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

ROSELANE GONÇALVES DOS SANTOS

MALDI -TOF MS PARA DIFERENCIAÇÃO DE ISOLADOS DE *Corynebacterium pseudotuberculosis* BIOVAR OVIS E EQUI, E CARACTERIZAÇÃO DE UMA NOVA ESPÉCIE DE *DIETZIA*

Belo Horizonte
Dezembro - 2023

ROSELANE GONÇALVES DOS SANTOS

MALDI -TOF MS PARA DIFERENCIAÇÃO DE ISOLADOS DE *Corynebacterium pseudotuberculosis* BIOVAR OVIS E EQUI, E CARACTERIZAÇÃO DE UMA NOVA ESPÉCIE DE *DIETZIA*

Versão final de tese apresentada como requisito para a obtenção do Grau de Doutora pelo Programa de Pós-graduação em Bioinformática, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais.

Orientador: Prof. Dr. Vasco Azevedo
Coorientadores: Prof. Dr. Sandeep Tiwari

Belo Horizonte
Dezembro – 2023

043 Santos, Roselane Gonçalves dos.
MALDI-TOF MS para diferenciação de isolados de Corynebacterium pseudotuberculosis biovar Ovis e Equi, e caracterização de uma nova espécie de Dietzia [manuscrito] / Roselane Gonçalves dos Santos. – 2023.
124 f. : il. ; 29,5 cm.

Orientador: Vasco Azevedo. Coorientador: Sandeep Tiwari Sousa.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Bioinformática.

1. Bioinformática. 2. Actinomycetales. 3. Corynebacterium pseudotuberculosis. 4. Genômica. I. Azevedo, Vasco. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 573:004

Ficha Catalográfica elaborada pelo bibliotecário Rafael Gonçalves Dias (CRB: 6 - 3277)

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

ATA DA DEFESA DE TESE

ROSELANE GONÇALVES DOS SANTOS

Às dez horas do dia **14 de dezembro de 2023**, reuniu-se, através de videoconferência, a Comissão Examinadora de Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: **"MALDI-TOF MS para diferenciação de isolados de *Corynebacterium pseudotuberculosis* biovar Ovis e Equi, e caracterização de uma nova espécie de *Dietzia*"**, requisito para obtenção do grau de Doutora em **Bioinformática.** Abrindo a sessão, o Presidente da Comissão, **Dr. Vasco Ariston de Carvalho Azevedo**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	**Instituição**	**Indicação**
Dr. Vasco Ariston de Carvalho Azevedo - Orientador	Universidade Federal de Minas Gerais	**Aprovada**
Dr. Sandeep Tiwari - Coorientador	Universidade Federal da Bahia	**Aprovada**
Dra. Flavia Figueira Aburjaile	Universidade Federal de Minas Gerais	**Aprovada**
Dr. Mateus Matiuzzi da Costa	Universidade Federal do Vale do São Francisco	**Aprovada**
Dra. Paula Carvalhal Lage von Buettner Ristow	Universidade Federal da Bahia	**Aprovada**
Dr. Thiago Motta Venancio	Universidade Estadual do Norte Fluminense	**Aprovada**

Pelas indicações, a candidata foi considerada: **Aprovada**

O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 14 de dezembro de 2023.

Documento assinado eletronicamente por **Flávia Figueira Aburjaile**, **Professora do Magistério Superior**, em 14/12/2023, às 13:10, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Sandeep Tiwari**, **Usuário Externo**, em 14/12/2023, às 13:10, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Vasco Ariston de Carvalho Azevedo**, **Professor do Magistério Superior**, em 14/12/2023, às 23:39, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Thiago Venancio**, **Usuário Externo**, em 18/12/2023, às 10:25, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Paula Carvalhal Lage Von Buettener Ristow**, **Usuário Externo**, em 18/12/2023, às 12:58, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Mateus Matiuzzi da Costa**, **Usuário Externo**, em 19/12/2023, às 14:43, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2889591** e o código CRC **31D42F5D**.

Referência: Processo nº 23072.276391/2023-98 SEI nº 2889591

“E quando comecei a redigir, cheio de confiança,
verifiquei logo em que terreno escorregadio me aventurava.
(...), mas um pensamento me consola. Quem fez a experiência
de pensar em outro domínio sobrepuja sempre aquele que não
pensa de modo algum ou muito pouco.”

Albert Einstein (*14 março de 1879 †18 de abril de 1955)
Como Vejo o Mundo (Mein Weltbild), Zurich, 1953.

AGRADECIMENTOS

Meus sinceros agradecimentos a todos que foram fundamentais em minha jornada até este momento. Não foi um caminho fácil, mas não seria possível sem o apoio e a ajuda proporcionados por cada um de vocês.

Em primeiro lugar, agradeço a Deus, que me deu forças nos momentos em que pensei em desistir de tudo, e a quem recorria para encontrar consolo e coragem para continuar. Ele abençoou minha jornada, iluminou meu caminho e colocou pessoas tão especiais ao meu lado ao longo dessa trajetória.

Ao meu orientador, Vasco Azevedo, minha imensa gratidão por ter acreditado em mim desde o início, investido seu tempo e dedicação em meu projeto, me aconselhado nos momentos necessários, compartilhado seus conhecimentos e me dado todo o suporte para a realização desse sonho.

Aos meus coorientadores, agradeço pela paciência, apoio e sabedoria.

Ao meu esposo, expresso minha profunda gratidão por toda a paciência e apoio durante esse período, por suportar meu estresse e minha falta de tempo, e por estar sempre ao meu lado me dando amor.

À minha família, que me apoiou mesmo sem compreender completamente o que eu estava passando, que sempre se orgulhou de mim, suportou minha ausência e me ofereceu apoio emocional, especialmente minha sobrinha(Keke) e meus irmãos.

Ao LGCM, o laboratório onde fiz meu doutorado, por ter sido o cenário onde desenvolvi meu conhecimento e habilidades acadêmicas. A equipe e o ambiente acolhedor foram essenciais

para o meu crescimento como pesquisador e para a conclusão bem-sucedida deste importante capítulo da minha vida.

Aos amigos que fiz durante essa jornada, em especial a Tosta, que esteve ao meu lado de perto, sempre oferecendo ajuda e tornando o processo mais leve e proveitoso.

À minha companheira de quatro patas, Dandara, que me ajuda com apoio emocional e me distrai nos momentos certos.

Gostaria de estender meus sinceros agradecimentos a Fernanda, Sheila, Natália e Thiago (pessoal da secretária). Sua dedicação e eficiência na gestão administrativa foram fundamentais para o bom funcionamento do programa/curso e para que eu pudesse me concentrar em minha pesquisa. Sempre prestativos e solícitos, vocês tornaram minha experiência acadêmica ainda mais gratificante. Muito obrigado por todo o suporte e ajuda ao longo desse percurso. Vocês fizeram a diferença e sou muito grato por tê-los em minha jornada.

E, por fim, minha sincera gratidão a todos os funcionários do ICB que, de alguma forma, contribuíram para que esse sonho se tornasse realidade.

Agradeço a todos por fazerem parte da minha jornada, por acreditarem em mim e por tornarem essa conquista possível. Seu apoio e encorajamento foram inestimáveis, e serei eternamente grato pelo impacto que cada um de vocês teve em minha vida.

RESUMO

Corynebacterineae é uma subordem de bactérias bacilares Gram-positivas, ácido-álcool resistentes, geralmente aeróbias, da ordem dos *Actinomycetales*, que inclui diversos gêneros como *Corynebacterium e Dietzia*. *Corynebacterium pseudotuberculosis* é uma bactéria de grande importância veterinária que leva a grandes prejuízos ao agronegócio, necessitando de melhor conhecimento biológico para se alcançar um controle das doenças. Por outro lado, espécies do gênero *Dietzia* estão amplamente distribuídas no ambiente, apresentando um papel oportunista, com algumas espécies apresentando taxonomia ainda indeterminada. Métodos bioquímicos, sorológicos e moleculares têm sido desenvolvidos para o diagnóstico laboratorial de doenças causadas por *C. pseudotuberculosis*, entretanto a identificação do patógeno e a diferenciação dos biovares algumas vezes pode ser imprecisa, e levar a classificações taxonômicas com outras bactérias, ou vice-versa, de forma errônea. Neste estudo realizamos uma análise no MALDI TOF para otimizar a identificação e diferenciação de biovares de *C. pseudotuberculosis*. Nossos dados mostram que a análise MALDI TOF são úteis para identificar *C. pseudotuberculosis* em nível de espécie e biovares, bem como para identificar outras cepas pertencendo a outros gêneros/espécies, como foi encontrado neste trabalho. Durante essa otimização recebemos uma cepa identificada como *C. pseudotuberculosis*, entretanto o resultado do PCR quadruplex, comumente utilizada para identificação de *C. pseudotuberculosis*, foi negativa, e ao utilizarmos o MALDI TOF para essa identificação, foi possível identificar como uma cepa pertencendo ao gênero *Dietzia*. Baseado nesses achados, neste trabalho partimos então para a caracterização genômica realizando o sequenciamento do isolado, análises da similaridade genômica entre o nosso isolado e as espécies de *Dietzia* disponível em banco de dados públicos, análises fenotípicas da cepa, usando como padrão espécies já descritas para o gênero, bem como algumas análises bioquímicas. Os resultados das análises genômicas, testes fisiológicos e bioquímicos utilizados para a caracterização de uma nova espécie determinaram que essa cepa é diferente das espécies conhecidas do gênero *Dietzia*. As análises de genômica comparativa fornecem informações sobre a diversidade genômica, taxonomia e adaptação o gênero *Dietzia* e são um recurso valioso para estudos genômicos funcionais, proporcionando resultados mais abrangentes que o convencional, auxiliando na identificação correta e melhorando o entendimento de sua patologia oportunista. Além disso, os dados deste trabalho sugerem que o isolado seja classificado como representante de uma nova espécie do gênero *Dietzia*, denominada *Dietzia calabriae*.

Palavras Chaves: *Corynebacterium pseudotuberculosis*; *Dietzia*, Identificação bacteriana, MALDI TOF, genômica comparativa, diversidade genômica, taxonomia.

ABSTRACT

Corynebacterineae is a suborder of Gram-positive, acid-alcohol resistant, usually aerobic, bacillary bacteria of the order Actinomycetales, which includes several genera such as Corynebacterium and *Dietzia*. *Corynebacterium pseudotuberculosis* is a bacterium of great veterinary importance that leads to great damage to agribusiness, requiring better biological knowledge to achieve disease control. On the other hand, species of the genus *Dietzia* are widely distributed in the environment, presenting an opportunistic role, with some species presenting an undetermined taxonomy. Biochemical, serological, and molecular methods have been developed for the laboratory diagnosis of diseases caused by *C. pseudotuberculosis*, however, the identification of the pathogen and the differentiation of biovars can sometimes be imprecise, and lead to taxonomic classifications with other bacteria, or vice versa, erroneously. In this study, we performed a MALDI TOF analysis to optimize the identification and differentiation of *C. pseudotuberculosis* biovars. Our data show that the MALDI TOF analysis is useful to identify *C. pseudotuberculosis* at the species and biovars level, as well as to identify other strains belonging to other genera/species, as was found in this work. During this optimization, we received a strain identified as *C. pseudotuberculosis*, however, the result of the quadruplex PCR, commonly used to identify *C. pseudotuberculosis*, was negative, and when we used the MALDI TOF for this identification, it was possible to identify it as a strain belonging to the genus *Dietzia*. Based on these findings, in this work we set out for the genomic characterization, performing the sequencing of the isolate, analysis of the genomic similarity between our isolate and the species of *Dietzia* available in public databases, phenotypic analysis of the strain, using as standard species already described for the genus, as well as some biochemical analyses. The results of the genomic analyses and physiological and biochemical tests used for the characterization of a new species determined that this strain is different from the known species of the genus *Dietzia*. Comparative genomic analyzes provide information on genomic diversity, taxonomy, and adaptation to the genus *Dietzia* and are a valuable resource for functional genomic studies, providing more comprehensive results than conventional ones, helping in the correct identification and improving the understanding of its opportunistic pathology. In addition, data from this work suggest that the isolate be classified as a representative of a new species of the genus *Dietzia*, called *Dietzia calabriae*.

Keywords: *Corynebacterium pseudotuberculosis*; *Dietzia*, Bacterial identification, MALDI TOF, comparative genomics, genomic diversity, taxonomy

LISTA DE ABREVIATURAS E TERMOS EM INGLÊS

ANI - Average Nucleotide Identity

AM - Ampicillin

AMC - Amoxicillin-clavulanate

BLAST - Basic Local Alignment Search Tool

CLO - Cloranfenicol

CDS- Coding Sequence

dDDH - DNA–DNA hybridization

DNA - Deoxyribonucleic acid

ELISA - Enzyme-Linked Immunosorbent Assay

HGT - Horizontal Gene Transfer

IS - Insertion Sequences

MDR - Multidrug resistance

NGS - Next Generation Sequencing

ORF - Open Reading Frame

PCR - Polymerase Chain Reaction

16S rRNA - 16S ribosomal ribonucleic acid

VFDB - Virulence Factors Database

WHO - World Health Organization

APRESENTAÇÃO DO TRABALHO DE TESE

Colaboradores

Este trabalho foi conduzido sob a orientação do Prof. Dr. Vasco Azevedo e coorientação do Prof. Dr. Sandeep Tiwari e Dr. Thiago Sousa, Laboratório de Genética Celular e Molecular (LGCM), Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas (ICB), Universidade Federal de Minas Gerais (UFMG), Belo Horizonte e outro (s) pesquisador (es)/colaborador (es) parceiro (s) integrais (nacionais e internacional) e suas respectivas instituições, entre outras, são:

- Professora Dr. Anna Rita Attili, National Specialist in Control of Food of Animal Origin School of Biosciences and Veterinary Medicine Microbiology and Infectious Diseases, University of Camerino.

O grupo de pesquisa (LGCM) está ativamente envolvido em abordagens de bioinformática em genômica de bactéria. Pioneiro em pesquisas em bioinformática no Brasil, o grupo se engajou em projetos de pesquisa intensivos abrangendo diversas áreas da biologia como genômica, transcriptômica e o desenvolvimento de vacinas e diagnósticos tornou o grupo referência para o estudo de microrganismos.

Este trabalho apresentou apoio financeiro da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), da Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Estrutura da qualificação

O delineamento da qualificação está organizado nas seguintes seções:

I. INTRODUÇÃO: Busca trazer informações relevantes sobre o trabalho, introduzindo a espécie de *Corynebacterium pseudotuberculosis*, que pode causar diversas doenças e sintomas em

diferentes hospedeiros, mostrando técnicas desenvolvidas para superar o problema de identificação clínica da infecção por *C. pseudotuberculosis*. Por ouro lado, também traz informações sobre o gênero *Dietzia*, considerado um gênero oportunista, em que suas espécies foram isoladas de diversos nichos. Além disso, o gênero *Dietzia*, pode ser muitas das vezes confundido com outros determinados actinomicetos corineformes, pois exibem características morfofisiológicas muito similares, podendo dificultar a diferenciação entre gêneros mirobianos.

II. JUSTIFICATIVA: Devido à dificuldade de se diferenciar isolados de *C. pseudotuberculosis* quanto ao seu biovar, nós testamos o MALDI TOF para ver a precisão da técnica na diferenciação dos biovares equi e ovis de *C. pseudotuberculosis*. Entretanto durante a durante a realização dos experimentos com o MALDI para diferenciação de *C. pseudotuberculosis* recebemos dois isolados identificados primeiramente como *Corynebacterium amycolatum e C. pseudotuberculosis*, pois as lesões macroscópicas apresentadas pelo animal pareciam lesões típicas da infecção por *C. pseudotuberculosis.* Entretanto, notou-se pelas técnicas utilizadas que esses isolados não eram *Corynebacterium.* Sendo assim, procedemos para caracterização dessa nova espécie, dando continuidade a este trabalho de pesquisa de Tese, em que saímos da identificação de biovares de *Corynebacterium pesudotuberculosis* para uma identificação imprecisa de um isolado do gênero *Dietzia*.

III. CAPÍTULO I – Este artigo de pesquisa descreve a diferenciação dos biovares equi e ovis de *C. pseudotuberculosis* utilizando a técnica MALDI TOF; intitulado: **Exploring the MALDI Biotyper for the identification of *Corynebacterium pseudotuberculosis* biovar Ovis and Equi.** O artigo foi publicado no Journal of the American Society for Mass Spectrometry.

IV. CAPÍTULO II – Este artigo de pesquisa descreve uma genômica comparativa, com foco na correta classificação das espécies descrita para o gênero até o momento; intitulado: **Comparative genomic analysis of the *Dietzia* genus: an insight into genomic diversity, and adaptation**. O artigo foi publicado no Journal Research in Microbiology.

V. CAPÍTULO III – Este capítulo está na forma de um draft de artigo de pesquisa, que descreve a Caracterização de uma nova espécie do gênero *Dietzia* em língua inglesa intitulado "***Dietzia calabriae* sp. nov., a bacterial species isolated from a mare's abscesso"** submetido no jornal Environmental Microbiology.

VI. CONCLUSÃO GERAL: sumariza os principais achados da Tese.

VII. PERSPECTIVAS: traz as principais ideias a serem exploradas para a continuidade deste trabalho até o final do doutorado.

VIII. REFERÊNCIAS BIBLIOGRÁFICAS: Lista as referências que foram utilizadas no desenvolvimento deste trabalho.

IX. APENDICE: Lista os trabalhos realizados durante esse período de doutorado.

SUMÁRIO

1. INTRODUÇÃO

1.1. Subordem *Corynebacterineae*

Corynebacterineae é uma subordem de bactérias bacilares Gram-positivas, ácido-álcool resistentes, geralmente aeróbias, da ordem dos *Actinomycetales*, que inclui diversas espécies patogênicas, como *Mycobacterium*, *Rhodococcus* (CONVILLE; WITEBSKY, 2015; LIM *et al.*, 2019; STACKEBRANDT; RAINEY; WARD-RAINEY, 1997a).

Os gêneros atribuídos à subordem *Corynebacterineae* contêm ácidos graxos hidroxila de cadeia longa chamados ácidos micólicos. O comprimento das cadeias laterais do ácido micólico variam e as diferenças no número de carbonos encontrados nos micolatos estão incluídas entre os critérios que definem gêneros ou famílias encontrados na subordem *Corynebacterineae* (CONVILLE; WITEBSKY, 2015). Assim, a subordem *Corynebacterineae* inclui alguns gêneros como *Corynebacterium, Dietzia, Amycolicicoccus, Mycobacterium, Gordonia, Millisia, Nocardia, Parvopolyspora, Rhodococcus, Skermania, Smaragdicoccus, Williamsia, Segniliparus, Tsukamurella* dentre outros (CONVILLE; WITEBSKY, 2015; NOUIOUI *et al.*, 2018).

1.2 *Corynebacterium pseudotuberculosis:* Aspectos microbiológicos e características patológicas

Corynebacterium pseudotuberculosis em termos microbiológicos caracteriza-se por ser um patógeno intracelular facultativo, bacilo Gram-positivo curto e irregular (0,5 a 0,6 μm por 1 a 3 μm) tendo como medidas estruturais formas cocóides e bastões filamentosos (BINNS; GREEN; BAILEY, 2007; CONNOR *et al.*, 2000; DERCKSEN *et al.*, 2000; DORELLA; PACHECO; *et al.*, 2006; SCHREUDER; TER LAAK; DERCKSEN, 1994).

C. pseudotuberculosis pode causar diversas doenças e sintomas em diferentes hospedeiros. A Linfadenite caseosa (LC) é uma dessas doenças, acometendo diferentes ruminantes como equinos, bovinos, ovinos e caprinos RO et al., 2013; SCHR(WINDSOR; BUSH, 2016). LC em pequenos ruminantes pode apresentar manifestações que são caracterizadas principalmente por necrose dos linfonodos. Entretanto a forma mais comum da doença é descrita pela formação de lesões granulomatosas nos linfonodos superficiais e nos tecidos subcutâneos (DORELLA; PACHECO; *et al.*, 2006; SCHREUDER; TER LAAK; DERCKSEN, 1994). Já os equinos são diagnosticados com linfangite ulcerativa, e os bovinos infectados apresentam lesões

granulomatosas ulcerativas e mastite (GUEDES *et al.*, 2015). Em búfalos é observado a doença cutânea edematosa (OSD), que é caracterizada por um inchaço cutâneo, edematoso na barbela, na parte posterior ou nos membros anteriores e na barriga. Essas doenças podem causar perdas econômicas ao agronegócio, devido a infecção por *C. pseudotuberculosis* ocasionar redução da produção de leite e carne, bem como o aproveitamento da pele do animal (SELIM, 2001).

Dentro da espécie *C. pseudotuberculosis* existe uma subdivisão a respeito da capacidade de redução do nitrato, o que permite a distinção das mesmas em dois biovares, respectivamente: o biovar *Equi*, representado por cepas nitrato redutoras, encontrados majoritariamente em equinos e bovinos; e o biovar *Ovis*, representado por cepas isoladas principalmente de caprinos e ovinos, e que não são capazes de reduzir nitrato (BIBERSTEIN; KNIGHT; JANG, 1971; CONNOR et al., 2007).(BIBERSTEIN; KNIGHT; JANG, 1971; CONNOR et al., 2007).

1.3. Diagnóstico de identificação para C. *pseudotuberculosis*

Vários testes de diagnóstico sorológico têm sido desenvolvidos para superar o problema de identificação clínica da infecção por *C. pseudotuberculosis*. (BAIRD; FONTAINE, 2007). Entretanto, este diagnóstico pode nem sempre ser vantajoso ou possível, pois algumas lesões externas podem conter pouco material caseoso ou poucos microrganismos viáveis, bem como podem apresentar dificuldades no crescimento bacteriano e contaminação das culturas (BAIRD; FONTAINE, 2007; RIBEIRO *et al.*, 2013).

Um método diagnóstico que pode ser utilizado para *C. pseudotuberculosis* é o teste bioquímico. Entretanto, esses testes possuem desvantagens devido ao custo e longo tempo para a identificação em laboratório (ÇETINKAYA *et al.*, 2002; MUCKLE; GYLES, 1982). Neste contexto, técnicas de imunodiagnóstico foram elaboradas na tentativa de otimizar o diagnóstico. O ensaio imunoenzimático (Enzyme-Linked Immunosorbent Assay) (ELISA), tem se mostrado ser um método flexível (BINNS; GREEN; BAILEY, 2007; CARMINATI *et al.*, 2003; DERCKSEN *et al.*, 2000; RIBEIRO *et al.*, 2013; SCHREUDER; TER LAAK; DERCKSEN, 1994). A performance e aplicabilidade de testes sorológicos baseiam-se na seleção de um antígeno imunodominante que seja reconhecido pela maioria dos hospedeiros infectados por *C. pseudotuberculosis* (HUERTA *et al.*, 2013; REZENDE *et al.*, 2016). A primeira proteína recombinante de *C. pseudotuberculosis* aplicada em ensaio de imunodiagnóstico foi a PLD

(MENZIES; HWANG; PRESCOTT, 2004; STAPLETON; BRADSHAW; O'KENNEDY, 2009). Entretanto um estudo mais recente de Barral e colaboradores (2019), desenvolveram ensaios para caprinos e ovinos utilizando um painel de proteínas recombinantes de *C. pseudotuberculosis* (rPLD, rCP40, rPknG, rDtxR e rGrx), pois as ferramentas de diagnóstico comercialmente disponíveis apresentam problemas de especificidade, o que pode levar a resultados falso-negativos. Assim, neste estudo foi desenvolvido um ensaio imunoenzimático indireto (ELISA) para a detecção de imunoglobulinas específicas em caprinos e ovinos, determinando, assim, melhores parâmetros de sensibilidade e especificidade (BARRAL *et al.*, 2019).

Visando a identificação de *Corynebacterium pseudotuberculosis* para detecção rápida desse patógeno em amostras clínicas, (MENZIES; HWANG; PRESCOTT, 2004; POLLOCK; ANDERSEN, 1997; SUNIL *et al.*, 2008). Pacheco et al. (2007) desenvolveram um ensaio de (Polymerase chain reaction) PCR multiplex (mPCR) com os genes 16S rRNA, *rpoB* e *pld*. Esse PCR foi específico o suficiente para diferenciar *C. pseudotuberculosis* de *C. ulcerans*, pois já foi relatado que essas duas bactérias possuem 99,7% de similaridade entre seus genes 16S rRNA e 93,6% entre seus genes *rpoB* (KHAMIS; RAOULT; LA SCOLA, 2004; PACHECO et al., 2007b; RIEGEL et al., 1995).

Por outro lado, esta abordagem foi aprimorada por Almeida et al. (2017), o qual desenvolveram um novo ensaio de PCR, denominado PCR quadruplex, para espécies de *C. pseudotuberculosis* com foco na diferenciação/identificação de biovares. Anteriormente, a identificação de biovares de *C. pseudotuberculosis*, era possível através dos procedimentos estabelecidos que incluíam o isolamento e identificação do agente por meio de testes bioquímicos, como o teste de redução de nitrato, onde as cepas de biovar ovis são negativas para redução de nitrato, enquanto as cepas de biovar equi são positivas (ALMEIDA *et al.*, 2017). O PCR quadruplex tem a adição de um novo par de oligonucleotídeos iniciadores direcionados ao gene *narG* do nitrato redutase ao antigo ensaio de PCR multiplex (ALMEIDA *et al.*, 2017; PACHECO *et al.*, 2007) que já possuía os *genes 16S, rpoB* e *pld.* A partir do PCR quadruplex é possível distinguir em níveis de espécies e biovares cepas de *C. pseudotuberculosis*. As vantagens destes ensaios moleculares sobre os testes bioquímicos são a velocidade, desempenho e reprodutibilidade, e a capacidade de testar um grande número de isolados simultaneamente (DORNELES *et al.*, 2014; SUTHERLAND; HART; BULLER, 1996). Além disso, o uso de técnicas moleculares reduz a manipulação de

bactérias viáveis em laboratório e consequentemente o risco de infecção acidental, pois *C. pseudotuberculosis* pode eventualmente ser um agente zoonótico (DORELLA; CARVALHO PACHECO; *et al.*, 2006).

Novos métodos de identificação e diferenciação dos biovares de *C. pseudotuberculosis* também podem ser implementados a partir de abordagens ômicas. Nesse sentido, Silva et al. (2017) relataram o primeiro estudo proteômico comparativo dos biovares Equi e *Ovis* de *C. pseudotuberculosis*, no qual os autores, considerando o proteoma exclusivo, detectaram proteínas específicas para as cepas *1002_ovis* e 258_ *equi* (modelos para os biovares Equi e Ovis, respectivamente). Essas diferenças proteômicas foram relacionadas principalmente a proteínas envolvidas no metabolismo celular, armazenamento e processamento de informações, e sinalização (SILVA *et al.*, 2014). No proteoma exclusivo de 258_ *equi*, as ORFs que codificam vinte proteínas são anotadas como pseudogene em 1002_ovis. No entanto, seis ORFs que codificam proteínas não foram detectadas no genoma de *1002_ovis* (SILVA *et al.*, 2014, 2017).

É importante ressaltar que as detecções antecipadas e rápidas na identificação dos microrganismos envolvidos em processos infeccioso, como o caso de *C. pseudotuberculosis*, são essenciais para a resposta mais eficaz a um surto de doenças infecciosas no rebanho. Esta identificação rápida é necessária não apenas para tratar os animais de forma eficaz, mas também para estabelecer uma gestão de surto, o rastreamento de origem, e análise de ameaça contra a doença (LISTA *et al.*, 2011).

No momento atual, os laboratórios de microbiologia clínica estão passando por transformações de ferramentas clássicas para novas ferramentas de diagnóstico como PCR, PCR em tempo real, espectrometria de massa MALDI-TOF e sequenciamento, e (GONÇALVES *et al.*, 2014; RAOULT; FOURNIER; DRANCOURT, 2004), que, devido à diminuição dos custos, podem preferir usar testes moleculares em vez de bioquímicos. Além disso, os testes moleculares costumam ser mais rápidos, fáceis e têm menos subjetividade na interpretação, tornando-os focos de vários estudos recentes (LANGONI *et al.*, 2017)).

Dentro desse contexto, uma técnica que está sendo utilizada para a identificação de microrganismos é o MALDI-TOF (LISTA et al., 2011). A Ionização de proteínas pela energia laser (MALDI) seguida pela separação e detecção de íons pelo tempo de voo (TOF) é um método rápido

para analisar diferenças biológicas em microrganismos. A sua capacidade para analisar as células bacterianas inteiras com virtualmente nenhuma preparação da amostra melhorou drasticamente o tempo necessário para a identificação de uma cultura positiva (1-2 min, começando com uma colónia). As vantagens de MALDI-TOF em relação às técnicas convencionais é que ela é um método preciso, rápido, eficaz em termos de custos, o que é apropriado para identificação de alto rendimento de bactérias por pessoal de laboratório (ASSIS *et al.*, 2017; LISTA *et al.*, 2011). Esta técnica tem desempenhado um papel cada vez mais significativo, e devido às suas capacidades tem sido utilizada numa grande variedade de aplicações (LEWIS *et. al*., 2000; MANN *et al*., 2001).

Há dois métodos para o preparo da amostra, o método direto e o da extração celular. No método direto, uma pequena porção da colônia isolada de um microrganismo é espalhada diretamente sobre a placa-alvo e a matriz orgânica (substância que facilita a ionização da amostra) é adicionada sobre ela. Esta técnica vem sendo utilizada com maior frequência para identificação de espécies bacterianas como *Streptococcus agalactiae*, *Lactococcus garvieae*, *Streptococcus iniae* e *Streptococcus dysgalactiae* subsp. disgalactiae (ASSIS *et al.*, 2017). Por vezes, melhor resultados de identificação bacteriana são relatados com o uso da extração proteica que consiste em uma metodologia que rompe as células ou tecidos para conseguir acessar as proteínas-alvo, melhorando a qualidade do espectro de massa (AGUSTINI, 2014),

Dados obtidos a partir do MALDI-TOF podem ser analisados com o auxílio de softwares, como o Biotyper (Bruker Daltonics, Alemanha), que contém um banco de dados utilizado para identificação de microrganismos, uma vez que contém espectros de massa, obtidos principalmente de bactérias, fungos e leveduras. Os resultados das análises neste banco são liberados na forma de log escores que variam de 0 a 3,0, sendo estes valores gerados pela comparação entre a lista de sinais de um isolado desconhecido com um espectro de referência, levando em consideração informações referentes a relação massa/carga (m/z), a intensidade relativa do sinal e as frequências dos picos presentes(ASSIS *et al.*, 2017; KANAK; YILMAZ, 2018; LISTA *et al.*, 2011).

Para prevenir identificações incorretas, o MALDI-TOF Byotiper usa critérios bem rigorosos para aceitar uma determinada identificação. Assim, um resultado é considerado consistente no que diz respeito a uma classificação a nível de espécie do microrganismo, quando apresentar um valor de log escore $\geq 2,0$. Quando o log escore fica entre 1,7 e 2,0, apenas o gênero é confirmado. Por outro lado, quando o resultado for $\leq 1,7$ significa que não se pode atribuir

nenhuma identificação ao microrganismo analisado (KANAK; YILMAZ, 2018; LISTA et al., 2011). Para (KANAK; YILMAZ, 2018; LISTA *et al.*, 2011). Portanto, a utilização de novos métodos de diferenciação dos biovares de *C. pseudotuberculosis* pode auxiliar numa melhor caracterização clínica e microbiológica destes patógenos.

É de suma importância também destacar que uma diferenciação mais precisa entre os biovares de *C. pseudotuberculosis* faz com que esta espécie também não seja erroneamente classificada como uma outra espécie ou vice-versa, principalmente aquelas pertencentes à subordem *Corynebacterinaceae,* como por exemplo bactérias do gênero *Dietzia*, que muitas vezes podem apresentar semelhanças microbiológicas em comum.

1.4 Gênero *Dietzia*

O gênero *Dietzia* é considerado um táxon monoespecífico, proposto por Rainey (1995) (RAINEY et al., 1995) para dois organismos classificados como *Rhodococcus maris* (NESTERENKO; NOGINA; KASUMOVA, 1982), anteriormente conhecido como '*Flavobacterium maris*'. *Dietzia* spp. pertence à família *Dietziaceae*, subordem *Corynebacterineae*, ordem *Actinomycetales* (NATARAJAN *et al.*, 2005). Até o momento da escrita deste trabalho, o género *Dietzia* compreende dezenove espécies descritas publicadas na Lista de nomes procarióticos com posição na nomenclatura (LPSN) **(https://lpsn.dsmz.de/genus/dietzia),** mas apenas onze dessas espécies tem a nomenclatura correta publicada e validada (GHARIBZAHEDI; RAZAVI; MOUSAVI, 2013).

Vários membros do gênero *Dietzia* foram originalmente isoladas de vários ambientes, como um lago de soda na África Ocidental(DUCKWORTH *et al.*, 1998), deserto frio do Himalaia Indiano (MAYILRAJ *et al.*, 2006a), deserto oriental do Egito (LI *et al.*, 2009), ar de celeiro de pato (KÄMPFER *et al.*, 2010), árvores de mogno em Timor Ocidental (YAMAMURA *et al.*, 2010a), comida tradicional coreana (KIM *et al.*, 2011), campos de petróleo ou solo contaminado por petróleo (NESTERENKO; NOGINA; KASUMOVA, 1982; WANG, XING BIAO *et al.*, 2011), água do produto em campo petrolífero (FANG *et al.*, 2021; WANG, XING BIAO *et al.*, 2011), e também de amostras clínicas, como prótese da articulação do quadril, cultura de amostra da parede aórtica e líquido pericárdico, além de cultura de pele de um paciente com papilomatose (BEMER-MELCHIOR *et al.*, 1999; JONES *et al.*, 2008; PIDOUX *et al.*, 2001a).

As espécies de *Dietzia* são caracterizadas por microrganismos Gram-positivos, aeróbicos, não móveis, não formadores de endósporos, não ácido-resistentes, oxidase-positivos e catalase-positivos, que formam cocos que germinam em bastonetes curtos ou células em forma de bastonete, e geralmente apresentam diâmetro celular médio de aproximadamente 1,1-1,5 μm. As colônias são circulares, elevadas ou convexas, brilhantes, de coloração laranja a vermelho coral, com bordas inteiras sendo formadas em meio de ágar (KOERNER; GOODFELLOW; JONES, 2009).

Algumas espécies de *Dietzia* apresentam um crescimento na faixa de temperatura de 10-45° C. No entanto, a temperatura ótima de crescimento para várias cepas varia entre 25 e 30° C. A maioria dos isolados de *Dietzia* podem crescer em condições salinas e alcalinas elevadas (GHARIBZAHEDI; RAZAVI; MOUSAVI, 2013), sendo que a tolerância máxima de NaCl para o cultivo de cepas fica na faixa de 15%, podendo variar entre 7 - 15% de NaCl, no qual as cepas apresentam um bom crescimento (GHARIBZAHEDI; RAZAVI; MOUSAVI, 2013)

As cepas de *Dietzia* são quimiorganotrófias, apresentando metabolismo oxidativo, com arabinose e galactose sendo os principais açúcares que compõem a parede celular. Os ácidos graxos celulares de cadeia longa consistem em componentes saturados e insaturados predominantemente de cadeia linear. Os principais perfis de lipídios polares presentes nas cepas das espécies de *Dietzia* são difosfatidilglicerol, fosfatidilglicerol, fosfatidilinositol manosídeo, fosfatidilinositol e fosfatidiletanolamina. Os ácidos graxos de células inteiras de *Dietzia* consistem predominantemente, em componentes saturados e insaturados de cadeia linear.

As quinonas isoprenoídes são uma classe de lipídeos terpenóides com potencial inerente semelhante à quimiotaxonomia. Eles são constituintes das membranas plasmáticas bacterianas e desempenham papéis importantes no transporte de elétrons, fosforilação oxidativa e, possivelmente, transporte ativo (GOODFELLOW; MALDONADO, 2006). A menaquinona desidrogenada com oito unidades de isopreno [MK-8 (H 2)] é a predominante em cepas de *Dietzia*. Entretanto, uma quantidade menor de MK-7(H_2) também foi observada em algumas espécies de *Dietzia* (*D. cinnamea*, *D. papillomatosis*, *D. aerolata* e *D. aurantiaca*. Além disso, *D. aerolata* e *D. aurantiaca* possuem o sistema de quinona composto por MK-9(H_2) (KÄMPFER *et al.*, 2010, 2012a; RAMMER *et al.*, 2013). O conteúdo de G + C do DNA genômico para *Dietzia* spp varia de aproximadamente 64,7 a 73% (GHARIBZAHEDI; RAZAVI; MOUSAVI, 2013).

1.4.1. Fatores de virulência e potencial patogênico

Até o momento, existem poucos trabalhos tentando identificar os potenciais fatores de patogenicidade das espécies de *Dietzia*. Estas espécies, são considerdas como patógeno oportunistas, capazes de colonizar a pele humana (KOERNER; GOODFELLOW; JONES, 2009) como já descrito no tópico de relevância clínica do gênero.

Chun et al. (1996) mostraram que as paredes das espécies *Dietzia* contêm ácido tuberculosteárico (CHUN, J. *et al.*, 1996), um componente de lipoarabinomanano, que por sua vez faz parte da membrana bacteriana. Este bioproduto é um poderoso estimulador do fator de necrose tumoral-α em macrófagos humanos (COLSTON, 1996; HUNTER; GAYLORD; BRENNANS, 1986; RHOADES; ULLRICH, 2000). . Além disso, algumas *Dietzia* spp são capazes de oxidar uma variedade de *n* -alcanos e iso-alcanos (YASSIN; HUPFER; SCHAAL, 2006a; YUMOTO *et al.*, 2002), assim como *Corynebacterium* spp. (BROADWAY; DICKINSON; RATLEDGE, 1993) o que poderia indicar sinais de patogênese deste microrganismo. Algumas *Dietzia* spp foram claramente implicadas como um patógeno humano potencialmente invasivo (BEMER-MELCHIOR *et al.*, 1999; PIDOUX *et al.*, 2001b; REYES *et al.*, 2006). Embora o significado desses achados permaneça incerto, a busca por *Dietzia* spp. como potenciais patógenos devem ser incluídos em investigações sistemáticas de infecções humanas e em animais quando nenhum outro patógeno bem reconhecido puder ser encontrado.

Ao investigar o potencial patogênico de microrganismos, os reservatórios naturais são considerados como fontes de infecção, sendo a identificação de fatores de patogenicidade os primeiros passos no desenvolvimento de uma compreensão desses patógenos emergentes (KOERNER; GOODFELLOW; JONES, 2009).

1.4.2. Suscetibilidade antimicrobiana

Até o momento, nenhuma averiguação abrangente foi realizada sobre a susceptibilidade a antibióticos para representantes de espécies de *Dietzia*. O que se conhece até o momento é que os isolados de alguns trabalhos realizados mostram que as espécies eram susceptíveis a uma ampla gama de antibióticos, incluindo vancomicina, amicacina, amoxicilina-ácido clavulânico, ampicilina, ceftriaxona, claritromicina, ciprofloxacina, imipenem e linezolida (NIWA *et al.*, 2011; PILARES *et al.*, 2010). *Dietzia maris* foi relatada como suscetível ao aztreonam, ciprofloxacina, mezlocilina, oxacilina, penicilina G, perfloxacina e ticarcilina (NIWA *et al.*, 2011; PILARES *et*

al., 2010), enquanto *Dietzia maris* JCM 6166 *D. natronolimnaea* mostraram-se resistentes ao trimetoprim/sulfametoxazolt por testes de difusão de disco em conformidade com as diretrizes estabelecidas pelo Comité de l' Antibiogramme de la Société Française de Microbilogie (BIZET *et al.*, 1997; NIWA *et al.*, 2011; PILARES *et al.*, 2010). Por sua vez, tem sido relatado que *Dietzia maris* JCM 6166 é altamente suscetível aos antifúngicos azólicos, mas é resistente ao voriconazol (DABBS *et al.*, 2003).

1.4.3. Potenciais dificuldades no isolamento e reconhecimento de espécies de *Dietzia*

Determinados actinomicetos corineformes, como *Rhodococcus equi*, *Dietzia* sp., *Gordonia* sp. e *Nocardia* sp., exibem características morfofisiológicas muito similares, o que pode dificultar a diferenciação entre os gêneros microbianos (BIZET et al., 1997). Muitos laboratórios de microbiologia utilizam essencialmente os testes fenotípicos para a identificação de *R. equi* (PRESCOTT, 1991).

Dietzia e *R. equi* são actinomicetos intimamente relacionados. A diferenciação fenotípica entre *R. equi* e *Dietzia* baseado em sua morfologia de coloração de Gram e aparência de colônia são consideradas imprecisas pois apresentam uma grande semelhança, as propriedades de crescimento e os perfis bioquímicos não é confiável e requer métodos moleculares para diferenciá-las (DE ALEGRÍA PUIG et al., 2017). Além disso, o teste API CORYNE, provavelmente o kit de identificação comercial mais utilizado, falha em identificar *Dietzia* ao nível de gênero (KOERNER; GOODFELLOW; JONES, 2009). As espécies de *Dietzia* crescem relativamente devagar, pode demandar incubação a temperaturas de 37° C, e, portanto, possivelmente podem ser perdidas se as placas de cultura não forem incubadas por pelo menos 48–72 h em uma temperatura apropriada (KOERNER; GOODFELLOW; JONES, 2009; NIWA et al., 2011).

Estudos tem realizado novas reclassificações de algumas de suas bibliotecas de isolados antes identificados como *Rhodococcus* spp. e posteriormente reclassificado como *Dietzia* spp. e outros taxons. O trabalho de (NIWA et al., 2011), reclassificaram uma coleção que possuíam 62 isolados clínicos humanos identificados como de *R. equi*, cuja identificação tinha sido realizada apenas por meio de testes bioquímico. Um estudo de (PILARES et al., 2010), relataram que 15 isolados bacterianos de amostras clínicas, que foram coletadas no Hospital Marqués de Valdecilla, Santander, Espanha, haviam sido identificados como *R. equi* utilizando o teste API Coryne (PILARES et al., 2010). O trabalho de (DE ALEGRÍA PUIG et al., 2017) reclassificaram um

painel de 154 isolados inicialmente identificados como *R. equi* pelo teste API Coryne V2.0, sendo que um total de 144 isolados foram positivos para *R. equi* e os outros 10 isolados foram identificados como pertencentes ao gênero *Dietzia*. Além disso eles avaliaram o desempenho da espectrometria de massas para a diferenciação entre essas espécies usando o equipamento Biotyper e o Viter MS que permitem a diferenciação entre *R. equi* e *Dietzia* spp., mas a identificação de todos *Dietzia* sp. isolados no nível de espécie precisaram do sequenciamento do gene 16S rRNA (DE ALEGRÍA PUIG et al., 2017).

Até o momento não foi encontrando nenhum trabalho descrevendo reclassificação, tão pouco algum problema na identificação entre *Dietzia* e *Corynebacterium pseudotuberculosis.* Entretanto as características celulares micro morfológicas, bioquímicas, de crescimento e quimiotaxonômicas de um microrganismo são frequentemente aplicadas para isolar e identificar esses microrganismos (GHARIBZAHEDI; RAZAVI; MOUSAVI, 2013; KOERNER; GOODFELLOW; JONES, 2009). Porém, essas estratégias nem sempre são adequadas para a classificação e identificação. Assim, análises que incluam abordagens ômicas, como a genômica comparativa, pangenômica, filogenômica, entre outras, vêm sendo continuamente exploradas para esta finalidade. Por exemplo, a análise filogenética usando sequências do gene 16S rRNA, espectrometria de massas, Identidade Média de Nucleotídeos (ANI), Valores digitais Hibridação DNA-DNA (dDDH) têm sido aplicados com sucesso como um relógio molecular para estimar as relações entre bactérias, e mais recentemente se tornou importante como um meio de identificar uma bactéria desconhecida para o gênero ou espécie.

1.5 Genômica

1.5.1 Sequenciamento de genomas

As primeiras tecnologias de sequenciamento de genomas conhecidas na literatura como primeira geração são os métodos de Sanger e o de Maxam e Gilbert, sendo que o mais utilizado é o de Sanger (COSTA, 2015; SANGER; NICKLEN; COULSON, 1977). Em 1987 foi lançado o primeiro sequenciador automático conhecido como AB370, que foi comercializado pela Applied Biosystems, utilizando eletroforese por 6 capilaridade (LIU *et al.*, 2012). Este sendo muito utilizado nos anos seguintes propiciou o início da chamada corrida do sequenciamento genômico, sendo utilizado inclusive no projeto genoma humano (LANDER *et al.*, 2001; LIU *et al.*, 2012).

A partir de 2004 intensificou-se o desenvolvimento das tecnologias NGS, a introdução destas tecnologias reduziu o custo do sequenciamento e aumentou a capacidade de produção de dados por corrida (KAUR & MALIK, 2013), permitindo assim que um genoma bacteriano fosse sequenciado em questão de horas ou poucos dias (KOBOLDT *et al.*, 2013; LOMAN *et al.*, 2012). O primeiro sequenciador com tecnologia NGS a ser lançado foi o 454 da Life Sciences (Roche) em 2005, que utiliza o método de pirosequenciamento (LIU et al., 2012). Um ano após o lançamento do 454, a Solexa lança o seu sequenciador, que foi adquirido pela Illumina no ano seguinte, o qual utiliza o sequenciamento por síntese. A próxima tecnologia de sequenciamento foi o SOLiD, o qual utiliza a detecção por ligação de oligonucleotídeos e um código de cores para identificar as bases desenvolvido pela Applied Biosystems em 2007 (atualmente pertence à Thermo Fisher Scientific) (VAN DIJK *et al.*, 2014). Em 2010, surge o PacBio RS (Pacific Biosciences) que realiza o sequenciamento pelo processo conhecido como Single Molecule, Real-Time (SMRT) (PACBIO SYSTEMS, 2016). No ano de 2015 o Oxford Nanopore Technologies MinION foi lançado, este tem o tamanho de um pendrive o que possibilita ser facilmente carregado para fazer o sequenciamento do DNA genômico no campo ou em qualquer lugar no momento da coleta do mesmo (OXFORD NANOPORE TECHNOLOGIES, 2016). O grande diferencial dessas tecnologias é a capacidade de sequenciamento por molécula única e o tamanho das reads (VAN DIJK et al., 2014).

As tecnologias NGS têm contribuído para a realização de diferentes estudos dentro da biologia, tais como: sequenciamento de genoma completo (WGS), ressequenciamento genômico, metagenômica, desenvolvimento de marcadores (geralmente genes alvos para o tratamento de doenças/infecções), transcriptômica, análises epigenéticas (regulações celulares) e estudos de interação proteína-DNA (KAUR & MALIK, 2013).

1.5.2 Genômica comparativa

Estudos de genômica comparativa iniciaram-se em meados de 1995, quando o grupo de Craig Venter publicou o sequenciamento de dois genomas bacterianos (COSTA, 2015; FLEISCHMANN *et al.*, 1995). A genômica comparativa nos permite avaliar relacionamentos evolutivos entre diferentes indivíduos, através da comparação entre os genomas de dois ou mais

organismos (COSTA, 2015). A quantidade de dados genômicos depositados nos bancos de dados públicos vêm crescendo, devido ao avanço das novas plataformas de sequenciamento, que vem auxiliando para o desempenho de estudos na área da genômica comparativa (Field et al., 2006; Hu et al., 2011).

A provável função dos genes que se diferenciam entre as linhagens e os estudos direcionados à composição do genoma, nos permitem identificar mecanismos importantes para sobrevivência e a adaptação de microrganismos nos seus respectivos hospedeiros, bem como, em seus nichos preferenciais. As pesquisas de genômicas comparativas fornecem detalhes adicionais, que podem auxiliar na descoberta de genes alvos de interesse biomédico, ambiental, biotecnológico etc. (SHEPPARD; GUTTMAN; FITZGERALD, 2018; TOFT; ANDERSSON, 2010). Através da comparação de várias linhagens de uma mesma espécie ou de espécies diferentes, faz se possível reconhecer similaridades e diferenças existentes entre os genomas, bem como se possível esclarecer quais sequências são capazes de divergir em mudanças fenotípicas nos organismos e elucidar os mecanismos de virulência entre organismos patogênicos (HU *et al.*, 2011).

1.5.3. Filogenética e Filogenômica

Análise filogenética pode ser considerada como um sistema de classificação que leva em conta a evolução entre as espécies a partir de sequências de ácidos nucleicos ou aminoácidos (CALDART *et al.*, 2016). No entanto, os métodos baseados em um único ou um conjunto de genes universalmente conservados, como genes codificadores de proteínas ribossômicas 16S rRNA, muitas vezes não são aplicáveis a genomas incompletos (por exemplo, os genes não são montados), e esses genes geralmente mostram maior conservação de sequência do que a média do genoma. Consequentemente, a análise de genes universais não fornece resolução suficiente no nível de espécie, e frequentemente resultou na falta de descontinuidades genéticas claras entre táxons intimamente relacionados (JAIN *et al.*, 2018).

A disponibilidade crescente de sequências de genoma completo hoje em dia, permite uma inferência filogenética de organismos em escala de genoma chamada filogenômica (EISEN, 1998). A filogenômica pode ser caracterizada por ser uma análise comparativa em escala genômica para o estudo dos relacionamentos evolutivos entre espécies, buscando uma melhor compreensão dos mecanismos de evolução molecular (CHAN; RAGAN, 2013; EISEN, 1998). A filogenômica é

indispensável na avaliação de diversas hipóteses biológicas, por exemplo, a distribuição e disseminação da patogenicidade bacteriana, a convergência ou divergência da função gênica, a origem das organelas ou a resolução da árvore (ou rede) da vida. Geralmente utilizar o genoma inteiro ou uma grande quantidade de genes torna a filogenia mais confiável do que utilizando um ou alguns genes podendo, assim, obter informação detalhada das diferenças entre os genomas e não simplesmente uma classificação geral destes (ÅGREN *et al.*, 2012).

1.5.4. Similaridade nucleotídica

Um trabalho primordial na avaliação dos limites das espécies é a estimativa da relação genética entre dois genomas. Nos últimos anos, a identidade de nucleotídeos médios de todo o genoma (ANI) surgiu como um método robusto para esse trabalho (JAIN *et al.*, 2018). ANI representa a identidade nucleotídica média de todos os genes ortólogos compartilhados entre quaisquer dois genomas, e oferece resolução robusta entre cepas da mesma espécie ou de espécies intimamente relacionadas considerando uma identidade ≥95% como evidência de organismos pertencentes à mesma espécie (RODRIGUEZ-R *et al.*, 2018). ANI apresenta algumas vantagens importante em relação a outras técnicas já existentes como, DDH, sequenciamenteo 16S, pois apresenta maior resolução entre genomas intimamente relacionados, pode ser estimado entre sequências genômicas incompletas recuperadas do ambiente, usando técnicas metagenômicas, ou de célula única que não codificam genes universalmente conservados, mas codificam pelo menos algumas centenas de genes compartilhados, assim expandindo bastante o número de sequências que podem ser estudados e classificados em comparação com uma abordagem universal baseada em genes (AUCH *et al.*, 2010; JAIN *et al.*, 2018; RODRIGUEZ-R *et al.*, 2018).

ANI reflete o conceito microbiológico tradicional de relação de hibridização DNA-DNA (dDDH) para definir espécies, e apresentou um potencial para substituir a hibridização DNA-DNA como medida padrão de parentesco, pois é mais fácil de estimar e representa dados portáteis e reprodutíveis. A hibridização DNA-DNA é um método de laboratório experimental, atualmente ainda usado como padrão-ouro taxonômico para delimitação de espécies em *Archaea* e *Bacteria* (MEIER-KOLTHOFF *et al.*, 2013a). Se o DNA genômico de dois organismos respectivos revela uma similaridade DDH abaixo de 70% este é o principal argumento para considerá-los como espécies distintas e vice-versa. No entanto, existem calculadoras in-silico capazes de simular com sucesso os resultados de hibridização experimental, fornecendo correlações mais altas com um

conjunto empírico de valores de DDH do que métodos anteriores baseados em sequência de genoma. Além disso, estes métodos são capazes de lidar com genomas bastante incompletos, também aplicando limite mínimo de 70% para se considerar um organismo da mesma espécie. (AUCH *et al.*, 2010; MEIER-KOLTHOFF *et al.*, 2013a).

1.5.5 Pangenômica

O conceito de pangenoma se refere ao conjunto de todas as famílias de genes que se encontram nos genomas de um grupo de linhagens. O pangenoma é subdividido em três grupos. O "genoma central", que configura os genes essenciais que por sua vez estão presentes entre todas as linhagens; o "genoma acessório" que inclui os genes que as bactérias necessitam para sobreviver em um ambiente específico são compartilhados entre duas ou mais cepas; ainda temos os genes espécie-específicos pertencentes a uma única linhagem. (MIRA *et al.*, 2010; TETTELIN *et al.*, 2008). Desta forma, nos estudos pangenômicos podem-se caracterizar o pangenoma de uma espécie como "aberto" ou "fechado". O pangenoma "aberto" indica a possibilidade do aumento de novos genes no genoma core à medida que um novo genoma é inserido na análise. O mesmo não acontece no pangenoma "fechado", pois ele é caracterizado por apresentar pouca ou nenhuma aquisição de novos genes (MUZZI; MASIGNANI; RAPPUOLI, 2007; TETTELIN *et al.*, 2008).

1.5.6 Plasticidade genômica

Estudos de genômica comparativa também têm contribuído na melhor compreensão de características como evolução e patogenicidade bacteriana. O genoma bacteriano é um agrupamento de regiões estáveis e instáveis. A variação na estrutura do genoma, ou plasticidade genômica, pode se dar pelos processos de duplicações, inversões, translocações, aquisição (inserção) e perda (deleção) de sequências (PATEL, 2016). O estilo de vida e o tamanho do genoma estão diretamente relacionados. O genoma bacteriano em organismos patogênicos pode ser menor que organismos não patogênicos devido a sua habilidade de utilizar metabólitos produzidos pelo hospedeiro (BARBOSA *et al.*, 2014; OCHMAN; DAVALOS, 2006; TOFT; ANDERSSON, 2010).

Outra característica que auxilia no estudo da plasticidade genômica são a predição de Ilhas genômicas (GEIs), que são regiões do DNA que podem ter sido adquiridas por transferência

horizontal (THG), contribuindo para a plasticidade, evolução, e a adaptação de microrganismos. GEIs podem ser classificadas em quatro diferentes categorias: Ilhas de Patogenicidade (PAIs), quando elas carregam genes de virulência (GAL-MOR; FINLAY, 2006; SCHMIDT; HENSEL, 2004; SOARES et al; 2016) Ilhas metabólicas (MIs), quando possuem genes associados à biossíntese de metabólitos (SOARES *et al.*, 2016; TUMAPA *et al.*, 2008); ilhas de resistência (RIs), quando possuem genes que codificam a resistência, tipicamente contra antibióticos ; (KRIZOVA; NEMEC, 2010; SOARES *et al.*, 2016); e por fim as ilhas simbióticas (SIs), que facilitam as associações simbióticas do hospedeiro com outros microrganismos) (BARCELLOS *et al.*, 2007; SOARES *et al.*, 2016).

2. JUSTIFICATIVA DA PESQUISA DE TESE

Corynebacterium. pseudotuberculosis é o causador da Linfadenite Caseosa que causa grandes perdas econômicas para pequenos e grandes produtores de animais ruminantes (ALVES; PINHEIRO; PIRES, 1997). *C. pseudotuberculosis* pode ser dividida em dois biovares: *Equi,* encontrados majoritariamente em equinos e bovinos, são nitrato redutoras e *Ovis,* isolado principalmente de caprinos e ovinos, não reduz nitrato. No entanto, muitas vezes se torna difícil diferenciar microbiologicamente estes dois biovares. Nesse contexto, vários testes de diagnóstico sorológico, bioquímico e molecular têm sido desenvolvidos para superar o problema de identificação da infecção por *C. pseudotuberculosis.* No entanto, muitos destes testes apresentam limitações, necessitando-se, dessa forma, que novos métodos sejam implementados (ÇETINKAYA *et al.*, 2002; MUCKLE; GYLES, 1982). O MALDI-TOF vem se mostrando uma excelente ferramenta para identificação de novos isolados bacterianos. Assim, devido à dificuldade de se diferenciar isolados de *C. pseudotuberculosis* quanto ao seu biovar, nós testamos o MALDI TOF para ver a precisão da técnica na diferenciação dos biovares equi e ovis de *C. pseudotuberculosis*, devido ao seu custo menor, após a aquisição do equipamento, agilidade metodológica, e às limitações que muitos testes convencionais apresentam na diferenciação dos biovares desta espécie. Sendo assim, esse trabalho vai fornecer mais uma técnica adequada para a identificação rotineira de microrganismos com baixo custo, fornecendo resultados precisos e rápidos.

Por outro lado, durante a realização dos experimentos com o MALDI para diferenciação de *C. pseudotuberculosis,* nosso laboratório recebeu dois isolados bacterianos removidos do abscesso de uma égua em Reggio Calabria, Itália, identificados primeiramente como *Corynebacterium amycolatum e C. pseudotuberculosis.* No animal, os locais das lesões, tamanho e aparência macroscópica pareciam lesões típicas da infecção por *C. pseudotuberculosis*. No primeiro momento, a identificação dos isolados como *Corynebacterium amycolatum* e *Corynebacterium pseudotuberculosis* foram realizados utilizando o teste API- Coryne com teste de maltose e leucina-β-naftilamida atípicas para o primeiro isolado e urease atípico para o segundo isolado.

Com base nesses resultados, foi administrado um ciclo de antibioticoterapia com rifampicina, duas vezes ao dia, por via oral por três semanas. Devido a esse tratamento foi observado um desaparecimento total dos nódulos menores e redução do tamanho das massas maiores, juntamente com uma recuperação geral e melhora das condições corporais do animal. Oito meses após o tratamento, os nódulos, embora menores, ainda estavam presentes e o quadro clínico do animal se deteriorou: anorexia, letargia, perda de peso, desidratação e hipertermia. Sendo assim, realizamos o PCR quadruplex que é comumente utilizado em nosso laboratório para identificação de *C. pseudotuberculosis,* e a técnica de MALDI TOF. O PCR quadruplex foi negativa para *C. pseudotuberculosis* e a técnica de MALDI para o isolado 1 foi identificado como *Arthrobacter creatinolyticus*, recentemente reclassificada como *Glutamicibacter creatinolyticus*. Por sua vez, o isolado 2 foi identificado como *Dietzia* pelo MALDI Biotyper (Bruker Daltonics, Bremen, Alemanha), indicando uma identificação confiável apenas em nível de gênero. Com base no diagnóstico bacteriológico, a escolha de uma terapia alvo final com Sulfametoxazol + Trimetoprima (10 g/100 kg por via oral no primeiro dia e depois 5 g/100 kg por 10 dias) foi realizada, resultando em extensas reduções no número e tamanho das lesões detectadas e a égua fazendo progressos clínicos significativos.

Baseado nesses achados, nós então decidimos trabalhar durante o mestrado com a espécie identificada como *Arthobacter creatinolyticus,* o qual realizamos toda a caracterização da espécie, sendo está a primeira espécie isolada de um equino, realizamos a genômica comparativa do gênero e a publicação. Por conseguinte, a segunda amostra identificada com o possível gênero *Dietzia* está sendo trabalhada nesse doutorado, no qual então buscamos realizar a caracterização dessa nova espécie, dando continuidade a este trabalho de pesquisa de Tese, em que saímos da identificação

de biovares de *Corynebacterium pesudotuberculosis* para uma identificação imprecisa de um isolado do gênero *Dietzia*.

3. CAPÍTULO I

3.1 OBJETIVOS

3.2 OBJETIVO GERAL

- Identificar biovares Equi e Ovis de *C. pseudotuberculosis* utilizando a técnica de MALDI TOF;

3.3 OBJETIVOS ESPECÍFICOS

- Selecionar 230 cepas de biovares equi e ovis de *C. pseudotuberculosis* isoladas de uma variedade de hospedeiros naturalmente infectados;
- Extrair proteínas de quatro linhagens duas biovar ovis e duas biovar equi para utilização como referência dos Perfis de espectros principais ((Main Spectra Profiles) (MSPs));
- Classificar os biovares usando o método Máquina de vetor de suporte ((Support Vector Machine) (SVM)) utilizando o software Bionumerics e método Incorporação de vizinhos estocásticos distribuídos ((t-Distributed Stochastic Neighbor Embedding) (t-SNE)), utilizando a linguagem R e o pacote Rtsne;
- Classificar os Biovares usando o Índice Geral de Relação do Genoma (OGRI) para 20 linhagens de *C. pseudotuberculosis*;

3.4. ARTIGO CIENTÍFICO 1: Exploring the MALDI Biotyper for the identification of *Corynebacterium pseudotuberculosis* biovar Ovis and Equi

Publicado em: Journal of the American Society for Mass Spectrometry; Fator de impacto (atual): 3.262.

- **DOI:** https://doi.org/10.1021/jasms.2c00174

Corynebacterium pseudotuberculosis é uma bactéria patogênica intracelular facultativa, classificada em em dois biovares Ovis e Equi que pode infectar animais silvestres e domésticos e é potencialmente zoonótica para humanos. *C. pseudotuberculosis* podem causar doenças como linfadenite caseosa em caprinos e ovinos, linfangite ulcerativa em equinos, dermatose edematosa em búfalos e acne contagiosa em bovinos, essas doenças podem estar relacionadas perdas econiomicas significativas. Existem diversos testes sorológicos e moleculares para o diagnóstico laboratorial de doenças causadas por *C. pseudotuberculosis*, entretanto a identificação do patógeno e a diferenciação de biovares podem ser demoradas, caras e confusas em comparação com outras bactérias.

Portanto, este artigo científico relata a utilização da técnica de MALDI- TOF na tentatiza de otimizar a identificação e diferenciação de biovares de *C. pseudotuberculosis*, juntamente com com as análises de Support Vector Machine, o método t-Distributed Stochastic Neighbor Embedding e abordagem baseada no Índice Geral de Relacionamento do Genoma. Para realização do trabalho utilizamos 230 cepas de *C.pseudotuberculosis*.

As análises realizadas classificaram corretamente as cepas de *C. pseudotuberculosis* usando o MALDI Biotyper, provando que esta tecnologia é altamente sensível. Entretanto para a diferenciação dos biovares Equi e Ovis, se faz necessária análises estatísticas complementares para a ccorreta diferenciação.

O nosso trabalho contribuirá para um correto e rápido diagnóstico por doenças causadas por *C. pseudotuberculosis*, complementando os métodos já existe para a espécie. Além disso, o trabalho será possível a diferenciação de *C. pseudotuberculosis* em níveis biovares.

ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Research Article

Exploring the MALDI Biotyper for the Identification of *Corynebacterium pseudotuberculosis* biovar Ovis and Equi

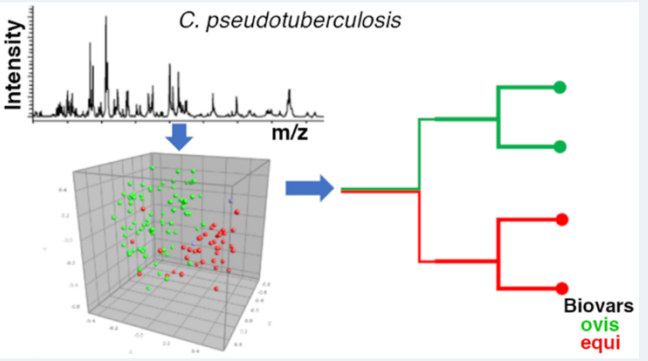
Roselane Gonçalves dos Santos,[†] Nubia Seyffert,[†] Elaine M. S. Dorneles, Eric R. G. R. Aguiar, Carolina P. Ramos, Dionei J. Haas, Gabriella B. N. Assis, Ricardo Dias Portela, Aristóteles Goes-Neto, Luis G. C. Pacheco, Henrique C. P. Figueiredo, Thiago de Jesus Sousa, Sandeep Tiwari, Arun Kumar Jaiswal, Andrey P. Lage, Thiago L. P. Castro,* and Vasco Azevedo*

Cite This: https://doi.org/10.1021/jasms.2c00174

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Biochemical, serological, and molecular methods have been developed for the laboratory diagnosis of diseases caused by *C. pseudotuberculosis* (CP), but the identification of the pathogen and biovars differentiation may be time-consuming, expensive, and confusing compared with other bacteria. This study aimed to evaluate MALDI Biotyper and Overall Genome Relatedness Index (OGRI) analysis to optimize the identification and differentiation of biovars of *C. pseudotuberculosis*. Out of 230 strains isolated from several hosts and countries, 202 (87.8%) were precisely classified using MALDI Biotyper and the BioNumerics platform. The classification accuracies for the Ovis and Equi biovars were 80 (88.75%) and 82 (92.68%), respectively. When analyzing a sampling of these strains by Average Nucleotide Identity based on BLAST and TETRA analyses using genomic sequence data, it was possible to differentiate 100% of the strains in Equi and Ovis. Our data show that MALDI Biotyper and OGRI analysis help identify *C. pseudotuberculosis* at the species and biovar levels.



KEYWORDS: *MALDI-TOF, C. pseudotuberculosis, Equi and Ovis biovars, Overall Genome Relatedness Index, Bionumerics software*

INTRODUCTION

Corynebacterium pseudotuberculosis is a facultative intracellular pathogenic bacterium that can infect wild and domestic animals and is potentially zoonotic to humans.[1] *C. pseudotuberculosis* (CP) isolates are catalase positive, oxidase negative, and β-hemolytic and can be classified into one of two biovars: Ovis or Equi.[2] The most common diseases caused by CP include caseous lymphadenitis (CLA) in goats and sheep, ulcerative lymphangitis in horses, oedematous skin disease in buffalo, and contagious acne in cattle.[3−6] The diseases caused by CP are widely distributed worldwide and responsible for significant economic losses and the risk of human infection related to occupational activities.[7]

Clinical diagnosis of CLA is mainly through macroscopic observation of lesions, such as caseous necrosis in the external lymph nodes of sheep and goats.[8] However, CLA may also affect the internal lymph nodes without causing evident clinical signs of the disease, hindering the identification of infected animals.[9] The diseases caused by CP in cattle and horses have different clinical manifestations and may present caseous lesions.[6,10]

The laboratory test considered the gold standard for detecting CP in clinical samples consists of isolating the agent by bacteriological culture and biochemical testing.[11] Although biochemical tests yield dependable identification of CP at the species level, there may be misidentification of biovars using the nitrate reduction assay.[12−14] This assay has long been used to distinguish the biovar Equi, which reduces nitrate to nitrite, from the biovar Ovis, which does not.[2,13]

In 2002, a PCR assay targeting the *16S rRNA* gene was proposed to identify CP in clinical isolates.[15] As this test was not able to differentiate CP from *Corynebacterium ulcerans*, a bacterium that is frequently associated with abscess formation in small ruminants, Pacheco et al.[16] developed a multiplex PCR (mPCR) test to successfully identify CP isolates using, as targets, the *16S rRNA* gene, alongside the *rpob* and *pld* genes.[16]

Received: June 24, 2022
Revised: October 4, 2022
Accepted: October 4, 2022

Phenotypic and molecular methodologies can differentiate the Ovis and Equi biovars of CP, with the nitrate reductase assay being the gold standard method.[17] The standard classification methods for CP are time-consuming, require significant manipulation of a potentially zoonotic organism, and may lead to highly subjective results. However, the potential of matrix-assisted laser desorption ionization (MALDI)-time-of-flight (TOF) mass spectrometry (MS) for CP classification has not yet been well explored. MALDI Biotyper has emerged as a powerful system for bacterial typing, improving the identification of microorganisms impacting public and animal health and providing important information for disease prevention and control.[15]

Genomic methods have also contributed to identifying and classifying microorganisms by inferring phylogenetic relationships. The Overall Genome Relatedness Index (OGRI) is a term coined for the first time by Chun and Rainey[18] and represents any measurements that indicate how similar the two sequences of the genome are. There are many algorithms for calculating OGRI values, but the most widely used algorithm for taxonomic studies is the Average Nucleotide Identity (ANI).[19] ANI measures the average nucleotide identity based on BLASTn[20] between two genome sequences where only matches with at least 30% general sequence identity are considered.[19]

Although the *in vitro* assay for nitrate reductase consists of the gold standard method, the MALDI Biotyper and OGRI analysis are likely to provide more accurate classifications of CP biovars. This study evaluated these omics methodologies to rapidly identify CP isolates at the species level and reliably classify the Ovis and Equi biovars.

MATERIAL AND METHODS

Bacterial Strains. All CP strains used in the present study are maintained in the culture collection of the Laboratory of Applied Bacteriology and the Laboratory of Cellular and Molecular Genetics, Federal University of Minas Gerais. These laboratories have previously classified the biovars of all strains using the nitrate reductase activity assay. In total, 111 strains of the biovar Ovis and 119 strains of the biovar Equi were selected to perform MALDI TOF-MS using the MALDI Biotyper (Bruker Daltonics). The strains were isolated from various naturally infected hosts, including horses, cattle, buffalo, goats, sheep, and humans (Table S1).

MALDI Biotyper Analysis. Main Spectra Profiles (MSPs) were created using two CP strains from biovar Ovis and two from biovar Equi as the references. Spectra of strains 1002, FRC41, 258, and E19 were initially evaluated according to Alatoom et al.[21] Briefly, the strains were subcultured on BHI agar plates (HiMedia, Mumbai, India) and incubated for 48 h at 37 °C. A bacterial colony was collected for each strain separately and transferred into a tube with 300 μL of deionized water delivered from Milli-Q Gradient A10 (Millipore, Burlington, VT). The samples were homogenized for 10 s, and 1200 μL of 70% ethanol (Merck, Darmstadt, Germany) was added for centrifugation at 13.000 × g for 2 min at 25 °C. Carefully, the supernatants were discarded, and the samples were centrifuged again under the same conditions. The precipitates were dried at 25 °C, and 70% formic acid (Merck, Darmstadt, Germany) was added sufficiently to homogenize and lyse the bacterial mass. After adding one volume of 100% acetonitrile (Merck, Darmstadt, Germany), the bacterial debris was pelleted by centrifugation for 2 min at 13,000*g* at 25 °C, and then the protein-containing supernatant was retrieved. The spots were dried for 10 min, and 1 μL of the α-cyano-4-hydroxycinnamic acid (CHCA) (Merck, Darmstadt, Germany) was added to the sample prior to loading the MALDI Biotyper assay plate (Bruker Daltonics, Bremen, Germany).[22] The spectra of all colonies were acquired three times and analyzed in FlexAnalysis software version 3.4 (Bruker Daltonics, Bremen, Germany) to assess reproducibility levels. Next, the spectra of each strain were uploaded to the MALDI Biotyper software version 3 (Bruker Daltonics, Bremen, Germany) and assembled to generate one Main Spectra Profile (MSP) using the BioTyper MSP creation standard method. All steps were performed according to the manufacturer's recommendations.

All other 226 CP strains were thawed, subcultured on BHI agar plates, and incubated for 48 h at 37 °C. Colonies of every bacterial strain were spotted on the target steel plate using sterile toothpicks. For each colony, 1 μL of formic acid 70% (Merck, Darmstadt, Germany) and 1 μL of the MALDI-TOF MS matrix, consisting of a saturated solution of CHCA (Merck, Darmstadt, Germany), were applied to the spot and allowed to air dry. A FlexControl MicroFlex LT mass spectrometer (Bruker Daltonics, Bremen, Germany) was used to verify the reproducibility of the identifications. Spectra were acquired using a FlexControl MicroFlex LT mass spectrometer (Bruker Daltonics, Bremen, Germany) with a 60-Hz nitrogen laser, in which up to 240 laser shots were fired in spiral movements to collect 40 shot steps for each strain spot, and the isolates were compared against each other to determine specific peaks. Spectra were recorded within the range of 2−20 kDa, and the bacterial species were confirmed by comparison with the mass-spectrum library using FlexAnalysis (Bruker Daltonics, Bremen, Germany) under standard conditions. All steps were performed according to the manufacturer's recommendations.[23,24]

Biovar Classification Using the Support Vector Machine (SVM) Method. Identification of the classes of peaks obtained from the MALDI Biotyper analysis capable of differentiating the biovar Ovis and the biovar Equi was performed using the Bionumerics software version 7.5 (Applied Maths, Belgium). The average similarity of the raw spectra from the biological replicates was calculated with Spearman's correlation coefficient. The spectra of the strains were divided into two databases; one was used to identify the peaks able to differentiate Ovis and Equi, and the other was used to evaluate the accuracy of this methodology. A reference set, which was further used to identify our unknown strains, was created using 70% of the isolates (n = 162), selected randomly within each biovar (82 Equi and 80 Ovis). For these strains, peak class matching was performed using a constant tolerance of 1.9, a linear tolerance of 550, and a peak detection rate of 10%. Then the selection of peaks showing significantly different intensities ($p \leq 0.05$) between Ovis and Equi was performed using the Mann−Whitney test. Next, the Principal Component Analysis (PCA) was used to reduce the complexity of the data set. This reference set was then trained by the linear Support Vector Machine classifying algorithm (SVM) to determine the ability of the peak classes to discriminate the different CP biovars. An internal method validation was carried out, considering all spectra used in training as unknown and classifying them with the trained SVM. The 30% of the remaining spectra were used in external validation by

https://doi.org/10.1021/jasms.2c00174
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

Table 2. *C. pseudotuberculosis* Strains Selected for OGRI Analysis

CP strains	Biovar	Host	Country of Origin	MALDI Biotyper score value[a]	Genomes BioProject
ATCC1941	Ovis	sheep	USA	2.215	PRJNA382169
MIC-6	Ovis	sheep	Brazil	2.583	PRJNA354849
12C	Ovis	sheep	Brazil	2.581	PRJNA282769
PAT10	Ovis	sheep	Argentina	2.332	PRJNA61115
N1	Ovis	sheep	Brazil	2.517	PRJNA301685
1002	Ovis	goat	Brazil	2.425	PRJNA40687
MEX1	Ovis	goat	Mexico	2.552	PRJNA348354
29156	Ovis	cattle	Israel	2.341	PRJNA273364
I19	Ovis	cattle	Israel	2.358	PRJNA52845
267	Ovis	llama	Peru	2.648	PRJNA73515
MEX30	Equi	horse	Mexico	2.528	PRJNA343017
MEX31	Equi	horse	Mexico	2.361	PRJNA341961
258	Equi	horse	Belgium	2.415	PRJNA157069
316	Equi	horse	USA	2.375	PRJNA71591
E19	Equi	horse	Chile	2.252	PRJNA271886
MB11	Equi	horse	USA	2.606	PRJNA256958
43	Equi	buffalo	Egypt	2.446	PRJNA317386
I37	Equi	cattle	Israel	2.293	PRJNA340871
262	Equi	cattle	Belgium	2.452	PRJNA271872
162	Equi	camel	UK	2.545	PRJNA89445

[a]Score values/best match (2.3−3.0 = highly probable species identification: 2.0−2.29 = secure genus identification and probable species identification).

comparing the biovars predicted by the SVM with the biovars classified according to biochemical testing.

Biovar Classification Using the t-Distributed Stochastic Neighbor Embedding (t-SNE) Method. We reduced the sampling to 10 strains of CP Equi and 10 strains of CP Ovis to perform the t-SNE analysis using the spectral data set (Table 2). One combined spectrum including strain replicates was generated for every CP biovar, and the 2,000 peptide/protein peaks with higher intensities were considered for the t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis. The R language and *Rtsne* packages[25] were used. The t-SNE algorithm was applied to two different matrix data files, one containing information ordered from higher to lower peptide/protein intensity and the other ordered from a higher to lower m/z ratio. The t-SNE parameters using the intensity-based ordering included *dims*, *perplexity*, and *max_iter* equal to 9, 30, and 500, respectively. The t-SNE parameters used with the m/z ratio-based ordering included *dims*, *perplexity*, and *max_iter* equal to 10, 20, and 500, respectively.

Biovar Classification Using OGRI. To construct one dendrogram for the 20 CP strains used in the t-SNE analysis (Table 2), we used an approach based on the Overall Genome Relatedness Index (OGRI). First, both Average Nucleotide Identity based on BLAST (ANIb) and Tetra-Nucleotide Analysis (TNA) calculations were performed in an all-versus-all fashion for 20 complete genomes recovered from the database of the National Center for Biotechnology Information (NCBI). The ANIb and TNA values were concatenated into one single matrix to generate a *Euclidean* distance matrix and one dendrogram with the *ward.D* hierarchical clustering method. The clades' bootstrap values (confidence) were calculated using the *pvclust* package.[26] The number of confident clades (k) was defined based on the number of clades in which the bootstrap confidence was higher than 80%. Finally, the dendrogram was drawn using the *factoextra*[27] and *ggdendro*[28] packages, with 2 being the minimum number of clusters (k = 2). All analyses were performed using the R language.

RESULTS AND DISCUSSION

Genus and Species Identification with the MALDI Biotyper. All 230 previously confirmed clinical CP isolates were correctly classified at both the genus and species levels using the MALDI Biotyper, proving this technology to be highly sensitive for detecting CP. The MALDI Biotyper provided highly probable species identification (scores between 2.3 and 3.0) for 202 CP isolates (87.82% of the total), while only 28 isolates (12.17%) presented secure genus and probable species identification (scores between 2.0 and 2.29) (Table S1). According to the manufacturer's recommended parameters, scores between 1.7 and 2.0 indicate genus-level identification, and those below 1.7 indicate unreliable identification. However, the MALDI Biotyper FlexAnalysis software could not automatically distinguish Ovis and Equi, which made us customize the analysis of the spectra to check whether these biovars could be reliably classified using this data set.

The gold standard method for CP identification relies on the biochemical testing of isolated bacterial colonies. When performed with the aid of commercialized biochemical panels such as the API Coryne System, testing of isolated colonies takes approximately 30 h.[29] An alternative method for identifying CP relies on the multiplexed detection of three validated gene markers, using the PCR technique and genomic DNA isolated from bacteria.[14] Both biochemical testing and PCR-based detection are expensive, time-consuming, and require highly technical skills.[16,30]

The MALDI Biotyper platform is well suited to routine low-cost identification of microorganisms, providing accurate results in only a few minutes.[29,31−34] Using one 96-well target plate, the MALDI Biotyper identifies dozens of bacterial colonies in less than 2 h.[35] In this work, species confirmation could be performed in as little as 7 min, from the loading of

https://doi.org/10.1021/jasms.2c00174
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

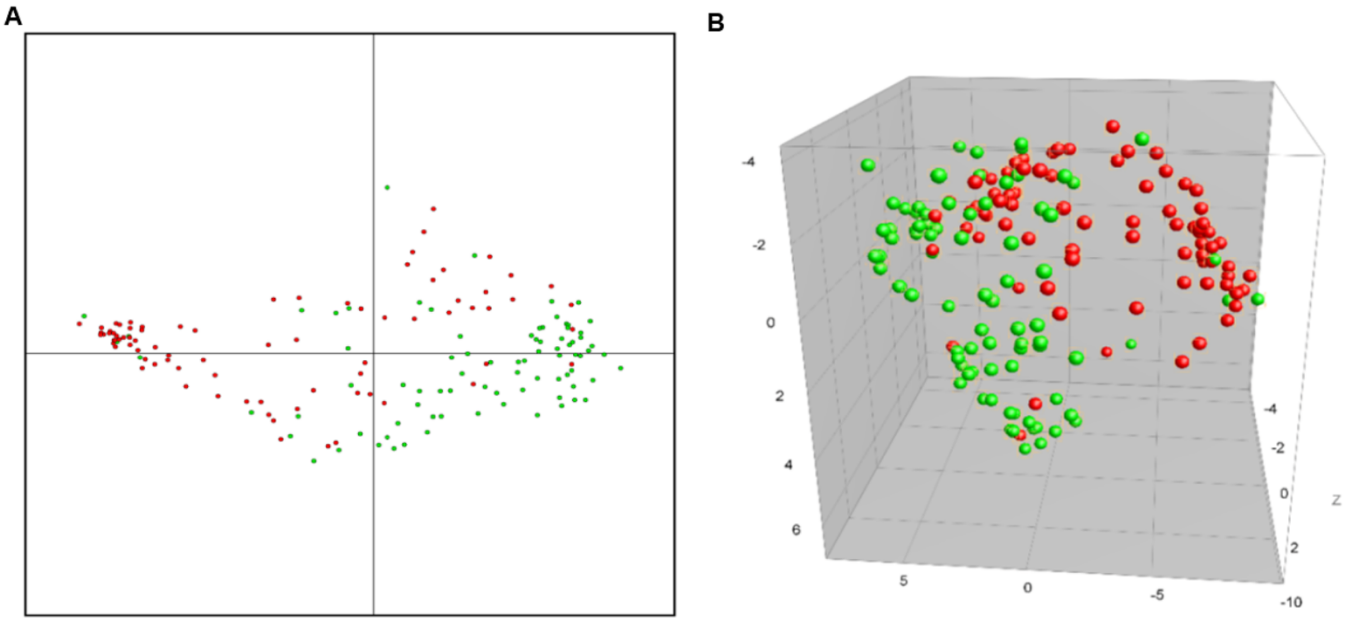


Figure 1. Principal Component Analysis of the 77 peak classes of *C. pseudotuberculosis* for the distinction of Ovis and Equi biovars considered statistically different by the Mann–Whitney test after Bionumerics analysis. One hundred sixty-two strains were evaluated, and the results are shown in a (A) two-dimensional image and (B) a three-dimensional image. The green dots correspond the biovar Ovis, and the red dots correspond to the biovar Equi.

colony samples until the result. In addition to the low 12.17% scores between only 2.0 and 2.29, which indicate probable species identification, the correct identification of all strains demonstrated how sensitive the MALDI Biotyper is for detecting CP.

Biovar Classification Using the SVM Method. The peptide/protein spectra data set for 162 randomly selected CP isolates were further evaluated to correlate the peak intensities with biovar distribution among samples. In total, 77 peptides/proteins peaks presented statistically relevant differential intensities between Ovis and Equi (Box S1). The most important peaks in the spectrum data for differentiation of the biovar Equi were 5466.35, 4015.1, 5984.06, 10932.56, and 3652.52, while those for the biovar Ovis were 3566 and 7130.81. The PCA results for these peaks are shown in Figure 1. Following the SVM algorithm training, the internal validation of the reference set presented 90.79% accuracy, with 88.75% true positives for the biovar Ovis strains and 92.68% true positives for the biovar Equi strains. The external validation resulted in biovar misclassification for only 5 out of the 67 CP strains evaluated, with 92.65% accuracy (Table S1). The classification of biovars might not interfere with the actions taken against CP in a clinical case; however, this classification may support epidemiological surveys in locations where the Ovis and Equi biovars coexist. The biovar classification in CP generally takes into consideration the host from which the bacterium was isolated—Ovis tends to infect goats, sheep, and camels, while Equi tends to infect horses, cows, and buffaloes—and, more importantly, the positive nitrate reductase activity restricted to the biovar Equi.[36]

Although the nitrate reductase assay remains broadly accepted, one study has pointed out exceptions to the rule that requires careful investigation.[14] The study initially proposed a multiplex PCR assay, including detecting the nitrate reductase gene (*narG*) to provide a facilitated classification assay over traditional biochemical testing. The authors have detected the *narG* gene in a few Equi strains that yielded negative nitrate reductase activity in biochemical tests. Also, the *narG* gene was not detected in a few strains that were able to reduce nitrate to nitrite.

Biovar Classification Using tSNE and OGRI. Peptide/protein spectra for the tSNE analysis included 10 known Ovis and 10 Equi strains to assess the effectiveness of this method in front of the OGRI method, which relies on the genomes of strains. When these analyses were conducted, a small number of genomes related to the CP strains submitted to the MALDI Biotyper in this work had been completely sequenced and available in the NCBI database.

First, the dendrogram generated for the concatenated ANIb and TNA results revealed two well-defined bacterial groups (Figure 2A). One of the groups was defined by the Ovis strains with bootstrap confidence of 100%, and the other group was defined by the Equi strains with bootstrap confidence of 80%. Within the Equi group, three strains (CpI37, Cp162, and Cp262) tended to group separately with bootstrap confidence of 80%. Once the dendrogram validated the previous biovar classifications, we proceeded to the tSNE analysis. When the spectra data set was ordered from higher to lower m/z ratio, the Ovis and Equi strains did not form two entirely distinct groups using tSNE (Figure 2B). When the spectra data set was ordered from higher to lower peak intensity, the tSNE analysis tended to form one larger and one smaller group of Ovis strains separated by one larger and one smaller group of Equi strains (Figure 2C).

Exploring the complete genome sequences for dozens of CP strains, another study attempted to find additional molecular markers for the Ovis and Equi biovars. Despite the high genomic conservation reported for CP, the authors identified 27 genes undergoing distinct adaptative changes in Ovis and Equi.[37] However, CP biovar classification had not been assured using the genome information, and the present study

https://doi.org/10.1021/jasms.2c00174
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

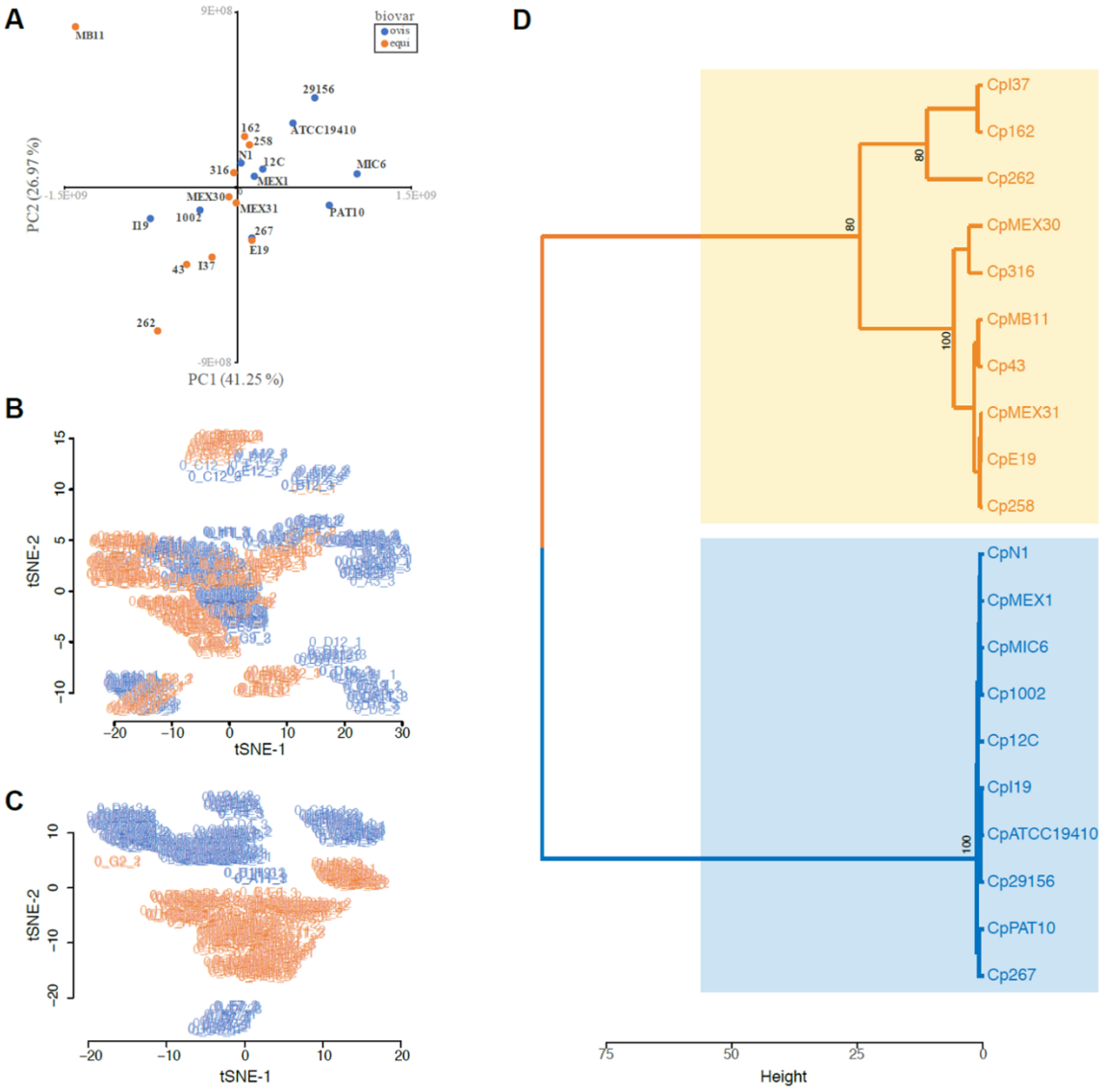


Figure 2. Applying MALDI-TOF and Overall genome relatedness index (OGRI) to classify 20 strains selected from *C. pseudotuberculosis* Equi and Ovis biovars. Mass/charge values obtained by MALDI-TOF were used to perform Principal Component Analysis (A) and tSNE (B). (C) tSNE was computed using the abundance of peptides/proteins assessed by MALDI-TOF. (D) Dendrogram based on relatedness of ANIb and TETRA nucleotide values obtained by evaluating bacterial genomes. In the dendrogram, numbers over the clades represent the confidence values estimated using pvclust.

effectively distinguished the CP biovars based on whole genome analysis. Concatenation of the ANI blast and TNA results generated genomic signatures that allowed accurate grouping of the Ovis and Equi biovars (Figure 2).

Even though high-throughput DNA sequencing technologies have evolved to provide complete genome sequences in a few hours, they are not within reach of every research/clinical microbiology laboratory. Additionally, ANIb and TNA are performed using whole-genome assemblies, which are often challenging to obtain. In this context, we investigated whether the peptide/protein spectra provided by the MALDI Biotyper might be explored to classify the CP strains into biovars.

Accurate classification of biovars in CP was not achieved using the standard parameters in the FlexAnalysis software. Similar results were reported for *C. diphtheriae*, one of the most representative pathogenic corynebacteria. Although the four *C. diphtheriae* biovars (Belfanti, Gravis, Mitis, and Intermedius) can be differentiated using classical microbiology techniques, the Microflex LT Biotyper did not provide reliable biovar classifications.[38] In another study, the MALDI Biotyper generated ambiguous peptide/protein spectra for bacteria from the *Brucella* genus. The identification of species was possible only after applying multivariate statistical analysis based on the SVM method, which has also allowed accurate classifications of biovars in the species *Brucella suis*.[39]

In this work, when the SVM method was applied to analyze the MS spectra generated for CP, the intensities of 77 peptides/proteins peaks were statistically different between the biovars. When considered for PCA, this dataset al.lowed highly accurate biovar classification (92.65%). This result confirms that the peak intensities in the spectra provided by the MALDI Biotyper are useful for biovar classification and epidemiological studies involving CP. However, given that these analyses were performed using the Bionumerics software, which requires a specific license, we developed another strategy for the biovar classification using only open-source software.

https://doi.org/10.1021/jasms.2c00174
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

In this new strategy, we performed tSNE analysis with a more significant number of peptide/protein peaks and disregarded the matching of peaks among the selected CP strains. The spectra matrices were ordered from higher to lower m/z ratio or intensity peak prior to the tSNE analysis. Considering the m/z ratio ordering, we noticed a highly overlapped tSNE profile for Ovis and Equi. This result suggests that the high genomic conservation observed in our ANIb and TNA results leads to a similar m/z ratio distribution in the two biovars. In contrast, the biovars were distinguished when the tSNE analysis considered peak intensity ordering. Altogether, our results suggest that the protein expression levels play an essential role in the differentiation of CP biovars.

CONCLUSION

We effectively demonstrated that the MALDI Biotyper could be used for biovar classification in CP as long as complementary statistical analyses are applied to the peptide/protein spectra generated by the equipment. This approach dramatically increases the complexity of analysis in a workflow designed initially for routine microorganism identification, and the standard technical approaches should remain as the reference for biovar classification in CP.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.2c00174.

MALDI Biotyper identification of *C. pseudotuberculosis* strains isolated from different hosts (Table S1), classes of protein peaks with different intensities between the Equi and Ovis biovars according to the Mann−Whitney test ($p < 0.05$) (Box S1) (PDF)

AUTHOR INFORMATION

Corresponding Authors

Vasco Azevedo − *Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil;* orcid.org/0000-0002-4775-2280; Email: vascoariston@gmail.com

Thiago L. P. Castro − *Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador 40231-300, Brazil; Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil;* Email: castrotlp@gmail.com

Authors

Roselane Gonçalves dos Santos − *Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Nubia Seyffert − *Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador 40231-300, Brazil; Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Elaine M. S. Dorneles − *Departamento de Medicina Veterinária, Universidade Federal de Lavras, Lavras 37200-000, Brazil*

Eric R. G. R. Aguiar − *Centro de Biotecnologia e Genética, Universidade Estadual de Santa Cruz, Ilhéus 45662-900, Brazil*

Carolina P. Ramos − *Departamento de Medicina Veterinária Preventiva, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Dionei J. Haas − *Departamento de Medicina Veterinária Preventiva, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Gabriella B. N. Assis − *Departamento de Medicina Veterinária Preventiva, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Ricardo Dias Portela − *Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador 40231-300, Brazil*

Aristóteles Goes-Neto − *Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Luis G. C. Pacheco − *Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador 40231-300, Brazil;* orcid.org/0000-0003-2500-0800

Henrique C. P. Figueiredo − *Departamento de Medicina Veterinária Preventiva, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Thiago de Jesus Sousa − *Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Sandeep Tiwari − *Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Arun Kumar Jaiswal − *Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Andrey P. Lage − *Departamento de Medicina Veterinária Preventiva, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil*

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.2c00174

Author Contributions

†R.G.d.S. and N.S. contributed equally to this work.

Funding

This study was financially supported by the following Brazilian research agencies: CAPES (grants 88887.091518/2014-00, 88881.068000/2014-01, and 8887.508856/2020-00), CNPq grant 433859/2018-0, and 402817/2019-2, and FAPEMIG grant APQ-02096-15.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Moore, R.; Miyoshi, A.; Pacheco, L. G. C.; Seyffert, N.; Azevedo, V. *Corynebacterium and Arcanobacterium. Pathogenesis of Bacterial Infections in Animals*, 4th ed.; Wiley, 2010; pp 133−147. DOI: 10.1002/9780470958209.CH8.

(2) Dorella, F. A.; Carvalho Pacheco, L.; Oliveira, S. C.; Miyoshi, A.; Azevedo, V. Corynebacterium pseudotuberculosis : microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Vet Res.* **2006**, *37*, 201−218.

(3) Shpigel, N. Y.; Elad, D.; Yeruham, I.; Winkler, M.; Saran, A. An outbreak of Corynebacterium pseudotuberculosis infection in an Israeli dairy herd. *Vet Rec.* **1993**, *133*, 89−94.

(4) Baird, G. J.; Fontaine, M. C. Corynebacterium pseudotuberculosis and its Role in Ovine Caseous Lymphadenitis. *J. Comp Pathol.* **2007**, *137*, 179−210.

(5) Blair, J. M. A.; Piddock, L. J. V.; Selim, S. A. Oedematous Skin Disease of Buffalo in Egypt. *Journal of Veterinary Medicine, Series B* **2001**, *48*, 241−258.
(6) Aleman, M.; Spier, S. J.; Wilson, W. D.; Doherr, M. Corynebacterium pseudotuberculosis infection in horses: 538 cases (1982−1993). *J. Am. Vet Med. Assoc.* **1996**, *209*, 804−809.
(7) Dorella, F. A.; Pacheco, L. G. C.; Seyffert, N.; Portela, R. W.; Meyer, R.; Miyoshi, A.; Azevedo, V. Antigens of Corynebacterium pseudotuberculosis and prospects for vaccine development. *Expert Rev. Vaccines.* **2009**, *8*, 205−213.
(8) de Sá Guimarães, A.; Borges Do Carmo, F.; Pauletti, R. B.; Seyffert, N.; Ribeiro, D.; Lage, A. P.; Heinemann, M. B.; Miyoshi, A.; Azevedo, V.; Guimarães Gouveia, A. M. Caseous Lymphadenitis: Epidemiology, Diagnosis, And Control. *Review: Veterinary Microbiology* **2011**, *2*, 33−43.
(9) Guimarães, A. S.; Seyffert, N.; Bastos, B. L.; Portela, R. W. D.; Meyer, R.; Carmo, F. B.; Cruz, J. C. M.; McCulloch, J. A.; Lage, A. P.; Heinemann, M. B.; Miyoshi, A.; Azevedo, V.; Gouveia, A. M. G. Caseous lymphadenitis in sheep flocks of the state of Minas Gerais, Brazil: Prevalence and management surveys. *Small Ruminant Research.* **2009**, *87*, 86−91.
(10) Yeruham, I.; Elad, D.; Van-Ham, M.; Shpigel, N. Y.; Perl, S. Corynebacterium pseudotuberculosis infection in Israeli cattle: clinical and epidemiological studies. *Vet Rec.* **1997**, *140*, 423−427.
(11) Seyffert, N.; Guimarães, A. S.; Pacheco, L. G. C.; Portela, R. W.; Bastos, B. L.; Dorella, F. A.; Heinemann, M. B.; Lage, A. P.; Gouveia, A. M. G.; Meyer, R.; Miyoshi, A.; Azevedo, V. High seroprevalence of caseous lymphadenitis in Brazilian goat herds revealed by Corynebacterium pseudotuberculosis secreted proteins-based ELISA. *Res. Vet Sci.* **2010**, *88*, 50−55.
(12) Hou, X. G.; Kawamura, Y.; Sultana, F.; Hirose, K.; Miyake, M.; Otsuka, Y.; Misawa, S.; Oguri, T.; Yamamoto, H.; Ezaki, T. Genetic identification of members of the genus Corynebacterium at genus and species levels with 16S rDNA-targeted probes. *Microbiol. Immunol.* **1997**, *41*, 453−460.
(13) Biberstein, E. L.; Knight, H. D.; Jang, S. Two biotypes of Corynebacterium pseudotuberculosis. *Vet Rec.* **1971**, *89*, 691−692.
(14) Almeida, S.; Dorneles, E. M. S.; Diniz, C.; Abreu, V.; Sousa, C.; Alves, J.; Carneiro, A.; Bagano, P.; Spier, S.; Barh, D.; Lage, A. P.; Figueiredo, H.; Azevedo, V. Quadruplex PCR assay for identification of Corynebacterium pseudotuberculosis differentiating biovar Ovis and Equi. *BMC Vet Res.* **2017**, 13.
(15) Çetinkaya, B.; Karahan, M.; Atil, E.; Kalin, R.; de Baere, T.; Vaneechoutte, M. Identification of Corynebacterium pseudotuberculosis isolates from sheep and goats by PCR. *Vet. Microbiol.* **2002**, *88*, 75−83.
(16) Pacheco, L. G. C.; Pena, R. R.; Castro, T. L. P.; Dorella, F. A.; Bahia, R. C.; Carminati, R.; Frota, M. N. L.; Oliveira, S. C.; Meyer, R.; Alves, F. S. F.; Miyoshi, A.; Azevedo, V. Multiplex PCR assay for identification of Corynebacterium pseudotuberculosis from pure cultures and for rapid detection of this pathogen in clinical samples. *J. Med. Microbiol.* **2007**, *56*, 480−486.
(17) Kostrzewa, M. Application of the MALDI Biotyper to clinical microbiology: progress and potential. *Expert Rev. Proteomics.* **2018**, *15*, 193−202.
(18) Chun, J.; Rainey, F. A. Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *Int. J. Syst. Evol Microbiol.* **2014**, *64*, 316−324.
(19) Richter, M.; Rosselló-Móra, R.; Oliver Glöckner, F.; Peplies, J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics.* **2016**, *32*, 929−931.
(20) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403−410.
(21) Alatoom, A. A.; Cunningham, S. A.; Ihde, S. M.; Mandrekar, J.; Patel, R. Comparison of direct colony method versus extraction method for identification of gram-positive cocci by use of Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin Microbiol.* **2011**, *49*, 2868−2873.
(22) Egli, A.; Tschudin-Sutter, S.; Oberle, M.; Goldenberger, D.; Frei, R.; Widmer, A. F. Matrix-assisted laser desorption/ionization time of flight mass-spectrometry (MALDI-TOF MS) based typing of extended-spectrum β-lactamase producing E. coli - A novel tool for real-time outbreak investigation. *PLoS One.* **2015**, 10.
(23) Agustini, B. C.; Silva, L. P.; Bloch, C.; Bonfim, T. M. B.; da Silva, G. A. Evaluation of MALDI-TOF mass spectrometry for identification of environmental yeasts and development of supplementary database. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 5645−5654.
(24) Assis, G. B. N.; Pereira, F. L.; Zegarra, A. U.; Tavares, G. C.; Leal, C. A.; Figueiredo, H. C. P. Use of MALDI-TOF Mass Spectrometry for the Fast Identification of Gram-Positive Fish Pathogens. *Front Microbiol.* **2017**, *8*, 1492.
(25) Blair, J. M. A.; Piddock, L. J. V.; Krijthe, J. *Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation version 0.16 from CRAN*, https://rdrr.io/cran/Rtsne/.
(26) Suzuki, R.; Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics.* **2006**, *22*, 1540−1542.
(27) Kassambara, A.; Mundt, F. *Extract and Visualize the Results of Multivariate Data Analyses·factoextra, R package version* 2016.13
(28) Blair, J. M. A.; Piddock, L. J. V.; de Vries, A. Using the ggdendro package for plotting dendrograms and tree diagrams, 2012.
(29) Bao, R.; Gao, X.; Hu, B.; Zhou, Z. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: a powerful tool for identification of Corynebacterium species. *J. Thorac Dis.* **2017**, *9*, 3239.
(30) 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. *Vet Res.* **2006**, *37*, 201−218.
(31) Sogawa, K.; Watanabe, M.; Sato, K.; Segawa, S.; Ishii, C.; Miyabe, A.; Murata, S.; Saito, T.; Nomura, F. Use of the MALDI BioTyper system with MALDI-TOF mass spectrometry for rapid identification of microorganisms. *Anal Bioanal Chem.* **2011**, *400*, 1905−1911.
(32) Seng, P.; Drancourt, M.; Gouriet, F.; La Scola, B.; Fournier, P. E.; Rolain, J. M.; Raoult, D. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis.* **2009**, *49*, 543−551.
(33) Prod'Hom, G.; Bizzini, A.; Durussel, C.; Bille, J.; Greub, G. Matrix-assisted laser desorption ionization-time of flight mass spectrometry for direct bacterial identification from positive blood culture pellets. *J. Clin Microbiol.* **2010**, *48*, 1481−1483.
(34) Bizzini, A.; Jaton, K.; Romo, D.; Bille, J.; Prod'hom, G.; Greub, G. Matrix-assisted laser desorption ionization-time of flight mass spectrometry as an alternative to 16S rRNA gene sequencing for identification of difficult-to-identify bacterial strains. *J. Clin Microbiol.* **2011**, *49*, 693−696.
(35) El-Bouri, K.; Johnston, S.; Rees, E.; Thomas, I.; Bome-Mannathoko, N.; Jones, C.; Reid, M.; Ben-Ismaeil, B.; Davies, A. P.; Harris, L. G.; Mack, D. Comparison of bacterial identification by MALDI-TOF mass spectrometry and conventional diagnostic microbiology methods: agreement, speed and cost implications. *Br J. Biomed Sci.* **2012**, *69*, 47−55.
(36) Soares, S. C.; Trost, E.; Ramos, R. T. J.; Carneiro, A. R.; Santos, A. R.; Pinto, A. C.; Barbosa, E.; Aburjaile, F.; Ali, A.; Diniz, C. A. A.; Hassan, S. S.; Fiaux, K.; Guimarães, L. C.; Bakhtiar, S. M.; Pereira, U.; Almeida, S. S.; Abreu, V. A. C.; Rocha, F. S.; Dorella, F. A.; Miyoshi, A.; Silva, A.; Azevedo, V.; Tauch, A. Genome sequence of Corynebacterium pseudotuberculosis biovar equi strain 258 and prediction of antigenic targets to improve biotechnological vaccine production. *J. Biotechnol.* **2013**, *167*, 135−141.
(37) Viana, M. V. C.; Sahm, A.; Neto, A. G.; Figueiredo, H. C. P.; Wattam, A. R.; Azevedo, V. Rapidly evolving changes and gene loss associated with host switching in Corynebacterium pseudotuberculosis. *PLoS One.* **2018**, *13*, No. e0207304.

https://doi.org/10.1021/jasms.2c00174
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

(38) Konrad, R.; Berger, A.; Huber, I.; Boschert, V.; Hörmansdorfer, S.; Busch, U.; Hogardt, M.; Schubert, S.; Sing, A. Matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry as a tool for rapid diagnosis of potentially toxigenic Corynebacterium species in the laboratory management of diphtheria-associated bacteria. *Euro Surveill.* **2010**, *15*, 1−5.
(39) Karger, A.; Melzer, F.; Timke, M.; Bettin, B.; Kostrzewa, M.; Nöckler, K.; Hohmann, A.; Tomaso, H.; al Dahouk, S. Interlaboratory Comparison of Intact-Cell Matrix-Assisted Laser Desorption Ionization -Time of Flight Mass Spectrometry Results for Identification and Differentiation of Brucella spp. *Journal of clinical microbiology* **2013**, *51* (9), 3123−3126.

https://doi.org/10.1021/jasms.2c00174
J. Am. Soc. Mass Spectrom. XXXX, XXX, XXX−XXX

4. CAPÍTULO II

4.1. OBJETIVOS

4.2 OBJETIVO GERAL

Identificar e caracterizar uma nova espécie do gênero *Dietzia*, isolada de abscesso de égua e avaliar a diversidade genômica do gênero;

4.3 OBJETIVOS ESPECÍFICOS

- Sequenciar, montar e depositar o primeiro genoma de *Dietzia* sp. B32;
- Conhecer a diversidade genética através de análises filogenéticas entre os genomas disponível nos bancos de dados do gênero *Dietzia*;
- Realizar análises para avaliação dos limites/ estimativa da relação genética entre dois genomas;
- Realizar análise de pangenoma para melhor compreensão da relação entre as cepas do gênero *Dietzia;*
- Analisar e identificar as regiões de plasticidade genômica;

4.5 ARTIGO CIENTÍFICO 2: Comparative Genomic Analysis of the *Dietzia* genus: an insight into genomic diversity, and adaptation

Publicado em: Research in Microbiology; Fator de impacto (atual): 3.946

DOI: https://doi.org/10.1016/j.resmic.2022.103998

Membros do género *Dietzia* são patógenos oportunistas, ubíquas, de importância na indústrias química, alimentícia e médica. A identificação taxonômica das espécies do género *Dietzia* são realizadas através dos métodos microbiológico, bioquímico e fenotípicos. Porém esses sistemas de classificação apresentam uma deficiência na predição a nível de espécie para o gênero *Dietzia*, apresentando baixa sensibilidade e especificidade para identificar e o gênero não apresenta testes populares de sistemas de classificação. Frequentemente, na ausência de métodos simples e precisos para sua identificação o gênero *Dietzia* tem sido desconsiderado na deteção clínica e veterinária, pelo qual existe uma falta de reportes, identificação, e acesso público de sistemas de classificação taxonômica, levando ao fato que algumas espécies podem ter sido identificadas erroneamente como um *Rhodococcus equi* e outros actinomicetos. As espécies de *Dietzia* podem ser facilmente distinguidas de membros de outros gêneros usando dados de sequência do gene 16S rRNA. Contudo, dentro do próprio gênero, às pode ser difícil determinar o status taxonômico de algumas espécies, pois certas espécies de *Dietzia* são idênticas com base no sequenciamento e amplificação por PCR das sequências do gene 16S rRNA. Devido á dificuldade para a identificação taxonômica a nível tradicional e conhecimento do género *Dietzia* é importante a procura de novas estratégias de classificação baseado em estratégias ômicas. Diante desse cenário, nos utilizamos um conjunto de abordagens genômicas que nos permite elucidar a nível taxonômico, diversidade e adaptação. Portanto, este artigo científico relata a genômica comparativa do gênero *Dietzia*, dos genomas disponíveis no banco de dados NCBI, elucidando a existência de erros na classificação das espécies deste gênero. Realizamos análises de ANI, dDDH, pangenoma e plasticidade genômica para melhor elucida as relações filogenómicas entre estirpes de *Dietzia.* As análises realizadas apoiaram o agrupamento de diferentes espécies de *Dietzia* em quatro grupos distinto apoiando a ideia de que algumas espécies de *Dietzia* poderiam ser reatribuídas em uma nova possível classificação.

O nosso trabalho contribuirá na avaliação da predição taxonômica do gênero *Dietzia* utilizando genomas, trazendo uma melhor identificação para as espécies do gênero, mostrando que muitas

delas deveriam ser reatribuídas, devido a confusão na sua caracterização, além disso nós permitiu elucidar a diversidade genética e filogrupos associados a seu potencial adaptativo e de nicho, juntamente com a caracterização de uma nova espécie do gênero, sendo esse a primeira espécie depositada no NCBI.

Research in Microbiology 174 (2023) 103998

Contents lists available at ScienceDirect

Research in Microbiology

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

Original Article

Comparative genomic analysis of the *Dietzia* genus: an insight into genomic diversity, and adaptation

Roselane Gonçalves dos Santos [a], Raquel Hurtado [a], Diego Lucas Neres Rodrigues [a], Alessandra Lima [a], William Ferreira dos Anjos [h], Claudia Rifici [d], Anna Rita Attili [e], Sandeep Tiwari [a, b, c], Arun Kumar Jaiswal [a], Sharon J. Spier [f], Giuseppe Mazzullo [d], Francielly Morais-Rodrigues [a], Anne Cybelle Pinto Gomide [a], Luís Cláudio Lima de Jesus [a], Flavia Figueira Aburjaile [a], Bertram Brenig [g], Vincenzo Cuteri [e], Thiago Luiz de Paula Castro [b, c, j], Núbia Seyffert [a, b], Anderson Santos [h], Aristóteles Góes-Neto [i], Thiago de Jesus Sousa [a, 1, **], Vasco Azevedo [a, 1, *]

[a] Cellular and Molecular Genetics Laboratory, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil
[b] Postgraduate Program in Microbiology, Institute of Biology, Federal University of Bahia, Salvador, BA, Brazil
[c] Postgraduate Program in Immunology, Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil
[d] Department of Veterinary Science, University of Messina (Italy), Polo Universitario dell'Annunziata, 98168 Messina (ME), Italy
[e] School of Biosciences and Veterinary Medicine, University of Camerino (Italy), Via Circonvallazione 93/95, 62024 Matelica (MC), Italy
[f] Department of Veterinary Medicine and Epidemiology, University of California, Davis, CA, USA
[g] Institute of Veterinary Medicine, University of Göttingen, Burckhardtweg 2, Göttingen, Germany
[h] Department of Computer Science, Federal University of Uberlandia, Uberlandia, Brazil
[i] Molecular and Computational Biology of Fungi Laboratory Department of Microbiology, Institute of Biological Sciences, Federal University of Minas Gerais Brazil
[j] Department of Biotechnology, Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil

ARTICLE INFO

Article history:
Received 13 June 2022
Accepted 2 November 2022
Available online 11 November 2022

Keywords:
Comparative genomics
Core genome
Dietzia
Phylogenetics
Pangenome
Reassign species
Phylogenomic

ABSTRACT

Dietzia strains are widely distributed in the environment, presenting an opportunistic role, and some species have undetermined taxonomic characteristics. Here, we propose the existence of errors in the classification of species in this genus using comparative genomics. We performed ANI, dDDH, pangenome and genomic plasticity analyses better to elucidate the phylogenomic relationships between *Dietzia* strains. For this, we used 55 genomes of Dietzia downloaded from public databases that were combined with a newly sequenced. Sequence analysis of a phylogenetic tree based on genome similarity comparisons and dDDH, ANI analyses supported grouping different Dietzia species into four distinct groups. The pangenome analysis corroborated the classification of these groups, supporting the idea that some species of *Dietzia* could be reassigned in a possible classification into three distinct species, each containing less variability than that found within the global pangenome of all strains. Additionally, analysis of genomic plasticity based on groups containing *Dietzia* strains found differences in the presence and absence of symbiotic Islands and pathogenic islands related to their isolation site. We propose that the comparison of pangenome subsets together with phylogenomic approaches can be used as an alternative for the classification and differentiation of new species of the genus Dietzia.

1. Introduction

The *Dietzia* genus was proposed by [1] to accommodate actinomycetes, previously known as *Flavobacterium maris*, and later classified as *Rhodococcus maris* [2]. In the corrected description, the *Dietziaceae* family is within the order of *Corynebacteriales* found in different habitats, notably in the soil ecosystem. Some strains are

* Corresponding author.
** Corresponding author.
E-mail addresses: roselanegr@gmail.com (R.G. Santos), raquelgen1@gmail.com (R. Hurtado), claudiarifici35@gmail.com (C. Rifici), annarita.attili@unicam.it (A.R. Attili), sjspier@ucdavis.edu (S.J. Spier), giuseppe.mazzullo@unime.it (G. Mazzullo), bbrenig@gwdg.de (B. Brenig), vincenzo.cuteri@unicam.it (V. Cuteri), castrotlp@gmail.com (T.L.P. Castro), nbseyffert@gmail.com (N. Seyffert), arigoesneto@gmail.com (A. Góes-Neto), thiagojsousa@gmail.com (T. de Jesus Sousa), vascoariston@gmail.com (V. Azevedo).
[1] Contributed equally.

https://doi.org/10.1016/j.resmic.2022.103998
0923-2508/

severe pathogens of humans (eg, *Mycobacterium tuberculosis*), animals (e.g., *Rhodococcus equi* = *Rhodococcus hoagii*), and plants (e.g., *Rhodococcus fascians*). Along with the *Dietziaceae* family, the order also comprises the families *Corynebacteriaceae*, *Gordoniaceae*, *Lawsonellaceae*, *Mycobacteriaceae*, *Nocardiaceae*, *Segniliparace, ae* and *Tsukamurellaceae* and the type of the genus is *Corynebacterium* [2–4].

At the time of writing, there were thirteen species described for the genus: *Dietzia aerolata* [1], *Dietzia alimentaria* [5], *Dietzia aurantiaca* [6], *Dietzia Cercidiphyllum* [7], *Dietzia cinnamea* [8], *Dietzia kunjamensis* [9], *Dietzia lutea* [10], *Dietzia maris* [4,11], *Dietzia natronolimnaea* [12] *Dietzia papillomatosis* [13], *Dietzia psychralcaliphila* [14], *Dietzia schimae* [7] and *Dietzia timorensis* [15], all these published validly with the correct name in the List of Prokaryotic names with Standing in Nomenclature [16]. Since its first description in 1995, *Dietzia* strains have been gradually isolated from natural habitats, such as moderately saline and alkaline East African soda lake [12], surface-sterilized plant tissues [7], soil [10], under mahogany trees in West [15], there are at least eight descriptions of *Dietzia* species that cause infections in humans [13,17–25], dog bite wound in an adult patient [26], Korean food [5], an air of a duck barn [1] and the cold desert of the Indian Himalaya [9].

Many of *Dietzia*'s isolation sites can be found in the BioSample/BioProject of each available sample at the National Center for Biotechnology Information taxonomy (NCBI) homepage (https://www.ncbi.nlm.nih.gov/genome/?term=Dietzia). These bacteria also have many applications in various industries, particularly medical, chemical, and food [27]. Some works have shown that *Dietzia* species can degrade aliphatic hydrocarbons, such as n-alkanes, and some strains can degrade aromatic compounds [4,27,28]. Most *Dietzia* species can grow in high saline and alkaline conditions, suggesting their potential applicability in diverse environments). Furthermore, as *Dietzia* species can grow in low-temperature environments, they can be used for bioremediation in cold environments [28].

Dietzia is aerobic, Gram-positive, immobile, catalase-positive actinomycetes forming cocci germinating into short rods. The GC content of the genus ranges from 66 to 73 mol% [2]. Formerly classified as *Rhodococcus* species, *Dietzia* continues to be misidentified in this genus with many commercially available biochemical testing methods due to their phenotypic similarity, including traditional and commercial biochemical tests (Coryne API), as the micromorphological and clinical appearance of *Dietzia* strains, for example, are very similar to those of *R. hoagii*, formerly *R. equi*, an organism frequently encountered by medical microbiologists in their daily practice [27,28].

Dietzia strains can be easily distinguished from members of other genera using sequence data from the 16S rRNA gene. However, within the genus itself, it is occasionally impossible to determine some species' taxonomic status, as certain species of Dietzia are identical based on sequencing and PCR amplification of 16S ribosomal RNA (rRNA) gene sequences [2,29]. Furthermore, the genus Dietzia is still considered new, and so far, few studies have been aimed at the comparative genomic analysis of the genus.

16S rRNA gene sequencing is often not applicable to incomplete genomes (e.g, genes are not assembled), and these genes generally show higher sequence conservation than the genome average. Consequently, analysis of universal genes does not provide sufficient resolution at the species level and has often resulted in the lack of apparent genetic discontinuities between closely related taxa. In recent years, a methodology that offers several significant advantages, such as higher resolution between closely related genomes, is genome-wide mean nucleotide identity (ANI). The ANI draws an estimate of the genetic relationship between two organisms, informing if they belong to the same species if they present a value of $\geq$95% of ANI between them [30].

Another method to obtain an in-depth view of the genome of a species or genus is the pan-genomic approach, which provides us with the complete and non-redundant collection of genes of a species or genus and is composed of three subsets (central genome, shared genome, and singletons): the core genome, which is the collection of all genes commonly shared among all genomes used as a dataset; the shared genome, which contains only genes shared between two or more lineages, which are not present in all strains in the dataset; and, singletons, which are present in only one strain and are referred to as strain-specific genes [31].

We present a new phylogenetic/phylogenomic and comparative study of 55 sequenced genomes of *Dietzia* species available in public databases, combined with a novel genomic sequence isolated from a mare abscess in the region of Italy. The phylogenetic and phylogenomic relationships between these isolates of the *Dietzia* genus were compared, along with the pangenome. The complete inventory of genes found in any member of the species: the "core genome", which is important for the basic processes of life; and singletons, which are usually related to environmental fitness and host adaptation. Finally, we provide information on specific subsets (pan and core genomes) of 56 *Dietzia* genomes and correlate these groups with the plasticity of genomic islands. These data will allow new insights into the contributions of the *Dietzia* genus to a correct characterization/differentiation of these species.

2. Material and methods

2.1. Genomes of the genus Dietzia used for comparative genomic analysis

All the genome sequences used in this study are listed in (Additional file: Table S1). The genome data of two samples, Dietzia sp. MeA6-2017 and *D. maris* 97 were retrieved from the BV-BRC: Bacterial and Viral Bioinformatics Resource Center public database, and 54 recovered genomes from the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/). These were combined with data from the *Dietzia* sp. B32 was isolated from diffuse subcutaneous nodules of a mare in Reggio Calabria, Italy; this is the first genome of the *Dietzia* genus isolated from an animal. The complete genome sequence of the *Dietzia* sp. B32 was obtained on an Hiseq 2500 platform (Illumina, San Diego, CA, USA) using paired-end libraries (2 $\times$ 150 bp). The quality of the sequencing reads was verified using the FastQC program (v.0.11.5) [32] and assembled by The Edena software (version 3.0) [33].

The Sspace and GapFiller algorithms were used to assess the contigs' order, distance, and orientation and fill the gap in paired readings, respectively [34,35]. In addition, gap closing procedures were conducted using the GFinisher tools [36], BLASTn [37], and the CLC Genomics Workbench 7 (Qiagen Inc.) was used for visualization and mapping of reads. The genome information can be found under Bio-project number (PRJNA802820), Biosample (SAMN25598616) GenBank accession number (CP093845.1). A workflow of the methodology is described in Fig. 1. It contains all the steps of the methods of this work (Fig. 1).

2.2. Phylogenetic/phylogenomyc

The phylogenomic analysis was performed using Gegenees software (Version 2.1) [38], which calculated the percentage of similarity among the genomes of all strains. Gegenees separates genomes into small sequences and stipulates the minimum content shared by all genomes. Subsequently, the minimum shared contents obtained were subtracted from all genomes resulting in variable contents, which were eventually compared with all other strains to calculate similarity percentages [38]. Data generated by

Gegenees were extracted as a distance matrix file in nexus format (.nex), and SplitsTree software (version 4.17.1) [39] was used to build a dendrogram using the method of Neighbor-joining with the generated distance matrix. Finally, these percentages were plotted on a heat map graph ranging from low similarity (red) to high similarity (green). Two other trees were inferred using the 16S rRNA gene sequence and another using the *gyrB* gene. The 16S rRNA and *gyrB* complete sequences were extracted from the genomes and aligned using the MUSCLE program provided in MEGAX v.10.1.8 [40]. After alignment, a phylogenetic tree was inferred using the maximum likelihood method with MEGAX software v.10.1.8 [40]. The topography of the reconstructed tree was evaluated by bootstrap analysis with 1000 repetitions [41]. *Tsukamurella paurametabola* ATCC8368 was used as an external group. Finally, figures of all phylogenetic trees were generated using the Interactive Tree of Life (iTOL) [42].

2.3. Average nucleotide identity and DNA–DNA hybridization

We performed an Average Nucleotide Identity (ANI) analysis using the whole-genome sequences. ANI represents a mean of identity/similarity values between homologous genomic regions shared by two genomes. ANI was estimated by the JSpecies Web Server (available at http://imedea.uib-csic.es/jspecies/about.html) [43], based on BLAST [37]. Typically, the threshold frontier to consider two organisms to belong to the same species could be set at > 95% identity for ANI [30,43].

The digital values of DNA–DNA hybridization (dDDH) were calculated using the Genome-to-Genome Distance Calculator 2.1 provided by the Leibniz Institute DSMZ website (http://ggdc.dsmz.de/distcalc2.php) with recommended parameters and/or default settings, respectively [44]. If the genomic DNA of two respective organisms reveals a DDH similarity of above 70%, this is the main argument to regard them as the same species and vice versa [44,45]. For this analysis, we performed the calculation just to the results obtained in the analysis of the ANI, for the groups formed in the ANI that had a type species, it was used as the query to search for the dDDH values.

2.4. Pangenome calculation

The core genome (shared by all strains), shared genome (genes present in two or more than two strains, and singletons (strain-specific genes) were identified using OrthoFinder software (Version 2.4.1) through the default parameters [46]. Furthermore, we used an *in-house* script in the results provided by the program for the separation of core, shared and singleton [47]. OrthoFinder uses the. faa amino acid sequence file for each genome to perform *all*-vs-*all* BLASTp for the Orthologous analysis [46].

To calculate the pangenome development, we apply Heap's Law, with the formula $n = n = k*N-\alpha$, where n is the expected number of genes for a given number of genomes, n is the number of genomes, and the other terms are constants defined to fit the specific curve [47,48]. According to Heap's law, a value of $\alpha \leq 1$ is representative of an open pangenome; this means that each added genome will contribute some new genes and the pangenome will increase. Instead, an α value > 1 represents a closed pangenome, in which the addition of new genomes will not significantly affect the size of the pangenome [47,48]. For this analysis, the same groups found in the ANI were defined in three different environments. Group A is classified as an environment in general for presenting several different environments, group B was classified as a pathogenicity group for being formed with most clinical samples, and group C was considered as a group of Industrial water for presenting the

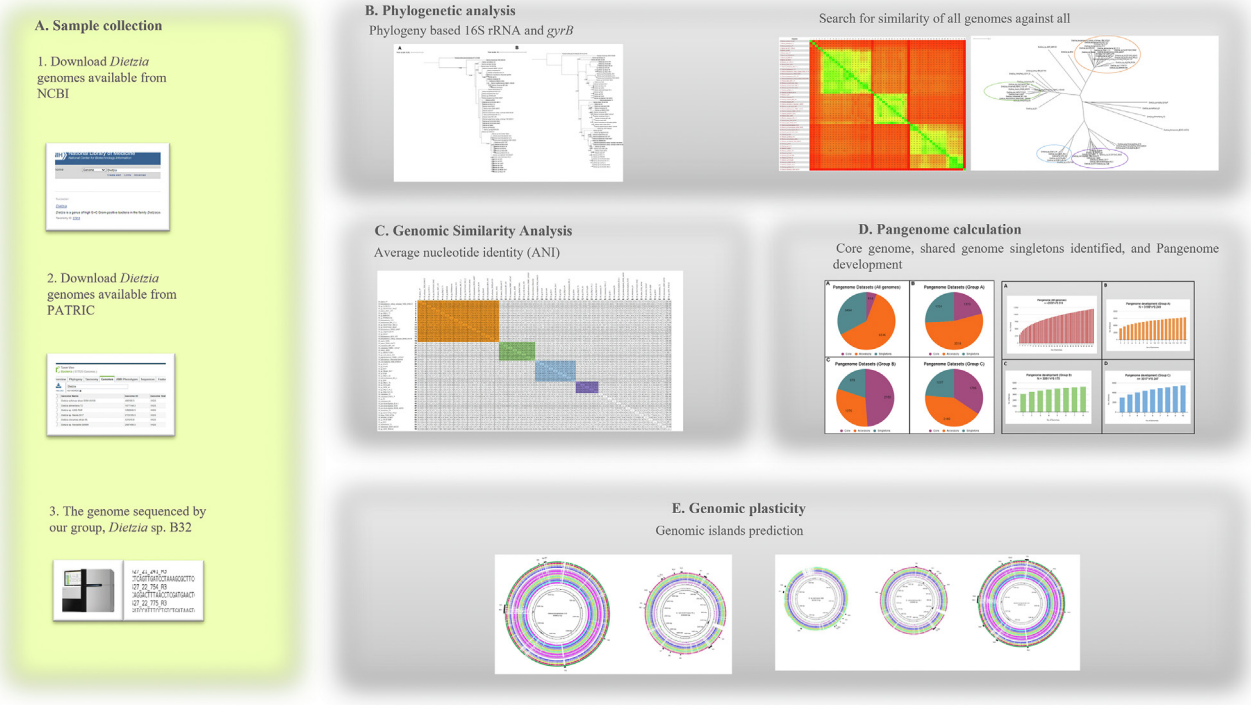


Fig. 1. Workflow analysis. (A) (1, 2, and 3) Samples used for analysis. (B) Phylogenetic and phylogenomic analysis. (C) Genomic Similarity Analysis. (D) Pangenome calculation. (E) Genomic islands prediction.

majority samples isolated from an environment containing water; group D was not used, as it does not present any described species, only designated with the genus (Additional file: Table S3).

2.5. Identification of genes encoding virulence factors and antibiotic resistance genes

In addition, a search for resistance and virulence genes was performed using the core genome obtained by pan-genomic analysis considering all strains in the dataset. The database used to search for resistance and virulence genes were CARD and VICTORS, respectively [49,50]. To infer presence, a minimum identity of 70% was considered.

2.6. Analysis of genomic plasticity

Predictions of genomic islands (GEIs) as metabolic islands (MIs), pathogenicity islands (PAIs), and symbiotic islands (SIs) were predicted using GIPSy (Genomic Island prediction) software (version 1.1) [51]. We used the same groups according to the environment defined in the pangenome analyses (Additional file: Table S3). We used the same groups according to the environment defined in the pangenome analysis (Additional file: Table S3). For group A, *D. kunjamensis* 313 was used as a reference; for group B *D.* sp. Oral taxon 368 was used as a reference, and in the group, C *D. natronolimnaea* HS-1 was used as a reference. The strain D. sp. oral taxon was used to predict the SIs and PAIs of groups A and C and the D. sp. JS16-p6b was used as a non-pathogenic species for the prediction of PAIs in group B. After curation of the PAIs, genes from all islands in each lineage were evaluated for presence/absence in all other lineages in the group. The island's SIs and PAIs were plotted using BRIG software [52].

3. Results

3.1. General features

The genome lengths varied by 3,4 Mb in size. The GC content ranged from 70.43% and around 3435 predicted CDS. The genome information can be found under Bio project number (PRJNA802820), Biosample (SAMN25598616) GenBank accession number (CP093845).

3.2. Phylogenomic and Phylogenetic analysis of Dietzia

We built two phylogenetic trees based on a single gene using the 16S rRNA genes and the *gyrB* gene (Additional file: Fig. S1 A and B). In the trees of 16S and *gyrB,* it is possible to observe that the species of *Dietzia* are mixed in the clades, without a correct distinction of the species, the species were present in several different clades. *D. maris* can be found in two clades with other species of the genus, making it difficult to distinguish which species belongs (Additional file: Fig. S1 A and B).

The distance matrix generated using Gegenees software was plotted as a heatmap (Fig. 2A), in which the similarity among genomes varied from ~4% (*D. timorensis* ID05-A0528^T) to ~98% (between the genomes of *D. psychralcaliphila* ILA-1 and *D. psychralcaliphila* DSM 44820^T). The similarity analysis using Gegenees revealed that the genomes of Dietzia can be clustered into at least four groups A, B, C, and D (Fig. 2B). However other genomes were only distantly related to the other strains, such as *D. timorensis* ID05-A0528T, *D. aurantiaca* J3, *D. aerolata* Sj14aT, and *D. alimentaria* 72T, these genomes did not cluster with any other sequence, *D. psychralcaliphila* groups correctly with its three strains described (Fig. 2B).

3.3. Dietzia genus has four major groups distinguished by ANI and dDDH without taxonomic distinction

We performed an Average Nucleotide Identity (ANI) analysis using genome sequence data (Additional file: Table S1). Using a 95% identity cutoff, this analysis revealed the presence of the same four groups presented in the Gegenees analyses. In cluster A it is possible to observe that the species *D. maris* 97, *D. kunjamensis* subsp. *schimae* YIM65001^T, *D. alimentaria* BP-27-1, and *D. kunjamensis* DSM44907 cluster together (Fig. 3). It is possible to find this same pattern in groups B and C, in which *D. maris* DSM 43672^T, *D. cinnamea* BP-168, *D. papillomatosis* NBRC-105045 and *Dietzia massiliensis* Marseille -Q0999^T are grouped in group B and group C two distinct species *Dietzia cercidiphylli* YIM 65002^T and *D. natronolimnaea* HS-1 are shown grouped (Fig. 3). However, in group D we did not obtain any species, only lineages designated with the genus (Fig. 3).

We also obtained some paired comparisons that resulted in values < 95%, indicating that they can be considered distinct bacterial species, such as *D. aurantiaca* J3 *a* (just a representative), *D. timorensis* ID05-A0528^T (just a representative), *D. aerolata* Sj14a^T (just a representative), *D. lutea* YIM 80766^T (just a representative) and *D. psychralcaliphila* (three representatives of the same species: shared more than 95% ANI and >70% dDDH with each other). The genome of *Dietzia* sp. B32 did not show an ANI value greater than 95%, indicating that it is a distinct species (Fig. 3).

Furthermore, the results of the dDDH calculation corroborate those of the ANI approaches in finding that the sequences of the genomes that clustered together present a dDDH value greater than 70%. However, *D. massiliensis* Marseille -Q0999^T and *D. kunjamensis* subsp schimae DSM 45139, samples showed a dDDH lower than allowed it belongs to the same species (69.3% and 69.5% respectively) (Additional file: Table S2). It is generally accepted that ANI values of 95−96% equate to a DDH value of 70% and can be used as a threshold for species delineation [53,54]. The analysis of similarity, ANI, and dDDH shows that some species of *Dietzia* deserve attention in their valid classification, and some of these species could be reassigned, as shown in Table 1. Although the ANI estimates are the gold standard for bacterial species determination, the ANI and phylogenomic analysis did not allow taxonomic classification, however, it will enable us to identify at least four major groups within the *Dietzia* genus, showing species that could be reassigned.

3.4. Pangenome calculation for 3 groups of Dietzia

To take a global view of the *Dietzia* genome and further explore the genomic diversity of this genus, we calculated the size of the pangenome based on different datasets. When we examined all the genomes together, orthologs analysis showed that the pangenome contained a total of 614 core genes, 6536 accessories, and 3494 singletons (Fig. 4A). When we consider only group A genomes, we found a total of 1370 core genes, 3518 accessory genes, and 1701 singletons (Fig. 4B). The group B pangenome had 2150 core genes, 1376 accessories, and 878 singletons (Fig. 4C) and Group C had 1796 core genes, 2180 accessories, and 1237 singletons (Fig. 4D). Using the formula $\alpha = 1-\gamma$, we inferred that the α value of the pangenome of all genomes was 0.69, indicating that the pangenome is probably open and increasing. Similarly, the extrapolation of the pangenome size calculated for groups A, B, and C generated α values of 0.74, 0.88, 0.82, and 0.75 respectively (Fig. 5A, B, C, and D). In the tree formed with the core genome, the results corroborate those found in the analysis of ANI, dDDH, Gegenees, being possible to visualize the formation of the 4 groups (Additional file: Fig. S2).

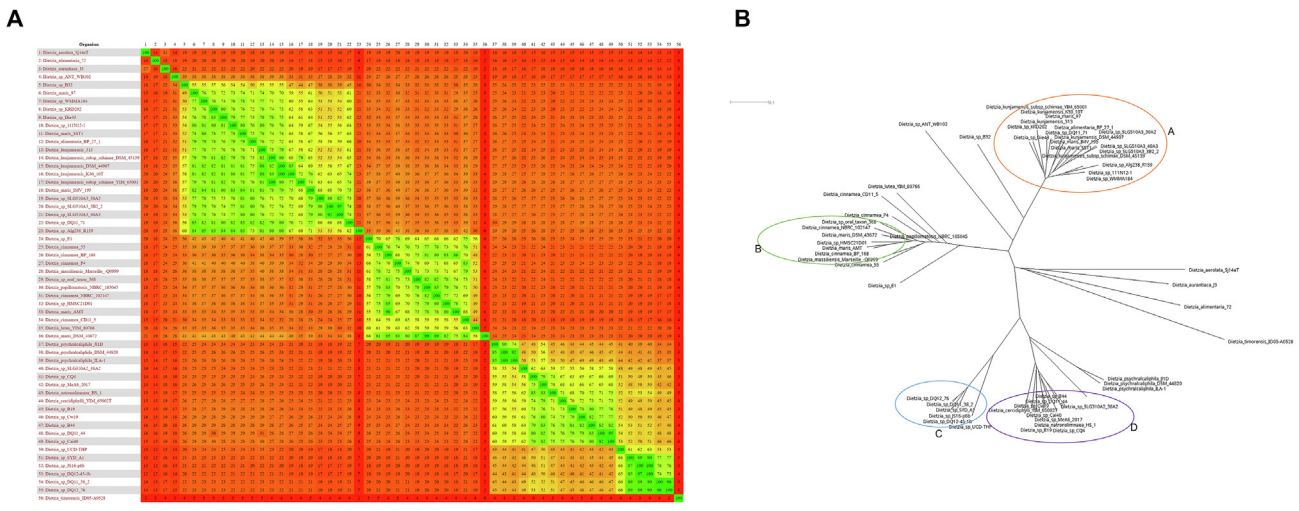


Fig. 2. Heat map analysis of 56 genomes of the *Dietzia* genus. (A) Heatmap and (B) distance-matrix-based phylogenetic network of *Dietzia* genus. The numbers in the heatmap show the percentage of similarity between genomes; the colors vary from red (low similarity) to green (high similarity). The network was constructed using SplitsTree software with NeighborNet and equal angle methods, based on a distance matrix from Gegenees software.

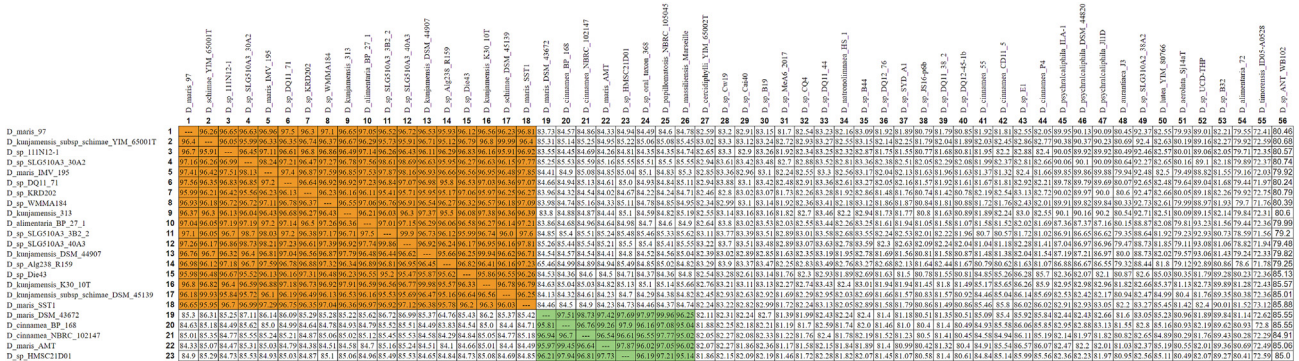


Fig. 3. Average nucleotide identity. The colors orange, green, blue and purple corresponds to clusters A, B, C, and D respectively.

3.5. Antibiotic resistance genes and virulence factors (VFs) predicted in the genus Dietzia within the core genome

In the search for resistance genes and virulence factors, the core genome obtained by pan-genomic analysis was used, considering all strains in the dataset. The database used was CARD and VICTORS respectively. To infer presence, a minimum identity of 70%. Two resistance genes were found, namely (I) mtrA, the gene encoding the "DNA-binding response regulator MtrA" protein, an efflux pump capable of generating resistance to macrolides; (II) rpoB2, an analog of rpoB that shows the same function as its paralog and that confers resistance to rifampicin by acting as a competitive inhibitor (https://www.uniprot.org/uniprotkb/?query=*).

Twelve virulence genes were found, as follows (I) mtrA – resistance gene mentioned above; (II) panD - coding for the aspartate alpha-decarboxylase protein that determines the rate of the gamma-aminobutyric acid reaction (III) icl - encoding the enzyme isocitrate lyase that acts on the malate biosynthesizer complex and on energy production; (IV) Rv3489 - encoding a hypothetical protein (V) accD1 - gene encoding the beta subunit of the carboxylase carboxyl transferase enzyme and acts in the production of energy through the manipulation of Acetyl-CoA; (VI) relay -

Table 1
Summary of results found in the analysis of ANI, dDDH, and pangenome.

Phylogeny *gyrB* and Phylogenomic	Species	DDH	ANI	Pangenome values	Strains that can be reassigned	References
Group A	*Dietzia maris* *Dietzia kunjamensis* *Dietzia kunjamensis* subsp *schimae* *Dietzia alimentaria*	>70%	>95%	$\alpha = 0.74$	97, YIM 65001T, IMV 195, 313, BP-27-1, DSM 44907, K30-10T, DSM 45139	[4,5,7,9,11]
Group B	*Dietzia maris* *Dietzia cinnamea* *Dietzia papillomatosis* *Dietzia massiliensis*	>70%	>95%	$\alpha = 0.82$	SST1, DSM-43672, BP_168, NBRC-102147, AMT, NBRC-105045, Marseille-Q0999T	[4,11,13,25]
Group C	*Dietzia cercidiphylli* *Dietzia natronolimnaea*	>70%	>95%	$\alpha = 0.75$	YIM-65002T, HS-1	[7,12,75]
Group D	Species no assigned	>70%	>95%	$\alpha = 0.88$	D. sp DQ12_76, D. sp JS16-p6b, D. sp SYD_A1, D.sp DQ11_38_2, D. sp DQ12-45-1b	

The table shows the results of the species that had an ANI value > 95% and dDDH >70%, the alpha value of the pangenome, and the strains that may need reassignment. All other species described for the genus showed ANI <95%, and dDDH <70%, corroborating the description of the species.

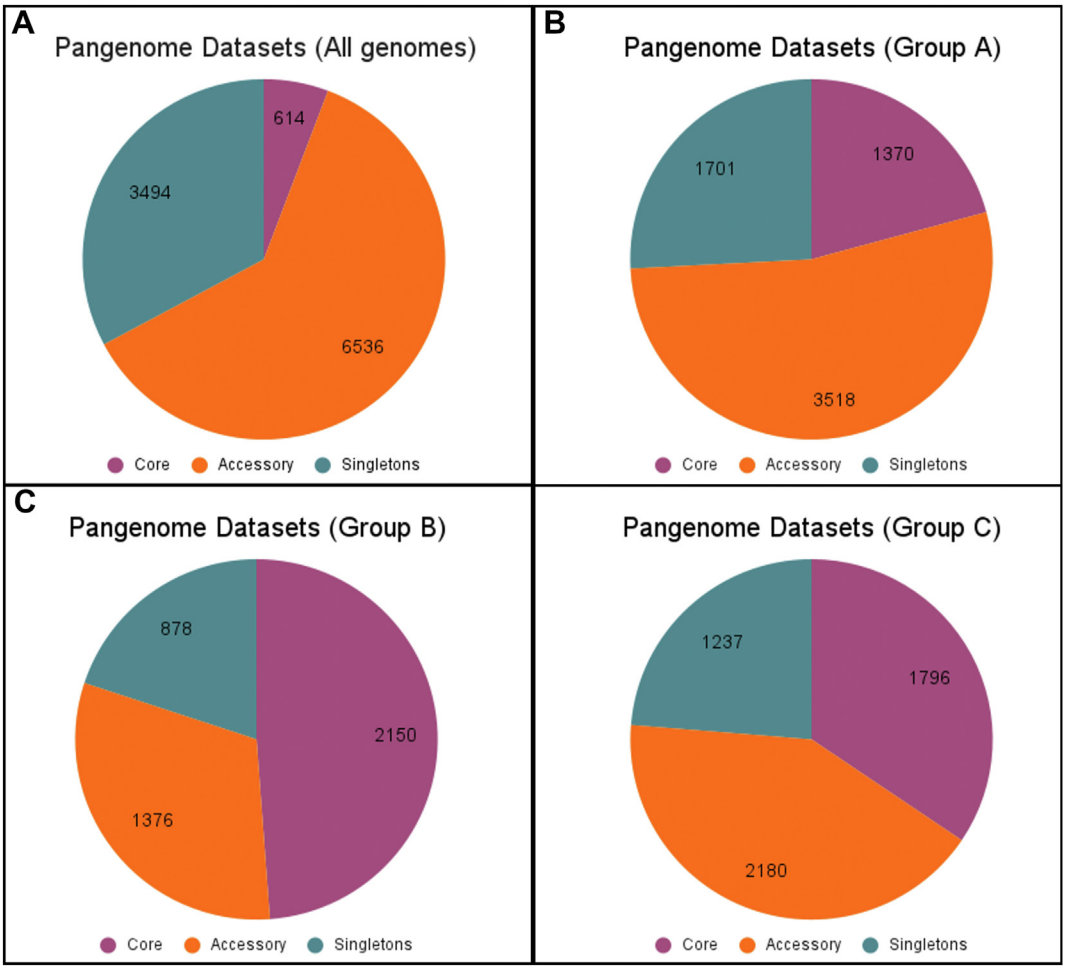


Fig. 4. Diagram depicting the subsets of the *Dietzia* pangenome (A): Pangenome subsets from an analysis based on all 56 *Dietzia* genomes. Graph (B): subset based on the analysis of eighteen genomes from group A. Graph (C): subset based on the analysis of eight genomes from group B. Graph (D): subset based on the analysis of nine genomes from group C.

a gene associated with the success of *Pseudomonas aeruginosa* as an extracellular pathogen (VII) sigH – coding factor H of RNA polymerase; (VIII) mprA – a transcriptional repressor and also acts as a response to antibiotics (IX) ino1 - encoded inositol-3-phosphate synthase and acts on the conversion of glucose into inositol; (X) coaA-a structural gene that acts in the biosynthesis of coenzyme A (XI) ftsY - Involved in targeting and insertion of nascent membrane proteins into the cytoplasmic membrane. Acts as a receiver for the complex formed by the signal recognition particle (SRP) and the ribosome-nascent chain (RNC). Interaction with SRP-RNC leads to the transfer of the RNC complex to the Sec translocase for insertion into the membrane, the hydrolysis of GTP by both Ffh and FtsY, and the dissociation of the SRP-FtsY complex into the individual components (UNIPROT, https://www.uniprot.org/uniprotkb/?

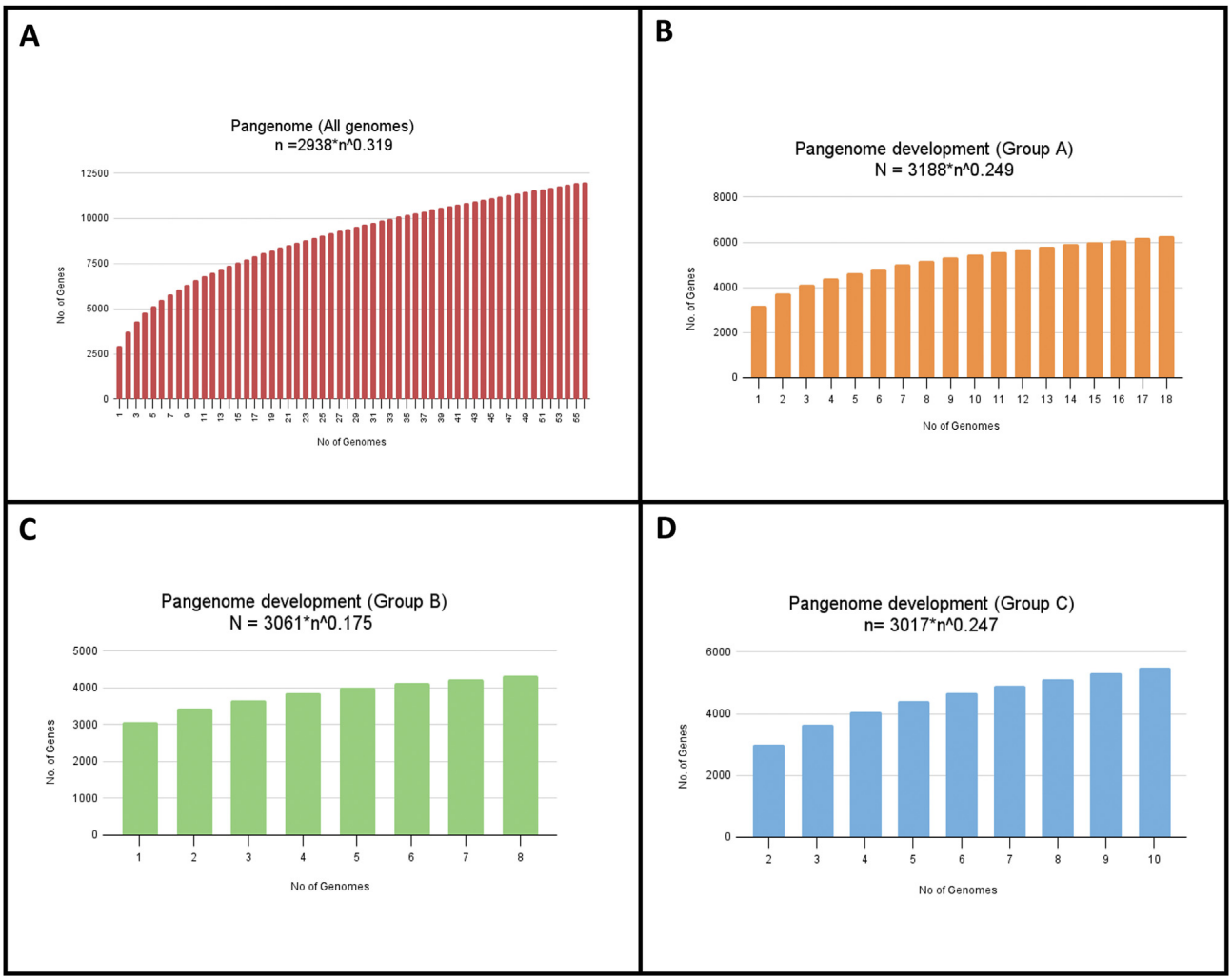


Fig. 5. Pangenome development. Top left graph: pangenome development using 56 *Dietzia* genomes. Upper right graph: development based on 18 genomes of group A. Lower left graph: development based on 8 genomes of group B, and the lower right graph: development based on 9 genomes of group C.

query=*); (XII) sigE — necessary factor for the differentiation and development of heterocysts.

3.6. Detection of islands in Dietzia's genome and variations in symbiotic islands between group A and group C

The genome plasticity of the three groups was determined using GIPSy (Genomic Island Prediction Software). BRIG software (BLAST Ring Image Generator) [25] was used to visualize the circular comparison of the genome, using only the samples present in the group. We found differences in the presence/absence of symbiotic islands in a study based on groups: seven symbiotic islands (SIs) in group A; and fourteen SIs in group C (Fig. 6A and B).

Regarding the presence of genes in SIs, we compared genes in all strains of group A with each other. When we compared the genes related to the SIs of group A with those of group C, there were differences in several genes found that were present only in group A and others only in group C. Among these genes, we can highlight the cupin domain-containing protein present only in the C group. Cupin domain-containing proteins have been found in organisms from thermophilic bacteria to plants and animals that inhabit a whole spectrum of environments. When the domain is conserved, it can provide a stable scaffold that allows these proteins to survive and function under various extreme conditions [55]. In addition, the clusters of genes related to SI7 of group A were similar to SI3 of group C, where there were differences in three genes found that were present only in A (SDR family oxidoreductase, TetR/AcrR family transcriptional regulator, phosphotransferase family protein) and three other genes present only in group C (C40 family peptidase, precorrin-6A/cobalt-precorrin-6A reductase, cobyrinic acid a,c-diamide synthase) (Fig. 6A and B).

Interestingly, we found that the genomic region related to SI7 and SI13 of group C was not present in any of the SIs of group A. This may reflect that the genomic signature of these regions can already be adapted to specific isolation sites. The list of genes related to SI7 and SI13 of group C is mentioned in Additional file: Table S4.

3.7. Detection of islands in Dietzia's genome and variations in pathogenic islands (PAIs) between groups A, B, and group C

The pathogenic islands were analyzed for the three groups A, B, and C: seven PAIs in group A; eight PAIs in group B, and ten PAIs in group C (Fig. 7A, B, and C, respectively). Nonetheless, aimed at finding specific genes in group B, as it is the group that presents isolated cases from clinical samples. Regarding PAIs, it was not possible to find any island for any of the three groups, however,

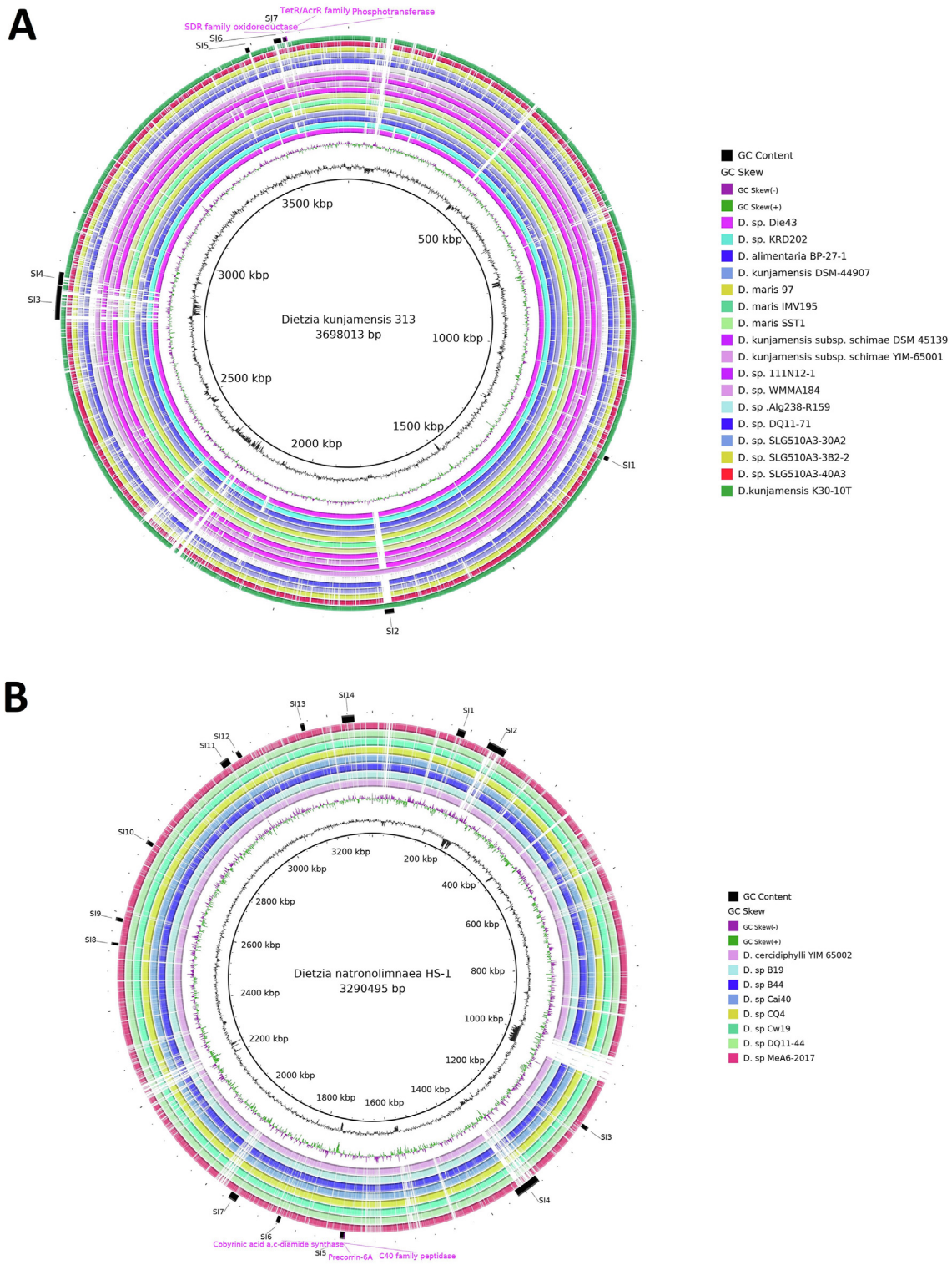


Fig. 6. Circular genomic representation of islands (SIs) in group A and C genomes. A) All genomes were aligned using *Dietzia kunjamensis* 313 as a reference. B) All genomes were aligned using *Dietzia natronolimnaea* HS-1. From the inner to the outer circle, the strains were represented in different colors. The figure represents the coding sequences (CDS) and SIs (Symbiotic islands).

several genes were specific to each group: group A (72), group B (47) and group C (97). However, group B presented important proteins such as type II toxin-antitoxin system Phd/YefM family antitoxin, antibiotic biosynthesis monooxygenase, ferric iron reductase, and universal stress protein, which may be associated with its pathogenic role (Fig. 7B).

4. Discussion

The micromorphological, biochemical, growth, and cellular chemotaxonomic characteristics of a microorganism are often applied to isolate and identify that microorganism [27]. However, these strategies are not always adequate for classification and

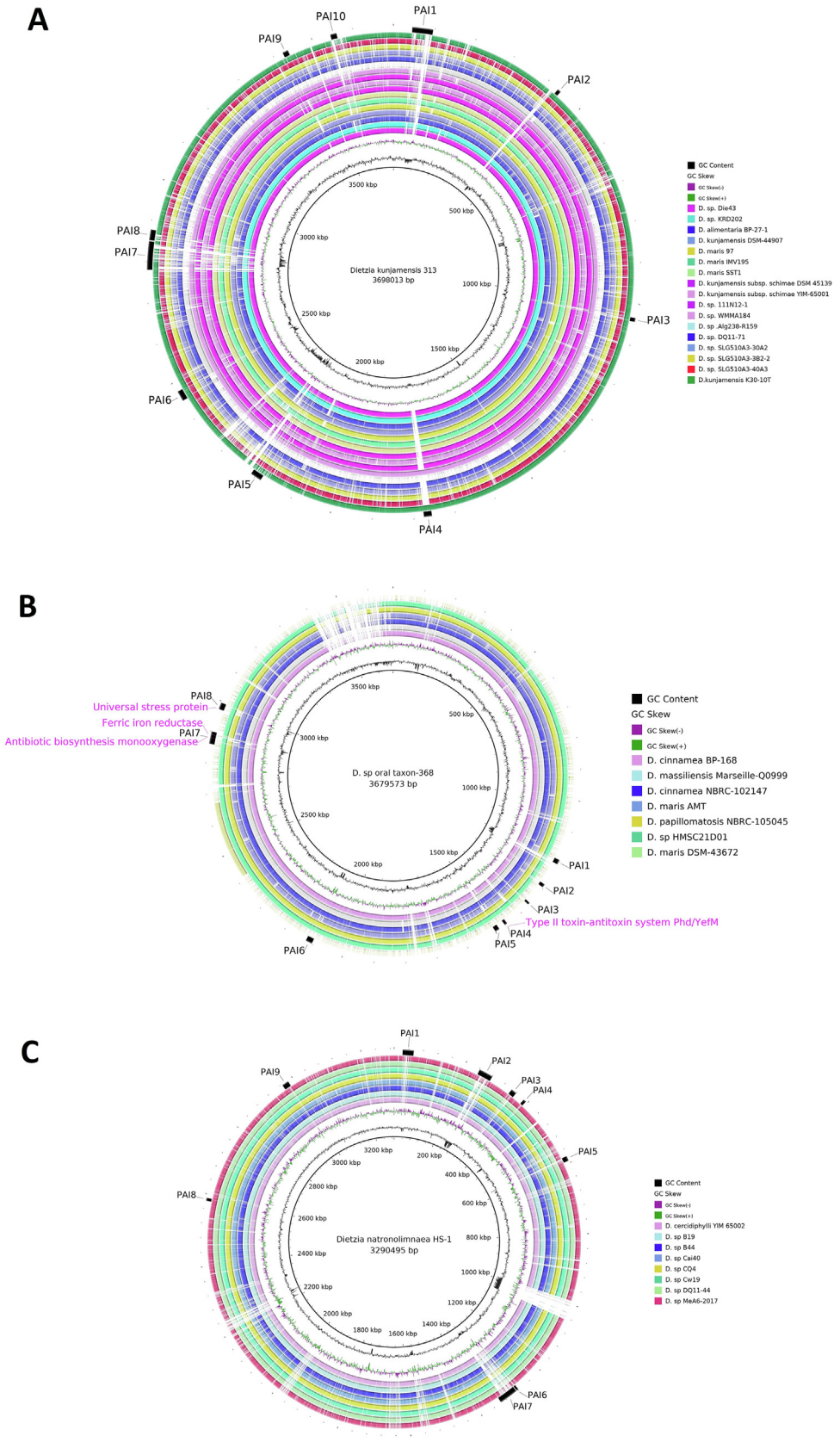


Fig. 7. Circular genomic representation of islands (PAIs) in group A b and C genomes. A) All genomes were aligned using *Dietzia kunjamensis* 313 as a reference. B) All genomes were aligned using *Dietzia* sp. oral taxon 368. C) All genomes were aligned using *Dietzia natronolimnaea* HS-1. From the inner to the outer circle, the strains were represented in different colors. The figure represents the coding sequences (CDS) and PAIs (Pathogenic islands).

identification, some studies have carried out the reclassification of some *Dietzia* species that had previously been misidentified as *R. equi*, as well as other actinomycetes [27].

Species definition and the best methods for taxonomic classification in prokaryotes continue to be debated [56,57]. Measures of phenotypic similarity have defined many bacterial species and many of these methods may be inadequate or unable to keep up with the levels of diversity being discovered in nature [58]. 16S rRNA genes have been widely used in bacteria to study phylogenetic relationships. Despite that, this approach is difficult since several forces that configure the evolution of bacterial genomes act with different strengths on different parts of the genome and bacterial strains [59,60]. Therefore, it is important to consider not only 16S rRNA sequences but also the whole genome to determine the diversity within a bacterial genus or species.

The study of the evolutionary history of the genus *Dietzia* has been primarily carried out through 16S rRNA sequence analysis and phenotypic tests, including traditional biochemical tests [2,27]. Some authors have reported difficulty in differentiating *Dietzia* species. For instance, in the study of Niwa et al., 2012, the sequence of the 16S rRNA gene from the type strains of *D. schimae* and *D. cercidiphylli* were identical to *D. maris* and *D. natronolimnaea*, respectively, and it was not possible to identify which species the strain belonged only to using the 16S rRNA gene [61]. Sudhindra et al., 2016 showed in their work that *Dietzia* spp. isolated from cardiac tissue showed that *D. cercidiphylli* has an identical 16S rRNA sequence and cannot be distinguished from *D. natronolimnaea* using the MicroSEQ 500 sequencing kit [62].

Phylogenetic analysis using sequences from the 16S rRNA gene has been applied and has also become important as a means of identifying an unknown bacterium at the genus or species level, but for the genus *Dietzia,* the 16S rRNA gene has not been a good marker for the identification of bacteria. species, as the 16S rRNA gene sequences, can sometimes be identical. The 16S rRNA gene is often not applicable to incomplete genomes (e.g., genes are not assembled, and have partial genes), as species characterization is based on sequencing and PCR amplification of 16S ribosomal RNA (rRNA) gene sequences [30]. Consequently, analysis of universal genes does not provide sufficient resolution at the species level and has often resulted in the lack of clear genetic discontinuities between closely related taxa.

Here in this work, using the 16S rRNA and *gyrB* gene, it was not possible to have a clear view of the distinction of the species of the genus, the species are mixed in the broths formed in the tree. However, the *gyrB* gene managed to separate these species with better precision, presenting a better bootstrap value than the 16S, most of the bootstrap values for the 16S were less than 50%. Despite better clade formation using the *gyrB* gene, many *Dietzia* strains do not cluster with their species. *D. maris* appears in two distinct clades, along with other species such as *D. cinnamea*, and *D. papillomatosis*, making it difficult to distinguish precisely which species is the real one.

These ambiguities motivated our use of other techniques to better understand diversity within the *Dietzia* genus. A genome-wide similarity matrix was obtained with the Gegenees software and used for a phylogenomic analysis showed 4 large groups: A, B, C and D. What we show once again that the 16S rRNA gene is not a good marker for classify *Dietzia* species. To improve the resolution of our analysis, ANI analysis showed new relationships identified, in group A, *D. maris* 97, *D. maris* IMV-195, *D. maris* SST1, *D. alimentaria* BP-27-1, *D. kunjamensis* K30-10^T, *D. kunjamensis* 313, *D. kunjamensis* DSM-44907, *D. kunjamensis* subs. schimae DSM-45139 and *D. kunjamensis* subs. schema YIM-65001^T, grouped, showing to be very close, and can be considered distinct strains and not distinct species; however, there are two types of strains in this group, which makes it difficult to know which species this group really belongs to. In group B, *D. cinnamea* NBRC 102147, *D. papillomatosis* NBRC105045, *D. maris* DSM-43672, *D. maris* AMT, and *D. massiliensis* Marseille-Q0999^T could belong to a single species, and not distinct species. Likewise, in the genomes of group C, the classification of *D. natronolimnaea* HS-1 could be a strain of *D. cercidiphylli* YIM 65002^T considering that *D. cercidiphylli* YIM 65002^T is a strain type. In group D, only genomes designated with the genus, but which have a high ANI value, share more than 98.64%.

We also applied a strategy based on the calculation of DNA−DNA Hybridization. Here, for the four groups (A, B, C, and D) the dDDH value presented a value greater than 70%, except *D. kunjamensis* subsp. schimae DSM-45139 and *D. massiliensis* Marseille-Q0999^T presented a DDH lower than 70%, which is the limit to consider that two samples belong to the same species, as shown in the results of the AN analysis, except for these two isolates. These results corroborate the results found in the work of Nouioui et al., 2018 in which in their DDH analyses it was not possible to distinguish some validly named species, even at the subspecies level, since they presented dDDH values > 79%, these included *Cellulosimicrobium aquatile* [63]) vs. *Cellulosimicrobium funkei* [64]; *D. cinnamea* [65] vs. *D. maris* [11]; several species of *Bifidobacterium, Nocardiopsis* and *Streptomyces* [3].

To further corroborate the existence of these potential "misclassifications of *Dietzia* species", we performed a pangenome analysis, which confirmed the new relationships identified in previous analyses. Using our grouping of the *Dietzia* genus, supported by the results of the Gegenees, ANI, and DDH, investigated genomic diversity through pangenome calculation analysis. The same four datasets were used to perform pangenome calculations. As would be expected, the number of genes in the core genome was greater in each cluster than in the dataset containing all 56 genomes, which is consistent with the idea that the genomes within a cluster are of the same species. Extrapolations from pangenome development also supported this assumption. The α value generated from an analysis of all genomes indicated that the *Dietzia* genus has an open pangenome ($\alpha = 0.69$), as well as each of the groups ($\alpha = 0.74$, $\alpha = 0.82$, $\alpha = 0.75$ and $\alpha = 0.88$, respectively). However, intragroup α values reveal that these latter pangenomes are increasing in size more slowly than the pangenome of all species (as indicated by higher α values). This means that if we consider all genomes as part of the *Dietzia* genus, each new genome sequenced will substantially increase the number of non-redundant genes in this genus. This phylogenetic approach to pangenome analysis revealed patterns fully in agreement with the results of our other analyses. In the tree of the core genome, it is also possible to visualize the separation of the four groups defined in the previous analyses, corroborating the idea that some species of *Dietzia* are strains of some species and not distinct species.

Interspecies genome plasticity can result from several events, of which horizontal gene transfer is particularly important because it can cause the acquisition of gene blocks (genomic islands, or GIs), producing evolution by quantum leaps [66]. Generally, these genes are often flanked by transposases (insertion elements) and have altered G + C content and skew, suggesting their acquisition through Horizontal Gene Transfer (HGT), mediated by phages or recombination [67].

The analysis of genome plasticity reveals differences in the presence and absence of some regions of the genome when compared at the group level. Symbiotic islands provide the bacterium with a genomic repertoire to sustain a symbiotic host−bacterium relationship [68] and characterize a class of Genomic Island [51]. Comparative analysis of SIs showed the absence of genes at the group level. We found clusters of genes

belonging to SI7 and SI13 of group C that was not identified in any SI of group A. However, only SI7 was present in all strains of group C. A protein present in SI7 is ribbon-helix-helix protein, *CopG* family, this protein regulates the transcription of genes that may be involved in metal uptake, amino acid biosynthesis, cell division, plasmid copy number control, bacteriophage lytic cycle [69]. Another protein found in SI7 is a trypsin-like serine protease. Prokaryotic serine proteases are involved in several physiological processes associated with cell signaling, defense response, and development. Proteases belonging to the trypsin family have been implicated in the heat shock response, as well as being important in the growth and defense response in various bacteria [70].

The type II protein toxin-antitoxin system Phd/YefM family antitoxin found for group B are small genetic elements composed of a toxic protein and its cognate antitoxin protein, the latter neutralizing the toxicity of the former. This protein has been implicated in several biological processes, including plasmid maintenance, phage inhibition, persistence, stress response, and biofilm formation [71]. Another important protein for the group was the antibiotic biosynthesis monooxygenase (ABM) domain, which is found in proteins involved in several biological processes, including metabolism, transcription, translation, and secondary metabolite biosynthesis, and in *Streptomyces coelicolor* ActVA-Orf6, monooxygenase, a role in the biosynthesis of aromatic polyketides, specifically the antibiotic actinorhodin, oxidizing phenolic groups to quinones [72]. Universal stress protein participates in a range of cellular responses to biotic and biotic stresses. Its physiological functions are associated with ion elimination, hypoxia responses, cell mobility, and regulation of cell growth and development, which may consist of its roles in resistance to multiple stresses [73]. Finally, ferric reductases are widely distributed electron transport proteins that are expressed by numerous infectious organisms and have been identified as a key virulence factor for the establishment of infection by single-celled pathogens [74]. The true pathogenic profile of *Dietzia* is not yet known, nevertheless, these proteins may contribute to the true pathogenicity of some species described for the genus.

5. Conclusion

Taken together, our analyzes suggest that there is more phylogenetic complexity in classifying these species than has been previously shown in other studies. In short, each of the analyzes we performed suggested the existence of an error in the classification of the species of the genus *Dietzia*, which leads us to the fact that notably, some species probably need to be reassigned. In addition, considering the genes in SIs and PAIs, we identified some absence of symbiotic islands in all strains present in group A. The genes present in the pathogenicity islands of group B are absent in the strains of groups A and C, which reveals that the presence/absence of genomic islands may be related to the isolation location.

5.1. Final remarks

Here, we used a variety of methods to analyze whole genome data, which together showed that: (i) the current application for differentiating between *Dietzia* species should be reviewed; (ii) this genus contains at least four separate groups; and (iii) sequencing of the *gyrB* gene can be used efficiently to reliably identify *Dietzia* species. Collectively, our results provide evidence for reconsidering the phylogenetic and genomic relationship between species currently assigned to *Dietzia*.

Furthermore, to the best of our knowledge, this is the first work to combine an analysis of pangenome development with an analysis of ANI and dDDH to support the attribution that some *Dietzia* species could be reassigned. Here, we aim that subsets of pangenomes and the α value generated by these analyzes can be used as an alternative approach, together with ANI, and dDDH, for the *in-silico* classification of species of the genus *Dietzia*. A high intragroup α value can be considered a good indicator of a new, more clonal species within the genus, although low α values can be discovered within a species group due to the high degree of variation between genomes resulting from horizontal gene transfer events.

Author Contributions

Wrote the manuscript (RGS); designed the study (RGS); performed the experimental work (ARA, RGS, NS); conducted the *in-silico* analyses and interpreted the results (RGS, RH, AS); critically reviewed and revised the manuscript (RHC, DLNR, AL, WFA, CR, ARA, ST, AKJ, SS, GM, FMR, ACPG, LCLJ, FFA, BB, VC, TLPC, NS, AS, AGN, TJS, VA); supervised the study (VA, TJS). All authors approved this manuscript for publication.

Funding

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil), and FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais).

Declaration of Competing Interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

The authors acknowledge the collaboration and assistance of all team members and the Brazilian funding agencies CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil) Finance Code 001, CNPq(Conselho Nacional de Desenvolvimento Científico e Tecnológico), and FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais) for providing financial support and the UFMG Microscopy Center, and finally, We thank the Network of Omic Sciences (RECOM) for the support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.resmic.2022.103998.

Accession codes *Dietzia* sp. B32 genome is publicly deposited in GenBank under accession number CP093845, Bio project number PRJNA802820, and Biosample SAMN25598616.

References

[1] Kämpfer P, Langer S, Martin E, Jäckel U, Busse H-J. Dietzia aerolata sp. nov., isolated from the air of a duck barn, and emended description of the genus Dietzia Rainey et al. 1995. Int J Syst Evol Microbiol 2010;60:393–6. https://doi.org/10.1099/IJS.0.012807-0.

[2] Koerner RJ, Goodfellow M, Jones AL. The genus Dietzia: a new home for some known and emerging opportunist pathogens. FEMS Immunol Med Microbiol 2009;55:296–305. https://doi.org/10.1111/j.1574-695X.2008.00513.x.

[3] Nouioui I, Carro L, García-López M, Meier-Kolthoff JP, Woyke T, Kyrpides NC, et al. Genome-based taxonomic classification of the phylum actinobacteria. Front Microbiol 2018;9:2007. https://doi.org/10.3389/FMICB.2018.02007/BIBTEX.

[4] Nesterenko OA, Nogina TM, Kasumova SA. Rhodococcus luteus nom. nov. and Rhodococcus maris nom. nov. Int J Syst Bacteriol 1982;32:1–14. https://doi.org/10.1099/00207713-32-1-1.

[5] Kim J, Roh SW, Choi J-H, Jung M-J, Nam Y-D, Kim M-S, et al. Dietzia alimentaria sp. nov., isolated from a traditional Korean food. Int J Syst Evol Microbiol 2011;61:2254–8. https://doi.org/10.1099/IJS.0.021501-0.
[6] Kämpfer P, Falsen E, Frischmann A, Busse HJ. Dietzia aurantiaca sp. nov., isolated from a human clinical specimen. Int J Syst Evol Microbiol 2012;62:484–8. https://doi.org/10.1099/IJS.0.032557-0.
[7] Li J, Zhao G-Z, Zhang Y-Q, Klenk H-P, Pukall R, Qin S, et al. Dietzia schimae sp. nov. and Dietzia cercidiphylli sp. nov., from surface-sterilized plant tissues. Int J Syst Evol Microbiol 2008;58:2549–54. https://doi.org/10.1099/IJS.0.2008/000919-0.
[8] Yassin AF, Hupfer H, Schaal KP. Dietzia cinnamea sp. nov., a novel species isolated from a perianal swab of a patient with a bone marrow transplant. Int J Syst Evol Microbiol 2006;56:641–5. https://doi.org/10.1099/IJS.0.63863-0.
[9] Mayilraj S, Suresh K, Kroppenstedt RM, Saini HS. Dietzia kunjamensis sp. nov., isolated from the Indian Himalayas. Int J Syst Evol Microbiol 2006;56:1667–71. https://doi.org/10.1099/IJS.0.64212-0.
[10] Li J, Chen C, Zhao GZ, Klenk HP, Pukall R, Zhang YQ, et al. Description of Dietzia lutea sp. nov., isolated from a desert soil in Egypt. Syst Appl Microbiol 2009;32:118–23. https://doi.org/10.1016/J.SYAPM.2008.11.007.
[11] Rainey FA, Klatte S, Kroppenstedt RM, Stackebrandt E. Dietzia, a new genus including Dietzia maris comb. nov., formerly Rhodococcus maris. Int J Syst Bacteriol 1995;45:32–6. https://doi.org/10.1099/00207713-45-1-32.
[12] Duckworth AW, Grant S, Grant WD, Jones BE, Meijer D. Dietzia natronolimnaios sp. nov., a new member of the genus Dietzia isolated from an East African soda lake. Extremophiles 1998;2:359–66. https://doi.org/10.1007/s007920050079.
[13] Jones AL, Koerner RJ, Natarajan S, Perry JD, Goodfellow M. Dietzia papillomatosis sp. nov., a novel actinomycete isolated from the skin of an immunocompetent patient with confluent and reticulated papillomatosis. Int J Syst Evol Microbiol 2008;58:68–72. https://doi.org/10.1099/IJS.0.65178-0/CITE/REFWORKS.
[14] Yumoto I, Nakamura A, Iwata H, Kojima K, Kusumoto K, Nodasaka Y, et al. Dietzia psychralcaliphila sp. nov., a novel, facultatively psychrophilic alkaliphile that grows on hydrocarbons. Int J Syst Evol Microbiol 2002;52:85–90. https://doi.org/10.1099/00207713-52-1-85/CITE/REFWORKS.
[15] Yamamura H, Lisdiyanti P, Ridwan R, Ratnakomala S, Sarawati R, Lestari Y, et al. Dietzia timorensis sp. nov., isolated from soil. Int J Syst Evol Microbiol 2010;60:451–4. https://doi.org/10.1099/IJS.0.012229-0.
[16] Parte AC. LPSN—list of prokaryotic names with standing in nomenclature. Nucleic Acids Res 2014;42:D613. https://doi.org/10.1093/NAR/GKT1111.
[17] Brown WD, Feinberg N, Stedman E, Dejace J, Hale AJ. Dietzia cinnamea: an increasingly recognized human pathogen. IDCases 2022;29:e01539. https://doi.org/10.1016/J.IDCR.2022.E01539.
[18] Bemer-Melchior P, Haloun A, Riegel P, Drugeon HB. Bacteremia due to Dietzia maris in an immunocompromised patient. Clin Infect Dis 1999;29:1338–40. https://doi.org/10.1086/313490.
[19] Cawcutt KA, Bhatti MM, Nelson DR. Pleural fluid infection caused by Dietzia cinnamea. Diagn Microbiol Infect Dis 2016;85:496–7. https://doi.org/10.1016/J.DIAGMICROBIO.2016.04.007.
[20] Rammer P, Calum H, Moser C, Björnsdóttir MK, Smedegaard H, Høiby N, et al. Dietzia papillomatosis bacteremia. J Clin Microbiol 2013;51:1977. https://doi.org/10.1128/JCM.03313-12.
[21] Navaratnam J, Dedi L, et al Reports in infectious, 2018 undefined. Identification of Dietzia Species in a Patient with Endophthalmitis Following Penetrating Injury with Retained Intraocular Metallic Foreign Body. HindawiCom n.d.
[22] Perkin S, Wilson A, Walker D, McWilliams E. Dietzia species pacemaker pocket infection: an unusual organism in human infections. BMJ Case Rep 2012;2012. https://doi.org/10.1136/BCR.10.2011.5011.
[23] Reyes G, Navarro JL, Gamallo C, de Las Cuevas MC. Type A aortic dissection associated with Dietzia maris. Interact Cardiovasc Thorac Surg 2006;5:666–8. https://doi.org/10.1510/ICVTS.2006.135640.
[24] Pidoux O, Argenson JN, Jacomo V, Drancourt M. Molecular identification of a Dietzia maris hip prosthesis infection isolate. J Clin Microbiol 2001;39:2634–6. https://doi.org/10.1128/JCM.39.7.2634-2636.2001.
[25] Olowo-okere A, Ibrahim YKE, Lo CI, Olayinka BO, Yimagou EK, Yacouba A, et al. Bhargavaea massiliensis sp. nov. And Dietzia massiliensis sp. nov., novel bacteria species isolated from human urine samples in Nigeria. Curr Microbiol 2021;79. https://doi.org/10.1007/S00284-021-02721-4.
[26] Hirvonen JJ, Lepistö I, Mero S, Kaukoranta SS. First isolation of Dietzia cinnamea from a dog bite wound in an adult patient. J Clin Microbiol 2012;50:4163–5. https://doi.org/10.1128/JCM.01939-12.
[27] Gharibzahedi SMT, Razavi SH, Mousavi SM. Characterization of bacteria of the genus Dietzia: an updated review. Ann Microbiol 2013;64(1):1–11. https://doi.org/10.1007/S13213-013-0603-3.
[28] Fang H, Xu J-B, Nie Y, Wu X-L. Pan-genomic analysis reveals that the evolution of Dietzia species depends on their living habitats. Environ Microbiol 2021;23:861–77. https://doi.org/10.1111/1462-2920.15176.
[29] Nazina TN, Shumkova ES, DSh Sokolova, Babich TL, Zhurina Mv, Xue Y-F, et al. Identification of hydrocarbon-oxidizing Dietzia bacteria from petroleum reservoirs based on phenotypic properties and analysis of the 16S rRNA and gyrB genes. Microbiology 2015;84(3):377–88. https://doi.org/10.1134/S0026261715030145.
[30] Jain C, Rodriguez -RLM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 2018;9(1):1–8. https://doi.org/10.1038/s41467-018-07641-9.
[31] Outten J, Warren A. Methods and developments in graphical pangenomics. J Indian Inst Sci 2021;101:485. https://doi.org/10.1007/S41745-021-00255-Z.
[32] Andrews S. FastQC A quality control Tool for high throughput sequence data. 2010. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. [Accessed 25 May 2020]. accessed.
[33] Hernandez D, François P, Farinelli L, Østerås M, Schrenzel J. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. Genome Res 2008;18:802–9. https://doi.org/10.1101/gr.072033.107.
[34] Boetzer M, Henkel C v, Jansen HJ, Butler D, Pirovano W. Scaffolding preassembled contigs using SSPACE. BIOINFORMATICS APPLICATIONS NOTE 2011;27:578–9. https://doi.org/10.1093/bioinformatics/btq683.
[35] Nadalin F, Vezzi F, Policriti A. GapFiller: a de novo assembly approach to fill the gap within paired reads. BMC Bioinf 2012;13. https://doi.org/10.1186/1471-2105-13-S14-S8.
[36] Guizelini D, Raittz RT, Cruz LM, Souza EM, Steffens MBR, Pedrosa FO. GFinisher: a new strategy to refine and finish bacterial genome assemblies. Sci Rep 2016;6:1–8. https://doi.org/10.1038/srep34963.
[37] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.
[38] Ågren J, Sundström A, Håfström T, Segerman B. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PLoS One 2012;7. https://doi.org/10.1371/JOURNAL.PONE.0039107.
[39] Kloepper TH, Huson DH. Drawing explicit phylogenetic networks and their integration into SplitsTree. BMC Evol Biol 2008;8:22. https://doi.org/10.1186/1471-2148-8-22.
[40] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 2018;35:1547. https://doi.org/10.1093/MOLBEV/MSY096.
[41] Felsenstein J. Confidence limits ON phylogenies: an approach using the bootstrap. Evolution 1985;39:783–91. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.
[42] Letunic I, Bork P. Interactive Tree of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 2019;47:W256–9. https://doi.org/10.1093/NAR/GKZ239.
[43] Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 2016;32:929–31. https://doi.org/10.1093/bioinformatics/btv681.
[44] Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 2007;57:81–91. https://doi.org/10.1099/ijs.0.64483-0.
[45] Stackebrandt E, Goebel BM. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Bacteriol 1994;44:846–9. https://doi.org/10.1099/00207713-44-4-846/CITE/REFWORKS.
[46] Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol 2019;20(1):1–14. https://doi.org/10.1186/S13059-019-1832-Y.
[47] Soares SC, Silva A, Trost E, Blom J, Ramos R, Carneiro A, et al. The pan-genome of the animal pathogen Corynebacterium pseudotuberculosis reveals differences in genome plasticity between the biovar ovis and equi strains. PLoS One 2013;8. https://doi.org/10.1371/JOURNAL.PONE.0053818.
[48] Carlos Guimaraes L, Benevides de Jesus L, Vinicius Canario Viana M, Silva A, Thiago Juca Ramos R, de Castro Soares S, et al. Inside the pan-genome - methods and software overview. Curr Genom 2015;16:245–52. https://doi.org/10.2174/1389202916666150423002311.
[49] Sayers S, Li L, Ong E, Deng S, Fu G, Lin Y, et al. Victors: a web-based knowledge base of virulence factors in human and animal pathogens. Nucleic Acids Res 2019;47:D693–700. https://doi.org/10.1093/NAR/GKY999.
[50] Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. Card 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 2020;48:D517–25. https://doi.org/10.1093/NAR/GKZ935.
[51] Soares SC, Geyik H, Ramos RTJ, de Sá PHCG, Barbosa EGV, Baumbach J, et al. GIPSy: genomic island prediction software. J Biotechnol 2016;232:2–11. https://doi.org/10.1016/j.jbiotec.2015.09.008.
[52] Alikhan N-F, Petty NK, ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genom 2011;12:402. https://doi.org/10.1186/1471-2164-12-402.
[53] Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 2005;102:2567. https://doi.org/10.1073/PNAS.0409727102.
[54] Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 2014;64:346–51. https://doi.org/10.1099/IJS.0.059774-0.
[55] Khuri S, Bakker FT, Dunwell JM. Phylogeny, function, and evolution of the cupins, a structurally conserved, functionally diverse superfamily of proteins. Mol Biol Evol 2001;18:593–605. https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A003840.

[56] Phillippe H, CJ D. Horizontal gene transfer and phylogenetics. Curr Opin Microbiol 2003;6:498–505. https://doi.org/10.1016/J.MIB.2003.09.008.
[57] Bruen TC, Philippe HH, Bryant D. A simple and robust statistical test for detecting the presence of recombination. Genetics 2006;172:2665–81. https://doi.org/10.1534/genetics.105.048975.
[58] Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, et al. Re-evaluating prokaryotic species. Nat Rev Microbiol 2005;3(9):733–9. https://doi.org/10.1038/nrmicro1236.
[59] Chan JZM, Halachev MR, Loman NJ, Constantinidou C, Pallen MJ. Defining bacterial species in the genomic era: insights from the genus Acinetobacter. BMC Microbiol 2012;12. https://doi.org/10.1186/1471-2180-12-302.
[60] Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J Clin Microbiol 2007;45:2761–4. https://doi.org/10.1128/JCM.01228-07.
[61] Niwa H, Lasker BA, Hinrikson HP, Franzen CG, Steigerwalt AG, Whitney AM, et al. Characterization of human clinical isolates of Dietzia species previously misidentified as Rhodococcus equi. Eur J Clin Microbiol Infect Dis 2012;31:811–20. https://doi.org/10.1007/s10096-011-1379-7.
[62] Sudhindra P, Wang G, Nadelman RB. Identification of Dietzia spp. from cardiac tissue by 16S rRNA PCR in a patient with culture-negative device-associated endocarditis: a case report and review of the literature. Case Rep Infect Dis 2016;2016:1–5. https://doi.org/10.1155/2016/8935052.
[63] Sultanpuram VR, Mothe T, Chintalapati S, Chintalapati VR. Cellulosimicrobium aquatile sp. nov., isolated from Panagal reservoir, Nalgonda, India. Antonie Leeuwenhoek 2015;108(6):1357–64. https://doi.org/10.1007/S10482-015-0588-Y.
[64] Brown JM, Steigerwalt AG, Morey RE, Daneshvar MI, Romero LJ, McNeilt MM. Characterization of clinical isolates previously identified as Oerskovia turbata: proposal of Cellulosimicrobium funkei sp. nov. and emended description of the genus Cellulosimicrobium. Int J Syst Evol Microbiol 2006;56:801–4. https://doi.org/10.1099/IJS.0.63882-0/CITE/REFWORKS.
[65] Yassin AF, Hupfer H. Williamsia deligens sp. nov., isolated from human blood. Int J Syst Evol Microbiol 2006;56:193–7. https://doi.org/10.1099/IJS.0.63856-0/CITE/REFWORKS.
[66] Karaolis DKR, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR. A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains. Proc Natl Acad Sci U S A 1998;95:3134–9. https://doi.org/10.1073/PNAS.95.6.3134.
[67] Tettelin H, Riley D, Cattuto C, Medini D. Comparative genomics: the bacterial pan-genome. Curr Opin Microbiol 2008;11:472–7. https://doi.org/10.1016/J.MIB.2008.09.006.
[68] Barcellos FG, Menna P, da Silva Batista JS, Hungria M. Evidence of horizontal transfer of symbiotic genes from a Bradyrhizobium japonicum inoculant strain to indigenous diazotrophs Sinorhizobium (Ensifer) fredii and Bradyrhizobium elkanii in a Brazilian Savannah soil. Appl Environ Microbiol 2007;73:2635–43. https://doi.org/10.1128/AEM.01823-06.
[69] Schreiter ER, Drennan CL. Ribbon-helix-helix transcription factors: variations on a theme. Nat Rev Microbiol 2007;5:710–20. https://doi.org/10.1038/NRMICRO1717.
[70] Tripathi LP, Sowdhamini R. Genome-wide survey of prokaryotic serine proteases: analysis of distribution and domain architectures of five serine protease families in prokaryotes. BMC Genom 2008;9:1–28. https://doi.org/10.1186/1471-2164-9-549/FIGURES/5.
[71] Fraikin N, Goormaghtigh F, Melderen L van. Type II toxin-antitoxin systems: evolution and revolutions. J Bacteriol 2020:202. https://doi.org/10.1128/JB.00763-19.
[72] Sciara G, Kendrew SG, Miele AE, Marsh NG, Federici L, Malatesta F, et al. The structure of ActVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. EMBO J 2003;22:205. https://doi.org/10.1093/EMBOJ/CDG031.
[73] Kvint K, Nachin L, Diez A, Nyström T. The bacterial universal stress protein: function and regulation. Curr Opin Microbiol 2003;6:140–5. https://doi.org/10.1016/S1369-5274(03)00025-0.
[74] Cain TJ, Smith AT. Ferric iron reductases and their contribution to unicellular ferrous iron uptake. J Inorg Biochem 2021;218:111407. https://doi.org/10.1016/J.JINORGBIO.2021.111407.
[75] Rostami F, Razavi SH, Sepahi AA, Gharibzahedi SMT. Canthaxanthin biosynthesis by Dietzia natronolimnaea HS-1: effects of inoculation and aeration rate. Braz J Microbiol 2014;45:447. https://doi.org/10.1590/S1517-83822014005000046.

5. CAPÍTULO III

5.1 OBJETIVOS

5.2 OBJETIVO GERAL

Caracterizar uma nova espécie do gênero *Dietzia*, isolada de abscesso de égua;

5.3 OBJETIVOS ESPECÍFICOS

- Caracterizar genomicamente a espécie identificada (sequenciar, montar, anotar e depositar o genoma);
- Caracterizar quimiotaxonomicamente a nova espécie identificada (classificação biológica baseado em semelhanças e diferenças na estrutura de certos compostos entre os organismos que estão sendo classificados).
- Caracterizar fenotipicamente a nova espécie identificada (observar e descrever características observáveis ou caracteres de um organismo ou população, como: morfologia, desenvolvimento, propriedades bioquímicas ou fisiológicas e comportamento).

5.4 ARTIGO CIENTÍFICO 3: *Dietzia calabriae* sp. nov., a bacterial species isolated from a mare's abscess

Artigo submetido ao journal em: Journal of Basic Microbiology; Fator de impacto (atual): 3.1

O gênero *Dietzia* é considerado um táxon monoespecífico e proposto por Rainey et al 1995. Este gênero pertence a subordem *Corynebacterineae*. O gênero apresenta 19 espécies descritas, entretanto apenas 13 dessas espécies apresenta o nome publicado validamente, incluindo sinônimos. Espécies de *Dietzia* já foram isoladas de diferentes habitantes, vários ambientes naturais, amostras clínicas, tecido de plantas dentre outros. Até o momento não havia nenhuma amostra de *Dietzia* isolada de um animal.

Neste trabalho isolamos e identificamos uma nova espécie do gênero, isolada de nódulos subcutâneos de uma égua de 12 anos de idade em Reggio Calabria, Itália. A amostra foi identificada através da técnica de MALDI - TOF, pertencendo ao gênero *Dietzia*, entretanto sem espécie distinguida. Sendo assim a amostra foi sequenciada, e depositada no NCBI. Por ser uma nova espécie todos os procedimentos necessários para descrição da espécie estão sendo realizados.

A nova espécie foi designada com o nome *Dietzia calabriae* B32, e apresenta na cor laranja, cresce de 4 a 40°C, catalase positiva, não reduz nitrato e é urease negativa. Nosso trabalho contribuirá na disponibilidade de dados do gênero, sendo essa a primeira espécie isolada de um animal sendo descrita, assim trará informações relevantes que podem ajudar no entendimento do gênero ser considerado oportunista.

Journal of Basic Microbiology

Dietzia calabriae sp. nov. is a bacterial species isolated from a mare's abscess

Journal:	*Journal of Basic Microbiology*
Manuscript ID	Draft
Wiley - Manuscript type:	Research Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	dos Santos, Roselane ; UFMG Rita , Attili ; University of Camerino, School of Biosciences and Veterinary Medicine de Jesus, Luís ; UFMG Neres Rodrigues, Diego; UFMG Lima, Alessandra ; UFMG Hurtado, Raquel ; UFMG Claudia, Rifici ; University of Messina Sharon, Spier ; University of California Davis Giuseppe , Mazzullo ; University of Messina Morais-Rodrigues , Francielly ; UFMG Gomide , Anne ; UFMG Brenig, Bertram; Georg-August-Universitat Gottingen Vincenzo , Cuteri ; University of Camerino, School of Biosciences and Veterinary Medicine Castro , Thiago ; UFBA Seyffert , Núbia ; UFBA Jaiswal, Arun ; UFMG Batista, Ana ; UFMG de Magalhaes , Mariana ; UFMG Figueira Aburjaile, Flávia; UFMG Góes-Neto, Aristóteles ; UFMG Tiwari, Sandeep; UFBA Azevedo, Vasco; UFMG
Keywords:	Dietzia, taxogenomics, genome analysis, mare, phylogeny

Dietzia calabriae sp. nov. is a bacterial species isolated from a mare's abscess

Roselane Gonçalves dos Santos[1], Attili Anna Rita[5] , Luís Cláudio Lima de Jesus[1], Diego Lucas Neres Rodrigues[9], Alessandra Lima[1], Raquel Hurtado[1],Rifici Claudia[4], Spier Sharon J.[6], Mazzullo Giuseppe[4], Francielly Morais-Rodrigues[1], Anne Cybelle Pinto Gomide[1], Bertram Brenig[7], Cuteri Vincenzo[5], Thiago Luiz de Paula Castro[1.3], Núbia Seyffert[1.2], Arun Kumar Jaiswal[1], Ana Carolina Silva Batista[8], Mariana T. Q. de Magalhaes[8], Flávia Figueira Aburjaile[9], Aristóteles Góes-Neto[10], Sandeep Tiwari[2,3]*, Vasco Azevedo[1]*

[1]Cellular and Molecular Genetics Laboratory, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil, vascoariston@gmail.com

[2]Institute of Biology, Federal University of Bahia, Salvador, BA, Brazil, nbseyffert@gmail.com

[3]Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil, castrotlp@gmail.com

[4]Department of Veterinary Science, University of Messina (Italy), Polo Universitario dell'Annunziata, 98168 Messina (ME) claudiarifici35@gmail.com, giuseppe.mazzullo@unime.it

[5]School of Biosciences and Veterinary Medicine, University of Camerino (Italy), Via Circonvallazione 93/95, 62024 Matelica (MC), Italy, annarita.attili@unicam.it, vincenzo.cuteri@unicam.it

[6] Department of Veterinary Medicine and Epidemiology, University of California, Davis, California, USA, institutional address, sjspier@ucdavis.edu

[7] Institute of Veterinary Medicine, University of Göttingen, Burckhardtweg 2, Göttingen, Germany, bbrenig@gwdg.de

[8] Macromolecular Biophysics Laboratory (LBM), Department of Biochemistry and Immunology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais 31270-901, Brazil. MQuezado@icb.ufmg.br

[9]Veterinary School of the Federal University of Minas Gerais, Department of Preventive Veterinary Medicine, Federal University of Minas Gerais – UFMG, Belo Horizonte, MG, Brazil, faburjaile@gmail.com

[10]Molecular and Computational Biology of Fungi Lab/ Department of Microbiology) / Institute of Biological Sciences, Federal University of Minas Gerais – UFMG, Belo Horizonte, MG, Brazil,

* **Correspondence:**

Corresponding Author: vascoariston@gmail.com and sandip_sbtbi@yahoo.com

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil), and FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais).

Abstract

This study provides a taxonomic characterization of a bacterial strain isolated from the abscess of a 12-year-old mare in Regio de Calabria, Italy. The strain of the new species shared the 16S rRNA gene sequence with species representing the genus *Dietzia*. However, the DNA-DNA *in silico* hybridization (dDDH) values indicate that they belong to a different genomic species. The estimated in silico value of dDDH between the species of the genus and the isolated B32^T strain was only 20.8% and 38.2%, respectively. The average nucleotide identity value (ANIb) of the B32^T strain among all genomes used in the study had a maximum value of 88.44%, which was not within the recommended threshold of 95-96% for bacterial species delimitation. In addition, the DNA G + C content of the B32^T strain was 70.5 mol%, and the growth temperature ranged from 4 to 40 °C, which is the expected limit for the genus. The results of this polyphasic approach support the membership of these strains in a new species within the genus *Dietzia,* for which we propose the name *Dietzia calabriae* sp. nov. The corresponding type of the strain is B32^T (B32^T= CBAS 986^T).

Keywords:

Dietzia; taxogenomics; genome analysis; mare; phylogeny.

The genus *Dietzia* was considered a monospecific taxon and was proposed by Rainey et al., to accommodate two organisms classified as *Rhodococcus maris* (Nesterenko et al., 1982), but formerly known as '*Flavobacterium maris*' (Rainey, KI, Kroppenstedt & Stackebrandt 1995; Nesterenko et al. 1982). *Dietzia* is a member of the suborder *Corynebacterineae* (Stackebrandt et al., 1997), which belongs to the family *Dietziaceae*, and includes 19 validly described names in the list of names of prokaryotes with the position in the nomenclature, but only 13 of these species have a validly published name, including synonyms (LPSN; https://lpsn.dsmz.de/genus/dietzia). *Dietzia* species occupy different niches and can be isolated from a wide variety of sources, including clinical specimens, different natural environments, such as hydrocarbons, alkaline soda ponds, soil, the skin of an immunocompromised patient, plant tissues (Bemer-Melchior et al., 1999; Duckworth et al., 1998; Jones et al., 2008a; Koerner et al., 2009; Mayilraj et al., 2006; Yamamura et al., 2010). Studies suggest that the genus may be underestimated, because the Gram morphology and colony appearance of *Dietzia* species share remarkable similarities with *Rhodococcus* and other related actinomycetes, and accurate and simple methods for identifying *Dietzia* species are still lacking (Fang et al., 2021; Koerner et al., 2009). The genus *Dietzia* includes aerobic, Gram-positive, non-motile, and catalase-positive actinomycetes that form cocci that germinate into short rods. The GC content of the genus ranges from 66-73 mol% (Koerner et al., 2009). In the present work, we isolated a strain of the genus *Dietzia* from subcutaneous nodules (abscess) of a 12-year-old mare in Reggio Calabria, Italy. The sequence of the 16S rRNA gene and the complete genome of the B32 strain showed no similarity to the strains described in the genus *Dietzia*. However, analysis of the complete genome sequence together with thin layer

chromatography (TLC) and other bioinformatic analyses provided evidence that these strains are a new species of *Dietzia*. The B32 strains are proposed as *Dietzia calabriae* sp. nov.

Material and methods

Bacterial isolation and identification

In 2018, an orange-pigmented strain, **B32^T** was isolated in Brain Heart Infusion (BHI) at 37 ° C from subcutaneous nodules (abscess) of a 12-year-old mare in Reggio Calabria, Italy (Rifici et al., 2020). A colony of the bacterial strain was plated onto a target steel plate using a sterile toothpick and mixed with a saturated matrix solution of α-cyano-4-hydroxycinnamic acid (HCCA). The preparation dried at room temperature and analyses were performed in a MALDI-TOF-MS (Bruker Daltonics, Bremen, Germany) equipped with a 337-nm nitrogen laser. Spectra were automatically recorded using the FlexControl Microflex LT mass spectrometer (Bruker Daltonics) according to the method of identification described by Assis et al (Assis et al., 2017) and the calibration performed using a bacterial standard test as described elsewhere (Seibold et al., 2010). For MALDI-TOF analysis, we proceed as recommended by the manufacturer, scores ≥2.000 indicate reliable identification at the species level; scores ≥1.700 and <2.000 indicate reliable identification at the gender level and scores <1.700 indicate unreliable identification.

Genome sequencing and analysis

The complete genome sequence of strain **B32^T** was obtained on an Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA), using paired-end libraries (2 × 150 bp). The quality of sequencing reads was verified using the FastQC program (v.0.11.5) (Andrews, 2010). The Edena software (version 3.0.) was used to reassemble the reads (Hernandez et al., 2008). The Sspace and

GapFiller algorithms were used to evaluate the order, distance, and alignment of contigs and to close the gap in paired reads, respectively (Boetzer et al., 2011; Nadalin et al., 2012). Gap-closing procedures were performed using the GFinisher (Guizelini et al., 2016), BLASTn (Altschul et al., 1990) tools, and CLC Genomics Workbench 7 (Qiagen Inc.) was used to visualize and map the reads with their own reads.

After we obtained the genome of strain **B32^T**, two trees were inferred based on the 16S rRNA gene sequence and another based on the *gyrB* gene. The 16S and *gyrB* sequences were extracted from the genomes and aligned using the program MUSCLE from MEGAX 10.1.8 software (Kumar et al., 2018). After alignment, the phylogenetic tree was inferred with the maximum likelihood method, Tamura-Nei model with gamma distribution, using MEGAX software v.10.1.8 (Kumar et al., 2018). The topography of the reconstructed tree was assessed by bootstrap analysis with 1000 repetitions (Felsenstein, 1985). *Tsukamurella paurametabola* ATCC8368 was used as an outgroup. All phylogenetic trees were plotted using the Interactive Tree of Life (iTOL) (Letunic and Bork, 2019).

Digital values of DNA-DNA hybridization (dDDH) and average nucleotide identity (ANI) were calculated using the Genome-to-Genome Distance Calculator 2.1 of the Leibniz Institute DSMZ website (http://ggdc.dsmz.de/distcalc2.php) and the JSpeciesWS web service (http://jspecies.ribohost.com/jspeciesws/#analyse), with recommended parameters and/or default settings, (Meier-Kolthoff et al., 2013; Richter et al., 2016).

Morphological, biochemical, and physiological exams

Colonial characteristics of the **B32^T** strain were recorded from a 37° Brain Heart Infusion (BHI) agar plate. Gram staining was performed using the RenyLab Kit according to the manufacturer's standard recommendations, according to the method of Gerhardt et al 1994 (Gerhardt P., 1994). Motility was examined by light microscopy in suspended drops, and cultures were grown on BHI agar plates. Cell morphology was observed by optical microscopy using SCANNING ELECTRON MICROSCOPE - JEOL JSM - 6360LV and transmission microscope Tecnai G2-12 - FEI SpiritBiotwin 120 kV using cells cultured in BHI agar for 48 hours at 37 °C. Colony color was determined by comparing color chips from standard samples from the ISCC-NBS color chart (Kenneth Kelly, 1965). Growth temperature was determined at 4-8, 10, 15, 20, 25, 30, 37, 40, 45, and 50 °C on BHI-agar plates. A growth curve using a temperature of 37°C was constructed by measuring the O.D. of the B32T strain every hour for 67 hours, using the Elisa reader Thermo Spectrophotometer - Multiskan SkyHi UV-Vis Reader. All parameters used were based on studies on the genus *Dietzia.*

NaCl tolerance was evaluated according to the method described by Xu et 2019 (XU *et al.*, 2019), with some modifications. For this purpose, 1 mL of strain culture was grown in BHI (Himedia) in a bacteriological incubator at 37 ºC for 48h, then centrifuged at 5000 rpm for 5 min and resuspended in 1mL of BHI. Then, 150 µL of the culture were inoculated in 15 mL of BHI previously supplemented with NaCl (Vetec) (previously autoclaved at 121 ºC for 15 min) at concentrations of 1-12%. As a control, 150 µL of strain culture was inoculated into 15 mL of BHI medium without NaCl supplementation. After 48 hours of growth at 37 ºC, the absorbance of the samples was measured at an O.D. (optical density) of 600 nm. Phenotypic identification was achieved by biochemical tests for catalase, glucose, sucrose, urea, and nitrate reduction, using commercial

gallery RapID™ CB Plus using commercial gallery RapID™ CB Plus (Remel, Thermo Fisher Scientific, Italy). According to the manufacturer's instructions, readings were taken using REMEL open software (http://www.remel.com/ERIC/IdentificationSingle.aspx) (MARKEY B *et al.*, 2013).

A previously established strain culture, in accordance with the methods detailed earlier, underwent a trilateral extraction procedure employing 70% ethyl alcohol (EtOH) at ambient conditions. The resultant ethanolic solution was differentially distributed between normal butanol (n-BuOH) and deionized water (H2O) in a 1:1 ratio, yielding both soluble constituents and an insoluble precipitate. The soluble n-BuOH fraction was subsequently subjected to chromatographic separation on a silica gel matrix, with elution being facilitated by a solvent mixture of dichloromethane (CH2Cl2) and ethyl acetate (EtOAc) in a 2:1 proportion. This process yielded a spectrum of seven distinct fractions. Selected fractions were subjected to further purification via thin layer chromatography (TLC) using a solvent mixture of dichloromethane (CHCl3) and methanol (MeOH) at a volumetric ratio of 13:1. Consequently, this purification strategy yielded the targeted compounds of significance, menaquinone as isoprenoid quinone lipids, and vanillic acid.

Hemolytic activity and antibiotic susceptibility

Hemolytic activity was assessed according to the methods of Xu and coworkers (Xu et al., 2019), using blood agar plates. Positive control was prepared using *Staphylococcus aureus* O46 cells. Antibiotic susceptibility of the **B32^{T}** strain was assessed using the disk diffusion method described by the Clinical and Laboratory Standards Institute (Clinical and Laboratory Standard Institute, 2009). The following antibiotics were tested (CECON): Gentamicin (GEN) (10mcg), Ampicillin

(AMP)(10mcg), Tetracycline (TET) (30mcg), Chloramphenicol (CLO) (30mcg), Vancomycin (VAN) (30mcg), Clindamycin (CLI) (2mcg), Strptomycin (EST) (10mcg), Erythromycin (ERI) (15mcg). We measured the diameters of the bacteria-free zones, and the results were expressed as resistance based on the Clinical & Laboratory Standards Institute design criteria (Clinical and Laboratory Standard Institute, 2009).

Results and discussion

The isolated and cultured microorganism was identified by protein profiling using MALDI-TOF MS obtained from BHI agar (37°C). The strain was identified with a reliable value for the genus *Dietzia* (score 1.791).

The genome of the species is available in the GenBank database under accession number CP093845. The strain was sequenced and has a length of 3.4 Mb, with G + C content of 70.5% and a total of 3435 CDSs, 9 rRNAS clusters (5S, 16S, and 23S) and 57 tRNA genes.

The sequence of 16S rRNA and *gyrB* gene of strain **B32^T** was compared with 11 sequences of recognized species of the genus *Dietzia*. The results of the reconstruction of the phylogenetic tree based on the maximum likelihood method and the Tamura-Nei model showed that the new strain formed a monophyletic clade with known species of this genus that had differences in bootstrap values **(Fig. 1 and Fig. 2**). In addition, we performed sequence similarity calculations based on pairwise comparison, which showed that the sequence similarities of the 16S rRNA gene with the known species of the genus *Dietzia* ranged from 93.24% to 97.07%. The closest phylogenetic relation was *Dietzia lutea* YIM 80766, which has 97.07% sequence similarity with the strain of this work, **strain B32^T** (Li et al., 2009)**.**

The G + C content of strain ***B32^T*** calculated from its genome, was 70.5 mol%, a value like species in the genus *Dietzia* **(Table 1)** (Kämpfer et al., 2012, 2010; Li et al., 2009)**.** As an alternative general index of genome relatedness (OGRI), genome-based taxonomic analysis based on *in silico* DDH in the TYGS database was performed (Meier-Kolthoff et al., 2022)**.** The dDDH value between strain B32^T and the species used in the work varied by 20.8% [18.6 - 23.2% as confidence interval] for *Dietzia timorensis* ID05-A0528 and 38.2 [35.7 - 40.7% as a confidence interval] for *Dietzia kunjamensis* DSM 44907 **(Table 1)**. In the analysis carried out to calculate ANI, the highest value was 88.44% for *Dietzia kunjamensis* DSM 44907. The *in silico* calculated dDDH and ANI of B32^T with species of the genus *Dietzia* was clearly below the recommended threshold of 70% (dDDH) and 95% (ANI) for species identification, indicating that isolate B32^T represents a distinct species **(Table 1)** (Felsenstein, 1985; Meier-Kolthoff et al., 2022; Richter et al., 2016)**.**

Colonies of strain B32^T were bright and had an orange color on BHI agar plates. The cells were Gram-positive, immobile, cocci-shaped that germinate in short rods or rod-shaped cells, that were up to 1.0 µm in size in agreement with previous studies (**Fig. 3 A, B, C, and D)** (Rainey et al., 1995)**.**

The temperature range for growth was 4–40 °C, with growth not possible at 45 and 50°C. Optimal growth occurred at 37 °C as shown by the growth curve performed at this temperature for 67 hours **(Fig.4 A).** Furthermore, the strain can grow in the presence of sodium chloride at a concentration of up to 12%. The B32^T strain has a positive catalase, does not reduce nitrate, and has a negative urease. Other physiological and biochemical results showed that the strain The strain had similar polar lipids, lipoquinones and major fatty acids profiles as listed in (**Table 2 and Table 3)** and in the species description, all these characteristics are consistent with previous

studies of the genus *Dietzia* (Duckworth et al., 1998; Kämpfer et al., 2012, 2010; Li et al., 2008; Mayilraj et al., 2006; F. A. Rainey et al., 1995; Rammer et al., 2013; Yumoto et al., 2002). The strain did not show any α- and β-hemolytic activity when cultivated on blood agar **(Figure 4 B).** Eight antibiotics from different families were investigated **(Table 4**). The B32^T strain showed sensitivity to all antibiotics tested **(Table 4).** Based on the physiological, biochemical, and genotypic data, strain B32^T represents a new species within the genus *Dietzia*, for which the name *Dietzia calabriae* B32 sp. nov. is proposed.

The characteristics of the proposed new species *D. calabriae* are listed in the protologue description in **Table 5**. Draft of the genome sequences of *Dietzia calabriae* sp B32^T been deposited in DDBJ/EMBL/GenBank under accession number GCA_024732245.1, Biosample (SAMN25598616) and Bioproject (PRJNA802820).

Description diagnostic characteristics of *Dietzia calabriae* sp. nov.

Cells are Gram-positive, aerobic, and non-motile, and form circular, smooth, bright orange colonies. Cells are coccus-shaped and germinate in short stalks or rod-shaped cells about 1.0 in diameter. Nitrate is not reduced, catalase is produced, and urea is not hydrolyzed. The temperature range for growth is 4 to 40°C, with optimal growth at 37°C. Above 45^o, no growth occurs. The bacterium grows on BHI agar in the presence of 0–12% NaCl. It uses D-glucose but not maltose, or sucrose, does not form hemolysis and is susceptible to eight different antibiotics (Gentamicin, Ampicillin, Tetracycline, Chloramphenicol, Vancomycin, Clindamycin, Streptomycin, and Erythromycin).

The strain type is B32^T (B32^T= CBAS 986^T) and was derived from an abscess sample from a 12-year-old mare in Reggio Calabria, Italy, in which the animal had diffuse subcutaneous nodules.

Dietzia calabriae (ca.la.bri'ae. NL fem. adj. *calabriae* belonging to Reggio Calabria, Italy, from where the organism was first isolated).

Figures:

Fig. 1. Maximum likelihood tree based on 1,506 positions of the multiple alignments of the 16S rRNA gene from **strain B32^T** and related species (accession numbers are given in parentheses). Phylogenetic analysis was performed using the Tamura-Nei model with gamma distribution. Bootstraps were calculated with 1000 repetitions. The *Tsukamurella paurometabola* ATCC8368 sequence was used as an outgroup. The scale bar indicates the number of replacements per site. The new species is shown in bold.

Fig.2. Phylogenetic analysis of the investigated *Dietzia* isolates based on the *gyrB* gene sequences (~1,324 to 1340 bp). Bootstraps were calculated with 1000 repetitions; only values above 50% are shown. The *Tsukamurella paurometabola* ATCC8368 sequence was used as an outgroup. The scale bar indicates the number of replacements per site. The new species is shown in red. The *gyrB* gene sequences were extracted from each of the genomes downloaded from the NCBI. *Dietzia aurantiaca* has no published genome and does not have the *gyrB* gene sequence deposited in public banks. The new species is shown in bold.

Fig.3: Cell morphology of the B32T strain observed by optical microscopy. **A** and **B**, cell morphology was observed by scanning electron microscopy using the Scanning Electron Microscope - JEOL JSM - 6360LV. **C** and **D**, cell morphology observed by transmission electron microscopy using the Tecnai G2-12 - FEI SpiritBiotwin 120 kV transmission microscope.

Fig. 4: **A** Growth curve of strain **B32^T** strain for 67 hours in BHI medium at 37°C, showing the O.D reached in each hour. **B.** Hemolysis test, strain **B32^T**, no reaction (γ-hemolysis); S. aureus O46, a clear zone of hydrolysis around the colony (β-hemolysis).

Tables:

Table 1. Values of dDDH and ANI between the genome of **strain B32^T** as a reference genome and that of closely related species. ANI values were calculated using JSpecies with the ANIb algorithm (mean nucleotide identity based on BLAST). The dDDH values were calculated using Genome-to-Genome Distance Calculator 2.1 provided by the Leibniz Institute DSMZ, numbers between parentheses after dDDH values are the confidence intervals.

Query genome	Reference genome	Bio project NCBI number	Size (MB)	G+C (mol %)	dDDH (%)	ANI %
Dietzia calabriae B32	*Dietzia lutea* YIM 80766^T	PRJNA319134	3,9	71	29.4 [27 - 31.9%]	84.64
Dietzia calabriae B32	*Dietzia alimentaria* 72^T	PRJNA72173	3.3	67.3	23.1 [20.8 - 25.6%]	80.27
Dietzia calabriae B32	*Dietzia alimentaria* BP 27/1	PRJNA338198	3.71	69.8	38.4 [35.9 - 40.9%]	80.68
Dietzia calabriae B32	*Dietzia aurantiaca J3*	PRJNA753838	3.77	69.7	23.6 [21.3 - 26.1%]	80.67
Dietzia calabriae B32	*Dietzia timorensis* ID05-A0528^T	PRJNA345110	3.6	65.6	20.8 [18.6 - 23.2%]	71.81
Dietzia calabriae B32	*Dietzia maris* 97	PRJNA500377	4.09	69.9	38 [35.6 - 40.6%]	89.21
Dietzia calabriae B32	*Dietzia maris* AMT	PRJNA454890	3.41	70.7	27.6 [25.2 - 30.1%]	84.27
Dietzia calabriae B32	*Dietzia maris* IMV 195	PRJNA607608	3.07	69.7	39.7 [37.2 - 42.2%]	89.71
Dietzia calabriae B32	*Dietzia maris* SST1	PRJNA478146	3.88	70	38.1 [35.7 - 40.6%]	89.03
Dietzia calabriae B32	*Dietzia maris* DSM 43672^T	PRJNA315745	3.5	70.9	28.3 [25.9 - 30.8%]	82.69
Dietzia calabriae B32	*Dietzia psychralcaliphila* ILA-1	PRJNA319132	3.9	69.67	24.8 [22.5 - 27.3%]	81.36
Dietzia calabriae B32	*Dietzia psychralcaliphila* JI1D	PRJNA686264	4.22	69.4	24.2 [21.9 - 26.7%]	81.16
Dietzia calabriae B32	*Dietzia psychralcaliphila* DSM 44820^T	PRJNA443690	3.9	69.6	24.3 [22 - 26.8%]	80.76

Dietzia calabriae B32	*Dietzia natronolimnaea* S-XJ-1	PRJNA391560	3.7	70	23.8 [21.5 - 26.3%]	80.34
Dietzia calabriae B32	*Dietzia papillomatosis* NBRC 105045	PRJNA224116	3.4	70.9	27.5 [25.1 - 30%]	83.25
Dietzia calabriae B32	*Dietzia kunjamensis K30 -10*^{T}	PRJNA607608	3.03	69.9	39.8 [37.3 - 42.3%]	89.66
Dietzia calabriae B32	*Dietzia kunjamensis* 313	PRJNA851992	4.02	70	38 [35.5 - 40.5%]	89.06
Dietzia calabriae B32	*Dietzia kunjamensis* DSM 44907	PRJNA363538	3.5	70.6	38.2 [35.7 - 40.7%]	88.33
Dietzia calabriae B32	*Dietzia kunjamensis* subsp schimae DSM 45139	PRJEB20683	3.65	70.5	38.1 [35.7 - 40.7%]	89.37
Dietzia calabriae B32	*Dietzia kunjamensis* subsp schimae YIM 65001^{T}	PRJNA607608	2.89	69.5	40.1 [37.6 - 42.6%]	89.67
Dietzia calabriae B32	*Dietzia cercidiphylli* YIM 65002^{T}	PRJNA607608	3.2	69.8	25.8 [23.4 - 28.2%]	82.05
Dietzia calabriae B32	*Dietzia aerolata* Sj14a^{T}	PRJNA607608	3.3	68.1	23.9 [21.6 - 26.4%]	80.29
Dietzia calabriae B32	*Dietzia cinnamea* 55	PRJNA500537	3.55	71	27.9 [25.5 - 30.3%]	84.65
Dietzia calabriae B32	*Dietzia cinnamea* BP 168	PRJNA701491	3.59	70.7	27.6 [25.3 - 30.1%]	84.29
Dietzia calabriae B32	*Dietzia cinnamea* CD11-5	PRJNA312284	4.14	67.6	28 [25.6 - 30.5%]	84.46
Dietzia calabriae B32	*Dietzia cinnamea* NBRC 102147	PRJDB449	3.6	70.8	27.7 [25.3 - 30.2%]	84.53
Dietzia calabriae B32	*Dietzia cinnamea* P4	PRJNA595013	3.56	71	28.6 [26.3 - 31.1%]	84.56
Dietzia calabriae B32	*Dietzia massiliensis* Marseille-Q0999^{T}	PRJNA727803	3.58	70.7	28 [25.6 - 30.5%]	84.49

Table 2. Differential characteristics of strain **B32^T** (a strain of this work) and type strains of *Dietzia* species. Strains: 1, D. spp. B32^T (strain of this work), 2, *D. aurantiaca* CCUG 35676 ^{T}; 3, *D. aerolata* Sj14a ^{T}; 4, *D. timorensis* ID05-A0528^T; 5, *D. cercidiphylli* YIM 65002 ^{T}; 6, *D. psychralcaliphila* ILA-1 ^{T}; 7, *D. natronolimnaea* CBS 107,95 ^{T}; 8, *D. kunjamensis* K30-10 ^{T}; 9, *D. maris* DSM 43672 ^{T}; 10, *D. lutea* DSM 45074 ^{T}; 11, *D. papilomatose* N 1280 ^{T}. The data for rates 2 to 3 are from (Kämpfer et al., 2012, 2010); the data for *D. timorensis* IDO5-A0528 are from (Yamamura et al., 2010), the data for de *D. papillomatosis* N 1280 ^{T} are from (Jones et al., 2008b; Rammer et al., 2013); the data of *D. lutea* are from (Li et al., 2009). + Positivo; - negativo; ND not determined.

Characteristics	1	2	3	4	5	6	7	8	9	10	11
Colony color	orange	Deep orange	Deep orange	nd	Reddish orange	Soft red	Soft pink	Red	Deep orange	Orange yellow	Orange
Urea hydrolysis	_	nd	–	–	+	+	+	–	+	–	+
Nitrate reduction	_	nd	nd	–	–	–	–	+	+	–	+
Maximum NaCl tolerance (%, w/v)	12	12	nd	nd	10	10	10	15	15	15	8
Growth temperature range (°C)	4 - 40	4–37	10–30	10–37	10–37	10–37	10–37	10–37	10–45	10–45	10–37
Utilization as a sole carbon source:											
d-Glucose	+	+	+	nd	+	–	+	+	+	+	+

Maltose	_	−	−	+	+	+	+	−	+	−	+
Sucrose	_	+	−	nd	+	+	+	+	−	−	+

Table 3. Polar lipids and isoprenoid menaquinone present in the strain type B32T. This table presents the composition of polar lipids and isoprenoid menaquinones identified in *Dietzia* strain Type B32T. The strain's lipid profile and quinone content provide insights into its potential characterization in the *Dietzia* genus.

Compound	Description
Majos polar lipids	Phosphatidylglycerol Diphosphatidylglycerol Phosphatidylethanolamine
Isoprenoid quinones*	MK8(H2) MK9(H2)

* Menaquinone (MK) are a type of quinone molecule essential for electron transport within the respiratory chain of bacteria. The number represents the length of the isoprene unit side chain attached to the quinone ring. Here we found the two variants with 9 and 8 isoprene units.

Table 4. Susceptibility to antibiotics of the strain B32^{T}.

Strain	Antibiotic tested *							
	AMP	CLI	CLO	ERI	EST	GEN	TET	VAN
Strain B32	S	S	S	S	S	S	S	S

* Antibiotics: Ampicillin (AMP), Clindamycin (CLI), Chloramphenicol (CLO), Erythromycin (ERI), Streptomycin (EST), Gentamicin (GEN), Tetracycline (TET) (30mcg), and Vancomycin (VAN). S, sensitive. Each experiment was repeated three times.

Table 5. Protologue description of *Dietzia calabriae* sp. nov.

Genus's name	*Dietzia*
Species name	*Dietzia calabriae*
Specific epithet	*carnosa*
Species status	sp. nov.
Species etymology	ca.la.bri'ae. NL fem. adj. *calabriae* belonging to Reggio Calabria, Italy, from where the organism was first isolated
Designation of the Type Strain	B32
Strain Collection Numbers	CBAS 986
Type Genome Nr. [INSDC databases]	GCA_024732245.1
Genome size	3.4 MB.
GC mol%	70.43 GC mol%.

Description of the new taxon and diagnostic traits	Cells are gram-positive, aerobic, and non-motile, and form circular, smooth, bright orange colonies. The cells are coccus-shaped and germinate in short stalks or rod-shaped cells about 1.0 in diameter. Nitrate is not reduced, catalase is produced, and urea is not hydrolyzed. The temperature range for growth is 4 to 40°C, with optimum growth at 37°C. No growth occurs above 45 °C. The bacterium grows on BHI agar in the presence of 0–12% NaCl. It uses D-glucose but not maltose, or sucrose, does not form hemolysis and is susceptible to eight different antibiotics (Gentamicin, Ampicillin, Tetracycline, Chloramphenicol, Vancomycin, Clindamycin, Streptomycin, and Erythromycin). The type of strain is B32 T (= CBAS 986T) and was from an abscess sample from a 12-year-old mare in Reggio Calabria, Italy, in which this animal had diffuse subcutaneous nodules.
Country of origin	Italy
Region of origin	Reggio Calabria
Date of isolation (dd/mm/yyyy)	05/03/2018
Source of isolation	Subcutaneous nodules (abscess) of a 12-year-old mare
Sampling date (dd/mm/yyyy)	05/03/2018
Latitude (xx°xx'xx"N/S)	not applicable

Longitude (xx°xx'xx"E/W)	not applicable
Altitude (meters above sea level)	not applicable
Number of strains in study	one
Information related to the Nagoya Protocol	not applicable

Abbreviations: ANI, average nucleotide identity; dDDH, digital DNA-DNA hybridization; MALDI-TOF MS, Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry, NaCl sodium chloride, BHI Brain Heart Infusion.

Author Contributions

Wrote the manuscript (RGS, ST); designed the study (RGS, ST, AKJ, VA, NS, TLPC, RC, SJ, MG, BB, CV); experimental work (RGS, MQM, DLNR, ACPG, AAR); conducted in silico analyses and interpreted the results (RGS, FMR, DLNR, AL); critically reviewed and revised the manuscript (ST, AKJ, VA); supervised the study (VA, ST). All authors approved this manuscript for publication.

Acknowledgments

The authors acknowledge the collaboration and assistance of all team members and the Brazilian funding agencies CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), FAPEMIG (Fundação de

Amparo à Pesquisa de Minas Gerais) for providing financial support and the UFMG Microscopy Center by the microscopic analysis performed and the RECOM.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J Mol Biol 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andrews, S., 2010. FastQC A Quality Control tool for High Throughput Sequence Data [WWW Document]. URL https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed 5.25.20).

Assis, G.B.N., Pereira, F.L., Zegarra, A.U., Tavares, G.C., Leal, C.A., Figueiredo, H.C.P., 2017. Use of MALDI-TOF Mass Spectrometry for the Fast Identification of Gram-Positive Fish Pathogens. Front Microbiol 8, 1492. https://doi.org/10.3389/fmicb.2017.01492

Bemer-Melchior, P., Haloun, A., Riegel, P., Drugeon, H.B., 1999. Bacteremia due to Dietzia maris in an immunocompromised patient. Clinical Infectious Diseases 29, 1338–1340. https://doi.org/10.1086/313490

Boetzer, M., Henkel, C. V, Jansen, H.J., Butler, D., Pirovano, W., 2011. Scaffolding pre-assembled contigs using SSPACE. BIOINFORMATICS APPLICATIONS NOTE 27, 578–579. https://doi.org/10.1093/bioinformatics/btq683

Clinical and Laboratory Standard Institute, 2009. Performance Standards for Antimicrobial Susceptibility; Approved Standard. Wayne, PA: Clinical and Laboratory Standard Institute Testing, 31st Edition [WWW Document]. URL https://clsi.org/standards/products/microbiology/documents/m100/ (accessed 9.12.21).

Duckworth, A.W., Grant, S., Grant, W.D., Jones, B.E., Meijer, D., 1998. Dietzia natronolimnaios sp. nov., a new member of the genus Dietzia isolated from an East African soda lake. Extremophiles 2, 359–366. https://doi.org/10.1007/s007920050079

Fang, H., Xu, J.-B., Nie, Y., Wu, X.-L., 2021. Pan-genomic analysis reveals that the evolution of Dietzia species depends on their living habitats. Environ Microbiol 23, 861–877. https://doi.org/10.1111/1462-2920.15176

Felsenstein, J., 1985. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. Evolution (N Y) 39, 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Gerhardt P., M.R.G.E., W.W.A., K.N.R., 1994. Methods for General and Molecular Bacteriology Washington, DC: American Society for Microbiology [WWW Document]. (editors) 1994. URL https://www.researchgate.net/publication/292432357_Methods_for_general_and_molecular_bacteriology_ASM_Press (accessed 9.12.21).

Guizelini, D., Raittz, R.T., Cruz, L.M., Souza, E.M., Steffens, M.B.R., Pedrosa, F.O., 2016. GFinisher: a new strategy to refine and finish bacterial genome assemblies. Sci Rep 6, 34963. https://doi.org/10.1038/srep34963

Hernandez, D., François, P., Farinelli, L., Østerås, M., Schrenzel, J., 2008. De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. Genome Res 18, 802–809. https://doi.org/10.1101/gr.072033.107

Jones, A.L., Koerner, R.J., Natarajan, S., Perry, J.D., Goodfellow, M., 2008a. Dietzia papillomatosis sp. nov., a novel actinomycete isolated from the skin of an immunocompetent patient with confluent and reticulated papillomatosis. Int J Syst Evol Microbiol 58, 68–72. https://doi.org/10.1099/ijs.0.65178-0

Jones, A.L., Koerner, R.J., Natarajan, S., Perry, J.D., Goodfellow, M., 2008b. Dietzia papillomatosis sp. nov., a novel actinomycete isolated from the skin of an immunocompetent patient with confluent and reticulated papillomatosis. Int J Syst Evol Microbiol 58, 68–72. https://doi.org/10.1099/IJS.0.65178-0/CITE/REFWORKS

Kämpfer, P., Falsen, E., Frischmann, A., Busse, H.J., 2012. Dietzia aurantiaca sp. nov., isolated from a human clinical specimen. Int J Syst Evol Microbiol 62, 484–488. https://doi.org/10.1099/IJS.0.032557-0

Kämpfer, P., Langer, S., Martin, E., Jäckel, U., Busse, H.-J., 2010. Dietzia aerolata sp. nov., isolated from the air of a duck barn, and emended description of the genus Dietzia Rainey et al. 1995. Int J Syst Evol Microbiol 60, 393–396. https://doi.org/10.1099/IJS.0.012807-0

Kenneth Kelly, 1965. ISCC-NBS color-name charts illustrated with centroid colors. - Yale University Library.

Koerner, R.J., Goodfellow, M., Jones, A.L., 2009. The genus Dietzia: A new home for some known and emerging opportunist pathogens. FEMS Immunol Med Microbiol. https://doi.org/10.1111/j.1574-695X.2008.00513.x

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35, 1547. https://doi.org/10.1093/MOLBEV/MSY096

Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47, W256–W259. https://doi.org/10.1093/NAR/GKZ239

Li, J., Chen, C., Zhao, G.Z., Klenk, H.P., Pukall, R., Zhang, Y.Q., Tang, S.K., Li, W.J., 2009. Description of Dietzia lutea sp. nov., isolated from a desert soil in Egypt. Syst Appl Microbiol 32, 118–123. https://doi.org/10.1016/J.SYAPM.2008.11.007

Li, J., Zhao, G.-Z., Zhang, Y.-Q., Klenk, H.-P., Pukall, R., Qin, S., Xu, L.-H., Li, W.-J., 2008. Dietzia schimae sp. nov. and Dietzia cercidiphylli sp. nov., from surface-sterilized plant tissues. Int J Syst Evol Microbiol 58, 2549–2554. https://doi.org/10.1099/IJS.0.2008/000919-0

Mayilraj, S., Suresh, K., Kroppenstedt, R.M., Saini, H.S., 2006. Dietzia kunjamensis sp. nov., isolated from the Indian Himalayas. Int J Syst Evol Microbiol 56, 1667–1671. https://doi.org/10.1099/IJS.0.64212-0

Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P., Göker, M., 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14, 1–14. https://doi.org/10.1186/1471-2105-14-60/TABLES/2

Meier-Kolthoff, J.P., Carbasse, J.S., Peinado-Olarte, R.L., Göker, M., 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acids Res 50, D801–D807. https://doi.org/10.1093/NAR/GKAB902

Nadalin, F., Vezzi, F., Policriti, A., 2012. GapFiller: A de novo assembly approach to fill the gap within paired reads. BMC Bioinformatics 13. https://doi.org/10.1186/1471-2105-13-S14-S8

Nesterenko, O.A., Nogina, T.M., Kasumova, S.A., 1982. Rhodococcus luteus nom. nov. and Rhodococcus maris nom. nov. Int J Syst Bacteriol 32, 1–14. https://doi.org/10.1099/00207713-32-1-1

Rainey, F A, Kl, S., Kroppenstedt, R.M., Stackebrandt, E., 1995. Formerly Rhodococcus maris, INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY.

Rainey, F. A., Klatte, S., Kroppenstedt, R.M., Stackebrandt, E., 1995. Dietzia, a new genus including Dietzia maris comb. nov., formerly Rhodococcus maris. Int J Syst Bacteriol 45, 32–36. https://doi.org/10.1099/00207713-45-1-32

Rammer, P., Calum, H., Moser, C., Björnsdóttir, M.K., Smedegaard, H., Høiby, N., Bjarnsholt, T., 2013. Dietzia papillomatosis Bacteremia. J Clin Microbiol 51, 1977. https://doi.org/10.1128/JCM.03313-12

Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., Peplies, J., 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32, 929–931. https://doi.org/10.1093/bioinformatics/btv681

Rifici, C., Attili, A.-R., De Biase, D., Gonçalves dos Santos, R., Seyffert, N., De Paula Castro, T.L., Pereira Figueiredo, H.C., Scaramozzino, C., Reale, S., Paciello, O., Cuteri, V., Spier, S.J., Azevedo, V., Mazzullo, G., 2020. Atypical Multibacterial Granulomatous Myositis in a Horse: First Report in Italy. Vet Sci 7, 47. https://doi.org/10.3390/vetsci7020047

Seibold, E., Maier, T., Kostrzewa, M., Zeman, E., Splettstoesser, W., 2010. Identification of Francisella tularensis by Whole-Cell Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry: Fast, Reliable, Robust, and Cost-Effective Differentiation on Species and Subspecies Levels. J Clin Microbiol 48, 1061–1069. https://doi.org/10.1128/JCM.01953-09

Stackebrandt, E., Rainey, F.A., Ward-Rainey, N.L., 1997. Proposal for a new hierarchic classification system, Actinobacteria classis nov. Int J Syst Bacteriol 47, 479–491. https://doi.org/10.1099/00207713-47-2-479

Xu, Y., Tian, Yuan, Cao, Y., Li, J., Guo, H., Su, Y., Tian, Yumin, Wang, C., Wang, T., Zhang, L., 2019. Probiotic Properties of Lactobacillus paracasei subsp. paracasei L1 and Its Growth Performance-Promotion in Chicken by Improving the Intestinal Microflora. Front Physiol 10. https://doi.org/10.3389/FPHYS.2019.00937

Yamamura, H., Lisdiyanti, P., Ridwan, R., Ratnakomala, S., Sarawati, R., Lestari, Y., Triana, E., Kartina, G., Widyastuti, Y., Ando, K., 2010. Dietzia timorensis sp. nov., isolated from soil. Int J Syst Evol Microbiol 60, 451–454. https://doi.org/10.1099/IJS.0.012229-0

Yumoto, I., Nakamura, A., Iwata, H., Kojima, K., Kusumoto, K., Nodasaka, Y., Matsuyama, H., 2002. Dietzia psychralcaliphila sp. nov., a novel, facultatively psychrophilic alkaliphile that grows on hydrocarbons. Int J Syst Evol Microbiol 52, 85–90. https://doi.org/10.1099/00207713-52-1-85/CITE/REFWORKS

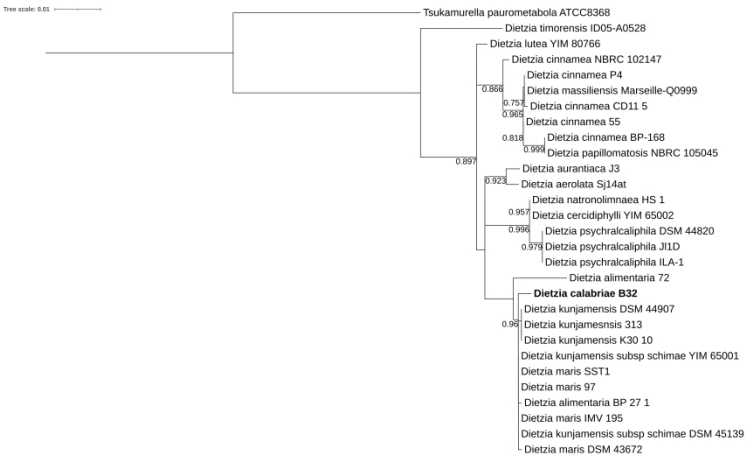


Fig. 1. Maximum likelihood tree based on 1,506 positions of the multiple alignments of the 16S rRNA gene from strain B32T and related species (accession numbers are given in parentheses). Phylogenetic analysis was performed using the Tamura-Nei model with gamma distribution. Bootstraps were calculated with 1000 repetitions. The Tsukamurella paurometabola ATCC8368 sequence was used as an outgroup. The scale bar indicates the number of replacements per site. The new species is shown in bold.

3242x1594mm (47 x 47 DPI)

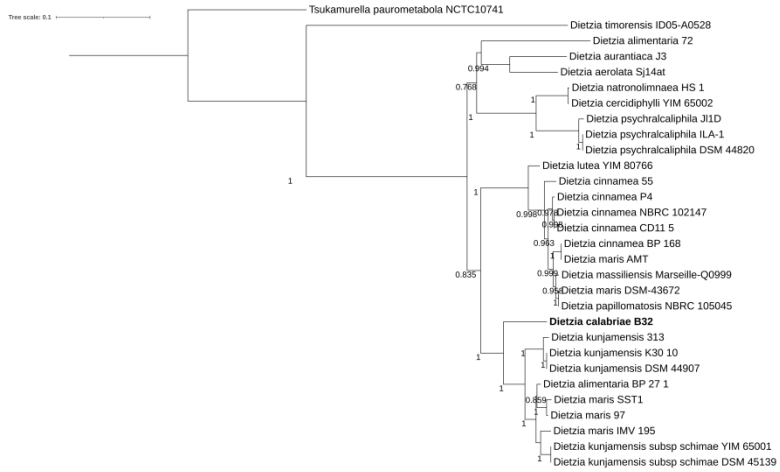


Fig.2. Phylogenetic analysis of the investigated Dietzia isolates based on the gyrB gene sequences (~1,324 to 1340 bp). Bootstraps were calculated with 1000 repetitions; only values above 50% are shown. The Tsukamurella paurometabola ATCC8368 sequence was used as an outgroup. The scale bar indicates the number of replacements per site. The new species is shown in red. The gyrB gene sequences were extracted from each of the genomes downloaded from the NCBI. Dietzia aurantiaca has no published genome and does not have the gyrB gene sequence deposited in public banks. The new species is shown in bold.

3242x1594mm (47 x 47 DPI)

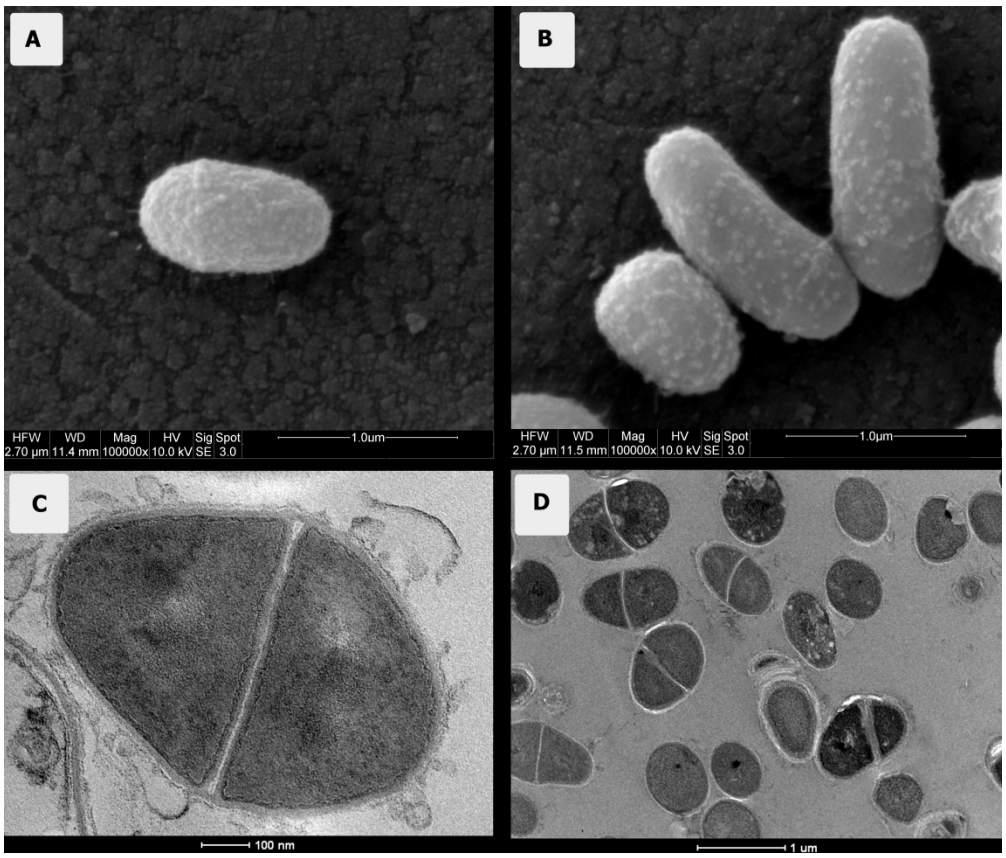


Fig.3: Cell morphology of the B32T strain observed by optical microscopy. A and B, cell morphology was observed by scanning electron microscopy using the Scanning Electron Microscope - JEOL JSM - 6360LV. C and D, cell morphology observed by transmission electron microscopy using the Tecnai G2-12 - FEI SpiritBiotwin 120 kV transmission microscope.

1410x1193mm (75 x 75 DPI)

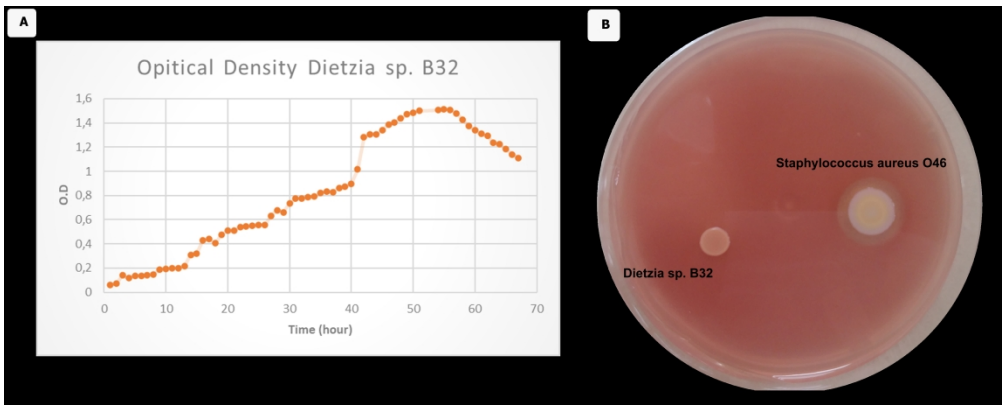


Fig. 4: A Growth curve of strain B32T strain for 67 hours in BHI medium at 37°C, showing the O.D reached in each hour. B. Hemolysis test, strain B32T, no reaction (γ-hemolysis); S. aureus O46, a clear zone of hydrolysis around the colony (β-hemolysis).

1547x618mm (72 x 72 DPI)

Submission Confirmation

Print

Thank you for your submission

Submitted to
Journal of Basic Microbiology

Manuscript ID
jobm.202300717

Title
Dietzia calabriae sp. nov. is a bacterial species isolated from a mare's abscess

Authors
dos Santos, Roselane
Rita , Attili
de Jesus, Luís
Neres Rodrigues, Diego
Lima, Alessandra
Hurtado, Raquel
Claudia, Rifici
Sharon, Spier
Giuseppe , Mazzullo
Morais-Rodrigues , Francielly
Gomide , Anne
Brenig, Bertram
Vincenzo , Cuteri
Castro , Thiago
Seyffert , Núbia
Jaiswal, Arun
Batista, Ana
de Magalhaes , Mariana
Figueira Aburjaile, Flávia
Góes-Neto, Aristóteles
Tiwari, Sandeep
Azevedo, Vasco

Date Submitted
03-Dec-2023

Author Dashboard

6. CONCLUSÃO GERAL

Em conclusão, este estudo ressalta a importância de aprimorar os métodos de identificação e classificação de cepas bacterianas, tanto para *Corynebacterium pseudotuberculosis* (CP) quanto para o gênero *Dietzia*. A infecção por CP representa um potencial risco zoonótico para humanos, tornando essencial a identificação precisa das cepas para o controle de doenças. Os métodos convencionais, como cultura bacteriana e testes bioquímicos, têm sido utilizados, mas a identificação incorreta de biovares pode ocorrer. Para superar esses desafios, técnicas moleculares, como o MALDI Biotyper e a análise genômica, mostraram-se promissoras.

O MALDI Biotyper é um poderoso sistema de tipagem bacteriana, sendo avaliado por sua capacidade de identificar microrganismos e fornecer informações valiosas para prevenção e controle de doenças. Além disso, métodos genômicos, especificamente a Identidade Nucleotídica Média (ANI), são discutidos como um meio de inferir relações filogenéticas e classificar microrganismos. Neste estudo, usamos os métodos MALDI Biotyper e Support Vector Machine (SVM). Os espectros das cepas de CP são adquiridos e os Perfis de Espectro Principal (MSPs) são criados. Os picos obtidos da análise MALDI Biotyper são então usados para diferenciar os biovares Ovis e Equi. O algoritmo SVM é empregado para classificar os picos e diferenciar com precisão entre os biovars. MALDI-TOF-MS, juntamente com métodos genômicos, pode ser usado para melhorar a identificação e classificação de cepas de CP. Esses avanços têm o potencial de fornecer classificações mais precisas e rápidas, o que pode auxiliar no gerenciamento de doenças e medidas de controle.

Da mesma forma, no gênero *Dietzia*, a distinção entre as cepas e a identificação de novas espécies podem ser desafiadoras devido à sua similaridade fenotípica. Embora o sequenciamento do gene 16S rRNA possa fornecer alguma ajuda na identificação, métodos adicionais, como

análise genômica, ANI e abordagens pan-genômicas, são necessários para determinar o status taxonômico de algumas espécies dentro do gênero. A análise comparativa de 56 genomas de *Dietzia* revelou informações valiosas sobre a diversidade genética e características genômicas do gênero, contribuindo para uma melhor compreensão e diferenciação dessas espécies. Essas análises envolveu relações filogenéticas/filogenômicas, caracterização do pangenoma e identificação de genes específicos e core. Os resultados fornecem informações valiosas sobre a diversidade genética e características genômicas do gênero *Dietzia*, contribuindo para uma melhor compreensão e diferenciação dessas espécies.

Além disso, o estudo descreve o isolamento de uma nova cepa, B32T, de nódulos subcutâneos de uma égua, que foi identificada como uma nova espécie de *Dietzia* por meio de análise genômica. A análise filogenética confirmou sua colocação dentro do gênero, e os valores de hibridização DNA-DNA e ANI apoiaram sua distinção como uma nova espécie. Exames morfológicos, bioquímicos e fisiológicos foram realizados para caracterizar a cepa, que apresentou características típicas do gênero *Dietzia*. A suscetibilidade a antibióticos e a atividade hemolítica foram avaliadas, fornecendo informações adicionais sobre suas características.

Os achados deste estudo contribuem para o entendimento da diversidade e taxonomia do gênero *Dietzia*. O isolamento de uma nova espécie enfatiza a potencial subestimação do gênero e destaca a importância de métodos precisos de identificação.

No geral, esta pesquisa expande nosso conhecimento sobre as espécies de *Dietzia* e ressalta a necessidade de métodos de identificação confiáveis. Mais estudos são necessários para explorar o significado ecológico, as aplicações biotecnológicas e o potencial patogênico das cepas de *Dietzia*.

7. PERSPECTIVAS

CAPITULO I

Testar uma subfunção do MALDI: Desenvolver um módulo de subtipagem que se concentra na análise dos picos altamente reproduzíveis de ambos os biovares. Isso se deve ao fato de que, devido ao log (pontuação), a diferenciação por si só não é viável. No entanto, é concebível realizar uma segunda rodada de classificação potencial exclusivamente com base no MALDI, sem a necessidade de utilizar ferramentas adicionais.

CAPÍTULO II

Será feito um artigo de revisão de como caracterizar uma nova espécie bacteriana, com ênfase no gênero *Dietzia*.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, Sintia *et al.* Quadruplex PCR assay for identification of Corynebacterium pseudotuberculosis differentiating biovar Ovis and Equi. *BMC Veterinary Research*, v. 13, n. 1, 25 set. 2017. Acesso em: 15 maio 2022.

ALTSCHUL, Stephen F. *et al.* Basic local alignment search tool. *Journal of Molecular Biology*, v. 215, n. 3, p. 403–410, 5 out. 1990. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0022283605803602?via%3Dihub>. Acesso em: 27 out. 2018.

ANDREWS, S. *FastQC A Quality Control tool for High Throughput Sequence Data*. Disponível em: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Acesso em: 25 maio 2020.

ASSIS, Gabriella B. N. *et al.* Use of MALDI-TOF Mass Spectrometry for the Fast Identification of Gram-Positive Fish Pathogens. *Frontiers in Microbiology*, v. 8, n. AUG, p. 1492, 9 ago. 2017. Disponível em: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01492/full>. Acesso em: 6 out. 2020.

BAIRD, G. J.; FONTAINE, M. C. Corynebacterium pseudotuberculosis and its role in ovine caseous lymphadenitis. *Journal of comparative pathology*, v. 137, n. 4, p. 179–210, nov. 2007. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17826790/>. Acesso em: 15 maio 2022.

BARRAL, Thiago Doria *et al.* A panel of recombinant proteins for the serodiagnosis of caseous lymphadenitis in goats and sheep. *Microbial Biotechnology*, v. 12, n. 6, p. 1313, 1 nov. 2019. Disponível em: </pmc/articles/PMC6801148/>. Acesso em: 15 maio 2022.

BEMER-MELCHIOR, P. *et al.* Bacteremia Due to Dietzia maris in an Immunocompromised Patient. *Clinical Infectious Diseases*, v. 29, n. 5, p. 1338–1340, 1 nov. 1999. Disponível em: <https://academic.oup.com/cid/article/29/5/1338/344883>. Acesso em: 3 ago. 2021.

BIBERSTEIN, E. L.; KNIGHT, H. D.; JANG, S. Two biotypes of Corynebacterium pseudotuberculosis. *The Veterinary record*, v. 89, n. 26, p. 691–692, 1971. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/5168555/>. Acesso em: 15 maio 2022.

BINNS, Sarah H.; GREEN, Laura E.; BAILEY, Mick. Development and validation of an ELISA to detect antibodies to Corynebacterium pseudotuberculosis in ovine sera. *Veterinary microbiology*, v. 123, n. 1–3, p. 169–179, 20 jul. 2007. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17363196/>. Acesso em: 15 maio 2022.

BOETZER, Marten *et al.* Scaffolding pre-assembled contigs using SSPACE. *BIOINFORMATICS APPLICATIONS NOTE*, v. 27, n. 4, p. 578–579, 2011. Disponível em: <www.baseclear.com/bioinformatics-tools/.>. Acesso em: 25 maio 2020.

CARMINATI, Renato *et al.* Determinação da sensibilidade e da especificidade de um teste de ELISA indireto para o diagnóstico de linfadenite caseosa em caprinos. *Revista de Ciências Médicas e Biológicas*, v. 2, n. 1, p. 88–93, 1 jan. 2003. Disponível em: <https://periodicos.ufba.br/index.php/cmbio/article/view/4256>. Acesso em: 15 maio 2022.

ÇETINKAYA, Burhan *et al.* Identification of Corynebacterium pseudotuberculosis isolates from sheep and goats by PCR. *Veterinary microbiology*, v. 88, n. 1, p. 75–83, 2 ago. 2002. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/12119139/>. Acesso em: 15 maio 2022.

CLINICAL AND LABORATORY STANDARD INSTITUTE. *Performance Standards for Antimicrobial Susceptibility; Approved Standard. Wayne, PA: Clinical and Laboratory Standard Institute Testing, 31st Edition*. Disponível em: <https://clsi.org/standards/products/microbiology/documents/m100/>. Acesso em: 12 set. 2021.

CONNOR, Kathleen M. *et al.* Characterization of United Kingdom Isolates of Corynebacterium pseudotuberculosis Using Pulsed-Field Gel Electrophoresis. *Journal of Clinical Microbiology*, v. 38, n. 7, p. 2633, 2000. Disponível em: </pmc/articles/PMC86984/>. Acesso em: 18 abr. 2022.

CONNOR, Kathleen M. *et al.* Molecular genotyping of multinational ovine and caprine Corynebacterium pseudotuberculosis isolates using pulsed-field gel electrophoresis. *Veterinary research*, v. 38, n. 4, p. 613–623, jul. 2007. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17565908/>. Acesso em: 15 maio 2022.

CONVILLE, Patricia S.; WITEBSKY, Frank G. Nocardia, Rhodococcus, Gordonia, Actinomadura, Streptomyces, and Other Aerobic Actinomycetes. *Manual of Clinical Microbiology* , p. 504–535, 26 maio 2015. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1128/9781555817381.ch29>. Acesso em: 18 abr. 2022.

COSTA, Daniela Arruda. *GENÔMICA COMPARATIVA DE LINHAGENS DE Saccharomyces E Kluyveromyces DE INTERESSE BIOTECNOLÓGICO*. . [S.l: s.n.], 2015. Disponível em: <http://www.locus.ufv.br/bitstream/handle/123456789/8473/texto%20completo.pdf?sequence=1&isAllowed=y>. Acesso em: 14 maio 2019.

DERCKSEN, Daan P. *et al.* A comparison of four serological tests for the diagnosis of caseous lymphadenitis in sheep and goats. *Veterinary microbiology*, v. 75, n. 2, p. 167–175, 31 jul. 2000. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/10889407/>. Acesso em: 15 maio 2022.

DORELLA, Fernanda Alves; PACHECO, Luis Gustavo Carvalho; *et al.* Corynebacterium pseudotuberculosis: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Veterinary research*, v. 37, n. 2, p. 201–218, mar. 2006. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/16472520/>. Acesso em: 18 abr. 2022.

DORELLA, Fernanda Alves; CARVALHO PACHECO, L.; *et al.* Corynebacterium pseudotuberculosis : microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Veterinary Research*, v. 37, n. 2, p. 201–218, mar. 2006.

DORNELES, Elaine M.S. *et al.* Evaluation of ERIC-PCR as Genotyping Method for Corynebacterium pseudotuberculosis Isolates. *PLOS ONE*, v. 9, n. 6, p. e98758, 5 jun. 2014. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098758>. Acesso em: 15 maio 2022.

DUCKWORTH, Andrew W. *et al.* Dietzia natronolimnaios sp. nov., a new member of the genus Dietzia isolated from an East African soda lake. *Extremophiles*, v. 2, n. 3, p. 359–366, ago. 1998. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/9783184/>. Acesso em: 13 out. 2020.

FELSENSTEIN, Joseph. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution*, v. 39, n. 4, p. 783–791, 1 jul. 1985. Disponível em: <http://doi.wiley.com/10.1111/j.1558-5646.1985.tb00420.x>. Acesso em: 25 maio 2020.

GERHARDT P., Murray R. G. E., Wood W. A., Krieg N. R. *Methods for General and Molecular Bacteriology Washington, DC: American Society for Microbiology*. Disponível em: <https://www.researchgate.net/publication/292432357_Methods_for_general_and_molecular_bacteriology_ASM_Press>. Acesso em: 12 set. 2021.

GONÇALVES, Juliano Leonel *et al.* Identification of Corynebacterium spp. isolated from bovine intramammary infections by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Veterinary microbiology*, v. 173, n. 1–2, p. 147–151, 2014. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25086477/>. Acesso em: 15 maio 2022.

GUEDES, Maria T. *et al.* Infecção por Corynebacterium pseudotuberculosis em equinos: aspectos microbiológicos, clínicos e preventivos. *Pesquisa Veterinária Brasileira*, v. 35, n. 8, p. 701–708, 2015. Disponível em: <http://www.scielo.br/j/pvb/a/gxLrmVttjZPQVLYMqS4xkKt/?lang=pt>. Acesso em: 18 abr. 2022.

GUIZELINI, Dieval *et al*. GFinisher: A new strategy to refine and finish bacterial genome assemblies. *Scientific Reports*, v. 6, n. 1, p. 1–8, 10 out. 2016. Acesso em: 25 maio 2020.

HERNANDEZ, David *et al*. De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. *Genome Research*, v. 18, n. 5, p. 802–809, 1 maio 2008. Disponível em: <www.genome.org/cgi/doi/10.1101/gr.072033.107.>. Acesso em: 24 jan. 2021.

HUERTA, Belén *et al.* Comparison of two biochemical methods for identifying Corynebacterium pseudotuberculosis isolated from sheep and goats. *Veterinary journal (London, England : 1997)*, v. 196, n. 3, p. 552–554, jun. 2013. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23352345/>. Acesso em: 15 maio 2022.

JONES, Amanda L. *et al.* Dietzia papillomatosis sp. nov., a novel actinomycete isolated from the skin of an immunocompetent patient with confluent and reticulated papillomatosis. *International Journal of Systematic and Evolutionary Microbiology*, v. 58, n. 1, p. 68–72, 1 jan. 2008. Disponível em: <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.65178-0>. Acesso em: 3 ago. 2021.

KÄMPFER, P. *et al.* Dietzia aerolata sp. nov., isolated from the air of a duck barn, and emended description of the genus Dietzia Rainey et al. 1995. *International Journal of Systematic and Evolutionary Microbiology*, v. 60, n. 2, p. 393–396, 1 fev. 2010. Disponível em: <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.012807-0>. Acesso em: 14 set. 2021.

KÄMPFER, P. *et al.* Dietzia aurantiaca sp. nov., isolated from a human clinical specimen. *International Journal of Systematic and Evolutionary Microbiology*, v. 62, n. Pt_3, p. 484–488, 1 mar. 2012. Disponível em: <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.032557-0>. Acesso em: 3 ago. 2021.

KANAK, Eda Kılıç; YILMAZ, Suzan Öztürk. Maldi-tof mass spectrometry for the identification and detection of antimicrobial activity of lactic acid bacteria isolated from local cheeses. *Food Science and Technology*, v. 39, p. 462–469, 29 nov. 2018. Disponível em: <http://www.scielo.br/j/cta/a/fNr6qnMTZs3nqHpw7vzcgrm/?lang=en>. Acesso em: 18 abr. 2022.

KENNETH KELLY. *ISCC-NBS color-name charts illustrated with centroid colors. - Yale University Library*. [S.l: s.n.], 1965. Disponível em: <https://findit.library.yale.edu/catalog/digcoll:2784235>. Acesso em: 12 set. 2021.

KHAMIS, Atieh; RAOULT, Didier; LA SCOLA, Bernard. rpoB Gene Sequencing for Identification of Corynebacterium Species. *Journal of Clinical Microbiology*, v. 42, n. 9, p. 3925, set. 2004. Disponível em: </pmc/articles/PMC516356/>. Acesso em: 15 maio 2022.

KOBOLDT, Daniel C. *et al.* The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell*, v. 155, n. 1, p. 27, 9 set. 2013. Disponível em: </pmc/articles/PMC3969849/>. Acesso em: 22 nov. 2022.

KOERNER, Roland J.; GOODFELLOW, Michael; JONES, Amanda L. *The genus Dietzia: A new home for some known and emerging opportunist pathogens*. *FEMS Immunology and Medical Microbiology*. [S.l: s.n.]. Disponível em: <https://academic.oup.com/femspd/article-lookup/doi/10.1111/j.1574-695X.2008.00513.x>. Acesso em: 13 out. 2020. , abr. 2009

KUMAR, Sudhir *et al.* MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, v. 35, n. 6, p. 1547, 1 jun. 2018. Disponível em: </pmc/articles/PMC5967553/>. Acesso em: 28 jul. 2021.

LANDER, Eric S. *et al.* Initial sequencing and analysis of the human genome. *Nature 2001 409:6822*, v. 409, n. 6822, p. 860–921, 15 fev. 2001. Disponível em: <https://www.nature.com/articles/35057062>. Acesso em: 22 nov. 2022.

LANGONI, Helio *et al.* Short communication: Identification of Corynebacterium bovis by MALDI-mass spectrometry. *Journal of dairy science*, v. 100, n. 6, p. 4287–4289, 1 jun. 2017. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28342611/>. Acesso em: 15 maio 2022.

LETUNIC, Ivica; BORK, Peer. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, v. 47, n. W1, p. W256–W259, 2 jul. 2019. Disponível em: <https://academic.oup.com/nar/article/47/W1/W256/5424068>. Acesso em: 28 jul. 2021.

LI, Jie *et al.* Description of Dietzia lutea sp. nov., isolated from a desert soil in Egypt. *Systematic and Applied Microbiology*, v. 32, n. 2, p. 118–123, 1 abr. 2009. Acesso em: 14 set. 2021.

LIM, Hoong Chuin *et al.* Identification of new components of the RipC-FtsEX cell separation pathway of Corynebacterineae. *PLOS Genetics*, v. 15, n. 8, p. e1008284, 2019. Disponível em: <https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1008284>. Acesso em: 18 abr. 2022.

LISTA, Florigio *et al.* Reliable identification at the species level of Brucella isolates with MALDI-TOF-MS. *BMC Microbiology*, v. 11, n. 1, p. 1–11, 23 dez. 2011. Disponível em: <https://bmcmicrobiol.biomedcentral.com/articles/10.1186/1471-2180-11-267>. Acesso em: 18 abr. 2022.

LIU, Lin *et al.* Comparison of next-generation sequencing systems. *Journal of Biomedicine and Biotechnology*, v. 2012, 2012. Acesso em: 22 nov. 2022.

LOMAN, Nicholas J. *et al.* High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nature reviews. Microbiology*, v. 10, n. 9, p. 599–606, set. 2012. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22864262/>. Acesso em: 22 nov. 2022.

MARKEY B *et al. Clinical Veterinary Microbiology - 2nd Edition 2nd ed.;Mosby Ltd.: Philadelphia, PA, USA*. 2nd ed ed. [S.l: s.n.], 2013. Disponível em: <https://www.elsevier.com/books/clinical-veterinary-microbiology/markey/978-0-7234-3237-1>. Acesso em: 14 set. 2021.

MAYILRAJ, S. *et al.* Dietzia kunjamensis sp. nov., isolated from the Indian Himalayas. *International Journal of Systematic and Evolutionary Microbiology*, v. 56, n. 7, p. 1667–1671, 1 jul. 2006. Disponível em: <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.64212-0>. Acesso em: 24 jan. 2021.

MEIER-KOLTHOFF, Jan P *et al.* Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics 2013 14:1*, v. 14, n. 1,

p. 1–14, 21 fev. 2013. Disponível em: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-60>. Acesso em: 29 ago. 2021.

MENZIES, P. I.; HWANG, Y. T.; PRESCOTT, J. F. Comparison of an interferon-gamma to a phospholipase D enzyme-linked immunosorbent assay for diagnosis of Corynebacterium pseudotuberculosis infection in experimentally infected goats. *Veterinary microbiology*, v. 100, n. 1–2, p. 129–137, 20 maio 2004. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/15135521/>. Acesso em: 15 maio 2022.

MUCKLE, C. A.; GYLES, C. L. Characterization of strains of corynebacterium pseudotuberculosis. *Canadian Journal of Comparative Medicine*, v. 46, n. 2, p. 206, 1982. Disponível em: </pmc/articles/PMC1320282/?report=abstract>. Acesso em: 15 maio 2022.

NADALIN, Francesca; VEZZI, Francesco; POLICRITI, Alberto. GapFiller: A de novo assembly approach to fill the gap within paired reads. *BMC Bioinformatics*, v. 13, n. SUPPL 1, 7 set. 2012. Acesso em: 25 maio 2020.

NESTERENKO, O. A.; NOGINA, T. M.; KASUMOVA, S. A. Rhodococcus luteus nom. nov. and Rhodococcus maris nom. nov. *International Journal of Systematic Bacteriology*, v. 32, n. 1, p. 1–14, 1 jan. 1982. Acesso em: 25 maio 2020.

NOUIOUI, Imen *et al.* Genome-based taxonomic classification of the phylum actinobacteria. *Frontiers in Microbiology*, v. 9, n. AUG, p. 2007, 22 ago. 2018. Acesso em: 19 set. 2022.

PACHECO, Luis G.C. *et al.* Multiplex PCR assay for identification of Corynebacterium pseudotuberculosis from pure cultures and for rapid detection of this pathogen in clinical samples. *Journal of Medical Microbiology*, v. 56, n. 4, p. 480–486, 1 abr. 2007. Disponível em: <https://www.microbiologyresearch.org/content/journal/jmm/10.1099/jmm.0.46997-0>. Acesso em: 15 maio 2022.

RAINEY, F. A. *et al.* Dietzia, a new genus including Dietzia maris comb. nov., formerly Rhodococcus maris. *International Journal of Systematic Bacteriology*, v. 45, n. 1, p. 32–36, 1995. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/7857805/>. Acesso em: 13 out. 2020.

RAOULT, Didier; FOURNIER, Pierre Edouard; DRANCOURT, Michel. What does the future hold for clinical microbiology? *Nature reviews. Microbiology*, v. 2, n. 2, p. 151–159, fev. 2004. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/15040262/>. Acesso em: 15 maio 2022.

REZENDE, Andrea de Fátima Silva *et al.* In silico identification of Corynebacterium pseudotuberculosis antigenic targets and application in immunodiagnosis. *Journal of medical microbiology*, v. 65, n. 6, p. 521–529, 1 jun. 2016. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/27071381/>. Acesso em: 15 maio 2022.

RIBEIRO, Dayana *et al.* Subclinical diagnosis of Caseous lymphadenitis based on Elisa in Sheep from Brazil. *Currículo Lattes*, v. 04, n. 03, p. 1–4, 2013. Disponível em: <https://repositorio.unesp.br/handle/11449/133093>. Acesso em: 15 maio 2022.

RICHTER, Michael *et al.* JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*, v. 32, n. 6, p. 929–931, 15 mar. 2016. Disponível em: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btv681>. Acesso em: 18 nov. 2018.

RIEGEL, P *et al.* Taxonomy of Corynebacterium diphtheriae and related taxa, with recognition of Corynebacterium ulcerans sp. nov. nom. rev. *FEMS Microbiology Letters*, v. 126, n. 3, p. 271–276, 1 mar. 1995. Disponível em: <https://academic.oup.com/femsle/article/126/3/271/496820>. Acesso em: 15 maio 2022.

RIFICI, Claudia *et al.* Atypical Multibacterial Granulomatous Myositis in a Horse: First Report in Italy. *Veterinary Sciences*, v. 7, n. 2, p. 47, 21 abr. 2020. Acesso em: 25 maio 2020.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, v. 74, n. 12, p. 5463–5467, 1 dez. 1977. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/271968>. Acesso em: 14 maio 2019.

SCHREUDER, B. E.; TER LAAK, E. A.; DERCKSEN, D. P. Eradication of caseous lymphadenitis in sheep with the help of a newly developed ELISA technique. *The Veterinary record*, v. 135, n. 8, p. 174–176, 1994. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/7992473/>. Acesso em: 15 maio 2022.

SEIBOLD, E. *et al.* Identification of Francisella tularensis by Whole-Cell Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry: Fast, Reliable, Robust, and Cost-Effective Differentiation on Species and Subspecies Levels. *Journal of Clinical Microbiology*, v. 48, n. 4, p. 1061–1069, 1 abr. 2010. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/20181907>. Acesso em: 7 dez. 2018.

SELIM, S. A. Oedematous Skin Disease of Buffalo in Egypt. *Journal of Veterinary Medicine, Series B*, v. 48, n. 4, p. 241–258, 24 maio 2001. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1046/j.1439-0450.2001.00451.x>. Acesso em: 18 abr. 2022.

SILVA, Wanderson M. *et al.* Label-free proteomic analysis to confirm the predicted proteome of Corynebacterium pseudotuberculosis under nitrosative stress mediated by nitric oxide. *BMC Genomics*, v. 15, n. 1, p. 1–15, 2014.

STACKEBRANDT, Erko; RAINEY, Fred A.; WARD-RAINEY, Naomi L. Proposal for a new hierarchic classification system, Actinobacteria classis nov. *International Journal of Systematic Bacteriology*, v. 47, n. 2, p. 479–491, 1 jan. 1997a. Disponível em: <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/00207713-47-2-479>. Acesso em: 18 abr. 2022.

STACKEBRANDT, Erko; RAINEY, Fred A.; WARD-RAINEY, Naomi L. Proposal for a new hierarchic classification system, Actinobacteria classis nov. *International Journal of Systematic Bacteriology*, v. 47, n. 2, p. 479–491, 1 abr. 1997b. Disponível em:

<https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/00207713-47-2-479>. Acesso em: 24 jan. 2021.

STAPLETON, Sharon; BRADSHAW, Bernard; O'KENNEDY, Richard. Development of a surface plasmon resonance-based assay for the detection of Corynebacterium pseudotuberculosis infection in sheep. *Analytica Chimica Acta*, v. 651, n. 1, p. 98–104, 28 set. 2009. Acesso em: 15 maio 2022.

SUTHERLAND, S. S.; HART, R. A.; BULLER, N. B. Genetic differences between nitrate-negative and nitrate-positive C. pseudotuberculosis strains using restriction fragment length polymorphisms. *Veterinary microbiology*, v. 49, n. 1–2, p. 1–9, 1996. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/8861638/>. Acesso em: 15 maio 2022.

VAN DIJK *et al.* Ten years of next-generation sequencing technology. *Trends in genetics : TIG*, v. 30, n. 9, p. 418–426, 1 set. 2014. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25108476/>. Acesso em: 22 nov. 2022.

WINDSOR, P. A.; BUSH, R. D. Caseous lymphadenitis: Present and near forgotten from persistent vaccination? *Small Ruminant Research*, v. 142, p. 6–10, 1 set. 2016. Acesso em: 18 abr. 2022.

XU, Yunhe *et al.* Probiotic Properties of Lactobacillus paracasei subsp. paracasei L1 and Its Growth Performance-Promotion in Chicken by Improving the Intestinal Microflora. *Frontiers in Physiology*, v. 10, n. JUL, 2019. Disponível em: </pmc/articles/PMC6670285/>. Acesso em: 12 set. 2021.

YAMAMURA, Hideki *et al.* Dietzia timorensis sp. nov., isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, v. 60, n. 2, p. 451–454, 1 fev. 2010a. Acesso em: 25 maio 2020.

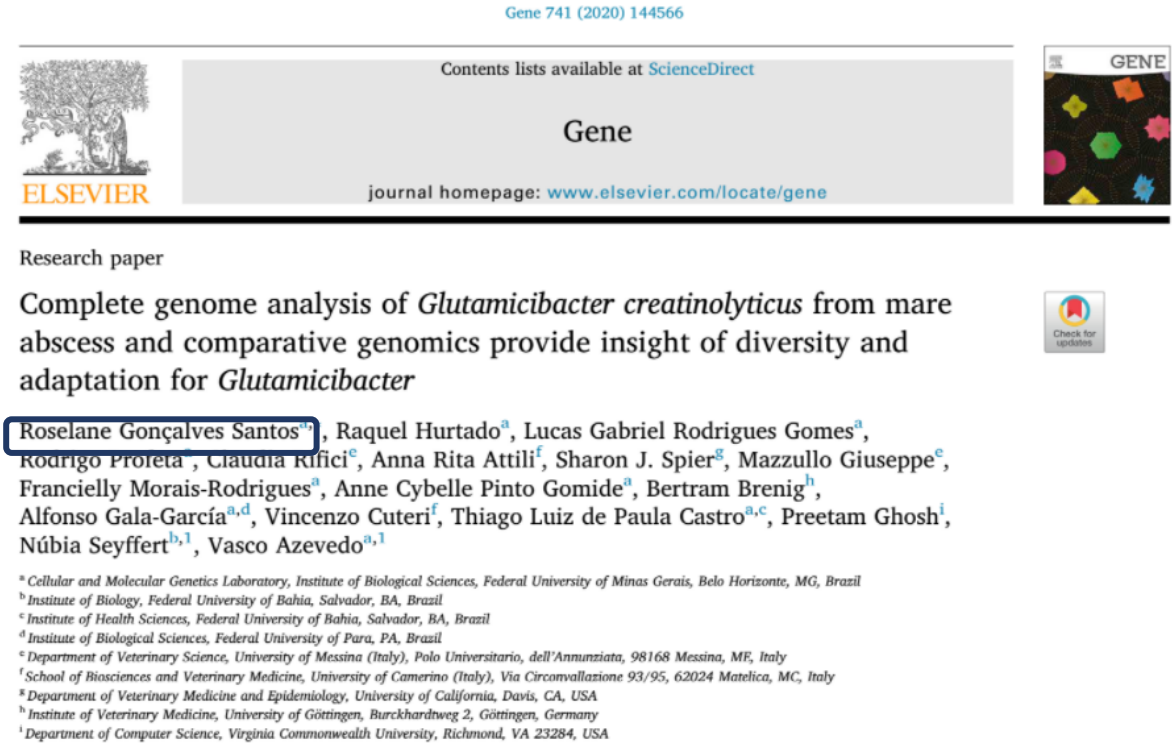
YAMAMURA, Hideki *et al.* Dietzia timorensis sp. nov., isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, v. 60, n. 2, p. 451–454, 1 fev. 2010b. Disponível em: <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.012229-0>. Acesso em: 29 jul. 2021.

9. APÊNDICE

Produção científica

Produção científica referente às atividades de pesquisa durante o período de 2019 a 2022. Um artigo como primeira autora, 9 como coautora, 1 capítulo livro publicados, e 1 artigo como coautora submetido.

Artigos Publicados como primeira autora:



Gene 741 (2020) 144566

Contents lists available at ScienceDirect

Gene

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

Research paper

Complete genome analysis of *Glutamicibacter creatinolyticus* from mare abscess and comparative genomics provide insight of diversity and adaptation for *Glutamicibacter*

Roselane Gonçalves Santos[a,*], Raquel Hurtado[a], Lucas Gabriel Rodrigues Gomes[a], Rodrigo Profeta[a], Claudia Rifici[e], Anna Rita Attili[f], Sharon J. Spier[g], Mazzullo Giuseppe[e], Francielly Morais-Rodrigues[a], Anne Cybelle Pinto Gomide[a], Bertram Brenig[h], Alfonso Gala-García[a,d], Vincenzo Cuteri[f], Thiago Luiz de Paula Castro[a,c], Preetam Ghosh[i], Núbia Seyffert[b,1], Vasco Azevedo[a,1]

[a] Cellular and Molecular Genetics Laboratory, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil
[b] Institute of Biology, Federal University of Bahia, Salvador, BA, Brazil
[c] Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil
[d] Institute of Biological Sciences, Federal University of Para, PA, Brazil
[e] Department of Veterinary Science, University of Messina (Italy), Polo Universitario, dell'Annunziata, 98168 Messina, ME, Italy
[f] School of Biosciences and Veterinary Medicine, University of Camerino (Italy), Via Circonvallazione 93/95, 62024 Matelica, MC, Italy
[g] Department of Veterinary Medicine and Epidemiology, University of California, Davis, CA, USA
[h] Institute of Veterinary Medicine, University of Göttingen, Burckhardtweg 2, Göttingen, Germany
[i] Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA

DOI https://doi.org/10.1016/j.gene.2020.144566

Este artigo foi referente ao trabalho do meu mestrado. Neste artigo, sequenciamos e analisamos o genoma de *Glutamicibacter creatinolyticus* LGCM 259, isolado de uma égua, que apresentava vários nódulos subcutâneos difusos, sendo esse o primeiro genoma dessa espécie isolada de um animal. Realizei todas as análises genômicas comparativas entre 4 representantes do gênero, de diferentes habitats, escrita do artigo, montagem e anotação do genoma.

Artigos Publicados como coautora

veterinary sciences

MDPI

Case Report

A Typical Multibacterial Granulomatous Myositis in a Horse: First Report in Italy

Claudia Rifici [1,†], Anna-Rita Attili [2,†,*], Davide De Biase [3], Roselane Gonçalves dos Santos [4], Núbia Seyffert [5], Thiago Luiz De Paula Castro [6], Henrique Cesar Pereira Figueiredo [7], Carmelo Scaramozzino [8], Stefano Reale [9], Orlando Paciello [3], Vincenzo Cuteri [2], Sharon Jane Spier [10], Vasco Azevedo [4] and Giuseppe Mazzullo [1]

DOI https://doi.org/10.3390/vetsci7020047

Este artigo também é um trabalho que está relacionado ao meu trabalho de mestrado e doutorado. Neste artigo, fiz a identificação de duas amostras isoladas uma égua italiana de 12 anos de idade, em que o animal apresentou nódulos subcutâneos difusos**,** foram realizadas investigações como PCR. MALDI-TOF, bem como também trabalhei na escrita do artigo.

Gene Reports 31 (2023) 101777

Contents lists available at ScienceDirect

Gene Reports

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

Genomic investigation of antimicrobial resistance in *Brucella abortus* strains isolated from cattle in Brazil

Carine Rodrigues Pereira [a], Rodrigo Bentes Kato [b], Fabrício Almeida Araújo [b], Alessandra Lima da Silva [c], Roselane Gonçalves dos Santos [c], Thiago de Jesus Sousa [c], Raquel Costa Neia [d], Saulo Britto da Silva [e], Charles H.D. Williamson [f], John Gillece [f], Andrey Pereira Lage [g], David O'Callaghan [h], Derek Pickard [i], Rommel Thiago Juca Ramos [b], Vasco Ariston de Carvalho Azevedo [c], Jeffrey T. Foster [f], Elaine Maria Seles Dorneles [a,*]

[a] Faculdade de Zootecnia e Medicina Veterinária, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil
[b] Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Pará, Brazil
[c] Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[d] Faculdade de Ciências Básicas, Universidade Federal Fluminense, Nova Friburgo, Rio de Janeiro, Brazil
[e] Escola de Ciências da Saúde, Universidade do Vale do Itajaí, Itajaí, Santa Catarina, Brazil
[f] Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, AZ, United States of America
[g] Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[h] Bacterial Virulence and Infectious Disease, University of Montpellier, Nimes, France
[i] Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom

Neste artigo, trabalhamos com a montagem de 52 genomas de *Brucella abortus* isoladas no Brasil. Trabalhei na montagem, anotação, deposito do genomas no NCB, analises de ilhas de patogenicidade, pangemona, dentre outras analises em que tinha a capacidade de fazer.

Gene
Volume 855, 1 March 2023, 147131

Comparative genomic analysis of ovine and other host associated isolates of *Staphylococcus aureus* exhibit the important role of mobile genetic elements and virulence factors in host adaptation

Alessandra Lima [a], Ana Carolina Barbosa Caetano [a], Raquel Hurtado Castillo [a], Roselane Gonçalves dos Santos [a], Diego Lucas Neres Rodrigues [a], Thiago de Jesus Sousa [a], Rodrigo Bentes Kato [a], Marcus Vinicius Canário Viana [a], Anne Cybelle Pinto Gomide [a], Flavia Figueira Aburjaile [f], Sandeep Tiwari [a b h], Arun Jaiswal [a], Alfonso Gala-García [c], Núbia Seyffert [b], Thiago Luiz de Paula Castro [b], Bertram Brenig [g], Mateus Matiuzzi da Costa [i], Elaine Maria Seles Dorneles [e], Yves Le Loir [d], Vasco Azevedo [a]

Neste artigo, trabalhamos com a montagem de genomas de *Stapylococcus aureus*. Trabalhei na montagem, anotação, deposito do genomas no NCB, bem como nas analises de bioinformática como busca por fatores de virulência, pangenoma.

International Journal of Scientific Research and Management (IJSRM)
||Volume||08||Issue||12||Pages||B-2020-85-95||2020||
Website: www.ijsrm.in ISSN (e): 2321-3418
DOI: 10.18535/ijsrm/v8i12.b01

Genes selected after application modified logistic regression in the microarrays gene expression for breast cancer.

Francielly Morais-Rodrigues[1*%], Diego Lucas Neres Rodrigues[1*], Rita Silvério-Machado[1], Lucas Gabriel Rodrigues Gomes[1], Rodrigo Bentes Kato[1], Roselane Gonçalves dos Santos[2], Vasco A. C. Azevedo[1], Marcos A dos Santos[2].

1 Institute of Biological Sciences, Federal University of Minas Gerais, Brazil. Av. Antônio Carlos, 6627, Belo Horizonte, MG 31270-901,Brazil.

2 Department of Computer Science, Federal University of Minas Gerais, Brazil Av Antônio Carlos, 6627 Belo Horizonte, MG 31270-901,Brazil.

DOI https://doi.org/10.18535/ijsrm/v8i12.b01

Neste trabalho utilizamos a regressão logística modificada foi usada para classificar os subtipos de câncer de mama usando todas as amostras de banco de dados de microarray. Neste artigo trabalhei junto com a autora na redação do artigo e revisão.

Gene 726 (2020) 144168

Contents lists available at ScienceDirect

Gene

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

Analysis of the microarray gene expression for breast cancer progression after the application modified logistic regression

Francielly Morais-Rodrigues[a,*,1], Rita Silvério-Machado[a], Rodrigo Bentes Kato[a], Diego Lucas Neres Rodrigues[a], Juan Valdez Baez[a], Vagner Fonseca[a,b], Emmanuel James San[b], Lucas Gabriel Rodrigues Gomes[a], Roselane Gonçalves dos Santos[a], Marcus Vinicius Canário Viana[a,c], Joyce da Cruz Ferraz Dutra[a], Mariana Teixeira Dornelles Parise[a], Doglas Parise[a], Frederico F. Campos[d], Sandro J. de Souza[e], José Miguel Ortega[a], Debmalya Barh[f], Preetam Ghosh[g], Vasco A.C. Azevedo[a], Marcos A. dos Santos[d]

[a] Institute of Biological Sciences, Federal University of Minas Gerais, Brazil. Av. Antônio Carlos, 6627, Belo Horizonte, MG 31270-901, Brazil
[b] KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), College of Health Sciences, University of KwaZulu-Natal, Durban 4001, South Africa
[c] Federal University of Pará, UFPA, Brazil
[d] Department of Computer Science, Federal University of Minas Gerais, Brazil Av Antônio Carlos, 6627, Belo Horizonte, MG 31270-901, Brazil
[e] Brain Institute, Federal University of Rio Grande d oNorte, Brazil
[f] Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur, West Bengal 721172, India
[g] Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA

DOI https://doi.org/10.1016/j.gene.2019.144168

Neste artigo, apresentamos um novo modelo baseado em regressão logística para classificar amostras de tumor de câncer de mama com base em dados de expressão de microarray, incluindo todos os recursos de expressão gênica e sem reduzir a matriz de dados de microarray. Trabalhei na escrita e revisão do artigo.

Data in Brief
Volume 34, February 2021, 106746

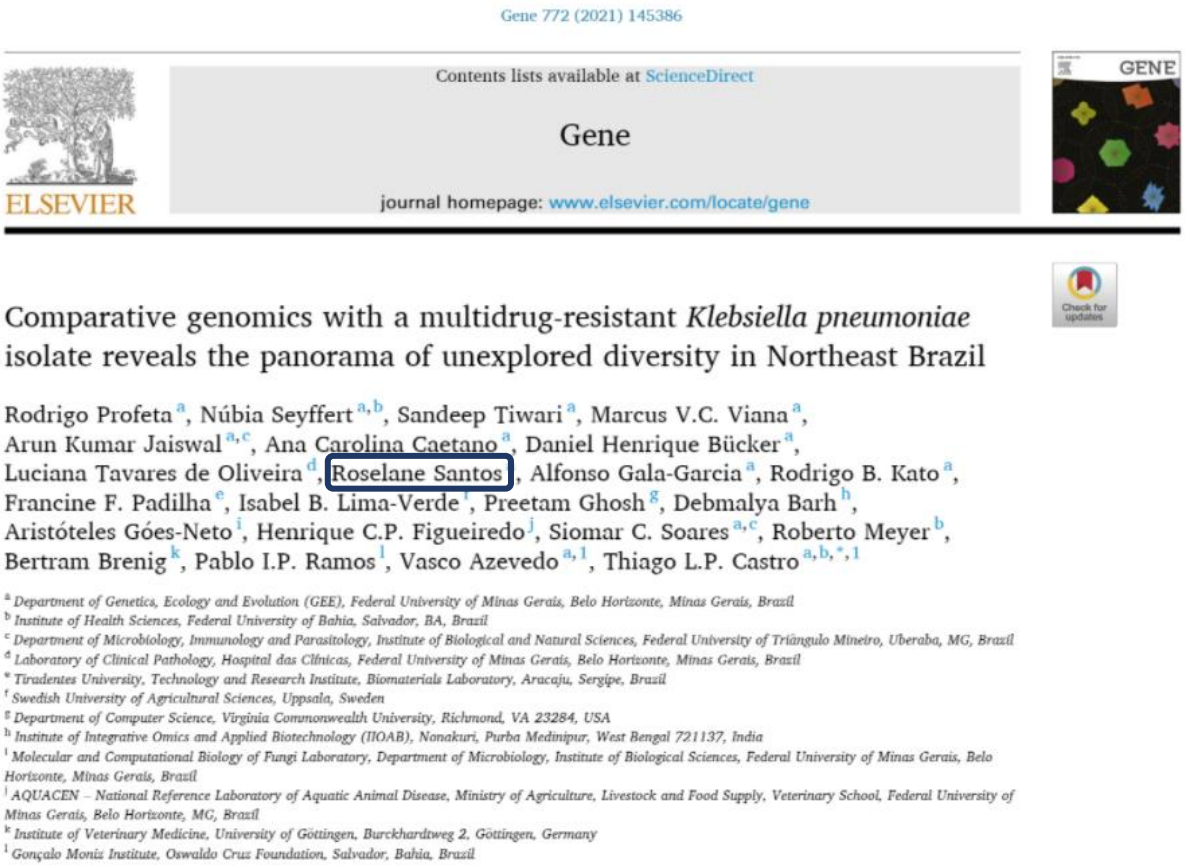
Data Article

Characterization of a new multidrug-resistant Brazilian *K. pneumoniae* isolate and 172 *Klebsiella* spp. sequenced strains: Genomic island, multilocus sequence typing and capsule locus dataset

Rodrigo Profeta [a], Núbia Seyffert [a, b], Sandeep Tiwari [a], Marcus V.C. Viana [a], Arun Kumar Jaiswal [a, c], Ana Carolina Caetano [a], Daniel Henrique Bücker [a], Luciana Tavares de Oliveira [d], Roselane Santos [a], Alfonso Gala-Garcia [a], Rodrigo B. Kato [a], Francine F. Padilha [e], Isabel B. Lima-Verde [f], Preetam Ghosh [g], Debmalya Barh [h], Aristóteles Góes-Neto [i], Henrique C.P. Figueiredo [j], Siomar C. Soares [a, c] ... Thiago L.P. Castro [a, b]

DOI https://doi.org/10.1016/j.dib.2021.106746

Neste artigo um isolado brasileiro de *K. pneumoniae* foi fenotipicamente caracterizada e submetida ao sequenciamento do genoma completo e análise genômica comparativa. Minha participação neste artigo foi nas analises de genômica compartiva.



Gene 772 (2021) 145386

Contents lists available at ScienceDirect

Gene

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

Comparative genomics with a multidrug-resistant *Klebsiella pneumoniae* isolate reveals the panorama of unexplored diversity in Northeast Brazil

Rodrigo Profeta [a], Núbia Seyffert [a,b], Sandeep Tiwari [a], Marcus V.C. Viana [a], Arun Kumar Jaiswal [a,c], Ana Carolina Caetano [a], Daniel Henrique Bücker [a], Luciana Tavares de Oliveira [d], Roselane Santos [a], Alfonso Gala-Garcia [a], Rodrigo B. Kato [a], Francine F. Padilha [e], Isabel B. Lima-Verde [f], Preetam Ghosh [g], Debmalya Barh [h], Aristóteles Góes-Neto [i], Henrique C.P. Figueiredo [j], Siomar C. Soares [a,c], Roberto Meyer [b], Bertram Brenig [k], Pablo I.P. Ramos [l], Vasco Azevedo [a,1], Thiago L.P. Castro [a,b,*,1]

[a] Department of Genetics, Ecology and Evolution (GEE), Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[b] Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil
[c] Department of Microbiology, Immunology and Parasitology, Institute of Biological and Natural Sciences, Federal University of Triângulo Mineiro, Uberaba, MG, Brazil
[d] Laboratory of Clinical Pathology, Hospital das Clínicas, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[e] Tiradentes University, Technology and Research Institute, Biomaterials Laboratory, Aracaju, Sergipe, Brazil
[f] Swedish University of Agricultural Sciences, Uppsala, Sweden
[g] Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA
[h] Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur, West Bengal 721137, India
[i] Molecular and Computational Biology of Fungi Laboratory, Department of Microbiology, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[j] AQUACEN – National Reference Laboratory of Aquatic Animal Disease, Ministry of Agriculture, Livestock and Food Supply, Veterinary School, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil
[k] Institute of Veterinary Medicine, University of Göttingen, Burckhardtweg 2, Göttingen, Germany
[l] Gonçalo Moniz Institute, Oswaldo Cruz Foundation, Salvador, Bahia, Brazil

DOI: 10.1016/j.gene.2020.145386

Neste artigo, um isolado de *K. pneumoniae* multirresistente, isolado de um paciente infectado na comunidade e internado em uma unidade de terapia intensiva no Nordeste do Brasil foi avaliado quanto a susceptibilidade antimicrobiana e as informações do genoma foram investigadas para caracterizar o isolado em frente a 172 cepas sequenciadas de diferentes países. Trabalhei com os dados de genômica, ajudando nas análises comparativa, e extração DNA.

antibiotics

MDPI

Article

Pan-Resistome Insights into the Multidrug Resistance of *Acinetobacter baumannii*

Diego Lucas Neres Rodrigues [1,†], Francielly Morais-Rodrigues [1,†], Raquel Hurtado [1], Roselane Gonçalves dos Santos [1], Daniela Camargos Costa [2], Debmalya Barh [1,3], Preetam Ghosh [4], Khalid J. Alzahrani [5], Siomar Castro Soares [6], Rommel Ramos [7], Aristóteles Góes-Neto [1], Vasco Azevedo [1,*,†] and Flávia Figueira Aburjaile [1,†]

1 Laboratory of Cellular and Molecular Genetics, Universidade Federal de Minas GeraisBelo Horizonte, Belo Horizonte 31270-901, MG, Brazil; dlnrodrigues@ufmg.br (D.L.N.R.); franrodriguesdacosta@ufmg.br (F.M.-R.); raquelgen1@gmail.com (R.H.); roselanegr@gmail.com (R.G.d.S.); dr.barh@gmail.com (D.B.); arigoesneto@icb.ufmg.br (A.G.-N.); faburjaile@gmail.com (F.F.A.)
2 FAMINAS-BH, Belo Horizonte 31744-007, MG, Brazil; daniela.costa@faminasbh.edu.br
3 Institute of Integrative Omics and Applied Biotechnology, Nonakuri West Bengal 721172, India
4 Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA; preetam.ghosh@gmail.com
5 Department of Clinical Laboratories Sciences, College of Applied Medical Sciences, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia; ak.jamaan@tu.edu.sa
6 Department of Research Development and Technological Innovation, Universidade Federal do Triângulo Mineiro, Uberaba 38025-180, MG, Brazil; siomars@gmail.com
7 Faculty of Biotechnology, Universidade Federal de Pará, Belém 66075-110, PA, Brazil; rommelramos@ufpa.br
* Correspondence: vasco@icb.ufmg.br
† These authors contributed equally to this work.

Citation: Rodrigues, D.L.N.;

DOI: **https://doi.org/10.3390/antibiotics10050596**

Neste artigo, trabalhei nas análises realizadas para busca do pan-resitoma de Acinetobacter baumannii.

frontiers in Bioinformatics

ORIGINAL RESEARCH
published: 07 September 2021
doi: 10.3389/fbinf.2021.711463

Organizing Events in Bioinformatics World: Insights Into Organizing Events in Bioinformatics

*Alessandra Lima da Silva[1†], Ana Paula de Abreu[2,3], Diego Mariano[3†], Felipe Caixeta[1], Fenícia Brito Santos[1†], Fernanda Stussi D. Lage[4†], Gabriel Quintanilha-Peixoto[1†], Heron. O. Hilário[5†], Joicymara. S. Xavier[6,7†], Lucio. R. Queiroz[1†], Nayara Evelin de Toledo[1†], Raphael Tavares[1], Rodrigo Bentes Kato[1†], Roselane Gonçalves dos Santos[1†], Stellamaris Soares[1,2†], Wanessa. M. Goes[1†], Wylerson. G. Nogueira[1†], Thiago. M. Batista[8], José Miguel Ortega[1], Vasco Ariston Azevedo De Carvalho[1], Glória. Regina Franco[1], Raquel. C. de Melo-Minardi[3] and Aristóteles Góes-Neto[1]**

[1]Institute of Biological Sciences, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil, [2]Department of Clinical and Toxicological Analysis, Faculty of Pharmacy, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [3]Department of Computer Science, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil, [4]Genomics for Climate Change Research Center, Universidade Estadual de Campinas (Unicamp), Campinas, Brazil, [5]Conservation Genetics Laboratory, Pontifícia Universidade Católica de Minas Gerais (PUC Minas), Belo Horizonte, Brazil, [6]Rene Rachou Institute (Fiocruz Minas), Belo Horizonte, Brazil, [7]Institute of Agricultural Sciences, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Unaí, Brazil, [8]Environmental Science Training Center, Universidade Federal do Sul da Bahia, Porto Seguro*

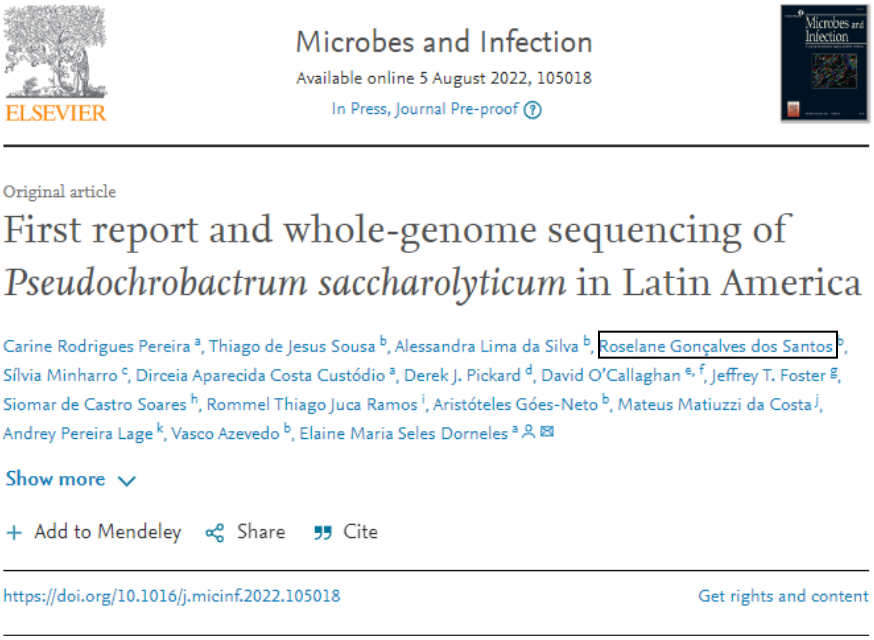
OPEN ACCESS

Edited by:
Sebastian Kmiecik, University of Warsaw, Poland

Reviewed by:
Nancy Wilkins-Diehr, University of California, San Diego, United States

DOI: https://doi.org/10.3389/fbinf.2021.711463

Este artigo é relacionado a organização do Curso de Verão em Bioinformática da UFMG, em que uma comissão organizadora composta por alunos de graduação e pós-doutorados da Universidade Federal de Minas Gerais (Brasil) promove um evento de uma semana denominado Curso de Verão em Bioinformática (CVBioinfo). Este evento tem como objetivo divulgar princípios, novidades e métodos de bioinformática voltados principalmente para o público de alunos de graduação. Aqui, apresentamos e discutimos os insights obtidos com a promoção do Workshop Online em Bioinformática (WOB) organizado em novembro de 2020. Eu fui uma dos membros da organização do curso, trabalhando também na escrita do artigo, realização do curso.



Microbes and Infection
Available online 5 August 2022, 105018
In Press, Journal Pre-proof

Original article

First report and whole-genome sequencing of *Pseudochrobactrum saccharolyticum* in Latin America

Carine Rodrigues Pereira [a], Thiago de Jesus Sousa [b], Alessandra Lima da Silva [b], Roselane Gonçalves dos Santos [b], Sílvia Minharro [c], Dirceia Aparecida Costa Custódio [a], Derek J. Pickard [d], David O'Callaghan [e, f], Jeffrey T. Foster [g], Siomar de Castro Soares [h], Rommel Thiago Juca Ramos [i], Aristóteles Góes-Neto [b], Mateus Matiuzzi da Costa [j], Andrey Pereira Lage [k], Vasco Azevedo [b], Elaine Maria Seles Dorneles [a]

Show more

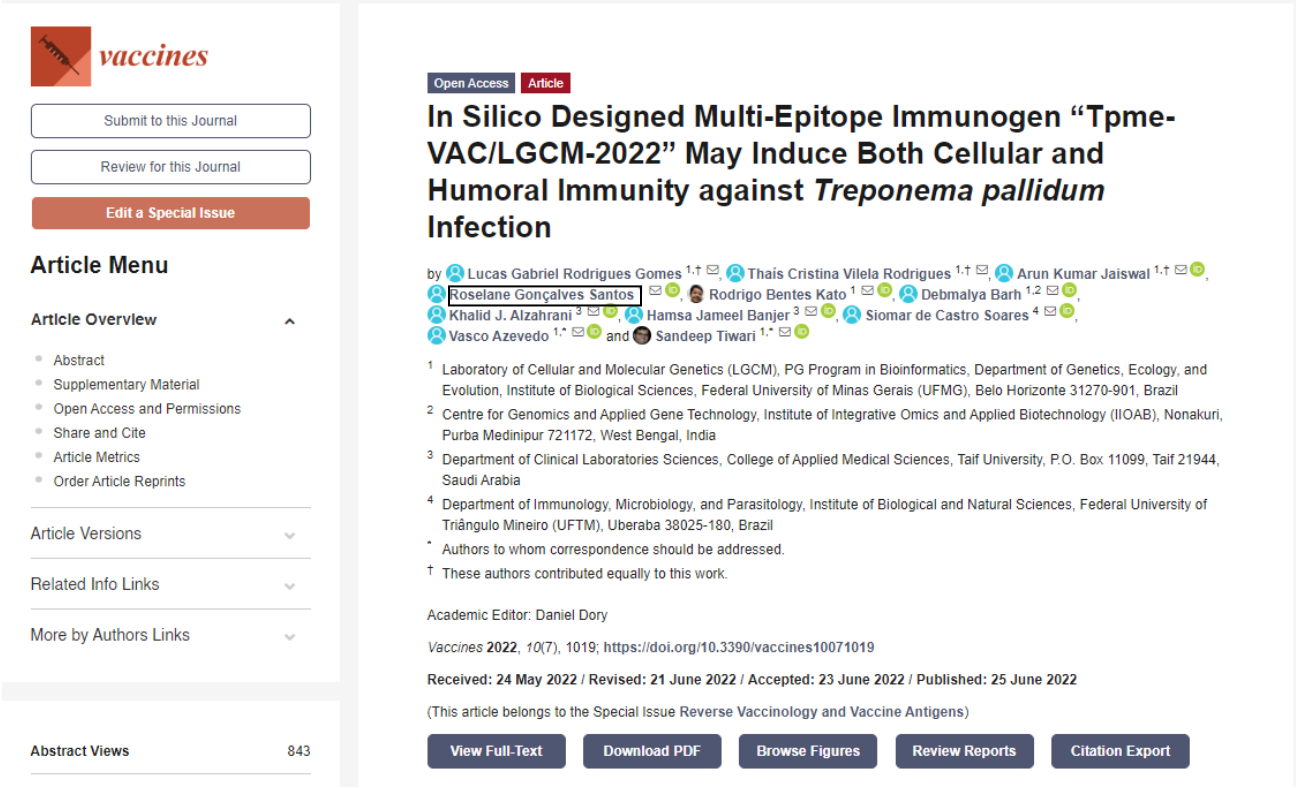
+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.micinf.2022.105018

Get rights and content

DOI: https://doi.org/10.1016/j.micinf.2022.105018

Neste artigo, relatamos o primeiro isolamento, identificação e caracterização de *Pseudochrobactrum saccharolyticum* da América Latina. Trabalhei na montagem, anotação, deposito do genoma, bem como na ajuda das analises realizadas.



vaccines

Submit to this Journal

Review for this Journal

Edit a Special Issue

Article Menu

Article Overview

- Abstract
- Supplementary Material
- Open Access and Permissions
- Share and Cite
- Article Metrics
- Order Article Reprints

Article Versions

Related Info Links

More by Authors Links

Abstract Views 843

Open Access Article

In Silico Designed Multi-Epitope Immunogen "Tpme-VAC/LGCM-2022" May Induce Both Cellular and Humoral Immunity against *Treponema pallidum* Infection

by Lucas Gabriel Rodrigues Gomes [1,†], Thaís Cristina Vilela Rodrigues [1,†], Arun Kumar Jaiswal [1,†], Roselane Gonçalves Santos, Rodrigo Bentes Kato [1], Debmalya Barh [1,2], Khalid J. Alzahrani [3], Hamsa Jameel Banjer [3], Siomar de Castro Soares [4], Vasco Azevedo [1,*] and Sandeep Tiwari [1,*]

[1] Laboratory of Cellular and Molecular Genetics (LGCM), PG Program in Bioinformatics, Department of Genetics, Ecology, and Evolution, Institute of Biological Sciences, Federal University of Minas Gerais (UFMG), Belo Horizonte 31270-901, Brazil

[2] Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur 721172, West Bengal, India

[3] Department of Clinical Laboratories Sciences, College of Applied Medical Sciences, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

[4] Department of Immunology, Microbiology, and Parasitology, Institute of Biological and Natural Sciences, Federal University of Triângulo Mineiro (UFTM), Uberaba 38025-180, Brazil

* Authors to whom correspondence should be addressed.

† These authors contributed equally to this work.

Academic Editor: Daniel Dory

Vaccines **2022**, *10*(7), 1019; **https://doi.org/10.3390/vaccines10071019**

Received: 24 May 2022 / Revised: 21 June 2022 / Accepted: 23 June 2022 / Published: 25 June 2022

(This article belongs to the Special Issue Reverse Vaccinology and Vaccine Antigens)

View Full-Text | Download PDF | Browse Figures | Review Reports | Citation Export

Este artigo utiliza uma abordagem baseada em imunoinformática e projeta uma candidata potencial a vacina para a sífilis. Trabalhei na revisão do artigo, entretanto fui sua coorientadora durante o desenvolvimento do trabalho.

Capítulo de livro publicado (Ensino)

Uma estratégia para engajamento de participantes de eventos online

Diego Mariano, Wylerson G. Nogueira, Wanessa M. Goes, Roselane G. dos Santos, Rodrigo Bentes Kato, Nayara Toledo, Lucio R. Queiroz, Heron O. Hilário, Gabriel Quintanilha-Peixoto, Fernanda S. D. Lage, Fenícia Brito Santos, Felipe Caixeta, Ana Paula de Abreu, Alessandra Lima da Silva, Joicymara S. Xavier

Revisão: Filipe Zimmer

BIOINFO – Revista Brasileira de Bioinformática. Edição #01. Julho, 2021.

DOI: 10.51780/978-6-599-275326-19

Neste capítulo de livro, apresentamos uma estratégia de gamificação usada para ampliar o engajamento dos participantes durante o I Workshop Online de Bioinformática (WOB20)., a fim de aumento o engajamento e interação entre participantes. Fui organizadora do evento e tdos juntos trabalhamos na escrita e desenvolvimento do trabalho.

Artigos submetidos

I – PeerJ. **Comparative genomics of persistent Staphylococcus warneri from recurrent goat mastitis associated with the use of enrofloxacin**. Fev 2021. Chirles Araújo de França, Alessandra da Silva, Danillo Sales Rosa, Gutiele do Nascimento do É, Diego Lucas Neres Rodrigues, Raquel Emna Hurtado Castillo, Thiago de Jesus Sousa, Roselane Gonçalves dos Santos, Anne Cybelle Pinto Gomide, Maria Aparecida Scatamburlo Moreira, Jackson Roberto Guedes da Silva Almeida, Gisele Veneroni Gouveia, João José de Simoni Gouveia, Renata de Faria Silva Souza, Bertram Brenig, Aristóteles Góes-Neto, Siomar Castro Soares, Rommel Thiago Juca Ramos, Vasco Ariston de Carvalho Azevedo, Mateus Matiuzzi da Costa. PeerJ. Comparative genomics of persistent Staphylococcus warneri from recurrent goat mastitis associated with the use of enrofloxacin.

Participação em eventos/ Palestras/ Ministração de cursos

I Apresentação de trabalho e palestra

1. **Gonçalves dos Santos, Roselane**
   Caracterização de uma nova espécie do gênero *Dietzia* e análise comparativa do gênero, 2021. (Seminário,Apresentação de Trabalho)
2. LUCAS NERES RODRIGUES, DIEGO; MORAIS-RODRIGUES, FRANCIELLY; **DOS SANTOS, ROSELANE GONÇALVES**; ABURJAILE, F. F.; AZEVEDO, VASCO
   ANÁLISE FUNCIONAL DO PANGENOMA DE Acinetobacter baumannii, 2020. (Simpósio,Apresentação de Trabalho)
3. LUCAS NERES RODRIGUES, DIEGO; MORAIS-RODRIGUES, F.; **Dos Santos Roselane Gonçalves**; ABURJAILE, F. F.; Azevedo, V
   Resistome profile of Acinetobacter baumannii, 2020. (Congresso,Apresentação de Trabalho)
4. **GONÇALVES, R. R.**; CARMO, F. L. R.; GALA-GARCIA, A.; JAN, G.; LE-LOIR, Y.; FIGUEIREDO, H. C. P.; GOMES, L. G. R COSTA, R. A.; Azevedo, V
   Analysis in silico of knockout genome mutation strain of Propionibacterium freudenreichii, 2019. (Conferência ou palestra,Apresentação de Trabalho)

II Participação em eventos

1. **Aula Magna da Pós-graduação do ICB 2021: Epidemiologia de viroses emergentes, com ênfase na disseminação do SARS-CoV-2 e suas variantes no mundo**, 2021. (Seminário)
   .
2. **Biologia Molecular de Ponta a Ponta: da PCR ao sequenciamento de 3° Geração**, 2021. (Outra)
   .

3. **Keynote Webinar with Geraldine Van der Auwera - Genomic analysis with GATK: Best Practices and latest innovations - X- meeting eXperience 2**, 2020. (Congresso)
.
4. Apresentação de Poster / Painel no(a) **X-meeting eXperience 2020 held online**, 2020. (Congresso)
Resistome profile of Acinetobacter baumannii.
5. Apresentação de Poster / Painel no(a) **"VII Simpósio de Microbiologia da UFMG - Conecta SIM",** 2020. (Simpósio)
ANÁLISE FUNCIONAL DO PANGENOMA DE Acinetobacter baumannii.

III Ministração de Curso curta duração

1. **GONÇALVES-SANTOS, R.**
Genômica Comparativa, 2020. (Outro, Curso de curta duração ministrado)

2. **Gonçalves dos Santos, Roselane**
Treinamento de Montagem, Anotação e Depósito de genomas de procarioto, 2020. (Outro, Curso de curta duração ministrado)

IV Organização de evento

1. **GONÇALVES, R. R.**
I Workshop Online de Bioinformática, 2020. (Outro, Organização de evento)

2. **GONÇALVES, R. R.**
IV Curso de Verão em Bioinformática da UFMG, 2020. (Outro, Organização de evento)

3. **GONÇALVES, R. R.**
III Curso de Verão em Bioinformática da UFMG, 2019. (Outro, Organização de evento)

V Revisora de um artigo na revista Large Animal Review:

Investigation of the effect of Parapoxvirus ovis, Corynebacterium cutis lysate and vitamin c on immunosuppression caused by long-term transport Stress in Morkaraman sheep.

VI Orientações e supervisões concluídas

Trabalho de conclusão de curso de graduação: Lucas Gabriel Rodrigues Gomes. Immunoinformatics-Aided Design and Evaluation of a Potential Multi-Epitope Vaccine against Treponema pallidum. 2021. Trabalho de Conclusão de Curso. (Graduação em Ciências Biológicas) - Universidade Federal de Minas Gerais. Coorientadora: Roselane Gonçalves dos Santos. CC de um aluno de graduação.

Submission Confirmation

Print

Thank you for your submission

Submitted to
Journal of Basic Microbiology

Manuscript ID
jobm.202300717

Title
Dietzia calabriae sp. nov. is a bacterial species isolated from a mare's abscess

Authors
dos Santos, Roselane
Rita , Attili
de Jesus, Luís
Neres Rodrigues, Diego
Lima, Alessandra
Hurtado, Raquel
Claudia, Rifici
Sharon, Spier
Giuseppe , Mazzullo
Morais-Rodrigues , Francielly
Gomide , Anne
Brenig, Bertram
Vincenzo , Cuteri
Castro , Thiago
Seyffert , Núbia
Jaiswal, Arun
Batista, Ana
de Magalhaes , Mariana
Figueira Aburjaile, Flávia
Góes-Neto, Aristóteles
Tiwari, Sandeep
Azevedo, Vasco

Date Submitted
03-Dec-2023

Author Dashboard