horizontal de genes é essencial para melhor apreender a evolução e adaptação de nicho das espécies, como demonstrou Fouts et al., (2016) em estudo de 20 espécies de *Leptospira* quando identificaram que a evolução é frequentemente impulsionada por elementos genéticos móveis via transferência horizontal de genes. Foram identificados novos genes e famílias de genes específicos para estas espécies, com previsão de possíveis adaptações de *Leptospira* infecciosa a mamíferos.

Reva et al., (2015) em estudo da evolução de micobactérias, identificaram que apesar da resistência a troca horizontal de genes e ser considerado um patógeno clonal, a *Mycobacterium tuberculosis* apresenta indícios que na fase inicial da evolução sua patogenicidade teve origem em aquisições horizontais de fatores de virulência. Nesse estudo os genomas de micobactérias possuíam vários genes comuns a todas as espécies, os quais foram associados a elementos genéticos móveis transferidos horizontalmente, como o *fadD22 – fadE*. As proteínas FadD22 mostraram uma conservação muito alta, implicando uma indispensabilidade dessa proteína para a patogênese.

1.2. Filogenia e o relógio molecular: uso em estudos microbiológicos

Devido às classificações filogenéticas iniciais serem baseadas principalmente em características morfológicas, classificar procariotos, por muitos anos, era difícil devido à morfologia simples e a difícil diferenciação desses organismos microscópicos (Rosselló-Mora & Amann, 2001).

Com o passar dos anos, já no século XX, o surgimento de metodologias como cultivo em meios artificiais, desenvolvimento de parâmetros bioquímicos e fisiológicos foram tentativas de classificação dos microrganismos (Schleifer, 2009). Contudo, as classificações fenotípicas sendo mesmo que úteis não estabelecem, de forma confiável, relações evolutivas nos organismos procariotos.

Em estudo, Zuckerkandl e Pauling (1965), afirmaram que a história evolutiva de um organismo poderia ser obtida a partir do genoma, devido à diferença do material genético que não são observadas nos polipeptídios, pois diferentes códonos podem ser traduzidos em um mesmo resíduo de aminoácido e, desta forma, o material genético possuía mais informações filogenéticas e poderia ser utilizado para o estudo de relações evolutivas entre os organismos.

Com esse novo entendimento iniciou-se uma abordagem molecular para melhor classificar procariotos (Gevers et al., 2005). Inicialmente uma comparação estabelecida pela composição de bases no genoma, permitiu somente uma classificação superficial (Krieg, 1988).

Com o advento de técnicas de hibridização DNA-DNA, possibilitou agrupamentos baseados na semelhança genética entre os indivíduos. Nesta técnica, em condições de pré-estabelecidas, as fitas de DNA dos diferentes microrganismos se associaram, de acordo o grau de similaridade entre os mesmos, permitindo a quantificação desta similaridade (Rosselló-Mora & Amann, 2001).

Para melhor identificação das bactérias passou-se a aplicar três métodos de classificação: fenótipo, que analisa características morfológicas, metabólicas, fisiológicas e química; genotípica que analisa o genoma dos microrganismos; e a filogenética, tentativa de melhor posicionamento dos organismos a partir dos dados de relação evolutiva baseado em sequências genéticas.

A maior disponibilidade de árvores filogenéticas de melhor resolução, associada ao aumento criterioso das análises baseadas em filogenias, permitiu acessar a influência de eventos de variância e dispersão na distribuição dos táxons, melhorando a capacidade técnica para lidar com a história biológica e geográfica (Lomolino *et al.*, 2006).

Moléculas como 16S rRNA estão presentes em todos os organismos e por se apresentarem conservada na homologia funcional e evolutiva (Wear *et al.*, 2018), se tornaram o marcador filogenético de maior importância nos estudos de diversidade e filogenia microbiana, possibilitando que relações filogenéticas distantes e próximas pudessem ser examinadas.

Outros genes conservados podem ser utilizados como marcadores filogenéticos no estudo da diversidade de procariotos, como genes que codificam proteínas ribossomais, toxinas, proteínas de ligação ao DNA e amino-acil sintetases (Kiel *et al.*, 2018; Mekadim *et al.*, 2018; Liang *et al.*, 2019).

Com o atual contexto científico de avanços no estudo da genômica, das ferramentas analíticas fornecidas pela bioinformática e a crescente disponibilidade de sequências completas de variadas espécies, o entendimento das condições e causas de evolução molecular tornaram-se cada vez mais acessíveis.

A aplicação da bioinformática diante da biologia evolutiva compreendem a análise e comparação de genomas de diferentes espécies, objetivando atingir um melhor entendimento da função de genes e regiões não codificantes do genoma, por meio de comparações que auxiliam a determinação das taxas de evolução molecular, assim como o estudo da relação entre estas taxas (Ali 2013; Choudhuri 2014).

Uma das particularidades da evolução molecular é a semelhança com um relógio devido à forma de progressão, isto a denominou de *relógio evolutivo molecular* (Zuckerkandl and Pauling 1965) esta abordagem tem sido amplamente utilizada para reconstruir filogenias e compreender a relação evolutiva ao nível dos ácidos nucleicos e proteínas, entre as espécies.

As modificações ocorridas, em relação a um ancestral relacionadas a mudanças evolutivas levam a modificações nos caracteres herdados em uma população, obtendo

como resultado final da evolução a formação de novas espécies. Essa evolução gera diversidade em todos os níveis possíveis de organização biológica, como ocorre em *Corynebacterium* spp. que apresenta grande diversidade de espécies com características distintas (patogênicas, patogênicas oportunistas e não patogênicas).

Alterações sequenciais podem ser utilizadas como relógio molecular, estimando o tempo desde que as duas linhagens de divergiram. O relógio molecular está baseado em alterações de nucleotídeos que se acumulam ao longo do tempo, são normalmente neutras e aleatórias. Estas estimativas são mais acuradas quando estão associadas a registros geológicos (MADIGAN et al., 2016).

Quando associados a divergências nucleotídicas e uma estimativa robusta da filogenia do grupo, os eventos geológicos podem ser usados para calibrar as taxas de evolução molecular, e fornecer a cronologia da diversificação de um táxon (Lundberg, 1998; Sanderson, 1998).

Louca et al., (2018) em estudo com grande diversidade de bactérias usando o tempo geológico e filogenia baseada no gene rRNA (RNA ribossômico 16S), de filos como *Firmecutes*, *Bacteroidetes*, *Proteobacterias* entre outras, encontraram evidências de que a diversidade bacteriana aumentou principalmente nos últimos um bilhão de anos, com taxas de especiação e extinção em geral aproximadamente constantes ou apenas lentamente alteradas.

1.3. Gênero *Corynebacterium* sp.

As bactérias constituem cerca de 23.458 espécies conhecidas, além de apresentarem uma distribuição de 7.372 gêneros, 1.969 famílias, 816 ordens, 296 classes e 112 filos. (<https://gtdb.ecogenomic.org/>). O gênero *Corynebacterium* está presente dentro de um subfilo de actinomicetos de eubactérias Gram-positivas (Stackebrandt et al., 1997), apresenta hastes que podem ser retas ou então com curvaturas ligeiras nas extremidades e dependendo da espécie, podem ter formas elipsoidais, ovoides ou “cabos de chicote” (Bernard e Funke, 2012), com seu DNA abrangendo um elevado conteúdo de guanina + citosina, além de ter uma representação na descendência do domínio Bactéria (Liebl, 2006). As espécies que têm uma maior relevância são as *Corynebacterium diphtheriae* e as não difteriais, que são, respectivamente, a principal causa da difteria e presentes na pele e mucosas (Alibi et al., 2017).

Com o intuito de descrever a estirpe tipo *Corynebacterium diphtheriae* Lehman e Neumann em 1896, descreveram o gênero *Corynebacterium*. Sendo que *C. diphtheriae* já havia sido descrita como sinônima de *Microsporon diphthericum* (Trujillo, 2016), *Bacillus diphtheriae* por Kruse (Corper e Sweni, 1918), *Pacinia loeffleri* (Trevisan, 1889) e *Mycobacterium diphtheriae* por Krasil'nikov em 1941 (<http://www.bacterio.net/corynebacterium.html>), antes de sua classificação taxonômica final. Posteriormente, o gênero passou a integrar outras espécies de bactérias que apresentavam forma e/ou patogenicidade semelhante (Pascual et al., 1995).

O gênero *Corynebacterium* conta atualmente com 110 espécies válidas, sendo 132 espécies contando com espécies sinônimas e 11 subespécies (<http://www.bacterio.net/corynebacterium.html>). Sendo as quatro espécies mais recentemente afiliadas ao gênero são *Corynebacterium senegalense*, descrita recentemente por Ndiaye et al., (2019) isolada da palma da mão de uma pessoa saudável, no Senegal, as espécie *Corynebacterium bouchesdurhonense* e *Corynebacterium provencense*, descritas por Lo et al., (2019) isoladas das fezes de um paciente obeso, e *Corynebacterium urinapleomorphum*, descrita por Niang et al., (2019) isoladas de uma amostra de urina pediátrica.

O gênero *Corynebacterium* está distribuído de maneira onipresente na natureza, sendo seus membros isolados dos mais variados ambientes (Alibi et al., 2017), compondo atualmente de:

- Espécie não patogênica como a *Corynebacterium glutamicum* e *Corynebacterium efficiens*, de importância biotecnológica (produção de aminoácidos) (KALINOWSKI et al., 2003; NISHIO et al., 2003), e *Corynebacterium variabile*, importante na indústria de queijos (SCHRÖDER et al., 2011).

- Espécies oportunistas como *Corynebacterium jeikeium*, *Corynebacterium urealyticum* e *Corynebacterium resistens* (SCHRÖDER et al., 2012), que são frequentemente associadas a infecções contraídas durante estadia em centros de saúde, *Corynebacterium aurimucosum*, que está associada a riscos à gravidez (TROST et al., 2010).

- Espécie patogênica como *Corynebacterium pseudotuberculosis*, *Corynebacterium diphtheriae* e *Corynebacterium ulcerans*, de importante relevância médica e veterinária (TROST & TAUCH, 2014).

As espécies mais importantes do ponto de vista clínico são, as de estirpe tipo *C. diphtheriae*, *C. ulcerans* e *C. pseudotuberculosis*, capazes de produzir a toxina diftérica

(TD) e urease positiva, patogênicas em humanos e/ou animais. Já do ponto de vista biotecnológico, a espécie *C. glutamicum* se destaca dentro do gênero (Hoffmann et al., 2018).

Estes gêneros compartilham certas características como alto conteúdo G+C e associados à parede celular, incluindo espessura, presença de ácidos micólicos, peptidoglicano e arabinogalactano e ácidos graxos saturados e insaturados (DORELLA et al., 2006; BELCHIOR et al., 2006).

1.3.1. *Corynebacterium glutamicum*

C. glutamicum é uma bactéria vastamente estudada devido à sua capacidade de secretar aminoácidos como L-glutamato e L-lisina, que são importantes industrialmente, como promotores de sabor e aditivos alimentares.

Essa espécie, de grande uso biotecnológico, foi isolada pela primeira vez em 1950 em solo japonês quando buscavam por um produtor natural de glutamato (Mimitsuka et al., 2007). *C. glutamicum* foi inicialmente identificada como *Micrococcus glutamicum*, posteriormente, com o progresso na taxonomia, principalmente com a utilização de sequenciamento do gene *rrs*, muitas das cepas foram reclassificadas como *C. glutamicum* (Jo et al., 2012).

Apresenta forma de bastonete, não produtor de esporos, catalase e nitrato positivo, anaeróbio facultativo, Gram-positiva, mesófilo e utiliza o metabolismo fermentativo para decompor os carboidratos. Possui parede celular com camada de peptidoglicano, ácidos micólicos de cadeia curta, ácidos meso-diaminopimélicos e polímeros de arabino-galactano (Riegel e Ruimy, 1997).

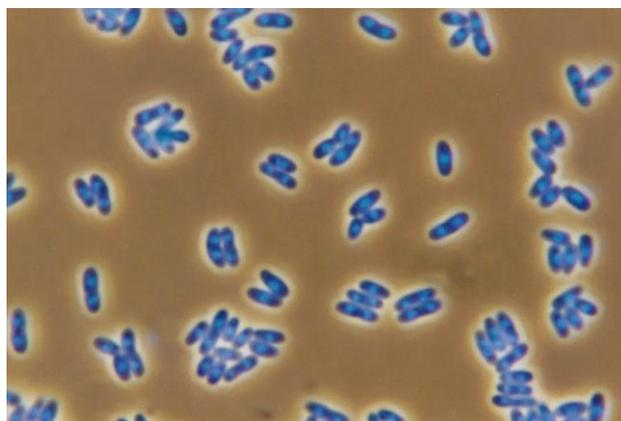


Figura 1. Fotomicrografia de *Corynebacterium glutamicum* em microscópio de luz após coloração de gram. FONTE: <http://www.actino.jp/DigitalAtlas/subwin.cgi?target=1-1>

Esse microrganismo apresenta cultivo e manuseio genético facilitados, sendo classificado como seguro e possui resistência a oscilações de suprimento de oxigênio e substratos, o que é comum em fermentações em grande escala (Buchholz et al. 2014).

Os alimentos utilizados para decomposição de carboidratos de forma industrial por *C. glutamicum* são amido de milho, mandioca ou trigo, melão de cana e melão de beterraba, comuns em países de alta produção agrícola como o Brasil. Este organismo pode converter cerca de 100 g/L de glicose em 50 g/L de ácido L-glutâmico em 2 a 4 dias

Em microrganismos a lisina é sintetizada por duas rotas possíveis, *C. glutamicum* é um dos poucos representantes capazes de operar com duas rotas biossintéticas de forma simultânea, a partir do 2-oxoglutarato e acetil-CoA pela rota do α -aminoadipato e do aspartato pela rota do diaminopimelato (An et al., 1999).

Um dos fatores ponderados na escolha de cepas é o rendimento máximo de uma célula em produzir lisina a partir de glicose (75%), sendo que as análises de fluxo metabólico, considerando as rotas metabólicas utilizadas por *C. glutamicum* para a produção de lisina, indicam que tal rendimento é crescente, podendo ser de até 82% (Shah, 2002; Hoffmann et al., 2018).

Em *C. glutamicum* existe apenas uma isoenzima de aspartato quinase, e é codificada por dois genes: *lysC* α e *lysC* β , representando as duas sequências codificadoras para as duas subunidades da enzima. A sua atividade é bioquimicamente regulada por inibição retroativa concertada por lisina e treonina, que se ligam à subunidade β (Becker et al., 2007).

Mais de 2,5 milhões de toneladas de lisina são produzidas anualmente, para alimentação animal, utilização nas indústrias farmacêutica, cosmética, de combustíveis, além de ocorrer um crescimento nesse valor em cerca de 7% ao ano, movimentando um mercado de aproximadamente US\$ 3,6 bilhões (Sebastian et al., 2019), e sendo a maior parte desta produção, feita utilizando espécies de *Corynebacterium*.

1.3.2. *Corynebacterium diphtheriae*

A *C. diphtheriae* é uma bactéria patogênica, gram-positiva, não esporulada, não encapsulada, não móvel, aeróbia com forma em bastonetes podendo medir até 0,5 a 1 μm de tamanho (CERDEÑO-TÁRRAGA et al., 2003).

Uma característica dos bacilos é ficarem arranados como letras chinesas ou em

paliçadas (Figura 2) devido aos movimentos logo após a divisão celular (TODAR, 2009).



Figura 2. Fotomicrografia de *Corynebacterium diphtheriae* após coloração de Gram. Bactérias Gram-positivas medindo de 0,5 a 1 µm de tamanho e em forma claviforme. FONTE: CDC (2014).

Baseado na gravidade da infecção junto com morfologia de colônias, padrões bioquímicos de fermentação de carboidratos, potencial de hemólise, as cepas são classificadas em 4 biovars: *Mitis*, *gravis*, *intermedius* e *belfanti* (LEGGETT et al., 2010).

Os biovars *mitis*, *gravis*, *intermedius* e *belfanti* se assemelham quanto reação positiva para produção de cistinase em meio Tinsdale, para glicose e maltose e negativo para atividade da pirazinamidase ureia e sacarose (EFSTRATIOU & GEORGE 1999).

Para redução de nitrato *mitis*, *gravis* e *intermedius* apresentam reação positiva e *belfanti* reação negativa, já para glicogênio *mitis*, *intermedius* e *belfanti* apresentam reação negativa e *gravis*, positiva (EFSTRATIOU & GEORGE 1999).

Os relatos de ocorrência de infecções têm sido causados pelos biótipos *gravis* e *mitis*, e apesar das particularidades dos biovars o tratamento e a prevenção são idênticos para todas as cepas toxigênicas (BONNET et al., 1999).

Esta bactéria é isolada principalmente em humanos, contudo já foi isolada em outros indivíduos, como equinos, gatos e cães (LEGGETT et al., 2010; HALL et al., 2010; DIXON, 2010).

Em humanos, *C. diphtheriae* causa a doença diftérica, associada principalmente

a ação da TD, sendo essa transmissível e aguda acometendo o trato respiratório superior (POPOVIC et al., 2000)

Na sua forma original a bactéria *C. diphtheriae* não apresenta patogenia, a patogenicidade está relacionada à infecção por um bacteriófago, tornando-a lisogênica. O bacteriófago transporta a informação genética de uma toxina, o gene *tox*, que é uma sequência altamente conservada e codificada pelo cromossomo bacteriano (FUNKE et al., 1997).

A TD atua em tecidos como o miocárdio, além do sistema nervoso, rins e pele. Em lesões na pele, onde ocorre à infecção, as pseudomembranas são o resultado da combinação dos efeitos do crescimento bacteriano, da produção de toxinas, da necrose tecidual e da resposta imune do hospedeiro, causando a difteria cutânea (SALYERS, 1994).

Houve redução nos casos de difteria no mundo após o desenvolvimento de uma vacina que atua na inativação da DT (POPOVIC et al., 2000). Entretanto devido o aumento da susceptibilidade de crianças e idosos, medidas de controle ineficientes, falta de disponibilidade de vacinas (MATTOS-GUARALDI et al., 2003), a emergência de novas cepas não toxigênicas as quais a vacina com toxóide não é eficiente (CERDEÑO-TÁRRAGA et al., 2003; VIGUETTI et al., 2012), possibilitaram a reemergência de casos associados a *C. diphtheriae*.

1.3.3. *Corynebacterium pseudotuberculosis*

A *C. pseudotuberculosis* é uma bactéria intracelular facultativa, gram-positiva, não-móvel, apresenta formas pleomórficas que variam desde cocóides a bastões filamentosos podendo medir de 0,6 µm a 3 µm de tamanho.

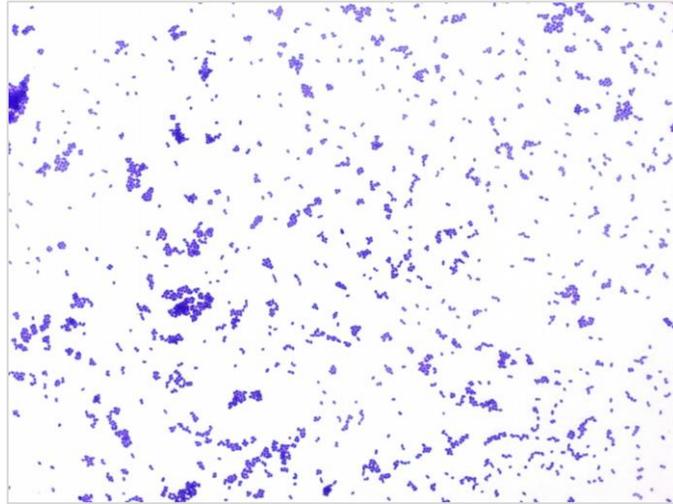


Figura 3. Fotomicrografia de *Corynebacterium pseudotuberculosis* após coloração de gram. As hastes gram-positivas, de formas pleomórficas que variam de cocóides a bastões filamentosos. FONTE: http://people.upei.ca/jlewis/html/lab_4_coryne.html

Apresenta grande quantidade de lipídeos na parede celular, particularmente o ácido micólico (COSTA, 2002). Esta bactéria é beta-hemolítica, produz esfingomielinase fosfolipase D e catalase, mas não produz oxidase (Buxton & Fraser, 1977).

Com reação positiva para produção de cistinase em meio Tinsdale, ureia, glicose e maltose, e negativa para atividade da pirazinamidase, sacarose e glicogênio (EFSTRATIOU & GEORGE 1999).

Possui dois biótipos ou biovars denominados *equi* e *ovis* que se diferenciam pela capacidade de reduzir ou não o nitrato e pela preferência do hospedeiro (BIBERSTEIN, KNIGHT, & JANG, 1971).

O biovar *equi* (redutor de nitrato a nitrito) infecta principalmente bubalinos, equinos, camélídeos e bovinos, podendo causar a linfangite ulcerativa (PRATT et al., 2005) e a doença edematosa da pele (SYAME et al., 2008).

Enquanto o biovar *ovis* (não redutor de nitrato) acomete ovinos e caprinos causando a linfadenite caseosa, e em humanos pode causar a linfadenite subaguda crônica (DORELLA et al., 2006; TROST et al., 2010a), recentemente este biovar foi isolada em suíno em Portugal (OLIVEIRA et al., 2014).

Os bovinos podem ser infectados pelos dois biovars, com predomínio do biovar *equi* (COSTA, 2002), o que acarreta vários sintomas variando de reações

piogranulomatosas, a mastite, dermatite ulcerativa, formação de abscessos e comprometimento de vísceras (YERUHAM et al., 2003).

As doenças associadas a esse microrganismo provocam a redução na produção de carne e leite, depreciação da lã, retardo no desenvolvimento dos animais, gastos com tratamento e honorários veterinários, deficiências nos índices reprodutivos do rebanho, condenação de carcaças, descarte precoce e morte ocasional de animais (RADOSTITS et al., 2007).

A linfadenite caseosa está entre as principais doenças infecciosas em países com tradicionalmente criadores de ovinos e caprinos como Austrália, Nova Zelândia, África do Sul, Estados Unidos, Canadá e Brasil (BELCHIOR et al., 2004).

No estado de Minas Gerais, sudeste do Brasil, 78,9% das cabras foi diagnosticado como soropositivas para *C. pseudotuberculosis* (SEYFFERT et al., 2010). Na região nordeste há relatos de grande número de animais infectados, sendo essa região a maior produtora de ovinos e caprinos do país (BROWN, OLANDER & ALVES, 1987; PINHEIRO, ALVES & HADDAD, 2000; SOUZA et al., 2011).

Os animais acometidos com linfadenite caseosa são tratados, normalmente, através de drenagem dos linfonódos, contudo, essa pratica não elimina totalmente as bactérias, devido à possibilidade de contaminar órgãos internos e também pelo grande potencial de contaminação do ambiente (DORELLA et al., 2006).

Além de que, o tratamento com antibióticos não gera resultados satisfatórios devido a difícil penetração na cápsula do abscesso, tornando o tratamento inviável (COLLETT, BATH & CAMERO, 1994), ressaltando que a profilaxia é o melhor método de combate à enfermidade.



(a)



(b)

Figura 4. Doença edematosa da pele (a) e linfadenite caseosa (b), lesões ocasionadas por *C. pseudotuberculosis*. Fonte: SYAME et al. (2008) e acervo próprio (2018).

A virulência de *C. pseudotuberculosis* está relacionada a três fatores principais; as estruturas da parede celular, sua capacidade para persistência intracelular em macrófago e a produção de fosfolipase D (PLD) como exotoxina, sendo esta considerada o principal fator de virulência (SELIM, 2001; HALL et al., 2010).

A PLD auxilia na disseminação da bactéria a partir do local da infecção causando hidrólise de paredes vasculares e linfáticas endoteliais, além disso, contribui com a sobrevivência de *C. pseudotuberculosis* no hospedeiro durante as 18 primeiras semanas, provocando assim forte estimulação do sistema imune (HODGSON et al., 1992).

Apesar do principal fator de virulência em *C. pseudotuberculosis* ser a exotoxina de fosfolipase D (PLD) (MCNAMARA et al., 1995), estipes toxigênicas dessa espécie também podem produzir toxinas de difteria (KRAEVA et al., 2007).

Mesmo que o gene PLD desempenhe um papel central na patogênese, *C. pseudotuberculosis* estirpe 31 apresenta uma mutação de frameshift perto da extremidade 3' neste gene, que pode contribuir para redução na capacidade desta estirpe se disseminar no hospedeiro. Contudo, esta mesma estirpe de *C. pseudotuberculosis* apresenta outro fator de virulência de grande importância, o gene da toxina diftérica (SOARES et al., 2013b).

Em estudo de *C. pseudotuberculosis* (SILVA, et al., 2012) foi identificada a ilha de patogenicidade, denominada PiCp12, que abriga o gene *tox* (SOARES et al., 2013b; VIANA et al., 2017), além da capacidade de cepas *C. pseudotuberculosis* linhagens 31, 992 e 993 isoladas de búfalo do Egito, produzirem TD (HEGGELUND, et al., 2015; SELIM et al., 2016).

REFERÊNCIA BIBLIOGRÁFICA

Alibi, S.; Ferjani, A.; Boukadida, J.; Cano, M. E.; Fernández-Martínez, M.; Martínez-Martínez, L.; e Navas, J. (2017). Occurrence of *Corynebacterium striatum* as an emerging antibiotic-resistant nosocomial pathogen in a Tunisian hospital. **Scientific reports**, 7(1), 9704.

ALVES, F. S. F.; SANTIAGO, L. B.; PINHEIRO, R. R. (2007). Linfadenite caseosa: o estado da arte. Sobral: **Embrapa Caprinos**.

An, G. H.; Song, K. B.; e Sinskey, A. J. (1999). Redirection of carbon flux to lysine in a recombinant of *Corynebacterium lactofermentum* ATCC 21799 by limited supply of pantothenate. **Journal of bioscience and bioengineering**, 88(2), 168-172.

Becker, J.; Klopprogge, C.; Herold, A.; Zelder, O.; Bolten, C. J.; e Wittmann, C. (2007). Metabolic flux engineering of l-lysine production in *Corynebacterium glutamicum*—over expression and modification of G6P dehydrogenase. **Journal of Biotechnology**, 132(2), 99–109.

BELCHIOR, S. E.; GALLARDO, A.; ABALOS, A.; JODOR, N.; JENSEN, O. Actualización sobre linfadenitis caseosa: el agente etiológico y la enfermedad. **Rev Vet Argent**. 2006; 23: 258-78.

Bernard, K. A., e Funke, G. (2015). *Corynebacterium*. **Bergey's Manual of Systematics of Archaea and Bacteria**, 1–70.

BIBERSTEIN, E.L., KNIGHT, H.D. & JANG, S. Two biotypes of *Corynebacterium pseudotuberculosis*. **Vet Rec**. 1971;89:691-692.

BROWN, C.C., OLANDER, H.J. & ALVES, S.F. Synergistic hemolysis-inhibition titers associated with caseous lymphadenitis in a slaughterhouse survey of goats and sheep in Northeastern Brazil. **Can J Vet Res**. 1987;51:46-49.

CERDEÑO-TÁRRAGA, A.M., EFSTRATIOU, A., DOVER, L.G., HOLDEN, M.T.G., PALLEN, M., BENTLEY, S.D., BESRA, G.S., CHURCHER, C., JAMES, K.D., DE ZOYSA, A. et al. The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. **Nucleic Acids Res**. 2003;31:6516-6523.

Corper, H. J., e Sweany, H. C. (1918). The enzymes of the tubercle bacillus. **Journal of bacteriology**, 3(2), 129.

COSTA, L. F. M. *Corynebacterium pseudotuberculosis*, o agente etiológico da linfadenite caseosa em caprinos. **Rev Cienc Med Biol**. 2002; 1: 105-15.

COLLETT MG, BATH GF, CAMERON CM (1994) *Corynebacterium pseudotuberculosis* infections. In: Infections diseases of livestock with special reference to Southern Africa. **Oxford University Press** 2: 1387–1395.

DIXON, B. Sick as a dog. **Lancet Infect Dis**. 2010;10:73.

DORELLA, F. A.; PACHECO, L. G. C.; OLIVEIRA, S. C.; MIYOSHI, A.; AZEVEDO, V. *Corynebacterium pseudotuberculosis*: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. **Veterinary research**, v. 37, n. 2, p. 201–18, 2006.

Drancourt, M.; Bollet, C.; Carlouz, A.; Martelin, R.; Gayral, J. P.; e Raoult, D. (2000). 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. **Journal of clinical microbiology**, 38(10), 3623-3630.

EFSTRATIOU, A.; GEORGE, R. C. Laboratory guidelines for the diagnosis of infections caused by *Corynebacterium diphtheriae* and *C. ulcerans*. **Communicable disease and public health**, v. 2, p. 250-257, 1999.

Ehresmann, C.; Fellner, P.; e Ebel, J. P. (1970). Nucleotide sequences of sections of 16S ribosomal RNA. **Nature**, 227(5265), 1321.

Fouts, D. E.; Matthias, M. A.; Adhikarla, H.; Adler, B.; Amorim-Santos, L.; Berg, D. E.; ... e Haake, D. A. (2016). What makes a bacterial species pathogenic?: comparative genomic analysis of the genus *Leptospira*. **PLoS neglected tropical diseases**, 10(2), e0004403.

FUNKE, G.; VON GRAEVENITZ, A.; CLARRIDGE, J. E.; BERNARD, K.A. Clinical microbiology of coryneform bacteria. **Clin Microbiol Rev.** 1997;10(1):125–159.

Gevers, D.; Cohan, F.M.; Lawrence, J.G.; Spratt, B.G.; Coenye, T.; Feil, E.J.; Stackebrandt, E.; Peer, Y.V.; Vandamme, P.; Thompson, F.L.; e Swings, J. (2005). Re-evaluating prokaryotic species. **Nature Reviews Microbiology**, 3 (9), 733-739.

HALL, A.J., CASSIDAY, P.K., BERNARD, K.A., BOLT, F., STEIGERWALT, A.G., BIXLER, D., PAWLOSKI, L.C., WHITNEY, A.M., IWAKI, M., BALDWIN, A. et al. Novel *Corynebacterium diphtheriae* in domestic cats. **Emerg Infect Dis.** 2010;16:688-691.

Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B.J., Gaigalat, L. et al. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of Laspartate- derived amino acids and vitamins. *J Biotechnol.* 2003;104:5-25.

Kiel, M.; Sagory-Zalkind, P.; Miganeh, C.; Stork, C.; Leimbach, A.; Sekse, C.; Mellmann, A.; Rechenmann, F.; e Dobrindt, U. (2018). Identification of Novel Biomarkers for Priority Serotypes of Shiga Toxin-Producing *Escherichia coli* and the Development of Multiplex PCR for Their Detection. **Frontiers in Microbiology**, 9.

Kraeva, L.A., Manina, Z., Tseneva, G.I., Radchenko, A.G., 2007. Etiologic role of *Corynebacterium nondiphtheriae* in patients with different pathology. **Zh. Mikrobiol. Epidemiol. Immunobiol.** 5, 3–7.

LEGETT, B.A.; ZOYZA, A.; ABBOTT, Y.E.; LEONARD, N.; MARKEY, B.; EFSTRATIOU, A. Toxigenic *Corynebacterium diphtheriae* isolated from a wound in a horse. **Vet Rec.**;166(21):656-7. DOI:10.1136/vr.b4846. 2010.

Liang, C.; Huang, Y.; e Wang, H. (2019). pahE, a functional marker gene for polycyclic aromatic hydrocarbon-degrading bacteria. **Appl. Environ. Microbiol.**, 85(3), e02399-18.

Liebl, W. (2006). *Corynebacterium*--Nonmedical. **The Prokaryotes: Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes**, 796-818.

Lo, C. I.; Morand, A.; Ndongo, S.; Raoult, D.; Fournier, P. E.; e Fenollar, F. (2019). *Corynebacterium urinapleomorphum* sp. nov., a new bacterial species isolated from human urine sample. **New Microbes and New Infections**, 100576.

Louca, S.; Shih, P. M.; Pennell, M. W.; Fischer, W. W.; Parfrey, L. W.; e Doebeli, M. (2018). Bacterial diversification through geological time. **Nature Ecology & Evolution**.

MATTOS-GUARALDI, A. L. M.; HIRATA, J. R.; CARVALHO AZEVEDO, V. A. *Corynebacterium diphtheriae*, *Corynebacterium ulcerans* and *Corynebacterium pseudotuberculosis*—General Aspects. In: *Corynebacterium diphtheriae* and Related Toxigenic Species. **Springer Netherlands**, 2014. p. 15-37.

MCNAMARA, P. J.; CUEVAS, W. A.; SONGER, J. G.; 1995. Toxic phospholipases D of *Corynebacterium pseudotuberculosis*, *C. ulcerans* and *Arcanobacterium haemolyticum*: cloning and sequence homology. **Gene** 156, 113–118.

Meier-Kolthoff, J. P.; Klenk, H. P.; e Göker, M. (2014). Taxonomic use of DNA G+ C content and DNA–DNA hybridization in the genomic age. **International journal of systematic and evolutionary microbiology**, 64(2), 352-356.

Mekadim, C.; Killer, J.; Pechar, R.; e Mrázek, J. (2018). Variable regions of the glyS, infB and rplB genes usable as novel genetic markers for identification and phylogenetic purposes of genera belonging to the family Propionibacteriaceae. **International journal of systematic and evolutionary microbiology**, 68(8), 2697-2705.

Ndiaye, C.; Fall, N. S.; Kuete, E.; Lo, C. I.; Raoult, D.; Sokhna, C.; e Lagier, J. C. (2019). *Corynebacterium senegalense* sp. nov. and *Arthrobacter senegalensis* sp. nov., two new Actinobacteria isolated from skin swab from the palm of hand. **New microbes and new infections**, 31, 100583.

NISHIO, Y.; NAKAMURA, Y.; KAWARABAYASI, Y.; USUDA, Y.; KIMURA, E.; SUGIMOTO, S.; MATSUI, K.; YAMAGISHI, A.; KIKUCHI, H.; IKEO, K. et al. Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. **Genome Res.** 2003;13:1572-1579.

OLIVEIRA, M.; BARROCO, C.; MOTTOLA, C.; SANTOS, R.; LEMSADDEK, A.; TAVARES, L.; SEMEDO-LEMSADDEK, T. First report of *Corynebacterium pseudotuberculosis* from caseous lymphadenitis lesions in Black Alentejano pig (*Sus scrofa domesticus*). **BMC veterinary research**, 10(1), 218. (2014).

PASCUAL, C.; LAWSON, P. A; FARROW, J. A; GIMENEZ, M. N.; COLLINS, M. D. Phylogenetic analysis of the genus *Corynebacterium* based on 16S rRNA gene sequences. **International journal of systematic bacteriology**, v. 45, n. 4, p. 724–8, out 1995.

PINHEIRO, R. R.; ALVES, F. S. F.; HADDAD, J. P. (2000). Aspectos epidemiológicos da caprinocultura cearense. **Arquivo Brasileiro Med Veterinaria Zootecnia**, 52(10).

POPOVIC, T.; MAZUROVA, I. K.; EFSTRATIOU, A.; VUOPIO-VARKILA, J.; REEVES, M. W.; DE ZOYSA, A.; GLUSHKEVICH, T.; GRIMONT, P. Molecular epidemiology of diphtheria. **J Infect Dis.** 2000;181 Suppl 1:S168-77.

PRATT, S. M.; SPIER, S. J.; CARROLL, S. P.; VAUGHAN, B.; WHITCOMB, M. B.; WILSON, W. D. Evaluation of clinical characteristics, diagnostic test results, and outcome in horses with internal infection caused by *Corynebacterium pseudotuberculosis*: 30 cases (1995-2003). **Journal of the American Veterinary Medical Association**, 227(3), 441-8. 2005.

RADOSTITS, O. M.; BLOOD, D. C.; GAY, C. C. Veterinary medicine: a textbook of the diseases of catlee, sheep, pigs, goats and horses. 9th ed. Philadelphia: **Bailliere Tindall**, 2007. p. 830- 9.

Reva, O.; Korotetskiy, I.; e Ilin, A. (2015). Role of the horizontal gene exchange in evolution of pathogenic Mycobacteria. **BMC evolutionary biology**, 15(1), S2.

SALYERS, A. A.; WHITT, D.D. Diphtheria. Bacterial pathogenesis: a molecular approach. **DC: American Society for Microbiology Press**, p. 113-121, Washington,1994.

SCHRÖDER, J.; MAUS, I.; TROST, E.; TAUCH, A. Complete genome sequence of *Corynebacterium variabile* DSM 44702 isolated from the surface of smear-ripened cheeses and insights into cheese ripening and flavor generation. **BMC Genomics**. 2011;12:545.

Sebastian, J.; Zheng, M.; Jiang, Y.; Zhao, Y.; Wang, H.; Song, Z.; Li, X.; Pang, J.; e Zhang, T. (2019). One-pot conversion of lysine to caprolactam over Ir/H-Beta catalysts. **Green Chemistry**, 21(9), 2462-2468.

SELIM, S. A. Review: Oedematous Skin Disease of Buffalo in Egypt. **Journal Vet. Med.** B 48, 241-258. (2001)

SEYFFERT, N., GUIMARÃES, A.S., PACHECO, L.G.C., PORTELA, R.W., BASTOS, B.L., DORELLA, F.A., HEINEMANN, M.B., LAGE, A.P., GOUEVIA, A.M.G., MEYER, R. ET al. High seroprevalence of caseous lymphadenitis in Brazilian goat herds revealed by *Corynebacterium pseudotuberculosis* secreted proteínas-based ELISA. **Res Vet Sci**. 2010;88:50-55.

Soares, S. C.; Silva, A.; Trost, E.; Blom, J.; Ramos, R.; Carneiro, A.; ... & Barbosa, E. G. 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, 2013b.

SOUZA, M.F., DE CARVALHO, A. Q., JR, F. G., & RIET-CORREA, F. (2011). Linfadenite caseosa em ovinos deslanados abatidos em um frigorífico da Paraíba. **Pesq. Vet. Bras**, 31(3), 224-230.

Stackebrandt, E.; Rainey, F. A.; e Ward-Rainey, N. L. (1997). Proposal for a new hierarchic classification system, Actinobacteria classis nov. **International Journal of Systematic and Evolutionary Microbiology**, 47(2), 479-491.

SYAME, S. M.; EL-HEWAIRY, H. M.; SELIM, S. A. (2008). Protection of buffaloes against oedematous skin disease by recombinent-bacterin and toxoid-bacterin vaccines. **Global Vet**, 2, 151-156.

TROST, E., GÖTKER, S., SCHNEIDER, J., SCHNEIKER-BEKEL, S., SZCZEPANOWSKI, R., TILKER, A., VIEHOEVER, P., ARNOLD, W., BEKEL, T., BLOM, J. et al. Complete genome sequence and lifestyle of black-pigmented *Corynebacterium aurimucosum* ATCC 700975 (formerly *C. nigricans* CN-1) isolated

from a vaginal swab of a woman with spontaneous abortion. **BMC Genomics**. 2010a;11:91.

TROST, E., & TAUCH, A. Comparative Genomics and Pathogenicity Islands of *Corynebacterium diphtheriae*, *Corynebacterium ulcerans*, and *Corynebacterium pseudotuberculosis*. In *Corynebacterium diphtheriae* and Related Toxigenic Species (pp. 39-65). **Springer Netherlands**. 2014. p. 39-65.

Trujillo, ME (2016). Actinobactérias. Em eLS, John Wiley & Sons, Ltd (Ed.). 1–16.

VIGUETTI, S.Z., PACHECO, L.G.C., SANTOS, L.S., SOARES, S.C., BOLT, F., BALDWIN, A., DOWSON, C.G., ROSSO, M.L., GUIZO, N., MIYOSHI, A. et al. Multilocus sequence types of invasive *Corynebacterium diphtheriae* isolated in the Rio de Janeiro urban area, Brazil. **Epidemiol Infect**. 2012;140:617-620.

Wear, E. K.; Wilbanks, E. G.; Nelson, C. E.; e Carlson, C. A. (2018). Primer selection impacts specific population abundances but not community dynamics in a monthly time- series 16S rRNA gene amplicon analysis of coastal marine bacterioplankton. **Environmental microbiology**, 20(8), 2709-2726.

Yang, Z. J.; Ramesh, V.; e Zhang, L. H. (2019). Synthesis and Evaluation of Novel Neamine–Nucleoside Conjugates as Potential Antibiotic Targets for *Escherichia coli* 16S Ribosomal RNA. In **Epitranscriptomics** (pp. 151-163). Humana Press, New York, NY.

YERUHAM, I.; ELAD, D.; FRIEDMAN, S.; PERL, S. *Corynebacterium pseudotuberculosis* infection in Israeli dairy cattle. **Epidemiol Infect**. 2003;131:947-955.

MADIGAN, Michael T. et al. Microbiologia de Brock-14ª Edição. Artmed Editora, 2016.

COLAUTTI, Robert I.; LAU, Jennifer A. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. **Molecular Ecology**, v. 24, n. 9, p. 1999-2017, 2015.

HAUCK, Sofia; MAIDEN, Martin CJ. Clonally Evolving Pathogenic Bacteria. In: **Molecular Mechanisms of Microbial Evolution**. Springer, Cham, 2018. p. 307-325.

2. OBJETIVOS

2.1. Objetivo Geral:

Caracterizar *Corynebacterium pseudotuberculosis* isolada no Pará e identificar, filogeneticamente, o surgimento da patogenicidade no gênero *Corynebacterium*.

2.2. Objetivos específicos:

- Coletar, isolar e identificar cepas de material caseoso de pequenos ruminantes com suspeita de linfadenite caseosa oriundos de propriedades no estado do Pará;
- Determinar o período temporal de surgimento da patogenicidade de *Corynebacterium pseudotuberculosis*;
- Caracterização da ontologia gênica em *Corynebacterium pseudotuberculosis*, *Corynebacterium diphtheriae* e *Corynebacterium glutamicum*;
- Correlacionar elementos estruturais: tamanho do genoma, quantidade de genes e densidade de genes de *Corynebacterium pseudotuberculosis*, *Corynebacterium diphtheriae* e *Corynebacterium glutamicum*;
- Determinar quais termos do *gene ontology* estão enriquecidos nos genes específicos da *Corynebacterium pseudotuberculosis*, *Corynebacterium diphtheriae* e *Corynebacterium glutamicum*.

Capítulo I

3. CAPÍTULO I: Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA05 Isolated from na Ovine Host in Pará State, Brazil

O primeiro capítulo trata do relato e apresentação da sequência do genoma preliminar a cepa de *Corynebacterium pseudotuberculosis*. Há uma carência tanto de dados epidemiológicos quanto de estudos publicados no estado do Pará no que diz respeito ao cenário da *C. pseudotuberculosis* na região Norte, tornando necessário a coletar de material da caseoso, para isolamento e identificação de cepas, através de métodos de microbiologia, bioquímica e de diagnóstico molecular, possibilitando a confirmação da ocorrência deste microrganismo nas propriedades rurais do estado e publicar essas informações nos banco de dados público.

Desconhecimento da enfermidade está relacionado a não notificação das doenças resultantes da infecção por este microrganismo, mesmo o Pará sendo um dos principais criadores de animais susceptivos a esta enfermidade (caprinos, ovinos, bovinos, bubalinos).

Desta forma, objetivou-se com este trabalho realizar a divulgação do genoma da cepa Cp05 *C. pseudotuberculosis*, isolado de um hospedeiro ovino no Estado do Pará, Brasil, com intuito de melhor apresentar esta espécie na sua apresentação molecular e disseminação no Brasil.

Para este trabalho foram utilizados análise de cultivo para identificação de padrão de hemólise, quando cultivado em agar sangue, de parâmetros morfológicos, quando cultivado em BHI (*Brain Heart Infusion* - Himedia) sólido e líquido, a partir de secreção caseosa obtiva da punção localizado nos gânglios linfáticos na região posterior da coxa de uma ovelha.

Ao atende os parâmetros coincidindo com o preconizado para *C. pseudotuberculosis*, foram aplicados testes de identificação microbiológica e bioquímica, coloração de Gram por kit (Larboclin, Brasil) que foi utilizada para a confirmação morfológica das células bacterianas, e a utilização do kit API®Coryne (API-bioMérieux, Inc) para identificação bioquímica.

Através de PCR Multiplex foi realizada a identificação molecular, para a confirmação das análises microbiológicas e bioquímicas. Os genes específicos avaliados foram: *rpoB*, *pld* e *rrs*. A extração de DNA foi executada de acordo o kit DNeasy® Blood & Tissue (Quiagen), sendo o genoma sequenciado com a plataforma Ion Torrent PGM.

Para avaliação da qualidade das leituras foi utilizado o software FastQC, o FASTX foi usado para realizar a filtragem e o recorte. As leituras tratadas foram montadas usando o SPAdes versão 3.9.0. A ferramenta Lasergene 11 Core Suite foi utilizada para retirada de redundância.

Para gerar os scaffold foi utilizado-se o software Mauve 2.4.0. Para fechamento de gaps utilizou-se GapBlaster e CLC Genomics Workbench. Por fim, para anotação automática, foi submetido ao servidor RAST (Anotações Rápidas usando Tecnologia de Subsistemas).

O genoma tem 2.435.137 pb, 2.295 sequências codificadoras, um conteúdo de G+C de 52,2%, cinco pseudogenes, 53 genes de tRNAs e seis genes de rRNAs. O genoma completo foi depositado no GenBank com o número de acesso CP019159.



AMERICAN
SOCIETY FOR
MICROBIOLOGY

genomeA^{announcements}TM

PROKARYOTES



Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA05 Isolated from an Ovine Host in Pará State, Brazil

Alyne Cristina Sodré Lima,^a Vitória Almeida Gonçalves de Moura,^a
Kenny da Costa Pinheiro,^a Carla Thais Moreira Paixão,^a
Wana Lailan Oliveira da Costa,^a Adriana Ribeiro Carneiro Folador,^a
Ana Luiza de Mattos Guaraldi,^b Rommel T. J. Ramos,^a Artur Silva,^a
Joana Montezano Marques^a

Laboratory of Genomic and Bioinformatics, Federal University of Pará, Center of Genomics and System Biology, Belém, Pará, Brazil^a; College of Medical Science, State University of Rio de Janeiro, Rio de Janeiro, Brazil^b

ABSTRACT We report here the draft genome sequence of *Corynebacterium pseudotuberculosis* PA05, isolated from an ovine host in Pará State, Brazil. *C. pseudotuberculosis* is an etiological agent of diseases with veterinary and medical importance. The genome contains 2,435,137 bp, a G+C content of 52.2%, 2,295 coding sequences, five pseudogenes, 53 tRNAs, and six rRNAs.

Corynebacterium pseudotuberculosis is a Gram-positive, nonmotile, nonsporulated, pleomorphic, intracellular, and aerobic-facultative bacterium (1). *C. pseudotuberculosis* is mainly known as the etiologic agent of caseous lymphadenitis (CLA), a chronic disease in small ruminants (goats and sheep) and is also considered an important agent to other diseases such as ulcerative lymphangitis and ulcerative dermatitis, among others. This pathogen affects several species, including sheep, goat, horse, cattle, llama, alpaca, buffalo, and human (2).

CLA is a contagious disease that is characterized by the presence of abscesses in the lymph nodes and internal organs of the host. After infection, the bacterium becomes encapsulated within walled-off lesions from which they escape immune system-mediated action, initiating a state of persistence (3). The infection presents itself as a disease difficult to eradicate. It is responsible for large economic losses resulting from the damage it inflicts on livestock, such as reducing the quality of wool, decreasing milk production, lowering animal weight, and causing death and damage to carcasses (4). The bacterium has numerous survival mechanisms and uses many strategies to adapt to its environment; the two toxic factors most described in the literature are a lipid factor that improves cellular wall resistance against digestion by cellular enzymes and exotoxin with hemolytic capacity that increases vascular permeability and facilitates invasion (5). Here, we report the genome sequencing of *C. pseudotuberculosis* PA05, a strain isolated from a punctured caseous abscess located in the lymph nodes in the posterior thigh region of a sheep host (Santa Inês breed) in Pará, Brazil.

The genome was sequenced with the Ion Torrent PGM platform using a fragment library that generated 1,251,198 reads. FastQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) was used to evaluate the quality of the reads, and the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit) was used to perform filtering and trimming, in which reads with a Phred quality score of ≥ 20 were removed. The trimmed reads were then assembled using SPAdes version 3.9.0 (<http://bioinf.spbau.ru/spades>) with *k*-mers 31, 33, 35, 61, 63, and 65. The assembly generated 24 contigs. The number of contigs was reduced to eight using the Lasergene 11 Core Suite tool

Received 24 January 2017 Accepted 25
January 2017 Published 30 March 2017

Citation Lima ACS, de Moura VAG, Pinheiro KDC, Paixão CTM, da Costa WLO, Folador ARC, Guaraldi ALDM, Ramos RTJ, Silva A, Marques JM. 2017. Draft genome sequence of *Corynebacterium pseudotuberculosis* strain PA05 isolated from an ovine host in Pará State, Brazil. *Genome Announc* 5:e00082-17. <https://doi.org/10.1128/genomeA.00082-17>.

Copyright © 2017 Lima et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Joana Montezano Marques, jomontezanomarkes@gmail.com.

Downloaded from <http://genomea.asm.org/> on November 6, 2017 by guest

(<http://www.dnastar.com/t-dnastar-lasergene.aspx>). The scaffold was generated using Mauve version 2.4.0 (6) with *C. pseudotuberculosis* strain 1002B (CP012837.1) as the reference genome. Gap closure was done using GapBlaster (7) and CLC Genomics Workbench (<http://www.clcbio.com>). The final scaffolding with six contigs was submitted to the RAST (Rapid Annotations using Subsystems Technology) server (<http://rast.nmpdr.org>) for automatic annotation. The draft genome has 2,435,137 bp, 2,295 coding sequences, a G+C content of 52.2%, five pseudogenes, 53 tRNAs genes, and six rRNAs genes.

Accession number(s). This whole-genome shotgun project has been deposited in GenBank under the accession number [CP019159](https://doi.org/10.1093/nar/gkz115).

ACKNOWLEDGMENTS

This study was supported by the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), the Foundation for Supporting Research in the State of Pará (FAPESPA), and the Genome and Proteome Network of the State of Pará (RPGP).

REFERENCES

- Dorella FA, Pacheco LG, Oliveira SC, Miyoshi A, Azevedo V. 2006. *Corynebacterium pseudotuberculosis* microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Vet Res* 37:201–218. <https://doi.org/10.1051/vetres:2005056>.
- Baird GJ, Fontaine MC. 2007. *Corynebacterium pseudotuberculosis* and its role in ovine caseous lymphadenitis. *J Comp Pathol* 137:179–210. <https://doi.org/10.1016/j.jcpa.2007.07.002>.
- Tashjian JJ, Campbell SG. 1983. Interaction between caprine macrophages and *Corynebacterium pseudotuberculosis*: an electron microscopic study. *Am J Vet Res* 44:690–693.
- D'Afonseca V, Moraes PM, Dorella FA, Pacheco LG, Meyer R, Portela RW, Miyoshi A, Azevedo V. 2008. A description of genes of *Corynebacterium pseudotuberculosis* useful in diagnostics and vaccine applications. *Genet Mol Res* 7:252–260. <https://doi.org/10.4238/vol7-1gm438>.
- Yeruham I, Elad D, Friedman S, Perl S. 2003. *Corynebacterium pseudotuberculosis* infection in Israeli dairy cattle. *Epidemiol Infect* 131:947–955. <https://doi.org/10.1017/S095026880300894X>.
- Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. <https://doi.org/10.1371/journal.pone.0011147>.
- de Sá PHCG, Miranda F, Veras A, de Melo DM, Soares S, Pinheiro K, Guimarães L, Azevedo V, Silva A, Ramos RTJ. 2016. GapBlaster—a graphical gap filler for prokaryote genomes. *PLoS One* 11:e0155327. <https://doi.org/10.1371/journal.pone.0155327>.

Capítulo II

4. CAPÍTULO II: A molecular timescale and comparative genomics of *Corynebacterium* genus

O segundo capítulo, trata da investigação através de relógio molecular de surgimento da patogenicidade de *Corynebacterium pseudotuberculosis* em conjunto com mais 62 espécies do gênero *Corynebacterium*, além de comparar geneticamente espécies patogênicas, *C. pseudotuberculosis*, *C. diphtheriae*, e não patogênicas, *C. glutamicum*. A investigação o período temporal de surgimento da patogenicidade e identificação da ontologia gênica em *Corynebacterium pseudotuberculosis*, *Corynebacterium diphtheriae* e *Corynebacterium glutamicum* possibilita melhor caracterização das variações associadas a tamanho do genoma, quantidade de genes, densidade de genes e função gênica dessas espécies.

O avanço da bioinformática diante da biologia evolutiva tornou possível reconstruir filogenias e compreender a relação evolutiva entre as espécies, possibilitando maior explanação quanto a estudo de gêneros com maior diversidade de espécies, como é o caso do gênero *Corynebacterium*. Dessa forma, objetivou-se investigar o período temporal de surgimento da patogenicidade de *Corynebacterium pseudotuberculosis* de maneira a explorar a plasticidade genômica dentro do gênero *Corynebacterium*.

Foram recuperados 64 genomas de espécies de *Corynebacterium*, com os outgroups *Mycobacterium tuberculosis* H37Rv, *Nocardia* sp. Y18, and *Rhodococcus jostii* RHA1. Para genômica comparativa, selecionou-se as espécies *C. pseudotuberculosis*, *C. diphtheriae* e *C. glutamicum*, que foram analisadas usando o software PGAP. Os genes espécie específicos, foram analisados no software OmicsBox para determinar os termos GO nas categorias compartilhamento celular, função molecular e processo biológico. Através das análises, foram agrupadas primeiramente as espécies *C. pseudotuberculosis*, *C. diphtheriae* e *C. ulcerans*, em outro aglomerado foram encontradas as espécies *C. kutscheri*, *C. matruchotii*, *C. mustelae* e *C. glutamicum* foi uma das espécies não patogênicas que mais se relacionou com o clado patogênico da *C. pseudotuberculosis*, indicando uma distância evolutiva significativa. A divergência entre espécies de *Corynebacterium* foram observadas, a cada 5.000 anos, além das primeiras espécies deste gênero ser bactérias de vida livre e sofreram adaptação aos diferentes hospedeiros.

A realização das análises genômicas comparativas nas três espécies de *Corynebacterium* permitiu um entendimento sobre a evolução genômica das corinebactérias patogênicas. A obtenção da sequência proteica dos genes e a análise dessas sequências apresentaram os principais termos de ontologia genética. Foi possível observar que grande parte dos genes que desempenham sua função na membrana celular e componente integral da membrana, implicando na adaptação a sobrevivência em ambientes adversos, sendo este um dos principais pontos de diferenciação destas espécies, patogênicas e não patogênicas.

-Artigo submetido à preprint BioRxiv

A molecular timescale and comparative genomics of *Corynebacterium* genus

Alyne Lima¹, Sávio de Souza Costa¹, Davi Josué Marcon¹, Edson Luiz Folador²,
Marcelo Vallinoto³, Rafael A. Baraúna^{1,4*}, Artur Silva^{1,4}

¹ Centro de Genômica e Biologia de Sistemas, Instituto de Ciências Biológicas, Universidade Federal do Pará, 66075-110, Belém, PA, Brazil.

² Centro de Biotecnologia, Universidade Federal da Paraíba, 58050-085, João Pessoa, PB, Brazil

³ Instituto de Estudos Costeiros, Universidade Federal do Pará, Campus Bragança, 68600-000, Bragança, Brazil.

⁴ Laboratório de Engenharia Biológica, Parque de Ciência e Tecnologia Guamá, 66075-750, Belém, Brazil.

* Author for correspondence: Rafael Azevedo Baraúna, Guamá Science and Technology

Park, Laboratory of Biological Engineering, Perimetral ave, 66075-750, Guamá, Belém,

Brazil. Email: rabarauna@ufpa.br

4.1. Abstract

The bacterial genus *Corynebacterium* is composed by pathogenic and non-pathogenic species. The capacity to interact with a host is related to the maintenance or acquisition of virulence genes under different survival conditions. With the advance of bioinformatics and genomics, it became possible to reconstruct phylogenies using genome sequences in order to understand the evolutionary relationship between bacterial species. The main goal of this study was to investigate the molecular clock of *Corynebacterium* genus in order to determine the emergence of pathogenic species such as *Corynebacterium pseudotuberculosis* and others. A comparative genomic approach was also performed to found the specie-specific genes of three species adapted to different habitats: *C. pseudotuberculosis*, *C. glutamicum* and *C. diphtheriae*. According to the Bayesian phylogenetic analysis, the species *C. pseudotuberculosis*, *C. diphtheriae* and *C. ulcerans* grouped in a clade formed only by pathogenic species. It was possible to observe polytomies in two clades, which may be related to the rapid molecular evolution of domain Bacteria. According to the molecular clock, the first species of *Corynebacterium* were free-living bacteria and were adapted to different habitats. Comparative genomic analysis showed that *C. pseudotuberculosis*, *C. diphtheriae* and *C. glutamicum* share 911 orthologous genes. The number of specie-specific genes were 207, 230 and 868 for *C. pseudotuberculosis*, *C. diphtheriae* and *C. glutamicum*, respectively. The analysis of gene ontology for these specie-specific genes showed that shifts in the enzymes involved in oxidation and reduction processes, transmembrane transport, and transcription regulation were important for the evolution of each species. Additionally, it was observed that most specie-specific genes play their role in the cell membrane, implying that survival of these species in a totally different environment requires changes in membrane proteins.

Keywords: *Corynebacterium*, genome evolution, pathogenicity, molecular clock.

4.2. Introduction

The *Corynebacterium* genus belongs to the Actinobacteria phylum. It is composed by several species that are highly relevant to the fields of systems biology, biotechnology, and human and animal medicine. Species of *Corynebacterium* share a similar guanine+cytosine content and some cell wall characteristics such as thickness, the presence of mycolic acids, peptidoglycan, arabinogalactan, and cellular membrane containing saturated and unsaturated fatty acids (Dorella et al., 2006).

Corynebacterium species are adapted to several types of habitat. *C. glutamicum* and *C. efficiens* are free-living non-pathogenic species commonly used in the industrial production of amino acids (Kalinowski et al. 2003; Nishio et al., 2003). *C. variabile* is used for cheese production (Dugat-Bony et al., 2016). *C. urealyticum* and *C. resistens* are opportunistic pathogens commonly associated to infections contracted in health centers (Schröder et al. 2012). *C. pseudotuberculosis*, *C. diphtheriae* and *C. ulcerans* are pathogenic species that infect animals and humans (Trost and Tauch, 2014).

The pathogenicity of some species is probably associated to the maintenance of virulence genes acquired by mobile genetic elements and horizontal gene transfer. Positive selection pressure over the whole genetic content also contribute with the capacity of some species to interact with a host (Resende et al., 2011; Soares et al., 2013; Soares et al., 2013; Viana et al., 2017). Technological advances in genomics and bioinformatics have led to an increase in the number of genome sequences available in public databases. Thus, researches aiming to evaluate the genome evolution processes have become increasingly accessible. Studying the evolution of the genomes of taxonomically related bacteria allows us to understand which genetic determinants are associated to the adaptation of these microorganisms to their habitats.

Holden and colleagues (2013) removed the mobile genetic elements from 193 genomes of *Staphylococcus aureus* and reconstructed the evolutionary history of the species using 8,905 single-nucleotide polymorphisms (SNPs). They found a mutation rate of 1.3×10^{-6} nucleotide substitutions per year ($p < 0,05$). Several other studies have used genome data to determine evolutionary relationships between bacterial species (Lebreton et al., 2017). Stoesser and colleagues (2016) determined the evolutionary history of the extraintestinal pathogen *Escherichia coli* ST131 using 215 genomes. They were able to temporally correlate the clonal expansion of two previously recognized antimicrobial-resistant clades (C1/H30R and C2/H30Rx) with the introduction of fluoroquinolones and extended-spectrum cephalosporins in clinical medicine (Stoesser et al., 2016). *Mycobacterium tuberculosis* is the closest bacterial species of the genus *Corynebacterium* to have its annual mutation rate defined in previous studies (Beyante et al., 2013; Menardo et al., 2019).

In this study, we aimed to evaluate the emergence period of *Corynebacterium* pathogenic species and to explore the genomic plasticity of some species of the genus in order to understand which genes are needed for corynebacteria to evolve to become pathogenic.

4.3. Methods

Data Collection

64 RefSeq genomes from different species that were used for phylogenetic and molecular clock analysis of the genus *Corynebacterium*. A total of 177 genomes of *Corynebacterium* species were recovered from GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), for comparative genomics, three species were further selected: *C. pseudotuberculosis* (74 genomes), *C. diphtheriae* (21 genomes) and

C. glutamicum (22 genomes). These three species were selected because they were closely related in phylogenetic analysis and molecular clock, although they colonize three distinct habitats and hosts.

Bayesian inference of phylogeny and molecular clock analysis

The genes *dnaA*, *dnaK*, *gyrB*, *recA*, *rpoB* and the gene of the 16S rRNA were recovered from the .gbk files downloaded from GenBank. Only the RefSeq genomes of 64 species of *Corynebacterium* were used in this analysis. The species *Mycobacterium tuberculosis* H37Rv (NC_000962.3), *Nocardia* sp. Y18 (CP018082), and *Rhodococcus jostii* RHA1 (CP000431) were used as outgroups. The online platform of MAFFT v.7 (Kato et al., 2017) was used to align the genes using the algorithm FFT-NS-i, then the alignments were filtered and corrected using the program Gblocks v.0.91b (Talavera and Castresana, 2007) with default parameters. The aligned and corrected genes were then concatenated into a phylip format file and submitted to MrBayes v.3.2 (Ronquist et al., 2012). The tree was edited using FigTree v.1.4.2 (Rambaut, 2014).

For molecular clock analysis we used the same six genes described above. Those genes are reported in literature as good phylogenetic markers (Wu et al., 2013). Alignment and gap removal was performed as previously described. A nexus format file containing the aligned genes were analyzed in BEAST v.1.10. (Suchard et al., 2018) using the pattern StarBeast. An uncorrelated lognormal relaxed molecular clock model was used. The population estimate value was set to an average rate of 10^{-10} up to 10^{-5} and a normal rate of 0,1 (Menardo et al., 2019). The molecular clock was calibrated using a mutation rate of 10^{-9} up to 10^{-7} nucleotide changes per year. This mutation rate was found for the phylogenetically-related genus *Mycobacterium* (Cohen et al., 2015; Merker et al., 2015; Lee et al., 2015; Liu et al., 2018; O'Neill et al., 2018).

Orthologs genes-based maximum likelihood-tree

Based on Bayesian analysis three *Corynebacterium* species adapted to different habitats and hosts were selected for phylogenomic and comparative genomic analysis: *C. pseudotuberculosis*, an animal pathogen that infect horses, cattle, buffaloes, camelids and goats; *C. diphtheriae*, a human pathogen; and *C. glutamicum*, a non-pathogenic specie closely-related to the other two pathogenic species cited above. Complete genomes were recovered from GenBank database and the orthologs genes were calculated in PGAP v.1.2.1 (Zhao et al., 2012) using the following parameters: Orthologs genes were used to construct an Unweighted Pair Group Method with Arithmetic Mean (UPGMA) tree using PGAP with coverage 0.5, identity 0.6 and method GF.

Comparative Genomics

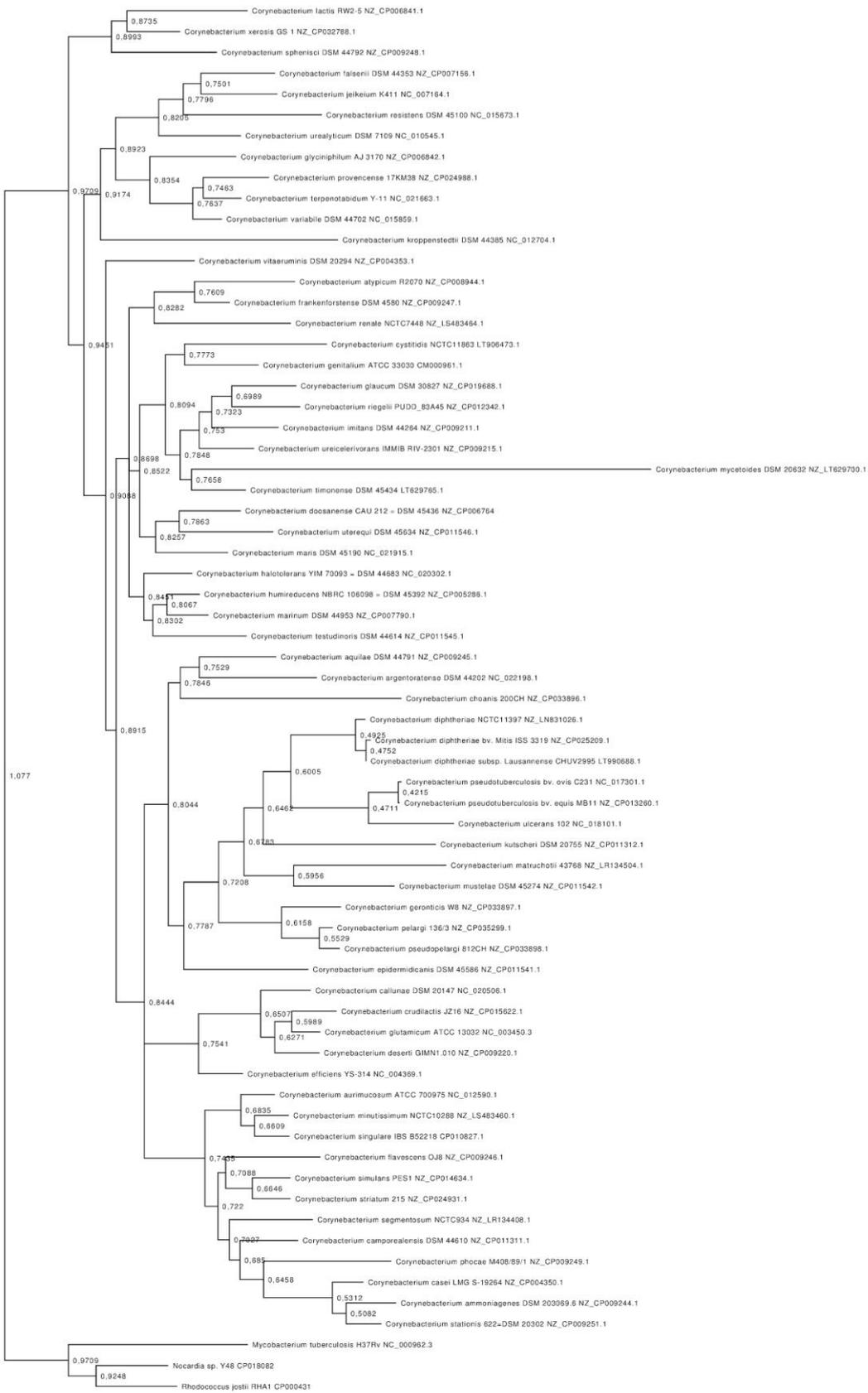
To determine differences among the genomic features of the three selected species the following characteristics were compared using analysis of variance (ANOVA) ($p < 0,05$): genome size, number of genes, gene density. To determine the specie-specific genes we calculated the pan-genome with PGAP using the same parameters as described above. The output file 1.Orthologs_cluster.txt was used to find the specie-specific genes. Subsequently, the multifasta file of these genes were analyzed in the software OmicsBox v.1.1.164 (Götz et al., 2008) in order to determine their GO terms in the categories of cellular compartment, molecular function and biological process. The enrichment analysis of GO terms was performed with fisher's test ($p < 0,05$).

4.4. Results and Discussion

Several studies have performed phylogenetic or phylogenomic analysis to resolve the relationships among species of *Corynebacterium*. Pascual and colleagues (1995) evaluate the phylogenetic relationship among 30 strains of *Corynebacterium* using the sequence of the 16S rRNA. Soares and colleagues (2013) analyzed a set of *Corynebacterium* genomes and determined their relationship based on phylogenomic analysis. However, there are no studies that attempt to predict the molecular clock of the genus. Our initial goal was to analyze the phylogenetic relationship among 64 species of *Corynebacterium* using seven phylogenetic markers and to calculate the molecular clock of the genus.

The pathogenic species *C. pseudotuberculosis*, *C. diphtheriae* and *C. ulcerans* grouped together according to the Bayesian phylogenetic analysis (Fig. 1). Compared to previous studies, the close relationship between these three species was maintained even after increasing the number of samples and genes analyzed (Pascual et al., 1995; Groman et al., 1984). A second cluster of pathogenic species was formed by *C. kutscheri*, *C. matruchotii* and *C. mustelae*, near to the cluster of *C. pseudotuberculosis*. These three species are related frequent parasite of mice and rats, occasional pathogen of humans or primates, and ferret (*Mustela putorius furo*), respectively (Bernard and Funke, 2015).

C. glutamicum forms a non-pathogenic clade which includes *C. crudilactis*, *C. allunae*, *C. deserti* and *C. efficiens*. It is worth noting that *C. glutamicum* do not clustered with other amino acid producing species. *C. glutamicum* was one of the non-pathogenic corinebacteria most closely related to the pathogenic clade of *C. pseudotuberculosis*. *C. mycetoides* showed the highest branch weight among the analyzed species indicating a significant evolutionary distance.



0.2

Figure 1. Rooted phylogenetic tree calculated using the six genes *dnaA*, *dnaK*, *gyrB*, *recA*, *rpoB* and 16S rRNA. Bayesian inference was used and the topology was tested using bootstrap analysis. *Nocardia* sp. sp. Y48 (NZ_CP018082), *Rhodococcus* WB9 (NZ_CP041769) e *Mycobacterium tuberculosis* H37Rv (NC_000962) were used as outgroups. Bootstrap values above 0.5 are shown in tree.

Different from that observed by Pascual and colleagues (1995), *C. vitarurmen* was grouped distant from *C. pseudotuberculosis* clade with a high bootstrap value (94.51%). Polytomies were observed in two clades (Fig. 1). This may be related to the rapid molecular evolution of the domain Bacteria. Despite our efforts we could not resolve this polytomies into dichotomies. Increasing the number of species or the number of phylogenetic markers and adding more outgroups to the analysis are actions that usually help to solve polytomies. In addition to the seven genes used in this study, *rpoS* and *rpoC* have also been commonly used in phylogenetic analysis of *Corynebacterium* (Khamis et al., 2004; Gao & Gupta, 2012).

The molecular clock in figure 2 represents the divergences among species of *Corynebacterium* every 5,000 years in the past. According to Brites and Gagneux (2015), the genus *Mycobacterium* diverged from the other CMNR species about 40,000 years ago. We observed the same result in our analysis (Fig. 2). *Nocardia* and *Rhodococcus* seams to diverge from *Corynebacterium* and *Mycobacterium* about 75,000 years ago (Fig. 2). Interestingly, *C. doosanense* is the oldest species of the genus, emerging about 60,000 years ago (Fig. 2). Thus, probably the first species of *Corynebacterium* were free-living bacteria that later adapted to different animal hosts. The first pathogenic species to emerge were *C. kroppenstedtii* and *C. matruchotii* (Fig. 2). Both are human pathogens (Wong et al., 2017; Barrett et al., 2001).

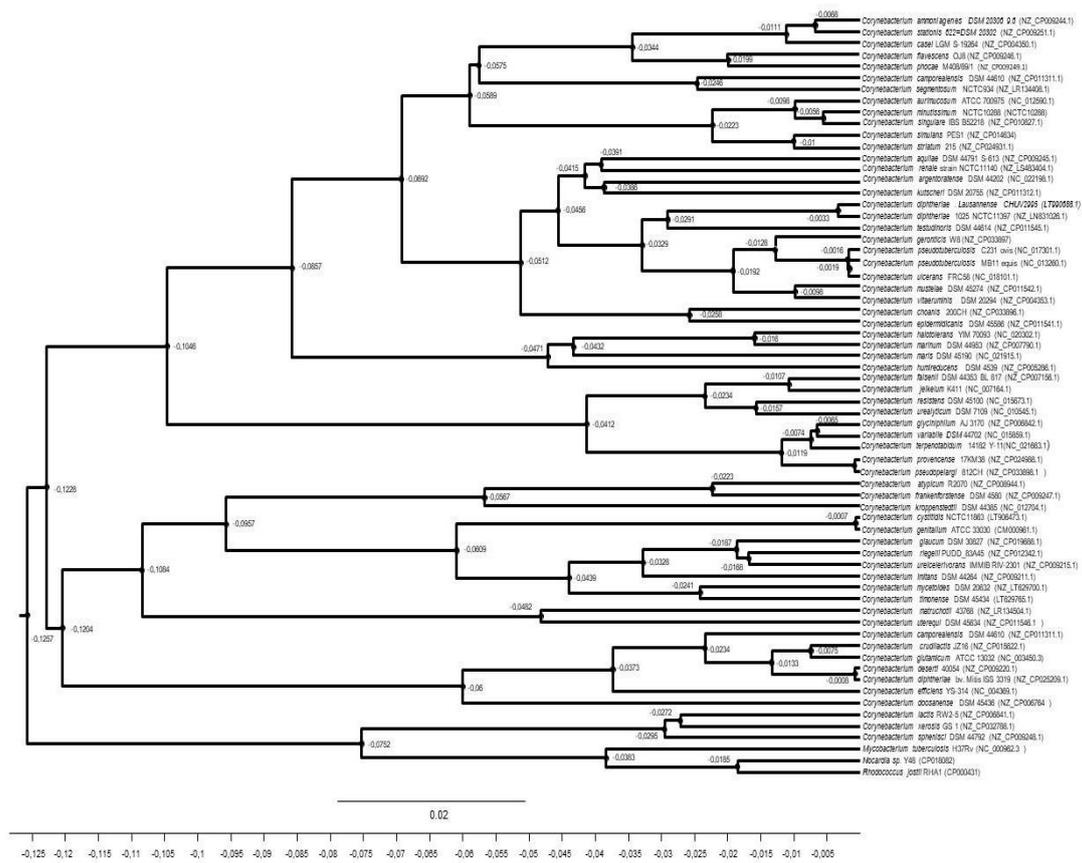


Figure 2. Molecular clock of the *Corynebacterium* genus. *Nocardia* sp. Y48 (NZ_CP018082), *Rhodococcus* WB9 (NZ_CP041769) e *Mycobacterium tuberculosis* H37Rv (NC_000962) were used as outgroups. Ruler values represent time at every 5,000 years in the past.

The majority of corinebacteria that affect animals emerged about 10,000 years ago during the Neolithic period which according to archaeological studies also corresponds to the period of the beginning of the domestication of cattle (*Bos taurus* and *Bos indicus*) (Driscoll et al., 2009). However, some species such as *C. pseudotuberculosis* have emerged more recently (Fig. 2).

To understand the adaptation to different habitats and hosts in terms of genome composition we selected three species for further comparative genomic analysis. *C. pseudotuberculosis* is an animal pathogen, *C. diphtheriae* is a human pathogen

phylogenetically-related to *C. pseudotuberculosis* (Fig. 1), and *C. glutamicum* is a non-pathogenic species. Comparing the genome of these three species allowed us to find some insights about the genomic evolution of pathogenic corinebacteria.

Pathogenic species have significantly smaller genomes ($p < 0,01$) (Table S1). *C. pseudotuberculosis*, *C. diphtheriae* and *C. glutamicum* showed an average genome size of 2.35 Mb, 2.47 Mb and 3.29 Mb, respectively. As expected, the same pattern is observed for the number of genes (Table S1). It is expected that pathogenic species have smaller genomes. These species undergo greater evolutionary pressure because smaller chromosomes replicate faster and therefore evolve faster. However, the reduction of a genome occurs by a slow and gradual process (Arnold et al., 2007).

These three species share 911 orthologs genes corresponding to 9% of all genes described in the three species (10,055 genes in total). *C. pseudotuberculosis*, *C. diphtheriae*, and *C. glutamicum* showed 207, 230 and 868 species-specific genes, respectively. A maximum-likelihood tree based on the 911 orthologous genes was calculated (Fig. 3). *C. pseudotuberculosis* formed two distinct clades according to its biovar (*ovis* or *equi*). In case of *C. diphtheriae*, although the species have three distinct biotypes, the phylogenomic analysis was not able to group these biotypes in distinct clades. In addition, our data were not sufficient to determine the phylogenomic relationship among *C. glutamicum* strains since its clade showed polytomies. Polytomies are common in evolutionary analysis of bacterial species (Koch et al., 2013; Tsang et al., 2017). Since bacteria evolve at very fast rates (not only by single nucleotide polymorphisms, but also by horizontal gene transfer), it is likely that in most phylogenetic analysis there is not enough data for the polytomy to be resolved in dichotomy.

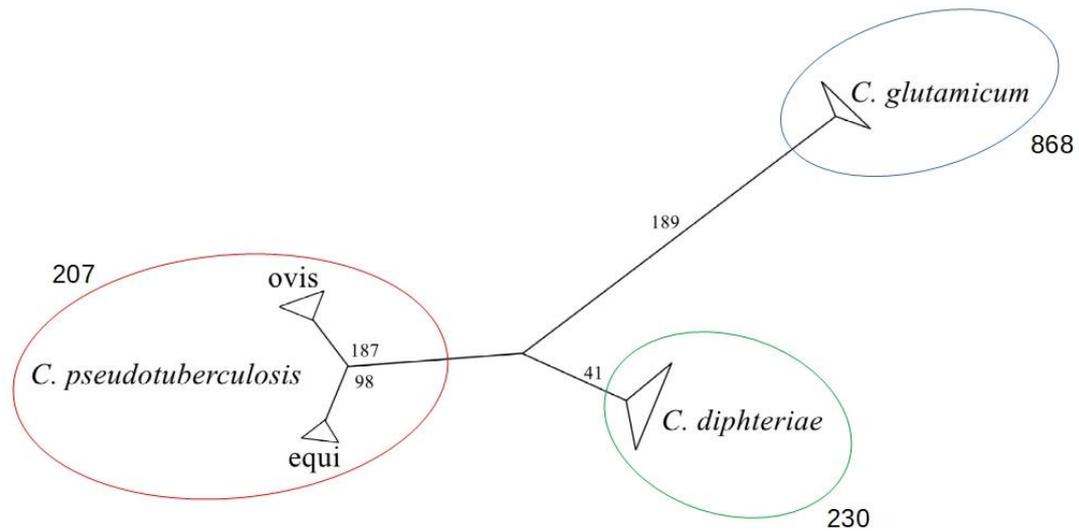


Figure 3. Unrooted phylogenomic tree calculated using the 911 orthologous genes of *C. pseudotuberculosis* (red clade), *C. diphtheriae* (green clade) and *C. glutamicum* (blue clade). Branch weight is shown between each resolved clade. Numbers next to each clade indicate the amount of species-specific genes identified by PGAP. A total of 117 genomes were used, 74 of *C. pseudotuberculosis*, 21 of *C. diphtheriae*, and 22 of *C. glutamicum*. Accession numbers are listed in table S2.

Species-specific genes have a valuable molecular information on how each species possibly adapts to their ecological niches. Thus, we obtained the protein sequence of all species-specific genes and analyzed these sequences using Blast2GO suite from OmicsBox in order to determine the gene ontologies in terms of biological process, molecular function and cellular component (Fig. S1, S2 and S3). Interestingly, for all species the main gene ontology terms (GO terms) were oxidation-reduction process, transmembrane transport, regulation of transcription and translation for the categories of biological process; ATP binding, DNA binding, ATPase activity and catalytic activity for molecular function and; integral component of membrane, membrane, cytoplasm, ribosome and integral component of plasma membrane for

cellular component (Fig. 4). Thus, most of the species-specific genes play their role in the cell membrane therefore the adaptation to different niches (hosts or environments) requires a considerable genotypic change in the membrane proteins.

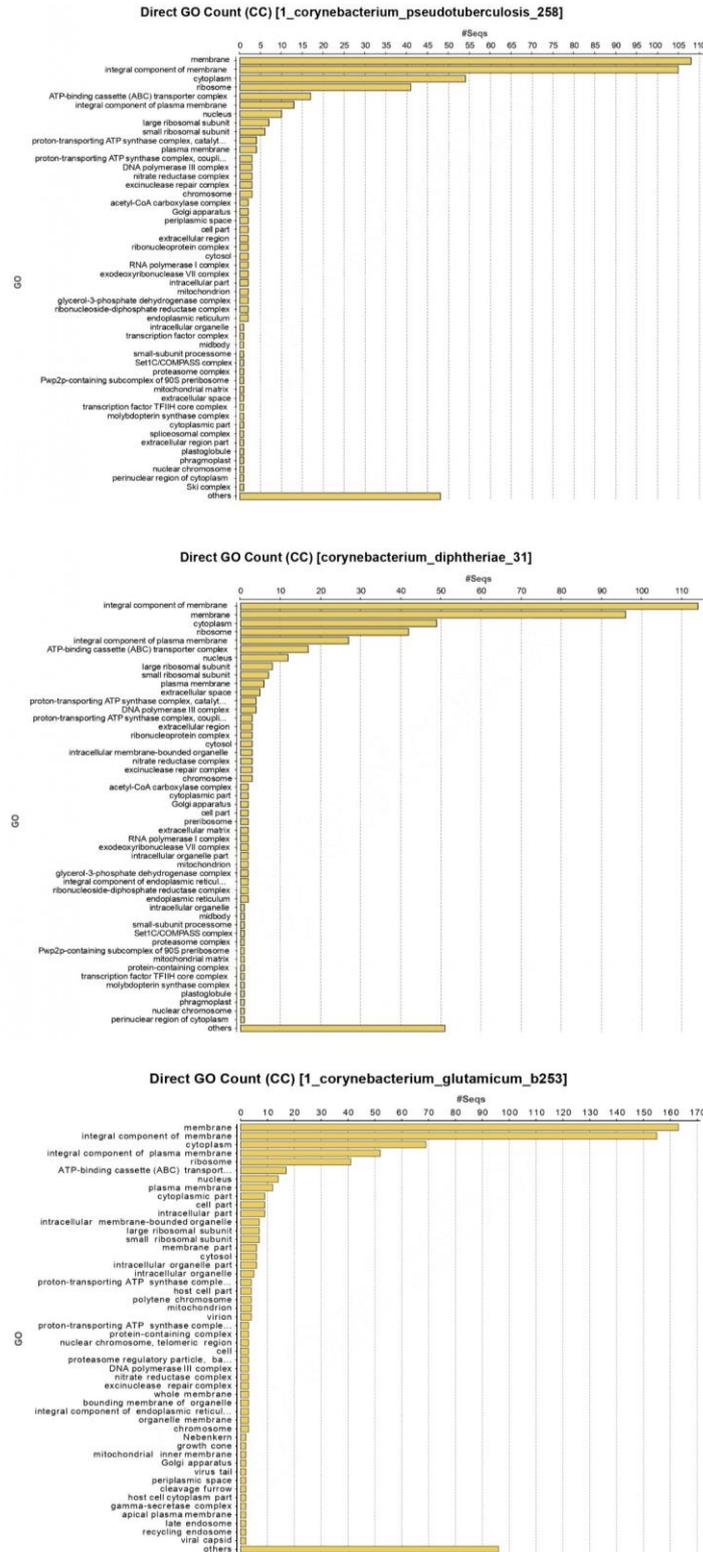


Figure 4. GO terms at level 5 of cellular component category for the specie-specific genes of *C. pseudotuberculosis* 258 (A), *C. diphtheriae* (B) and *C. glutamicum* (C). It is worth noting that for all three species, membrane proteins were the most representative in the dataset.

These membrane proteins are essential to the microbial physiology and serve as a target for drugs and may act as ion channels among other cellular function (Arinaminpathy et al., 2009; Henrich et al., 2015). Modification on the composition of the membrane proteins in the presence of different carbon sources have been previously demonstrated for *C. glutamicum* (Haußmann et al., 2009). This is a common feature of free-living bacteria as they must be able to use several nutrient sources to allow them to grown under the suboptimal conditions of the environment. In the GO enrichment analysis, membrane proteins were overrepresented in the dataset of genes specific of the free-living species *C. glutamicum* ($p < 0,05$) (Fig. 5).

The three bacterial species presented different enriched GO terms (Fig. 5). *C. pseudotuberculosis* seems to have acquired many genes involved in catabolic processes compared to the other two species (Fig. 5). In the dataset of *C. diphtheria* the enriched GO terms were nickel cation binding and photosynthesis-like genes.

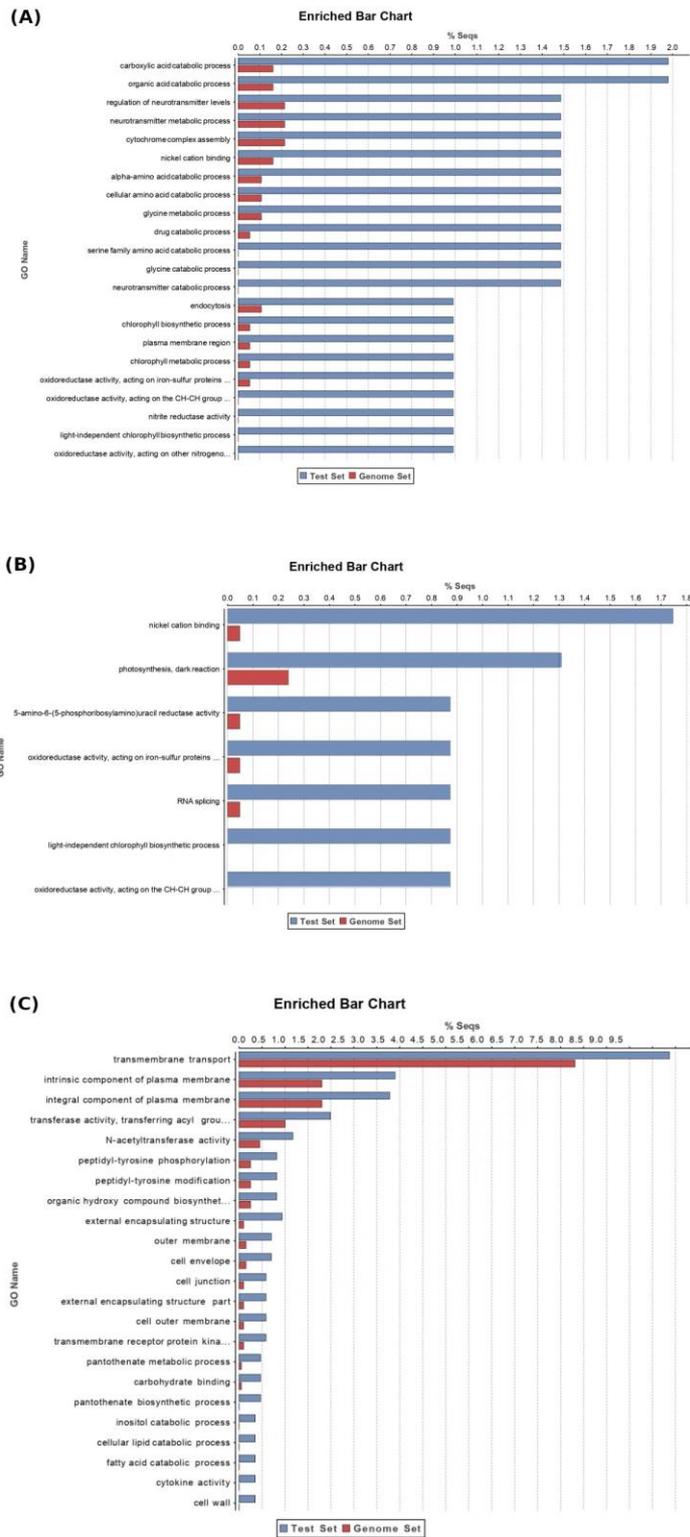


Figure 5. GO enrichment analysis. Bar graphs show the GO terms that were significantly ($p < 0.05$) enriched in the dataset of genes found only in *C. pseudotuberculosis* (A), *C. diphtheriae* (B), and *C. glutamicum* (C), compared to the set

of genes in their genomes. GO terms are ranked starting from the most significantly enriched term.

Acknowledgments

The authors would like to thank to the Brazilian agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the research grant and scholarship support.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Arinaminpathy, Y., Khurana, E., Engelman, D. M., & Gerstein, M. B. (2009). Computational analysis of membrane proteins: the largest class of drug targets. *Drug discovery today*, 14(23-24), 1130-1135. DOI: 10.1016/j.drudis.2009.08.006.
- Arnold, C. (2007). Molecular evolution of *Mycobacterium tuberculosis*. *Clinical microbiology and infection*, 13(2), 120-128. DOI: 10.1111/j.1469-0691.2006.01637.x
- Barrett, S. L. R., Cookson, B. T., Carlson, L. C., Bernard, K. A., & Coyle, M. B. (2001). Diversity within reference strains of *Corynebacterium matruchotii* includes *Corynebacterium durum* and a novel organism. *Journal of clinical microbiology*, 39(3), 943-948. DOI: 10.1128/JCM.39.3.943-948.2001
- Bernard, K. A., & Funke, G. (2015). *Corynebacterium*. *Bergey's Manual of Systematics of Archaea and Bacteria*, 1-70. DOI: 10.1002/9781118960608.gbm00026
- Brites, D., & Gagneux, S. (2015). Co-evolution of *Mycobacterium tuberculosis* and *Homo sapiens*. *Immunological reviews*, 264(1), 6-24. DOI:

<https://doi.org/10.1111/imr.12264>

- Bryant, J. M., Schürch, A. C., Van Deutekom, H., Harris, S. R., De Beer, J. L., De Jager, V., ... & Bentley, S. D. (2013). Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. *BMC infectious diseases*, 13(1), 110. DOI: 10.1186/1471-2334-13-110
- Cohen, K. A., Abeel, T., McGuire, A. M., Desjardins, C. A., Munsamy, V., Shea, T. P., & Chapman, S. B. (2015). Evolution of extensively drug-resistant tuberculosis over four decades: whole genome sequencing and dating analysis of *Mycobacterium tuberculosis* isolates from KwaZulu-Natal. *PLoS medicine*, 12(9), e1001880. DOI: 10.1371/journal.pmed.1001880
- Dorella FA, Carvalho Pacheco L, Oliveira SC, Miyoshi A, Azevedo V. (2006). *Corynebacterium pseudotuberculosis*: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Vet. Res. [Internet]* 37:201–218. Available from: <http://www.edpsciences.org/10.1051/vetres:2005056>. DOI: 10.1051/vetres:2005056
- Driscoll, C. A., Macdonald, D. W., & O'Brien, S. J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences*, 106(Supplement 1), 9971-9978. DOI: 10.1073/pnas.0901586106
- Dugat-Bony, E., Sarthou, AS, Loux, V., Vidal, M., Bonnarme, P., Irlinger, F., e Layec, S. (2016). Draft Genome Sequence of *Corynebacterium variabile* Mu292, Isolated from Munster, a French Smear-Ripened Cheese. *Genome Announc.* , 4 (4), e00669-16. DOI: 10.1128/genomeA.00669-16
- Gao B, Gupta RS. Phylogenetic framework and molecular signatures for the main clades of the phylum Actinobacteria. *Microbiol Mol Biol Rev.* (2012). 76(1):66–

112. doi:10.1128/MMBR.05011-11

Götz S., Garcia-Gomez JM., Terol J., Williams TD., Nagaraj SH., Nueda MJ., Robles M., Talon M., Dopazo J. and Conesa A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research*, 36(10), 3420-35. <https://doi.org/10.1093/nar/gkn176>

Groman, N., Schiller, J., & Russell, J. (1984). *Corynebacterium ulcerans* and *Corynebacterium pseudotuberculosis* responses to DNA probes derived from coryneophage beta and *Corynebacterium diphtheriae*. *Infection and immunity*, 45(2), 511-517.

Haußmann, U., Qi, S. W., Wolters, D., Rögner, M., Liu, S. J., & Poetsch, A. (2009). Physiological adaptation of *Corynebacterium glutamicum* to benzoate as alternative carbon source—a membrane proteome-centric view. *Proteomics*, 9(14), 3635-3651. DOI 10.1002/pmic.200900025

Henrich, E., Hein, C., Dötsch, V., & Bernhard, F. (2015). Membrane protein production in *Escherichia coli* cell-free lysates. *FEBS letters*, 589(15), 1713-1722. DOI: 10.1016/j.febslet.2015.04.045

Holden, M. T., Hsu, L. Y., Kurt, K., Weinert, L. A., Mather, A. E., Harris, S. R., ... & Skov, R. (2013). A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome research*, 23(4), 653-664. DOI: 10.1101/gr.147710.112

Kalinowski J, Bathe B, Bartels D, Bischoff N, Bott M, Burkovski A, Dusch N, Eggeling L, Eikmanns BJ, Gaigalat L, et al. (2003). The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.* 104:5–25. DOI: 10.1016/S0168-1656(03)00154-8

- Katoh, K., Rozewicki, J., & Yamada, K. D. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in bioinformatics*. DOI: 10.1093/bib/bbx108
- Khamis A, Raoult D, La Scola B. Rpo B gene sequencing for identification for *Corynebacterium* species. *J Clin Microbiol.* (2004);42(9):3925-31. DOI: 10.1128/JCM.42.9.3925-3931.2004
- Koch, H., Abrol, D. P., Li, J., Schmid-Hempel, P. (2013). Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. *Molecular Ecology*. 22(7): 2028-2044. DOI: 10.1111/mec.12209
- Lebreton, F., Manson, A. L., Saavedra, J. T., Straub, T. J., Earl, A. M., & Gilmore, M. S. (2017). Tracing the enterococci from Paleozoic origins to the hospital. *Cell*, 169(5), 849-861. DOI: 10.1016/j.cell.2017.04.027
- Lee, M. R., Sheng, W. H., Hung, C. C., Yu, C. J., Lee, L. N., & Hsueh, P. R. (2015). *Mycobacterium abscessus* complex infections in humans. *Emerging infectious diseases*, 21(9), 1638. DOI: 10.3201/2109.141634
- Liu, H., Zhang, Y., Liu, Z., Liu, J., Hauck, Y., Liu, J., ... & Jiang, Y. (2018). Associations between *Mycobacterium tuberculosis* Beijing genotype and drug resistance to four first-line drugs: a survey in China. *Frontiers of medicine*, 12(1), 92-97. DOI: 10.1007/s11684-017-0610-z
- Menardo, F., Duchene, S., Brites, D., & Gagneux, S. (2019). The molecular clock of *Mycobacterium tuberculosis*. *BioRxiv*, 532390. DOI: 10.1101/532390
- Merker, M., Blin, C., Mona, S., Duforet-Frebourg, N., Lecher, S., Willery, E., ... & Allix-Béguec, C. (2015). Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature genetics*, 47(3), 242. DOI: 10.1038/ng.3195

- Nishio Y, Nakamura Y, Kawarabayasi Y, Usuda Y, Kimura E, Sugimoto S, Matsui K, Yamagishi A, Kikuchi H, Ikeo K, et al. (2003). Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. *Genome Res.* 13:1572–1579. Available from: <http://genome.cshlp.org/content/13/7/1572.short>. DOI: 10.1101/gr.1285603
- O'Neill, M. B., Kitchen, A., Zarley, A., Aylward, W., Eldholm, V., & Pepperell, C. S. (2018). Global Phylogeography of *Mycobacterium Tuberculosis* Reveals Role of Recent Human History in Pathogen Dispersal. In *AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY* (Vol. 165, pp. 193-193). 111 RIVER ST, HOBOKEN 07030-5774, NJ USA: WILEY. DOI: 10.1016/S1473-3099(07)70108-1
- Pascual, C.; Lawson, P.A.; Farrow, J.A.; Gimenez, M.N.; & Collins, M.D. (1995) Phylogenetic analysis of the genus *Corynebacterium* based on 16S rRNA gene sequences. *International journal of systematic bacteriology*, v. 45, n. 4, p. 724-8. DOI: 10.1099/00207713-45-4-724
- Pinto, A. C., Ramos, R. T., Silva, W. M., Rocha, F. S., Barbosa, S., Miyoshi, A., ... & Azevedo, V. (2012). The core stimulon of *Corynebacterium pseudotuberculosis* strain 1002 identified using ab initio methodologies. *Integrative Biology*, 4(7), 789-794. DOI: 10.1039/c2ib00092j
- Rambaut, A. (2014). FigTree v1.4.2: A Graphical Viewer of Phylogenetic Trees. Available from <<http://tree.bio.ed.ac.uk/software/figtree/>>
- Resende BC, Rebelato AB, D'Afonseca V, Santos AR, Stutzman T, Azevedo VA, Santos LL, Miyoshi A, Lopes DO. 2011. DNA repair in *Corynebacterium* model. *Gene [Internet]* 482:1–7. <http://dx.doi.org/10.1016/j.gene.2011.03.008>

- Ronquist, F., M. Teslenko, P. van der Mark, DL Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, MA Suchard e JP Huelsenbeck. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, 61(3): 539-542. DOI: 10.1093/sysbio/sys029
- Rui, B., Shen, T., Zhou, H., Liu, J., Chen, J., Pan, X., ... & Shi, Y. (2010). A systematic investigation of *Escherichia coli* central carbon metabolism in response to superoxide stress. *BMC systems biology*, 4(1), 122. DOI: 10.1186/1752-0509-4-122
- Schröder J, Maus I, Meyer K, Wördemann S, Blom J, Jaenicke S, Schneider J, Trost E, Tauch A. (2012). Complete genome sequence, lifestyle, and multi-drug resistance of the human pathogen *Corynebacterium resistens* DSM 45100 isolated from blood samples of a leukemia patient. *BMC Genomics* [Internet] 13:141. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3350403&tool=pmcentrez&rendertype=abstract> DOI: 10.1186/1471-2164-13-141
- Soares SC, Silva A, Trost E, Blom J, Ramos R, Carneiro A, Ali A, Santos AR, Pinto AC, Diniz C, et al. (2013). The Pan-Genome of the Animal Pathogen *Corynebacterium pseudotuberculosis* Reveals Differences in Genome Plasticity between the Biovar *ovis* and *equi* Strains. *PLoS One* 8. DOI: 10.1371/journal.pone.0053818
- Soares SC, Trost E, Ramos RTJ, Carneiro AR, Santos AR, Pinto AC, Barbosa E, Aburjaile F, Ali A, Diniz CAA, et al. (2013). Genome sequence of *Corynebacterium pseudotuberculosis* biovar *equi* strain 258 and prediction of antigenic targets to improve biotechnological vaccine production. *J. Biotechnol.* [Internet] 167:135–141. Available from: <http://dx.doi.org/10.1016/j.jbiotec.2012.11.003>. DOI:

10.1016/j.jbiotec.2012.11.003

- Stoesser, N., Sheppard, A. E., Pankhurst, L., De Maio, N., Moore, C. E., Sebra, R., ... & Kos, V. (2016). Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *MBio*, 7(2), e02162-15. DOI: 10.1128/mBio.02162-15
- Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.*; 4 (1). DOI: 10.1093/ve/vey016
- Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic biology*, 56(4): 564-577. DOI: 10.1080/10635150701472164
- Trost E, Tauch A. (2014). Comparative Genomics and Pathogenicity Islands of *Corynebacterium diphtheriae*, *Corynebacterium ulcerans*, and *Corynebacterium pseudotuberculosis*. In: *Corynebacterium diphtheriae* and Related Toxigenic Species. Dordrecht: Springer Netherlands. p. 39–65. Available from: http://link.springer.com/10.1007/978-94-007-7624-1_3. DOI: 10.1007/978-94-007-7624-1_3
- Tsang, A. K. L., Lee, H. H., Yiu, S. M., Lau, S. K. P., Woo, P. C. Y. (2017). Failure of phylogeny inferred from multicolocus sequence typing to represent bacterial phylogeny. *Scientific Reports*. 7: 4536. DOI: 10.1038/s41598-017-04707-4
- Viana MVC, Figueiredo H, Ramos R, Guimarães LC, Pereira FL, Dorella FA, Selim SAK, Salaheldean M, Silva A, Wattam AR, et al. (2017). Comparative genomic analysis between *Corynebacterium pseudotuberculosis* strains isolated from buffalo. Lin B, editor. *PLoS One* [Internet] 12:e0176347. Available from: <http://dx.plos.org/10.1371/journal.pone.0176347>. DOI: 10.1371/journal.pone.0176347

- Wong, S. C., Poon, R. W., Chen, J. H., Tse, H., Lo, J. Y., Ng, T. K., ... & Luk, W. K. (2017). *Corynebacterium kroppenstedtii* is an emerging cause of mastitis especially in patients with psychiatric illness on antipsychotic medication. In *Open forum infectious diseases* (Vol. 4, No. 2). Oxford University Press. DOI: 10.1093/ofid/ofx096
- Wu, D., Jospin, G., & Eisen, J. A. (2013). Systematic identification of gene families for use as “markers” for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their major subgroups. *PloS one*, 8(10), e77033. DOI: 10.1371/journal.pone.0077033
- Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. 2012. PGAP: pan-genomes analysis pipeline. *Bioinformatics* [Internet] 28:416–418. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btr655>. DOI: 10.1093/bioinformatics/btr655

Supplementary material

Table S1. Genome size, number of genes and genome density for the three species analyzed by comparative genomics. Data was compared using ANOVA ($p < 0,05$).

Different letter in the same column indicates statistically significant differences.

Organism	#	Genome		
	genomes	Genome size	# genes	density
<i>C. pseudotuberculosis</i>	74	2,35±0,03 c	2163,16±31,35 c	919,502±3,49 c
<i>C. diphtheriae</i>	20	2,47±0,03 b	2388,79±44,89 b	964,57±9,41 a
<i>C. glutamicum</i>	22	3,29±0,12 a	3101,23±99,73 a	943,94±8,90 b
p-value		<.0001	<.0001	<.0001
DMS		0,04	39,55	4,67

*tukey distribuição normal

Table S2. List of genomes of *C. pseudotuberculosis*, *C. diphtheriae*, and *C. glutamicum* available in GenBank at March-2019. These genes were downloaded and used to calculate the unrooted phylogenomic tree of figure 3.

Lineage <i>Corynebacterium</i>	Acession number
<i>Corynebacterium diphtheriae</i>	241 NC 016782.1
<i>Corynebacterium diphtheriae</i>	31 NC 016799.1
<i>Corynebacterium diphtheriae</i>	B-D-16-78 NZ CP018331.1
<i>Corynebacterium diphtheriae</i>	BH8 NC 016800.1
<i>Corynebacterium diphtheriae</i>	BQ11 NZ CP029644.1
<i>Corynebacterium diphtheriae</i>	bv mitis str ISS 3319 NZ CP025209.1
<i>Corynebacterium diphtheriae</i>	C7 NC 016801.1
<i>Corynebacterium diphtheriae</i>	CDCE 8392 NC 016785.1
<i>Corynebacterium diphtheriae</i>	CHUV2995 LT990688.1
<i>Corynebacterium diphtheriae</i>	FDAARGOS 197 NZ CP020410.2
<i>Corynebacterium diphtheriae</i>	HC01 NC 016786.1
<i>Corynebacterium diphtheriae</i>	HC02 NC 016802.1
<i>Corynebacterium diphtheriae</i>	HC03 NC 016787.1

<i>Corynebacterium diphtheriae</i>	HC04	NC 016788.1
<i>Corynebacterium diphtheriae</i>	INC402	NC 016783.1
<i>Corynebacterium diphtheriae</i>	NCTC 13129	NC 002935.2
<i>Corynebacterium diphtheriae</i>	NCTC11397	NZ LN831026.1
<i>Corynebacterium diphtheriae</i>	NCTC3529	NZ LR134538.1
<i>Corynebacterium diphtheriae</i>	NCTC7838	NZ LR134537.1
<i>Corynebacterium diphtheriae</i>	PW8	NC 016789.1
<i>Corynebacterium diphtheriae</i>	VA01	NC 016790.1
<i>Corynebacterium glutamicum</i>	AR1	NZ CP007724.1
<i>Corynebacterium glutamicum</i>	ATCC 13032	NC 003450.3
<i>Corynebacterium glutamicum</i>	ATCC 13869	NZ CP016335.1
<i>Corynebacterium glutamicum</i>	ATCC 14067	NZ CP022614.1
<i>Corynebacterium glutamicum</i>	ATCC 21831	NZ CP007722.1
<i>Corynebacterium glutamicum</i>	B253	NZ CP010451.1
<i>Corynebacterium glutamicum</i>	C1	NZ CP017995.1
<i>Corynebacterium glutamicum</i>	CP	NZ CP012194.1
<i>Corynebacterium glutamicum</i>	DN N24	NZ AP017369.1
<i>Corynebacterium glutamicum</i>	DN AJ1511	NZ AP017557.1
<i>Corynebacterium glutamicum</i>	HA	NZ CP025534.1
<i>Corynebacterium glutamicum</i>	K051 ATCC 13032	NC 003450.3
<i>Corynebacterium glutamicum</i>	MB001	NC 003450.3
<i>Corynebacterium glutamicum</i>	R DN	NC 009342.1
<i>Corynebacterium glutamicum</i>	SCgG1	NC 021351.1
<i>Corynebacterium glutamicum</i>	SCgG2	NC 021352.1
<i>Corynebacterium glutamicum</i>	TQ2223	NZ CP020658.1
<i>Corynebacterium glutamicum</i>	USDA-ARS- USMARC-56828	NZ CP013991.1
<i>Corynebacterium glutamicum</i>	WM001	NZ CP022394.1
<i>Corynebacterium glutamicum</i>	XV	NZ CP018175.1
<i>Corynebacterium glutamicum</i>	YI	NZ CP014984.1
<i>Corynebacterium glutamicum</i>	ZL-6	NZ CP004062.1
<i>Corynebacterium pseudotuberculosis</i>	1002	NC 017300.1
<i>Corynebacterium pseudotuberculosis</i>	1002B	NZ CP012837.1
<i>Corynebacterium pseudotuberculosis</i>	1-06-A	NC 017308.1
<i>Corynebacterium pseudotuberculosis</i>	12C	NZ CP011474.1
<i>Corynebacterium pseudotuberculosis</i>	226	NZ CP010889.1
<i>Corynebacterium pseudotuberculosis</i>	258	NC 017945.2
<i>Corynebacterium pseudotuberculosis</i>	262	NZ CP012022.2
<i>Corynebacterium pseudotuberculosis</i>	267	NC 017462.1
<i>Corynebacterium pseudotuberculosis</i>	29156	NZ CP010795.1
<i>Corynebacterium pseudotuberculosis</i>	31	NC 017730.3
<i>Corynebacterium pseudotuberculosis</i>	316	NC 016932.1
<i>Corynebacterium pseudotuberculosis</i>	32	NZ CP015183.1

<i>Corynebacterium pseudotuberculosis</i>	33	NZ CP015184.1
<i>Corynebacterium pseudotuberculosis</i>	34	NZ CP015192.1
<i>Corynebacterium pseudotuberculosis</i>	35	NZ CP015185.1
<i>Corynebacterium pseudotuberculosis</i>	36	NZ CP015186.1
<i>Corynebacterium pseudotuberculosis</i>	38	NZ CP015187.1
<i>Corynebacterium pseudotuberculosis</i>	39	NZ CP015188.1
<i>Corynebacterium pseudotuberculosis</i>	3-99-5	NC 016781.1
<i>Corynebacterium pseudotuberculosis</i>	42-02-A	NC 017306.1
<i>Corynebacterium pseudotuberculosis</i>	43	NZ CP015189.1
<i>Corynebacterium pseudotuberculosis</i>	46	NZ CP015190.1
<i>Corynebacterium pseudotuberculosis</i>	48	NZ CP015191.1
<i>Corynebacterium pseudotuberculosis</i>	48252	NZ CP008922.1
<i>Corynebacterium pseudotuberculosis</i>	ATCC 19410	NZ CP021251.1
<i>Corynebacterium pseudotuberculosis</i>	C231	NC 017301.1
<i>Corynebacterium pseudotuberculosis</i>	Cap1W	NZ CP034411.1
<i>Corynebacterium pseudotuberculosis</i>	CAP3W	NZ CP026500.1
<i>Corynebacterium pseudotuberculosis</i>	CAPJ4	NZ CP026499.1
<i>Corynebacterium pseudotuberculosis</i>	CIP 5297	NC 017307.2
<i>Corynebacterium pseudotuberculosis</i>	Cp13	NZ CP014998.1
<i>Corynebacterium pseudotuberculosis</i>	Cp162 sequence	NC 018019.2
<i>Corynebacterium pseudotuberculosis</i>	CS 10	NZ CP008923.1
<i>Corynebacterium pseudotuberculosis</i>	E19	NZ CP012136.1
<i>Corynebacterium pseudotuberculosis</i>	E55	NZ CP014341.1
<i>Corynebacterium pseudotuberculosis</i>	E56	NZ CP013699.1
<i>Corynebacterium pseudotuberculosis</i>	FRC41	NC 014329.1
<i>Corynebacterium pseudotuberculosis</i>	Ft 2193-67	NZ CP008924.1
<i>Corynebacterium pseudotuberculosis</i>	I19 sequence	NC 017303.2
<i>Corynebacterium pseudotuberculosis</i>	I37	NZ CP017384.1
<i>Corynebacterium pseudotuberculosis</i>	KM01	NZ CP024995.1
<i>Corynebacterium pseudotuberculosis</i>	MB11	NZ CP013260.1
<i>Corynebacterium pseudotuberculosis</i>	MB14	NZ CP013261.1
<i>Corynebacterium pseudotuberculosis</i>	MB154	NZ CP024442.1
<i>Corynebacterium pseudotuberculosis</i>	MB20	NZ CP016829.1
<i>Corynebacterium pseudotuberculosis</i>	MB278	NZ CP023395.1
<i>Corynebacterium pseudotuberculosis</i>	MB295	NZ CP026501.1
<i>Corynebacterium pseudotuberculosis</i>	MB30	NZ CP013262.2
<i>Corynebacterium pseudotuberculosis</i>	MB302	NZ CP021982.1
<i>Corynebacterium pseudotuberculosis</i>	MB66	NZ CP013263.1
<i>Corynebacterium pseudotuberculosis</i>	MEX1	NZ CP017711.1
<i>Corynebacterium pseudotuberculosis</i>	MEX25	NZ CP013697.1
<i>Corynebacterium pseudotuberculosis</i>	MEX29	NZ CP016826.1
<i>Corynebacterium pseudotuberculosis</i>	MEX30	NZ CP017291.1
<i>Corynebacterium pseudotuberculosis</i>	MEX31	NZ CP017292.1

<i>Corynebacterium pseudotuberculosis</i>	MEX9	NZ CP014543.1
<i>Corynebacterium pseudotuberculosis</i>	MIC6	NZ CP019769.1
<i>Corynebacterium pseudotuberculosis</i>	N1	NZ CP013146.1
<i>Corynebacterium pseudotuberculosis</i>	OVI03	NZ CP026524.1
<i>Corynebacterium pseudotuberculosis</i>	OVI2C	NZ CP026374.1
<i>Corynebacterium pseudotuberculosis</i>	OVIAF1	NZ CP034410.1
<i>Corynebacterium pseudotuberculosis</i>	P54B96	NC 017031.1
<i>Corynebacterium pseudotuberculosis</i>	PA01	NZ CP013327.1
<i>Corynebacterium pseudotuberculosis</i>	PA02	NZ CP015309.1
<i>Corynebacterium pseudotuberculosis</i>	PA04	NZ CP019587.1
<i>Corynebacterium pseudotuberculosis</i>	PA07	NZ CP024457.1
<i>Corynebacterium pseudotuberculosis</i>	PA08	NZ CP024602.1
<i>Corynebacterium pseudotuberculosis</i>	PAT10	NC 017305.1
<i>Corynebacterium pseudotuberculosis</i>	phoP	NZ CP019768.1
<i>Corynebacterium pseudotuberculosis</i>	PO222-4-1	NZ CP013698.1
<i>Corynebacterium pseudotuberculosis</i>	PO269-5	NZ CP012695.1
<i>Corynebacterium pseudotuberculosis</i>	SigmaE	NZ CP020356.1
<i>Corynebacterium pseudotuberculosis</i>	T1	NZ CP015100.1
<i>Corynebacterium pseudotuberculosis</i>	VD57	NZ CP009927.1

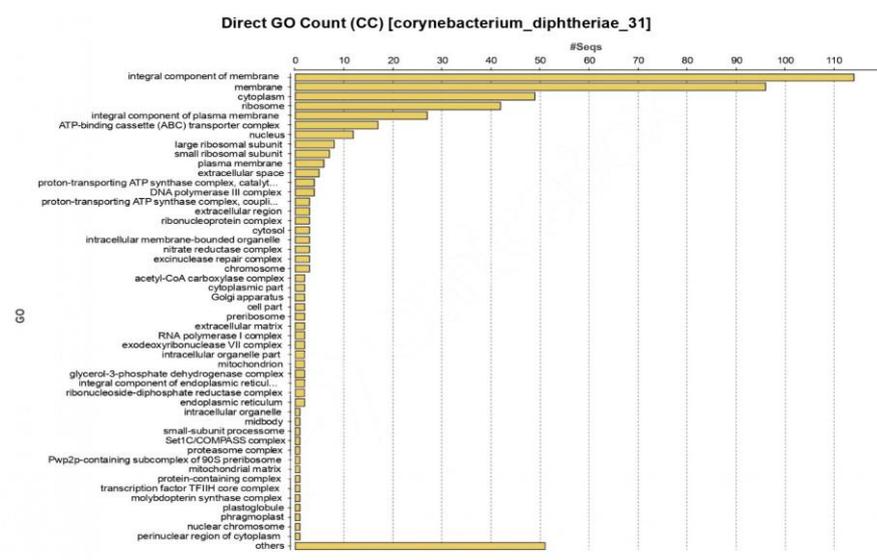
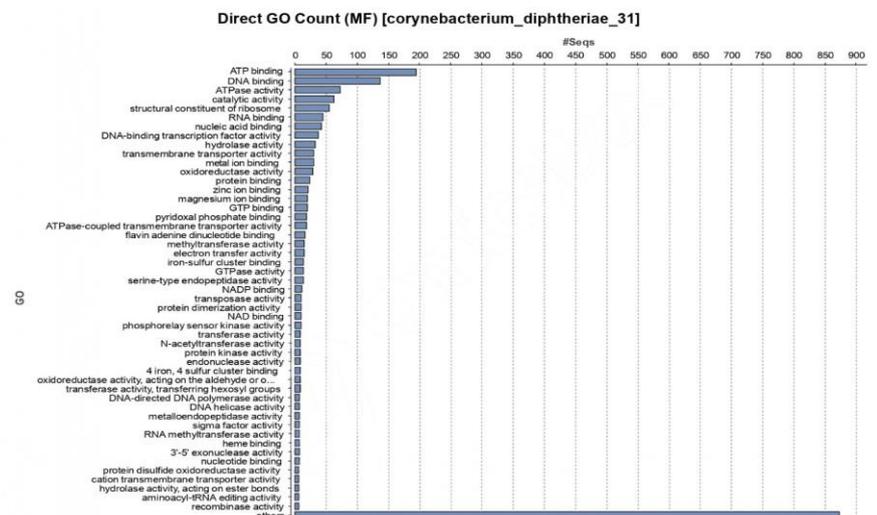
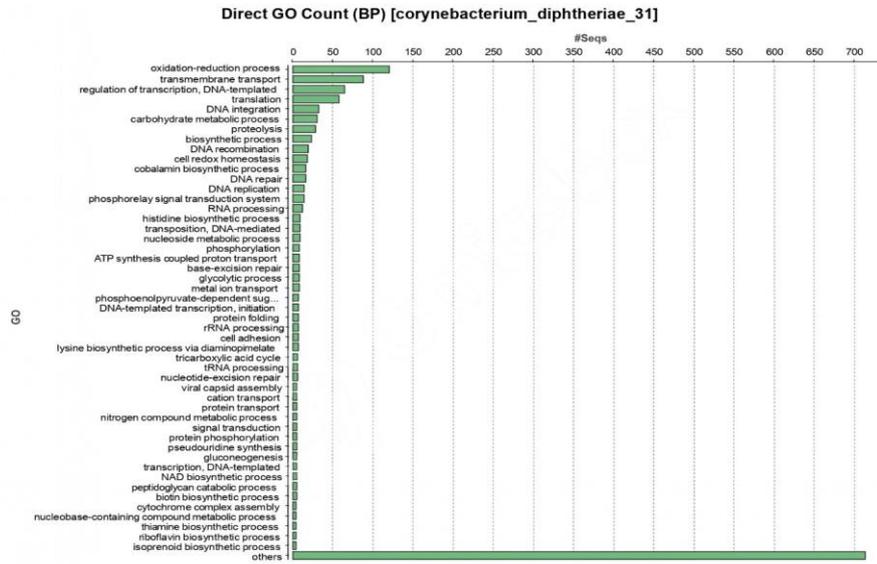


Figure S1. GO terms of genes found only in the genomes of *C. diphtheriae*.

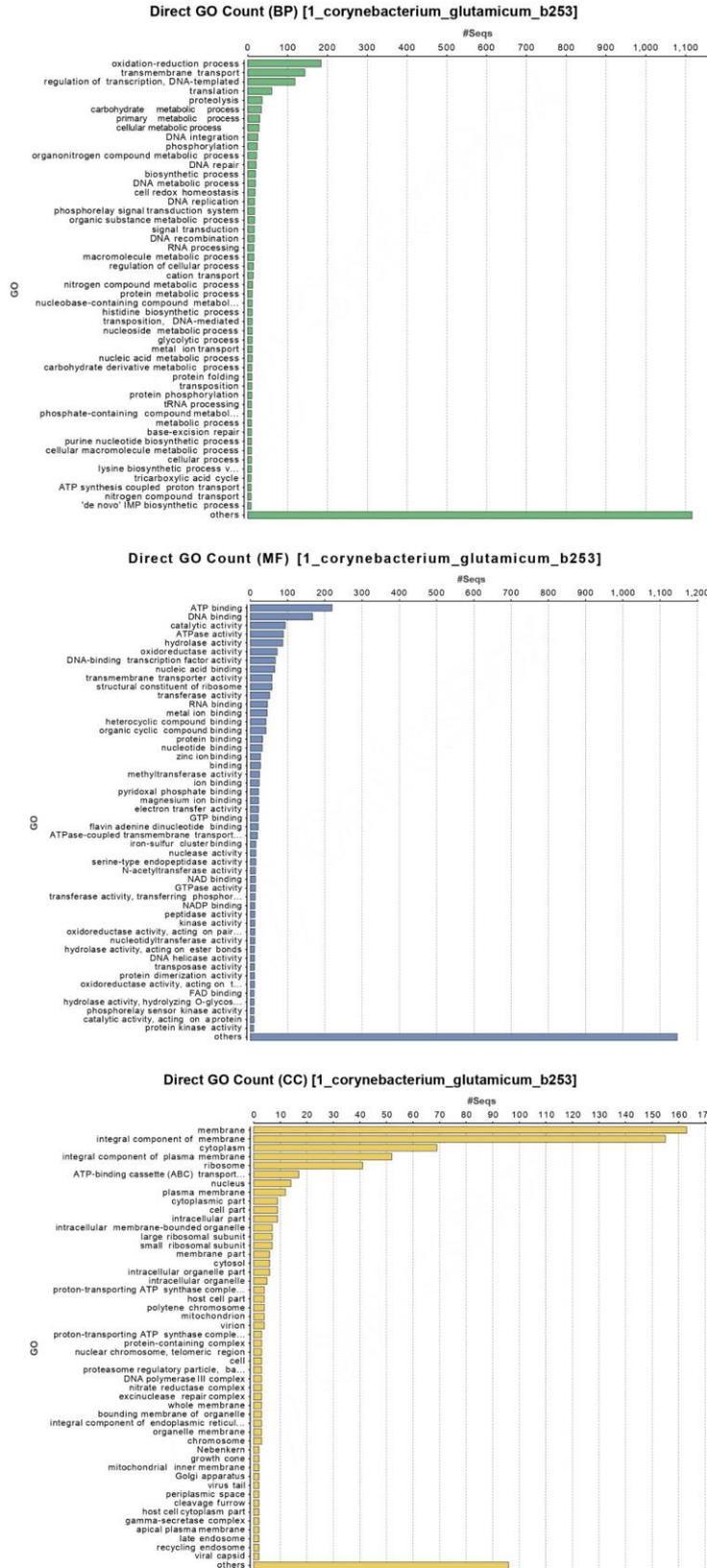


Figure S2. GO terms of genes found only in the genomes of *C. glutamicum*.

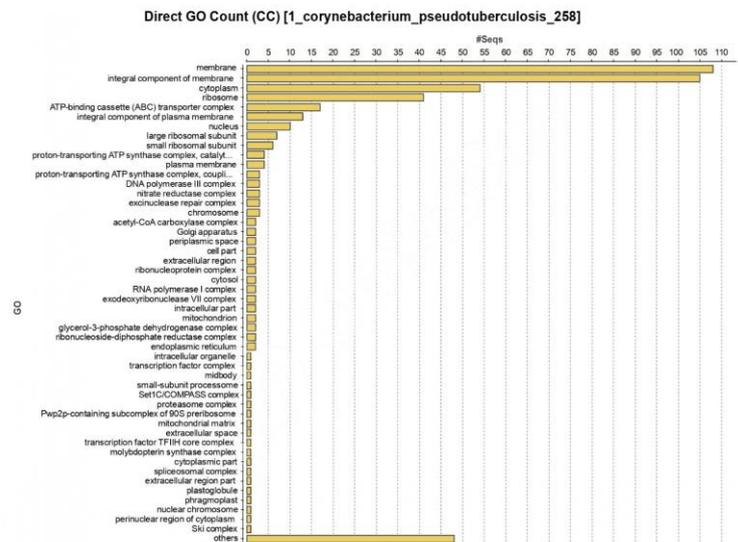
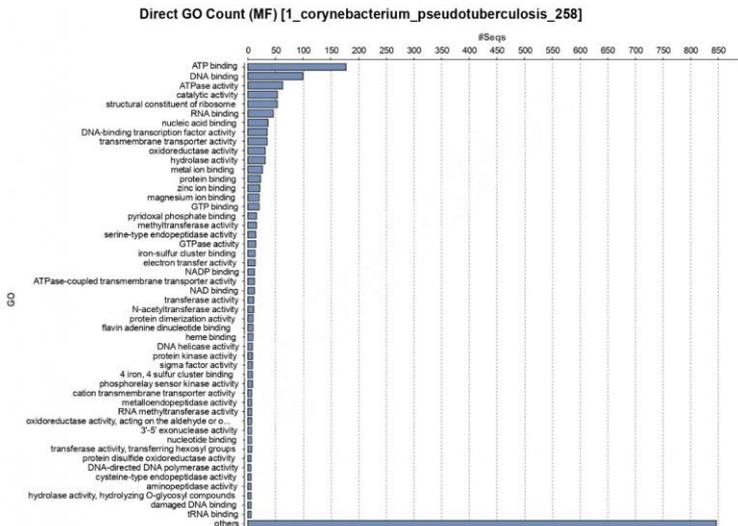
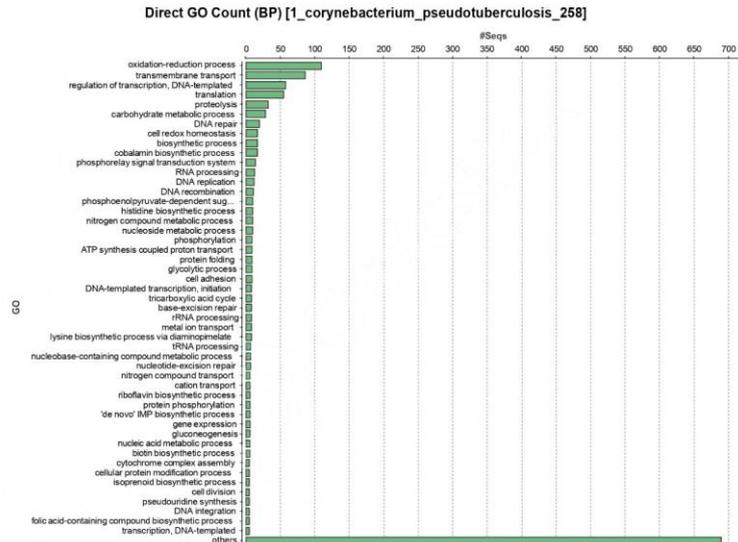


Figure S3. GO terms of genes found only in the genomes of *C. pseudotuberculosis*.

5. DISCUSSÃO GERAL

C. pseudotuberculosis PA05 apresentou elementos estruturais similares aos descrito na literatura (Muñoz-Bucio et al., 2016; Araújo et al., 2017; Marques et al., 2017; Gao et al., 2018) com 237 pseudogenes, 2,183 CDS, presença de *rpoB* *pld*, conteúdo G+C de 52,2%. Apresentou maior proximidade com *Corynebacterium pseudotuberculosis*, do ponto de vista estrutural, a cepa PA06 identificada como do biovar *ovis*, descrito na literatura como altamente clonal, mas com variações que permitem distinção, como identificado por Soares e colaboradores (2013), através de estudo filogenômico, pan-genômico, de core genoma e singleton observaram relações próximas entre as corinebactérias patogênicas, como comportamento clonal de *C. pseudotuberculosis* e com destaque de maior comportamento clonal em *C. pseudotuberculosis* biovar *ovis*.

O gênero *Corynebacterium* conta atualmente com 110 espécies válidas, sendo 132 espécies contando com espécies sinônimas e 11 subespécies (<http://www.bacterio.net/corynebacterium.html>). Este gênero está distribuído de maneira onipresente na natureza, sendo seus membros isolados dos mais variados ambientes (Alibi et al., 2017), com as espécies mais importantes do ponto de vista clínico: *C. diphtheriae*, *C. ulcerans* e *C. pseudotuberculosis*, capazes de produzir a toxina diftérica (TD) e uréase positiva patogênicas em humanos e/ou animais. Já do ponto de vista biotecnológico, a espécie *C. glutamicum* se destaca dentro do gênero (Hoffmann et al., 2018).

Para entendimento destas variações do gênero *Corynebacterium*, ferramentas filogenéticas e de evolução possibilitam maiores esclarecimentos, como observado em Liu et al., (2018), que ao estudar *Escherichia coli* ST131-H22, utilizando relógio

molecular, pode observar a aquisição de ColV plasmids por seis vezes de 1940 aos dias atuais, o que conferiu ao microrganismo maior resistência a humanos e aves.

O conceito de evolução molecular é relativamente recente e ganhou mais espaço com a maior disponibilidade de informações de sequência de DNA e de proteína e avanços na utilização da bioinformática. Modificações ocorridas em relação a um ancestral, relacionadas a mudanças evolutivas, levam a modificações nos caracteres herdados em uma população (Madigan et al., 2016), obtendo como resultado final da evolução a formação de novas espécies, resultando em maior diversidade como observado em *Corynebacterium* spp.

A partir de análise filogenética, observamos que *C. pseudotuberculosis* apresentou maior proximidade com *C. diphtheriae*, patogênica e com *C. glutamicum*, não patogênica. As corinebactérias são quimiorganotróficas, associados a diversos ambientes como solo (Mimitsuka et al. 2007), água (Wei et al. 2018), organismos animais (Lima et al. 2017) e humano (Hoskisson 2018). *C. pseudotuberculosis* foi isolada pela primeira vez em uma vaca com quadro de linfangite, na França, por Edmond Isidore Étienne Nocard (1888). *C. diphtheriae*, foi descrita por Lehman e Neumann em 1896 (Collins and Cummins, 1986) e *C. glutamicum* identificada inicialmente em 1950 em solo japonês na busca por um produtor natural de glutamato por pesquisadores da empresa Kyowa Hakko. Como esperando para uma espécie não patogênica de vida livre (Haußmann et. al., 2009), a *C. glutamicum* apresentou maior genoma e maior quantidade de genes. Entretanto foi observada maior densidade gênica em *C. diphtheriae*.

Os valores de Bootstrap variaram na árvore, os nós considerados mais antigos possuem um grau de apoio maior e esperado devido à grande complexidade do gênero, enquanto nós mais recentes possuem um grau menor devido o número não tão grande

de exemplares da mesma espécie. A Árvore demonstrou a distância filogenética entre diferentes espécies do gênero *Corynebacterium*, ficando clara a distância da *Corynebacterium mycetoides* para todo o gênero.

As espécies *C. diphtheria*, *C. ulcerans* e *C. pseudotuberculosis* possivelmente possuem um ancestral comum, que possui distância relativa do resto do gênero, e foram agrupadas próximas a *C. kutscheri* o que pode possuir correlação com o fato das quatro espécies estarem relacionadas a infecções em mamíferos (Hall I.C. e Fisher, 1915; Gubler et al., 1990; Hiromi et al., 2002; Dias et al., 2011) resultados que vão de acordo com os resultados propostos por Pascual et al., (1995) e reforçam a relação entre *C. ulcerans* e *C. pseudotuberculosis*, proposta por Groman, Schiller, e Russell (1984) . Esta relação não pode ser tão bem resolvida na árvore devido a limitações computacionais, seria possível avaliar essa distinção se mais linhagens fossem adicionadas no pool amostral.

Também foi observada uma relação entre espécies não patogênicas e de interesse biotecnológico, como *C. Efficiens* e *C. Glutamicum*, apesar de não possuírem um nó direto os separando, foram agrupadas em um único ramo, junto com *C. callunae*, e *C. deserti* que já haviam sido descritas como semelhantes por Zimmerman et al., (2016) as quais também não são patogênicas.

Destoando do resultado de Pascual et al. (1995), a *C. vitarurmen* foi agrupada em um ramo distante de *C. pseudotuberculosis*, e não foi agrupada com outras espécies, o valor de bootstrap suporta fortemente esse distanciamento (94,51%) entre a espécie e outros agrupamentos.

Existe uma politomia entre dois agrupamentos, através da metodologia proposta, não foi possível resolvê-lo com a capacidade computacional disponível. Pois seria necessário adicionar mais grupos de referência, como foi realizado no trabalho de

PASCUAL et al. (1995). Utilizamos o outgroup com apenas três espécies de outros gêneros, e seria adequado aumentar este número, pois estes dados auxiliam na melhor resolução da árvore.

Os estudos relacionados à filogenia do gênero *Corynebacterium* começaram a ser feitos a 24 anos atrás por Pascual (1995), onde foi utilizado uma quantidade limitada de linhagens existentes no momento e a árvore filogenética foi feita utilizando-se as sequências 16S rRNA. Posteriormente a este estudo, análises mais recentes para o gênero, foram feitos por Nouioui I (2018), utilizando apenas o gene 16s e o conteúdo G+C. Contudo, a quantidade de marcadores utilizados interfere diretamente nas análises filogenéticas (Wu D et al., 2013). Diversos outros trabalhos foram propostos para se utilizar mais marcadores filogenéticos para a obtenção de uma árvore bem resolvida. Em *Corynebacterium* já foram utilizados também os genes 16s, rpoS, rpoB, RpoC e gyrA (Khamis et al., 2004 e 2005; Gao e Gupta, 2012).

A utilização de relógio molecular com intuito de estabelecer uma possível árvore filogenética para o gênero *Corynebacterium* se torna inovador e necessário para melhor entendimento do período temporal de surgimento da patogenicidade, através da inferência bayesiana com aproximação, a posteriori utilizando Markov chain Monte Carlo (MCMC), que se mostrou mais adequado para calcular a probabilidade das árvores (Metropolis et al., 1953; Hastings, 1970)

Não existem estudos com a tentativa para prever o relógio molecular da espécie *Corynebacterium*, desta forma, para obter o relógio molecular com uma resolução confiável utilizamos seis genes (16s, dnaA, dnaK, rpoB, recA e gyrA), além de adaptar estudos feitos para a *Mycobacterium*, que pertencente ao mesmo grupo que a *Corynebacterium*, o grupo CMNR, que também engloba os gêneros *Nocardia* e *Rhodococcus*. A *Mycobacterium* entre este grupo é aquela que mais possui trabalhos

envolvendo a tentativa de traçar os tempos de divergência do gênero, por isto, neste trabalho adaptamos a metodologia utilizada neste organismo (Fabrizio et al 2019; Menardo et al., 2019).

Os resultados do relógio molecular para *Corynebacterium* apresentam dados de divergência a cada cinco mil anos no passado. Os dados já presentes na literatura para o tempo de divergência de *Mycobacterium* corroboram com os obtidos neste trabalho. Observamos neste relógio molecular, o tempo onde à linhagem de *Mycobacterium* surgiu e sua possível divergência para outras linhagens. Estudos apontam que este tempo de divergência a outros gêneros ocorreu há aproximadamente 40 mil anos atrás (Brites e Gagneux, 2015). Portanto, o que ocorreu possivelmente também com a *Nocardia* e *Rhodococcus*, que observamos estarem surgindo e se divergindo da *Mycobacterium* nesse mesmo período.

Observamos que o ramo correspondente ao ancestral comum da *M. tuberculosis* surgiu há 75 mil anos atrás. Corroborando com dados que mostram que os ancestrais do grupo CMNR emergiram aproximadamente 70 mil anos atrás (Gagneux & Small, 2007; Comas et al., 2013). A origem evolutiva das Actinobacterias é antiga, que remonta a aproximadamente 2,7 milhões de anos atrás, antes da oxigenação da atmosfera (~ 2,3 milhões de anos atrás). O sequenciamento do gene 16S rRNA, a filogenética e a filogenômica derivam de um ancestral em comum compartilhado para Actinobacteria, Cyanobacteria e Deinococcus que remete a mais de 2 milhões de anos atrás (Battistuzzi et al., 2004).

Para *Corynebacterium* levantamos a hipótese de que as primeiras espécies do gênero não eram patógenas, e sim microrganismos de vida livre, pois a espécie *C. doosanense* é relatada como uma bactéria de vida livre (Lee H. J., et al 2009) e foi a

espécie mais basal dentro da nossa análise (aproximadamente 60 mil anos atrás), levando a esta espécie ser a mais próxima do ancestral comum de todo o gênero.

A *C. humireducens* apresentou surgimento há aproximadamente 45 mil anos, e não é uma bactéria patogênica (Wu et al., 2011). Reiterando a possibilidade dos primeiros ancestrais da *Corynebacterium* tenham sido organismos de vida livre.

As primeiras espécies patogênicas de *Corynebacterium* estão associadas a infecção primeiramente aos seres humano, onde observamos a *C. kroppenstedtii* e a *C. matruchotii* sendo as espécies mais basais após os organismos de vida livre (Wong SCY et al 2017, Barrett SL et al 2001).

Estudos arqueológicos identificaram um fócil humano, de aproximadamente 24 000 anos de idade, uma porção substancial de microrganismos associados ao corpo humano, pertencendo predominantemente ao *Clostridia* (*Eubacterium*, *Pseudoramibacter*), Actinobacteria (*Propionibacterium*, *Corynebacterium*, *Actinomyces*) e Bacteroidia (*Tannerella*) (Ventura et al, 2007). Baseando em estudos arqueológicos, o fato do surgimento das linhagens de *Corynebacterium* patogênicas ao humano podem estar relacionados aos primeiros eventos de migração humana conhecida (Broschet al., 2002) que co-migraram dos humanos modernos e quando houve, se espalhou pela África e Eurásia (Gagneux & Small, 2007).

As espécies de *C. pseudotuberculosis* e a *C. ulcerans*, ambas patogênicas a humanos e animais, tiveram seu surgimento recentemente, entre os últimos mil anos.

Grande parte das linhagens de *Corynebacterium* que afetam animais surgiram aproximadamente a 10 mil anos atrás, no período neolítico, que segundo estudos arqueológicos também corresponde a data do início da domesticação do gado (*Bos taurus* e *Bos indicus*) (Driscoll CA et al 2009).

Após a obtenção das sequências dos genomas das 117 *Corynebacterium* spp., pertencentes as três espécies selecionadas, observamos tamanhos de genoma de *Corynebacterium* com um alcance ligeiramente amplo, tendo genomas em tamanho médio de 2,55 Mb, com o menor sendo observado em *C. pseudotuberculosis* com 2,35 Mb ($p < 0,01$), seguido de *C. diphtheriae* 2,47 ($p < 0,01$) e o maior sendo de *C. glutamicum* com 3,29 Mb ($p < 0,01$). O mesmo observado para o número de genes de 2163,16, 2388,79 e 3101,23, para *C. pseudotuberculosis*, *C. diphtheriae* e *C. glutamicum*, respectivamente. Como esperando para uma espécie não patogênica de vida livre, a *C. glutamicum* apresentou maior genoma e maior quantidade de genes. Entretanto foi observada maior densidade gênica em *C. diphtheriae*.

Do conjunto de genes do pangenoma, mostrou-se interessante caracterizar os genes espécie-específicos de *C. pseudotuberculosis*, *C. diphtheriae* e *C. glutamicum*, e através do Blast2GO obteve-se a funcionalidade destes genes. Identifica a ontologias gênicas relacionadas ao “processo biológico”, relacionada à atividade biológica com o qual o gene contribui; “função molecular”, associada à atividade bioquímica do gene; e o “componente celular”, que trata do compartimento celular onde o produto gênico desempenha sua função (Conesa et al., 2005).

Destaca-se também, no processo biológico, o reparo de DNA nos grupos de gene espécie específicos de todas as espécies. Este é responsável por manter a integridade do DNA quando exposto a várias condições de estresse, explicando a necessidade deste para estas espécies (Gorna et al., 2010). Dentre as funções metabólicas, destaca-se as atividades de endonuclease, bactérias patogênicas quando em ação infecciosa, atua sob estresse oxidativo, resultando na produção de lesões de DNA. A maioria dessas lesões é reparada pela via de reparo por excisão de base. As endonucleases desempenham um

papel importante no reparo do DNA danificado, compreendendo os locais abásicos (Puri et al., 2013).

Puri e colaboradores (2013) caracterizaram pela primeira vez duas endonucleases AP, a Endonuclease IV (End) e Exonuclease III (XthA), que desempenham funções distintas na *M. tuberculosis*, com a atividade de endonuclease AP de End e XthA estimulada por Mg²⁺ e Ca²⁺, sendo a endonuclease End mais eficiente do que XthA, desempenhando um papel crucial na defesa contra o estresse oxidativo.

Foi possível observar que grande parte dos genes desempenham sua função na membrana celular e componente integral da membrana, implicando que a sobrevivência destas espécies em um ambiente totalmente diferente, potencialmente propensas ao estresse tanto ambiental como de adaptação ao hospedeiro (Haußmann et. al., 2009), requerem alterações nas proteínas de membrana, sendo este um dos principais pontos de diferenciação destas espécies, patogênicas e não patogênicas.

Além disso, os genes espécie específico de *C. pseudotuberculosis* estão relacionados a várias categorias de processos biológicos, o que sugere esses genes como essenciais e alvos candidatos para tratamento, fator favorável ao desenvolvimento de novas drogas contra este microrganismo patogênico (Folador et al. 2016).

Dentre as funções moleculares destacam-se as atividades de hidrolase, ligação de ATP, atividade de transferase, ligação de DNA, atividade de oxi-redução.

Já para *C. glutamicum*, os resultados dos processos biológicos a nível 5 segue a ordem apresentada em *C. pseudotuberculosis*, com destaque para processo de oxi-redução, seguido por transporte de transmembrana e de regulação e transcrição do DNA, e também apresenta boa parte dos genes envolvido em processos metabólicos.

Dentre as funções moleculares destacam-se as ligações de ATP, de DNA e de

íon metal, atividades de hidrolase, atividade de transferase, atividade de fator de transcrição de ligação de DNA, atividade de oxi-redução (Figura 4B). Sendo que mais de 95% (320 genes) do componente celular é composto de componentes integrais da membrana, seguido de membrana de plasma (Figura 4C).

Destacam-se também, no processo biológico, o reparo de DNA nos grupos de gene espécie específicos de *C. pseudotuberculosis* e *C. glutamicum*. Este é responsável por manter a integridade do DNA quando exposto a várias condições de estresse, explicando a necessidade deste para estas espécies. Dentre as funções metabólicas as atividades de endonuclease, atuante no reparo de excisão de base, podem prevenir a perda da funcionalidade do DNA por ocorrências de lesão mutagênica ou por radiações.

Foi possível observar que grande parte dos genes, nas duas espécies, desempenha sua função na membrana celular, componente integral da membrana, o que implica que a sobrevivência destas espécies em um ambiente totalmente diferente, potencialmente propensas ao estresse tanto ambiental como de adaptação ao hospedeiro, requer alterações nas proteínas de membrana, tornado este um dos principais pontos de diferenciação destas espécies, patogênicas e não patogênicas.

6. CONCLUSÃO GERAL

- Foi possível isolar, identificar, montar e anotar cepa obtidas via material caseoso de ovino no estado do Pará, disponibilizando a comunidade científica mais um dado genômico da espécie *C. pseudotuberculosis*.
- Através da investigação da inferência bayesiana de filogenia e utilização do relógio molecular para estimar o período temporal de surgimento da patogenicidade, pode-se induzir que as primeiras cepas do gênero *Corynebacterium* eram de vida livre, que a patogenicidade surgiu a aproximadamente 10 mil anos atrás, mesmo período de domesticação de animais de interesse zootécnico. E para surgimento de *C. pseudotuberculosis* como patógeno, a estimativa é de cinco mil anos.
- Na análise estatística dos elementos estruturais, tamanho do genoma, número de genes e densidade gênica de *Corynebacterium pseudotuberculosis*, *Corynebacterium diphtheriae* e *Corynebacterium glutamicum*, observou-se diferença estatística significativa entre as espécies, sendo *C. pseudotuberculosis* com menores tamanhos de genoma e quantidade de genes, e *C. glutamicum*, com maiores. Confirmando que organismos patogênicos utilizam mecanismo de redução de genoma para otimização da infecção e por apresentar menor necessidade de adaptação a ambientes adversos.
- Na identificação de ontologias gênicas e o enriquecimento, as bactérias patogênicas mudaram principalmente genes relacionados a metabolismo celular, enquanto a não patogênica apresentou maior relevância em funções de membrana celular, processos essenciais para manutenção e sobrevivência da célula.

7. REFERÊNCIA BIBLIOGRÁFICA

ALIBI, S.; FERJANI, A.; BOUKADIDA, J.; CANO, M. E.; FERNÁNDEZ-MARTÍNEZ, M.; MARTÍNEZ-MARTÍNEZ, L.; E NAVAS, J. (2017). Occurrence of *Corynebacterium striatum* as an emerging antibiotic-resistant nosocomial pathogen in a Tunisian hospital. **Scientific reports**, 7(1), 9704.

AMAO, H.; AKIMOTO, T.; KOMUKAI, Y.; SAWADA, T.; SAITO, M.; & TAKAHASHI, K. W. (2002). Detection of *Corynebacterium kutscheri* from the oral cavity of rats. **Experimental animals**, 51(1), 99-102.

ARAÚJO, F. A., MARQUES, J. M., DE MOURA, V. A. G., SCHNEIDER, M. P. C., ANDRADE, S. S., LIMA, A. C. S., ... & RAMOS, R. T. (2017). Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA07 Biovar ovis, Isolated from a Sheep Udder in Amazonia. **Genome Announc.**, 5(12), e00040-17.

BARRETT, S. L. R., COOKSON, B. T., CARLSON, L. C., BERNARD, K. A., & COYLE, M. B. (2001). Diversity within reference strains of *Corynebacterium matruchotii* includes *Corynebacterium durum* and a novel organism. **Journal of clinical microbiology**, 39(3), 943-948.

BATTISTUZZI, F. U., FEIJAO, A. & HEDGES, S. B. (2004). A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. **BMC evolutionary biology**. 4: 44.

BRITES, D., & GAGNEUX, S. (2015). Co-evolution of *Mycobacterium tuberculosis* and *Homo sapiens*. **Immunological reviews**, 264(1), 6-24.

COLLINS, M.D., & CUMMINS, C.S. (1986). Genus *Corynebacterium* Lehmann and Neumann 1896, 350A1. In: *Bergey's Manual of Systematic Bacteriology*, Vol 2 (Sneath PHA, NS Mair, ME Sharpe and JG Holt, eds), pp 1266–1276, Williams and Wilkins, Baltimore, MD.

CONESA, A., GÖTZ, S., GARCÍA-GÓMEZ, J. M., TEROL, J., TALÓN, M., AND ROBLES, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, 21, 3674–3676.

DIAS, A. A. D. S. D., SANTOS, L. S., SABBADINI, P. S., SANTOS, C. S., JUNIOR, S., CORREA, F., ... & GUARALDI, A. L. M. (2011). *Corynebacterium ulcerans* diphtheria: an emerging zoonosis in Brazil and worldwide. **Revista de saúde pública**, 45(6), 1176-1191.

DRISCOLL, C. A., MACDONALD, D. W., & O'BRIEN, S. J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. **Proceedings of the National Academy of Sciences**, 106(Supplement 1), 9971-9978.

FOLADOR EL, DE CARVALHO PVSD, SILVA WM, FERREIRA RS, SILVA A, GROMIHA M, GHOSH P, BARH D, AZEVEDO V, RÖTTGER R. (2016). In silico identification of essential proteins in *Corynebacterium pseudotuberculosis* based on protein-protein interaction networks. *BMC Syst. Biol.* [Internet] 10:103.

GAO B, GUPTA RS. (2012). Phylogenetic framework and molecular signatures for the main clades of the phylum Actinobacteria. **Microbiol Mol Biol Rev.**, 76(1):66–112.

GAO, H., MA, Y., SHAO, Q., HONG, Q., ZHENG, G., & LI, Z. (2018). Genome Sequence of *Corynebacterium pseudotuberculosis* Strain KM01, Isolated from the Abscess of a Goat in Kunming, China. **Genome Announc.**, 6(11), e00013-18.

GROMAN, N., SCHILLER, J., & RUSSELL, J. (1984). *Corynebacterium ulcerans* and *Corynebacterium pseudotuberculosis* responses to DNA probes derived from coryneophage beta and *Corynebacterium diphtheriae*. **Infection and immunity**, 45(2), 511-517.

GUBLER, J. G.; WÜST, J.; KRECH, T.; & HANY, A. (1990). Classical pseudomembranous diphtheria caused by *Corynebacterium ulcerans*. **Schweizerische Medizinische Wochenschrift**, v. 120, n. 48, p. 1812-1816.

HALL I.C. & FISHER C.W. (1915). Suppurative lesions in horses and a calf of California due to the diphtheroid bacillus of Preisz-Nocard. **J Am Vet Med Assoc**, v.1, p.18-30.

HASTINGS, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, 57:97-109.

HAUBMANN, U., QI, S. W., WOLTERS, D., RÖGNER, M., LIU, S. J., & POETSCH, A. (2009). Physiological adaptation of *Corynebacterium glutamicum* to benzoate as alternative carbon source—a membrane proteome-centric view. **Proteomics**, 9(14), 3635-3651.

HOSKISSON, P.A. (2018). Microbe Profile: *Corynebacterium diphtheriae*—an old foe always ready to seize opportunity. **Microbiology**, v.164, n.6, p.865.

KHAMIS A, RAOULT D, LA SCOLA B. (2004). Rpo B gene sequencing for identification for *Corynebacterium* species. **J Clin Microbiol**, 42(9):3925-31.

LIMA ACS, DE MOURA VAG, PINHEIRO K DA C, PAIXÃO CTM, DA COSTA WLO, FOLADOR ARC, GUARALDI AL DE M, RAMOS RTJ, SILVA A, MARQUES JM. (2017). Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA05 Isolated from an Ovine Host in Pará State, Brazil. **Genome Announc.**, v.5, n.13, p.e00082-17.

LIU, H., ZHANG, Y., LIU, Z., LIU, J., HAUCK, Y., LIU, J., ..., & JIANG, Y. (2018). Associations between *Mycobacterium tuberculosis* Beijing genotype and drug resistance to four first-line drugs: a survey in China. **Frontiers of medicine**, 12(1), 92-97.

MADIGAN, Michael T. et al. *Microbiologia de Brock-14ª Edição*. Artmed Editora, 2016.

MARQUES, J. M., DE MOURA, V. A. G., LIMA, A. C. S., PAIXÃO, C. T. M., LOBATO, A. R. F., ALVES, J. T. C., ... & SILVA, A. (2017). Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA06 Isolated from a Subauricular Abscess in an Ovine Host. **Genome Announc.**, 5(13), e00083-17.

MENARDO, F., DUCHENE, S., BRITES, D., & GAGNEUX, S. (2019). The molecular clock of *Mycobacterium tuberculosis*. **BioRxiv**, 532390.

MUÑOZ-BUCIO, A. V., CORTÉS-PÉREZ, Y. A., ARELLANO-REYNOSO, B., HERNÁNDEZ-GIL, M., HERNÁNDEZ-CASTRO, R., & DÍAZ-APARICIO, E. (2017). Identification of *Corynebacterium pseudotuberculosis* isolated from muscular abscesses in two horses: First report in Mexico. **Equine Veterinary Education**, 29(8), 431-435.

METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., & TELLER, E. (1953). Equation of state calculations by fast computing machines. **The journal of chemical physics**, 21(6), 1087-1092.

PASCUAL, C.; LAWSON, P. A; FARROW, J. A; GIMENEZ, M. N.; COLLINS, M. D. Phylogenetic analysis of the genus *Corynebacterium* based on 16S rRNA gene sequences. **International journal of systematic bacteriology**, v. 45, n. 4, p. 724–8, out 1995

SOARES, SC, SILVA, A., TROST, E., BLOM, J., RAMOS, R., CARNEIRO, A., ... & BARBOSA, EG. (2013). O pan-genoma do patógeno animal *Corynebacterium pseudotuberculosis* revela diferenças na plasticidade do genoma entre as linhagens de biovar ovis e equi. **PLoS One**, 8(1), e53818.

WONG, S. C., POON, R. W., CHEN, J. H., TSE, H., LO, J. Y., NG, T. K., ... & LUK, W. K. (2017). *Corynebacterium kroppenstedtii* is an emerging cause of mastitis

especially in patients with psychiatric illness on antipsychotic medication. In: **Open forum infectious diseases** (Vol. 4, No. 2). Oxford University Press.

WU, D., JOSPIN, G., & EISEN, J. A. (2013). Systematic identification of gene families for use as “markers” for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their major subgroups. **PloS One**, 8(10), e77033.

ZIMMERMANN, J., RÜCKERT, C., KALINOWSKI, J., & LIPSKI, A. (2016). *Corynebacterium crudilactis* sp. nov., isolated from raw cow's milk. **International journal of systematic and evolutionary microbiology**, 66(12), 5288-5293.

8. ANEXOS

1. Artigo I



genomeAnnouncements™

PROKARYOTES



Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA06 Isolated from a Subauricular Abscess in an Ovine Host

Joana Montezano Marques,* Vitória Almeida Gonçalves de Moura,* Aiyne Cristina Sodré Lima,* Carla Thais Moreira Paixão,* Amélia Raiana Fonseca Lobato,* Jorlaine Thyessa Castro Alves,* Ana Lúcia de Mattos Guaraldi,* Adriana Ribeiro Carneiro Faldor,* Rommel T. J. Ramos,* Artur Silva*

Laboratory of Genetic and Bioinformatics, Federal University of Pará, Center of Genomics and System Biology, Belém, Pará, Brazil; College of Medical Science, State University of Rio de Janeiro, Rio de Janeiro, Brazil*

ABSTRACT We report here the draft genome sequence of *Corynebacterium pseudotuberculosis* PA06, isolated from a subauricular abscess in an ovine host. *C. pseudotuberculosis* is a worldwide pathogen of small and large ruminants. The genome comprises 2,320,074 bp, with a G+C content of 52.2%, 2,195 coding sequences, 48 tRNAs, and three rRNAs.

Corynebacterium pseudotuberculosis is a Gram-positive, nonmotile, non-spore-forming, pleomorphic, intracellular, and aerobic-facultative bacterium (1). This bacterium is the etiological agent of caseous lymphadenitis (CLA) disease in small ruminants. Besides its veterinary interest, this pathogen is a zoonotic agent and may be associated with many diseases distributed worldwide (2). The pathogenic effects of this agent have been observed in different hosts, including ulcerative lymphangitis in horses and ulcerative dermatitis in cattle. Recently, *C. pseudotuberculosis* has also been reported in pigs (3), and, by molecular genetic evidence, it was demonstrated that a veterinary student had pneumonia caused by this pathogen (4).

The primary pathology of this bacterium, caseous lymphadenitis, is a contagious disease that is characterized by the presence of abscesses in the lymph nodes and internal host organs (2). Considering the difficulties in identifying the disease's early stages and the presence of antigens produced by infectious agents (5), caseous lymphadenitis presents itself as a difficult disease to eradicate. Thus, it is responsible for large economic losses by reducing wool quality, decreasing milk production, lowering weight, and causing death and damage to carcasses (2).

C. pseudotuberculosis PA06 was obtained from a caseous secretion from a punctured lymph node in the subauricular region of a sheep host (Santa Inês breed) located on a farm in Pará State, Brazil. The draft genome was sequenced using the Ion Torrent Personal Genome Machine (PGM) platform (Thermo Fisher) with a fragment library. The genome assembly was performed by SPAdes version 3.9.0 (6) with k-mers of 31, 33, 35, 61, 63, and 65, resulting in 20 contigs. Additionally, MIRA version 4.0 (7) was used with default parameters to reduce the assembly to seven contigs. MAUVE version 2.4 (8) was used to generate the scaffold that was automatically annotated with RAST version 2.0 (9). The genome has 2,195 coding sequences, three rRNAs, 48 tRNAs, and a G+C content of 52.2%.

Accession number(s). This whole-genome shotgun project has been deposited in GenBank under the accession number CP019289.

Received 24 January 2017; Accepted 25 January 2017; Published 30 March 2017

Citation: Marques JM, de Moura VG, Lima ACS, Paixão CTM, Lobato AR, Alves JT, Guaraldi ALM, Faldor AR, Ramos RTJ, Silva A. 2017. Draft genome sequence of *Corynebacterium pseudotuberculosis* strain PA06 isolated from a subauricular abscess in an ovine host. *Genome Announcements* 5:e00083-17. <https://doi.org/10.1128/genomeA.00083-17>.

Copyright © 2017 Marques et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Joana Montezano Marques, jmontezano@ufpa.br.

2. Artigo II



Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA05 Isolated from an Ovine Host in Pará State, Brazil

Alyne Cristina Sodré Lima,* Vitória Almeida Gonçalves de Moura,*
Kenny da Costa Pinheiro,* Carla Thais Moreira Palácio,*
Wana Lailan Oliveira da Costa,* Adriana Ribeiro Carneiro Fozador,*
Ana Luitza de Mattos Guaraldi,* Rommel T. J. Ramos,* Artur Silva,*
Joana Montazano Marques*

Laboratory of Genomic and Bioinformatics, Federal University of Pará, Center of Genomics and System Biology, Belém, Pará, Brazil; College of Medical Science, State University of Rio de Janeiro, Rio de Janeiro, Brazil*

ABSTRACT We report here the draft genome sequence of *Corynebacterium pseudotuberculosis* PA05, isolated from an ovine host in Pará State, Brazil. *C. pseudotuberculosis* is an etiological agent of diseases with veterinary and medical importance. The genome contains 2,435,137 bp, a G+C content of 52.2%, 2,295 coding sequences, five pseudogenes, 53 tRNAs, and six rRNAs.

Corynebacterium pseudotuberculosis is a Gram-positive, nonmotile, nonsporulated, pleomorphic, intracellular, and aerobic-facultative bacterium (1). *C. pseudotuberculosis* is mainly known as the etiologic agent of caseous lymphadenitis (CLA), a chronic disease in small ruminants (goats and sheep) and is also considered an important agent to other diseases such as ulcerative lymphangitis and ulcerative dermatitis, among others. This pathogen affects several species, including sheep, goat, horse, cattle, llama, alpaca, buffalo, and human (2).

CLA is a contagious disease that is characterized by the presence of abscesses in the lymph nodes and internal organs of the host. After infection, the bacterium becomes encapsulated within walled-off lesions from which they escape immune system-mediated action, initiating a state of persistence (3). The infection presents itself as a disease difficult to eradicate. It is responsible for large economic losses resulting from the damage it inflicts on livestock, such as reducing the quality of wool, decreasing milk production, lowering animal weight, and causing death and damage to carcasses (4). The bacterium has numerous survival mechanisms and uses many strategies to adapt to its environment; the two toxic factors most described in the literature are a lipid factor that improves cellular wall resistance against digestion by cellular enzymes and exotoxin with hemolytic capacity that increases vascular permeability and facilitates invasion (5). Here, we report the genome sequencing of *C. pseudotuberculosis* PA05, a strain isolated from a punctured caseous abscess located in the lymph nodes in the posterior thigh region of a sheep host (Santa Inês breed) in Pará, Brazil.

The genome was sequenced with the Ion Torrent PGM platform using a fragment library that generated 1,251,198 reads. FastQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) was used to evaluate the quality of the reads, and the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit) was used to perform filtering and trimming, in which reads with a Phred quality score of ≥ 20 were removed. The trimmed reads were then assembled using SPAdes version 3.9.0 (<http://bioinf.spbau.ru/spades>) with *k*-mers 31, 33, 35, 61, 63, and 65. The assembly generated 24 contigs. The number of contigs was reduced to eight using the Lasergene 11 Core Suite tool

Received 24 January 2017 Accepted 25 January 2017 Published 30 March 2017

Citation Lima ACS, de Moura VAG, Pinheiro KDC, Palácio CTM, da Costa WLO, Fozador AR, Guaraldi ADM, Ramos RTJ, Silva A, Marques JM. 2017. Draft genome sequence of *Corynebacterium pseudotuberculosis* strain PA05 isolated from an ovine host in Pará State, Brazil. Genome Announc 5:e00002-17. <https://doi.org/10.1128/genomeA.00002-17>.

Copyright © 2017 Lima et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Joana Montazano Marques, jmontazano@ufpa.br.

Downloaded from <http://genome.asm.org/> on November 6, 2017 by guest

3. Artigo III



Draft Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA07 Biovar *ovis*, Isolated from a Sheep Udder in Amazonia

Fabrizio Almeida Araújo, Joana Montezano Marques, Vitória Almeida Gonçalves de Moura, Maria Paula Cruz Schneider, Soraya Silva Andrade, Aylene Cristina Sodré Lima, LUIS Carlos Guimarães, Adriana Ribeiro Carmelo Folador, Artur Silva, Rommel T. J. Ramos

University of Pará, Center of Genomics and System Biology, Laboratory of Genomic and Bioinformatics, Belém, Pará, Brazil

ABSTRACT In this work, we present the draft genome sequence of *Corynebacterium pseudotuberculosis* strain PA07 biovar *ovis*, isolated from a caseous secretion from a sheep udder in Pará, Brazil. The genome contains 2,320,235 bp, 52.2% G+C content, 2,191 coding sequences (CDSs), five pseudogenes, 48 tRNAs, and three rRNAs.

Corynebacterium pseudotuberculosis is a Gram-positive, facultative intracellular, pleomorphic, nonsporulating, noncapsulated, nonmotile bacterium that, together with *Mycobacterium*, *Nocardia*, and *Rhodococcus*, forms the CMNR group which shares specific characteristics, such as (i) high G+C content and (ii) specific cell wall organization, characterized by the presence of peptidoglycan, arabinogalactan, and mycolic acids (1, 2).

C. pseudotuberculosis is the etiological agent of caseous lymphadenitis (CLA), ulcerative lymphangitis, mastitis (clinical and subclinical), and ulcerative dermatitis. CLA is characterized by the formation of abscesses in the superficial lymph nodes and subcutaneous tissues (3, 4). CLA is a widespread disease that has been reported all over the world (5). This disease is associated with direct economic losses due to the progressive reduction in weight gain, depreciated wool and skin, reduced milk production, and eventually death caused by toxemia of the infected animals. Moreover, this bacterium presents significant zoonotic potential (5).

C. pseudotuberculosis presents some variable tools for persistence and escape from the immune system of the host (3). Thus, the molecular study of the genome of this pathogen and the search for new target genes is an important step for vaccine development (6, 7).

Herein, we present the draft genome of *C. pseudotuberculosis* strain PA07 biovar *ovis*, isolated from a Santa Inês sheep udder in Pará, Brazil. To the best of our knowledge, we report the first recovery of *C. pseudotuberculosis* from an abscess located in sheep udder in the state of Pará. This strain is part of the collection of Laboratory of Genomics and System Biology located at the Federal University of Pará, Belém, Pará, Brazil.

The genome sequencing was performed using an Ion Torrent personal genome machine (PGM) (Thermo, Fisher) with a fragment library. The *de novo* assembly strategy was performed using SPAdes genome assembler software version 3.9.0 (8). The assembly generated seven contigs with 2,320,235 bp. The contigs were automatically annotated using Rapid Annotations using Subsystems Technology (RAST) (9). The genome has 2,191 coding sequences (CDSs), five pseudogenes, 48 tRNAs, three rRNAs, and a G+C content of 52.2%.

Received 16 January 2017 Accepted 19 January 2017 Published 23 March 2017
Citation Araújo FA, Marques JM, de Moura VAG, Schneider MPC, Andrade SS, Lima ACS, Guimarães LC, Folador AR, Silva A, Ramos RTJ. 2017. Draft genome sequence of *Corynebacterium pseudotuberculosis* strain PA07 biovar *ovis*, isolated from a sheep udder in Amazonia. *Genome Announcements* 5:e00040-17. <https://doi.org/10.1128/genomeA.00040-17>.
Copyright © 2017 Araújo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.
Address correspondence to Rommel T. J. Ramos, rommel@iagpa.ufrpa.br.

4. Capitulo de livro I

The Genus *Corynebacterium* in the Genomic Era

Carlos Leonardo Araújo, Jorianne Alves, Alyne Lima,
Larissa Dias, Patricia Silva, Joana Marques,
Vasco Azevedo, Artur Silva and Adriana Folador

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.80445>

Abstract

This chapter will address the main omics approaches used in studies involving the genus *Corynebacterium*, Gram-positive microorganisms that can be isolated from many diverse environments. Currently, the genus *Corynebacterium* has more than 130 highly diversified species, many of which present medical, veterinary and biotechnological importance, such as *C. diphtheriae*, *C. pseudotuberculosis*, *C. ulcerans* and *C. glutamicum*. Due to the wide application in these fields, several omics methodologies are used to better elucidate the species belonging to this genus, such as genomics, transcriptomics and proteomics. The genomic era has contributed to the development of more advanced and complex approaches that enable the increase of generated data, and consequently the advance on the structural, functional and dynamic knowledge of biological systems.

Keywords: *Corynebacterium*, genomic era, genomics, transcriptomics, proteomics

1. Introduction

The genus *Corynebacterium* was proposed by Lehman and Neumann in 1896, to describe a type strain bacillus *Corynebacterium diphtheriae*. However, antecedent to its final taxonomic classification, *C. diphtheriae* had already been described as synonymous species such as *Microsporon diphthericum*, *Bacillus diphtheriae*, and *Paciua loeffleri*. After its classification, the species was again described as the synonym *Mycobacterium diphtheriae* by Krasil'nikov in 1941 [1]. Afterward, the genus came to accommodate other bacterial species that presented

IntechOpen

© 2018 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

CAPÍTULO 14

PERFIL FENOTÍPICO E GENOTÍPICO DE UMA CEPA DE *Escherichia coli* MULTIRRESISTENTE A ANTIBIÓTICOS, ISOLADA DO LAGO ÁGUA PRETA, BELÉM, PARÁ

Ícaro Rainyer Rodrigues de Castro

Universidade Federal Rural da Amazônia – UFRA
Belém – Pará

Jorianne Thyeska Castro Alves

Universidade Federal do Pará – UFPA
Belém – Pará

Alyne Cristina Sodré Lima

Instituto Federal de Ciência e Tecnologia do
Amapá
Porto Grande – Amapá

Vitória Almeida Gonçalves de Moura

Universidade Federal do Pará – UFPA
Belém – Pará

Carla Thais Moreira Paixão

Universidade Estadual de Campinas – UNICAMP
Campinas – São Paulo

Wana Lailan Oliveira da Costa

Universidade Federal do Pará – UFPA
Belém – Pará

Adriedson Jameson Chaves de Alcântara

Universidade Federal do Pará – UFPA
Belém – Pará

Carlos Leonardo de Aragão Araújo

Universidade Federal do Pará – UFPA
Belém – Pará

Larissa Maranhão Dias

Universidade Federal do Paraíba – UFPB
João Pessoa – Paraíba

Artur Luiz da Costa da Silva

Universidade Federal do Pará – UFPA
Belém – Pará

Adriana Ribeiro Carneiro Folador

Universidade Federal do Pará – UFPA
Belém – Pará

RESUMO: O presente trabalho teve como objetivo caracterizar o perfil fenotípico e genotípico de uma cepa de bactéria resistente a antibióticos de último recurso, isolada do lago Água Preta, manancial de abastecimento de Belém, Pará. Como métodos de caracterização fenotípica foram utilizados a coloração de Gram, microscopia direta e teste de sensibilidade *in vitro*. E para fins de caracterização genotípica foi realizada a extração do DNA genômico, diferenciação da espécie por meio de PCR Quadrex e PCR para identificação de genes de resistência. A cepa C40A foi identificada como *Escherichia coli* pertencente ao filogruppo D. A bactéria foi resistente a 87,5% dos antibióticos clinicamente disponíveis para o tratamento de infecções bacterianas e susceptíveis a apenas 12,5% dos mesmos (Imipenem e Canamicina). Foi identificada a presença do gene *bla*CTX e dos elementos móveis *Int1* e *Int2*, sendo estes componentes genéticos, marcadamente associados a expressão de perfis de resistência a antibióticos.

PALAVRAS-CHAVE: Genética molecular, Microbiologia e Resistência

6. Resumos de congresso



29º Congresso Brasileiro de Microbiologia

De 22 a 25 de Outubro de 2017 / Foz do Iguaçu-Paraná-Brasil
"60 anos da Sociedade Brasileira de Microbiologia"



29º Congresso Brasileiro de Microbiologia

De 22 a 25 de Outubro de 2017 / Foz do Iguaçu-Paraná-Brasil
"60 anos da Sociedade Brasileira de Microbiologia"



1614-1 - GENOTYPIC AND PHENOTYPIC IDENTIFICATION OF STRAINS OF CORYNEBACTERIUM PSEUDOTUBERCULOSIS ISOLATED FROM SMALL RUMINANTS IN BRAZILIAN AMAZON

Vitória Almeida G de Moura (UFPA - Universidade Federal do Pará) ; **Alyne Cristina Lima** (UFPA - Universidade Federal do Pará) ; **Adriana Ribeiro Carneiro Folador** (UFPA - Universidade Federal do Pará) ; **Arthur Silva** (UFPA - Universidade Federal do Pará) ; **Joana Marques** (UFPA - Universidade Federal do Pará)

Área: Genética de Microrganismos e Bioinformática

Sub-Área: Genética Microbiana

Data de Apresentação: 08/10/2019

Painel Número: 910

Palavras-chave: Corynebacterium pseudotuberculosis, Caseous lymphadenitis, Small ruminants, Amazon

Trabalho

972-1 - BAYESIAN PHYLOGENETIC ANALYSIS AND MOLECULAR CLOCK OF THE CORYNEBACTERIUM GENUS

José Lucas Souza dos Reis (CGBS - Centro de Genômica e Biologia de Sistemas) ; **Sávio de Souza Costa** (CGBS - Centro de Genômica e Biologia de Sistemas) ; **Davi Josué Marcon** (CGBS - Centro de Genômica e Biologia de Sistemas) ; **Alyne Cristina Sodré Lima** (IFAP - Instituto federal do Amapá - Campus Porto Grande, CGBS - Centro de Genômica e Biologia de Sistemas) ; **Marcelo Nazareno Vallinoto de Souza** (UFPA - Laboratório de Evolução - Universidade Federal do Pará) ; **Artur Luiz da Costa da Silva** (CGBS - Centro de Genômica e Biologia de Sistemas, ENGBIO - Laboratório de Engenharia Biológica) ; **Rafael Azevedo Baraúna** (CGBS - Centro de Genômica e Biologia de Sistemas, ENGBIO - Laboratório de Engenharia Biológica)

Área: Genética de Microrganismos e Bioinformática

Sub-Área: Genética Microbiana

Data de Apresentação: 08/10/2019

Painel Número: 924

Palavras-chave: Bayesian inference, Corynebacterium, Molecular clock, Phylogenetics

Trabalho