



UnB

Universidade de Brasília

Instituto de Ciências Biológicas

Departamento de Biologia Celular

Programa de Pós-Graduação em Biologia Molecular

Isabel Cristina Cunha Ferreira

Caracterização taxonômica e funcional de linhagens bacterianas isoladas a partir do uso do sobrenadante de *Paenibacillus elgii* como mecanismo de seleção.

Brasília, 2023

ISABEL CRISTINA CUNHA FERREIRA

Caracterização taxonômica e funcional de linhagens bacterianas isoladas a partir do uso do sobrenadante de *Paenibacillus elgii* como mecanismo de seleção.

Tese apresentada ao Programa
de Pós-graduação em Biologia Molecular
da Universidade de Brasília como requisito
para obtenção do Título de
Doutora em Biologia Molecular.
Orientador: Dr. Ricardo Henrique Krüger

Brasília, 2023

Agradecimentos

Ao meu orientador, Dr. Ricardo Krüger, pela confiança depositada em mim, e por compartilhar comigo seus conhecimentos, além de sempre se colocar disponível e sanar minhas dúvidas sem medir esforços.

A minha praticamente co-orientadora (pena que não deu pra formalizar), Dra. Carla Vizzotto, por ter acreditado em mim, pela paciência e compreensão nas minhas dificuldades, por se dispor a me ensinar e por todo tempo dispensado a mim desde o processo seletivo. Por segurar a minha mão desde o início e não ter largado nunca.

Ao José Maria, sem você eu não teria conhecido a Carla e nada disso seria possível. Obrigada por sempre se interessar e apoiar o andamento do meu trabalho.

À toda minha família por sempre acreditar em mim. Por todo apoio, carinho, torcida e incentivo em tudo na vida. E principalmente, por me dar a certeza de que se algo der errado vocês estarão sempre aí pra mim.

À Universidade de Brasília por toda infraestrutura fornecida para a execução deste trabalho. Ao Programa de Pós-Graduação em Biologia Molecular por nos proporcionar um curso de alto nível. À CAPES por me fornecer bolsa durante todo o curso, ao CNPq e à FAP-DF pelo apoio financeiro ao nosso laboratório. Ao Laboratório de Microscopia do IB-UnB pela estrutura para a obtenção das imagens em Microscopia Eletrônica de Varredura e Transmissão e às meninas que foram tão solícitas comigo sempre.

À todos os integrantes do Laboratório de Enzimologia pela amizade e por contribuírem de alguma forma para a realização deste trabalho. Muito amor, carinho e gratidão a todos vocês. Ao Raul, por ter me recebido tão bem no laboratório e me alfabetizado na microbiologia. Obrigada pela sua amizade do doutorado para vida! À Tayná, minha parceira em vários experimentos, sempre disposta a me ajudar e me ensinar e por me alfabetizar na Bioinformática. E ao Pedro, às Jéssicas, à Gabi e à Helena, amigos que a pós-graduação me proporcionou, sempre dispostos a ajudar e sempre com uma palavra amiga.

Ao Wladimir, amigo e irmão do mestrado para sempre, que me acompanhou, apoiou e ajudou muito em todas as fases dessa empreitada. À Flávia, minha amiga de infância, presente a vida toda independente de distância física, por todo amor, apoio e torcida sempre. Ao meu primo Douglas e meu amigo Londerson, programadores sempre disponíveis a "bater cabeça" comigo nas minhas análises de bioinformática. Sempre com muito carinho! À Mayanne, minha parceira da Antártica pra vida, sempre presente e solícita. Aos colegas do Jiu Jitsu por sempre me incentivarem e me apoiarem. Aos amigos de Arraial Bebele, Ana, Pati, Augusto, Fe, Hedinho e Lila, e minha prima de Curitiba, Vivi, por todo apoio e positividade sempre.

Resumo

Neste estudo, avaliou-se a prospecção de linhagens bacterianas com potencial para a produção de compostos antimicrobianos a partir do enriquecimento de meios de cultura com o sobrenadante de *Paenibacillus elgii*, que é abundante em peptídeos antimicrobianos e outras moléculas sinal. Amostras de solo foram inoculadas nesses meios a fim de selecionar apenas linhagens bacterianas resistentes àquele sobrenadante e, por isso, com potencial para a produção de compostos similares. A busca por estratégias direcionadas à obtenção de linhagens com potencial biotecnológico se fazem necessárias devido as limitações das técnicas clássicas de cultivo em prospectar novas espécies produtoras de antimicrobianos. Sabe-se também que o repertório de compostos bioativos estabelecidos para os micro-organismos já conhecidos ainda não foram totalmente desvendados. Há ainda a preocupação com a resistência a antibióticos desenvolvida por algumas bactérias, fato que tem sido uma ameaça crescente ao tratamento eficaz de infecções causadas por esses micro-organismos. Por isso, há uma demanda para descoberta de novos antibacterianos que substituam aqueles que se tornaram ineficazes. O desenvolvimento de antimicrobianos exige uma série de etapas, onde a primeira é a descoberta e caracterização de produtores potenciais. As linhagens bacterianas obtidas através desse protocolo foram avaliadas quanto a sua atividade antibacteriana contra *Escherichia coli*, *Bacillus subtilis* e *Pseudomonas aeruginosa* e foram previamente identificadas por sequenciamento do tipo Sanger do gene marcador molecular *rRNA* 16S. Todas as linhagens apresentaram atividade antibacteriana para *B.subitilis* e *E. coli* e apenas duas delas para *P. aeruginosa*. A análise do gene *rRNA* 16S mostrou que as linhagens selecionadas pertencem a gêneros já conhecidos por seu potencial biotecnológico, são eles: *Paenibacillus*, *Pseudomonas*, *Enterobacter*, *Kitasatospora*, *Sphingomonas*, *Xanthobacter*, *Burkholderia* e *Methylobacterium*. As linhagens bacterianas isoladas e com atividade antibacteriana contra *P. aeruginosa* foram denominadas linhagem K002 e linhagem K003 e tiveram seus genótipos e fenótipos caracterizados. Seus genomas foram sequenciados nas plataformas *Illumina* e *Oxford Nanopore Technologies*. A classificação taxonômica baseada em genomas mostrou que a espécie mais próxima da linhagem K002 é *Kitasatospora xanthocidica* (*dDDH* 32,8-37,8% e *ANI* 86,86%) indicando que pode se tratar de uma nova espécie. A linhagem K003 foi confirmada como uma *Methylobacterium radiotolerans* (*dDDH* 90,3-94% e *ANI* 97,51%). A anotação funcional dos genomas indicou a presença de sessenta *clusters* gênicos biossintéticos relacionados ao metabolismo secundário para a linhagem K002 e dez para a linhagem K003. A maioria deles pode ser relacionada com o potencial antimicrobiano das linhagens. A presença de determinados *clusters* gênicos biossintéticos em comum com *P.elgii* pode indicar a produção de compostos semelhantes entre ele e as linhagens K002 e K003 selecionadas a partir de seu sobrenadante. A metodologia descrita neste trabalho permitiu o isolamento de linhagens de gêneros distintos de bactérias. Os resultados mostram que as linhagens K002 e K003 abrigam vários genes responsáveis pela produção biossintética de metabólitos secundários confirmando seu potencial como uma valiosa fonte de compostos bioativos. Assim, este estudo mostrou a eficiência do uso do sobrenadante de *P. elgii* em selecionar micro-organismos com potencial para aplicação biotecnológica. Além disso, o fato da linhagem K002 se tratar de uma espécie ainda não descrita pode indicar que a metodologia favorece o isolamento de novas espécies bacterianas.

Palavras-chave: prospecção, sobrenadante, linhagens bacterianas, atividade antibacteriana

Abstract

This study evaluated the prospection of bacterial strains with potential for the production of antimicrobial compounds from the enrichment of culture media with the supernatant of *Paenibacillus elgii*, which is abundant in antimicrobial peptides and other signal molecules. Soil samples were inoculated in these media in order to select only bacterial strains resistant to that supernatant and, therefore, with potential for the production of similar compounds. The search for strategies aimed at obtaining strains with biotechnological potential is necessary due to the limitations of classic cultivation techniques in prospecting new species that produce antimicrobials. It is also known that the repertoire of bioactive compounds established for known microorganisms has not yet been fully unveiled. There is also concern that antibiotic resistance developed by some bacteria, a fact that has been a growing threat to the effective treatment of infections caused by these microorganisms. Therefore, there is a demand for the discovery of new antibacterials to replace those that have become ineffective. The development of antimicrobials requires a series of steps, the first of which is the discovery and characterization of potential producers. The bacterial strains obtained through this protocol were evaluated for their antibacterial activity against *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas aeruginosa* and were previously identified by Sanger sequencing of the 16S rRNA molecular marker gene. All strains showed antibacterial activity for *B. subtilis* and *E. coli* and only two of them for *P. aeruginosa*. Analysis of the 16S rRNA gene showed that the selected strains belong to genera already known for their biotechnological potential. They are: *Paenibacillus*, *Pseudomonas*, *Enterobacter*, *Kitasatospora*, *Sphingomonas*, *Xanthobacter*, *Burkholderia* and *Methylobacterium*. Bacterial strains isolated and with antibacterial activity against *P. aeruginosa* were named K002 strain and K003 strain and had their genotypes and phenotypes characterized. Their genomes were sequenced on the Illumina and Oxford Nanopore Technologies platforms. The taxonomic classification based on genomes showed that the species closest to the K002 strain is *Kitasatospora xanthocidica* (dDDH 32.8-37.8% and ANI 86.86%) indicating that it may be a new species. The K003 strain was confirmed to be a *Methylobacterium radiotolerans* (dDDH 90.3-94% and ANI 97.51%). The functional annotation of the genomes indicated the presence of sixty biosynthetic gene clusters related to secondary metabolism for the K002 strain and ten for the K003 strain. Most of them can be related to the antimicrobial potential of the strains. The presence of certain biosynthetic gene clusters in common with *P. elgii* may indicate the production of similar compounds between it and the K002 and K003 strains selected from its supernatant. The methodology described in this work allowed the isolation of distinct bacterial genera. The results show that K002 and K003 strains harbor several genes responsible for the biosynthetic production of secondary metabolites confirming their potential as a valuable source of bioactive compounds. Thus, this study showed the efficiency of using *P. elgii* supernatant in selecting microorganisms with potential for biotechnological application. Furthermore, the fact that the K002 strain is a species not yet described may indicate that the methodology favors the isolation of new bacterial species.

Keywords: prospecting, supernatant, bacterial strains, antibacterial activity

Lista de Ilustrações

Figura 1.1 Fluxograma das principais etapas deste trabalho apresentado em capítulos	14
Figura 2.1 Atividade antibiótica de <i>P. elgii</i> e seu sobrenadante, demonstrando atividade da cultura indicada pela seta vermelha (A) e do sobrenadante (300µl/L) indicada pela seta azul (B). Os ensaios foram feitos contra <i>B. subtilis</i> (indicado por setas amarelas) (Carvalho, 2013)	46
Figura 2.2 Crescimento de culturas a partir de amostras de solo em diferentes concentrações de sobrenadante de <i>P. elgii</i> . A placa A (300µl/L), a placa B (600 µl/L), a placa C (1ml/L) e a placa D (2ml/L). As linhagens de interesse coletadas por possuírem zonas de inibição são indicadas por setas vermelhas. (Carvalho, 2013)	46
Figura 2.3 Gel (TBE/agarose 1% corado com brometo de etídeo) com resultados da <i>PCR</i> das linhagens. O marcador de referência de tamanho de bandas utilizado foi o 1kb plus DNA <i>Ladder</i> indicado a esquerda com algumas marcações de tamanho de bandas (Carvalho, 2013)	47
Figura 2.4 Ensaios <i>overlay</i> de atividade antibacteriana de algumas das linhagens obtidas a partir do uso de sobrenadante de <i>P. elgii</i> . Os halos de inibição são bastante visíveis. (A) Linhagem A3, (B) Linhagem A9, (C) Linhagem C5 e (D) Linhagem C12. Todos os ensaios foram realizados contra <i>B. subtilis</i> (Carvalho, 2013)	49
Figura 2.5 Árvore filogenética baseada na sequência parcial do gene 16S ribossomal <i>rRNA</i> obtido por clonagem e sequenciamento de Sanger da linhagem <i>Methylobacterium</i> K003, obtido a partir da utilização do sobrenadante de <i>P. elgii</i> . A posição evolutiva da linhagem foi estimada pela construção de uma árvore filogenética usando o método <i>Maximum-likelihood</i> (ML) com 1000 repetições de amostragem (<i>Bootstrap</i>), usando sequências representativas de 16S ribossomal <i>rRNA</i> de linhagens válidas do gênero listadas na <i>List of Prokaryotic names with Standing in Nomenclature - LPSN</i>	53
Figura 2.6 Mapa circular da montagem híbrida do genoma da linhagem <i>Methylobacterium</i> K003, obtida a partir do sobrenadante de <i>P. elgii</i> . A ilustração foi gerada pelo recurso <i>DNAPlotter</i> do <i>software</i> Artemis demonstrando as regiões codificantes e o conteúdo GG. De fora para o centro: <i>CDS</i> direto (azul escuro), <i>CDS</i> reverso (vermelho), conteúdo de GC e inclinação de GC (cinza escuro e cinza claro claro)	54
Figura 2.7 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de <i>Methylobacterium</i> K003, linhagem obtida a partir do sobrenadante de <i>P. elgii</i> , baseada na sequência do gene ribossomal <i>rRNA</i> 16S	56

Figura 2.8 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de <i>Methylobacterium</i> K003, linhagem obtida a partir do sobrenadante de <i>P. elgii</i> , baseada em seu genoma total	56
Figura 2.9 Visualização circular do alinhamento dos genomas da linhagem <i>M. radiotolerans</i> K003 (em cinza) e sua linhagem mais próxima de acordo com TYGS-DSMZ, <i>M. radiotolerans</i> JCM2831 (em azul). O alinhamento e a ilustração foram gerados usando o Plot Nucmer. Apenas os blocos de alinhamento do genoma com pelo menos 1000 nucleotídeos e 90% de identidade são mostrados. A banda externa é a escala de tamanho dos <i>contigs</i> . Quando a banda azul tem duas camadas de profundidade, há duas sequências na linhagem mais próxima que se sobrepõem no mesmo local na linhagem <i>M. radiotolerans</i> K003	57
Figura 2.10 Árvore filogenética de máxima verossimilhança baseada em 233 genes ortólogos principais concatenados de sequências genômicas de 12 linhagens de <i>Methylobacterium</i> , incluindo <i>M. radiotolerans</i> K003, as linhagens intimamente relacionadas de acordo com o TYGS-DSMZ e o grupo externo <i>Enterovirga rhinocerotis</i> . O servidor <i>web</i> M1CR0B1AL1Z3R foi utilizado para concatenar os genes ortólogos e construir a árvore filogenética	58
Figura 2.11 Análise comparativa do pangenoma realizada no Roary. O genoma de <i>M. radiotolerans</i> K003 e os 10 genomas mais próximos a ele, de acordo com TYGS-DSMZ, foram utilizadas. Uma identidade percentual mínima para BLASTp de 80% foi definida para agrupar os genes que codificam sequências de proteínas completas em genomas centrais.....	59
Figura 2.12 Proteínas preditas no genoma da linhagem <i>M. radiotolerans</i> K003, obtidas a partir do sobrenadante de <i>P. elgii</i> , classificadas de acordo com as categorias COGs. A classificação foi realizada usando o mapeador eggNOG contra o banco de dados da <i>Kyoto Encyclopedia of Genes and Genomes (KEGG)</i>	60
Figura 2.13 Análise de similaridade entre as famílias de <i>clusters</i> de genes usando BiG-SCAPE. Cada união representa <i>clusters</i> gênicos semelhantes entre as linhagens <i>M. radiotolerans</i> K003 e JCM283. Os <i>clusters</i> específicos, ou seja, sem semelhanças entre as linhagens também são apresentados em forma única. Não há representação gráfica dos <i>clusters</i> gênicos semelhantes por se tratarem de <i>clusters</i> idênticos.....	63
Figura 2.14 Comparação entre as predições fenotípicas da montagem híbrida do genoma da linhagem <i>M. radiotolerans</i> K003, obtida a partir do uso do sobrenadante de <i>P. elgii</i> , e dos genomas de suas linhagens mais próximas de acordo com o TYGS-DSMZ obtidas no banco de dados do <i>NCBI</i> . O valor zero significa que a linhagem é negativa para o fenótipo, os valores 1 e 2 correspondem ao positivo para o fenótipo nos preditores <i>phypat</i> e <i>phypat+PGL</i> ,	

respectivamente. O valor 3 mostra que o fenótipo foi positivo nos dois preditores. A análise foi realizada pelo <i>software</i> Traitair	65
Figura 3.1 Micrografias eletrônicas de varredura (A e B) e de transmissão (C e D) de <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . Podem ser observados micélios vegetativos (forma alongada) e esporos submersos (forma arredondada)	84
Figura 3.2 Árvore filogenética baseada na seqüência parcial do gene 16S ribossomal <i>rRNA</i> obtido por clonagem e sequenciamento de Sanger de <i>Kitasatospora sp.</i> K002, obtida a partir da utilização do sobrenadante de <i>P. elgii</i> . A posição evolutiva da linhagem foi estimada pela construção de uma árvore filogenética usando o método <i>Maximum-likelihood (ML)</i> com 1000 repetições de amostragem (<i>Bootstrap</i>), usando seqüências representativas de 16S ribossomal <i>rRNA</i> de linhagens válidas do gênero listadas na <i>List of Prokaryotic names with Standing in Nomenclature - LPSN</i>	88
Figura 3.3 Mapa circular da montagem híbrida do genoma de <i>Kitasatospora sp.</i> K002, obtida a partir do sobrenadante de <i>P. elgii</i> . A ilustração foi gerada pelo recurso <i>DNAPlotter</i> do <i>software</i> Artemis demonstrando as regiões codificantes e o conteúdo GG. De fora para o centro: <i>CDS</i> direto (azul escuro), <i>CDS</i> reverso (vermelho), conteúdo de GC e inclinação de GC (cinza escuro e cinza claro claro)	90
Figura 3.4 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de <i>Kitasatospora</i> K002, linhagem obtida a partir do sobrenadante de <i>P. elgii</i> , baseada na seqüência do gene ribossomal <i>rRNA</i> 16S	91
Figura 3.5 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de <i>Kitasatospora</i> K002, linhagem obtida a partir do sobrenadante de <i>P. elgii</i> , baseada em seu genoma total	92
Figura 3.6 Visualização circular do alinhamento dos genomas de <i>Kitasatospora sp.</i> K002 (em cinza) e sua linhagem mais próxima de acordo com TYGS-DSMZ, <i>K. xanthocidica</i> JCM4862 (em azul). O alinhamento e a ilustração foram gerados usando o Plot Nucmer. Apenas os blocos de alinhamento do genoma com pelo menos 1000 nucleotídeos e 90% de identidade são mostrados. A banda externa é a escala de tamanho dos <i>contigs</i> . Quando a banda azul tem duas camadas de profundidade, há duas seqüências na linhagem mais próxima que se sobrepõem no mesmo local em <i>Kitasatospora sp.</i> K002.....	93
Figura 3.7 Árvore filogenética de máxima verossimilhança baseada em 520 genes ortólogos principais concatenados de seqüências genômicas de 12 linhagens de <i>Kitasatospora</i> , incluindo <i>Kitasatospora sp.</i> K002, as linhagens intimamente relacionadas de acordo com o	

TYGS-DSMZ e o grupo externo <i>Streptomyces violaceorubidus</i> . O servidor <i>web</i> M1CR0B1AL1Z3R foi utilizado para concatenar os genes ortólogos e construir a árvore filogenética	94
Figura 3.8 Análise comparativa de pangenoma realizada no Roary. O genoma de <i>Kitasatospora sp.</i> K002 e os 10 genomas mais próximos a ele, de acordo com TYGS-DSMZ, foram utilizadas. Uma identidade percentual mínima para BLASTp de 80% foi definida para agrupar os genes que codificam sequências de proteínas completas em genomas centrais	95
Figura 3.9 Proteínas preditas no genoma de <i>Kitasatospora sp.</i> K002, obtidas a partir do sobrenadante de <i>P. elgii</i> , classificadas de acordo com as categorias COGs. A classificação foi realizada usando o mapeador eggNOG contra o banco de dados da <i>Kyoto Encyclopedia of Genes and Genomes (KEGG)</i>	96
Figura 3.10 Análise de similaridade entre as famílias de <i>clusters</i> de genes usando BiG-SCAPE. Cada união representa <i>clusters</i> gênicos semelhantes entre as linhagens <i>Kitasatospora</i> K002 e <i>K. xanthocidica</i> JCM7862. Os <i>clusters</i> específicos, ou seja, sem semelhanças entre as linhagens também são apresentados em forma única. A representação gráfica dos <i>clusters</i> gênicos semelhantes também é apresentada.	101
Figura 3.11 Comparação entre as predições fenotípicas da montagem híbrida do genoma de <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> , e dos genomas de suas linhagens mais próximas de acordo com o TYGS-DSMZ obtidas no banco de dados do <i>NCBI</i> . O valor zero significa que a linhagem é negativa para o fenótipo, os valores 1 e 2 correspondem ao positivo para o fenótipo nos preditores <i>phypat</i> e <i>phypat+PGL</i> , respectivamente. O valor 3 mostra que o fenótipo foi positivo nos dois preditores. A análise foi realizada pelo <i>software</i> Traitair	104

Lista de Tabelas

Tabela 2.1 Resultados da comparação dos <i>amplicons</i> das linhagens obtidas a partir da utilização do sobrenadante de <i>P. elgii</i> com os bancos de dados do <i>NCBI</i> através do BLAST.	48
Tabela 2.2 Resultados de ensaios <i>overlay</i> de atividade antibacteriana das linhagens obtidas a partir do uso de sobrenadante de <i>P. elgii</i> . Os resultados foram observados após 48 h do inóculo de <i>B. subtilis</i>	49
Tabela 2.3 Propriedades fenotípicas comparativas entre a linhagem <i>Methylobacterium</i> K003, obtida a partir do sobrenadante de <i>P. elgii</i> , e a linhagem de referência <i>M. radiotolerans</i> CBMAI 824. +, Positivo, entre parênteses, crescimento ideal.	50
Tabela 2.4 Composição de ácidos graxos (%) de forma comparativa entre a linhagem <i>Methylobacterium</i> K003, obtida a partir do uso do sobrenadante de <i>P. elgii</i> , e sua linhagem de referência <i>M. radiotolerans</i> CBMAI 824. O perfil de ácidos graxos foi obtido por cromatografia gasosa. *O <i>Summed Feature</i> ocorre quando há separação de pico imperfeita. O <i>Summed Feature 2</i> representa possivelmente aldeído 12:0, o <i>Summed Feature 3</i> representa C16:1 ω6c e/ou C16:1 ω7c e o <i>Summed Feature 8</i> representa C18:1 ω7c e/ou C18:1 ω6c. –, ausente	51
Tabela 2.5 Resultados do teste de suscetibilidade a antibióticos da linhagem <i>Methylobacterium</i> K003, obtida a partir do sobrenadante de <i>P. elgii</i> , e sua linhagem de referência <i>M. radiotolerans</i> CBMAI 824. O método adotado foi o de Kirby-Bauer (difusão em disco), onde são utilizados discos de papel filtro impregnados com uma concentração padrão de antibiótico e observada a formação de um halo de inibição. +, Suscetível; –, Resistente. .	52
Tabela 2.6 Dados da montagem híbrida do genoma de <i>Methylobacterium</i> K003, obtida a partir do sobrenadante de <i>P. elgii</i> . A montagem foi realizada a partir dos sequenciamentos de leitura longa e curta através do programa Unicycler. A anotação foi realizada com o <i>software</i> Prokka e os parâmetros de qualidade foram avaliados nos <i>softwares</i> CheckM, Bowtie2 e Quast	55
Tabela 2.7 Identidade Média de Nucleotídeos (<i>ANI</i>) calculadas para a linhagem <i>M. radiotolerans</i> K003, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . <i>ANIB</i> : <i>ANI</i> com base no cálculo BLAST calculada na plataforma JSpeciesWS. <i>ANIm</i> : <i>ANI</i> com base no cálculo MUMmer calculada na plataforma JSpeciesWS. <i>ANI</i> : calculada no <i>software</i> GTDB-Tk	57
Tabela 2.8 Anotação de genes potenciais responsáveis pela resistência a antibióticos da linhagem <i>M. radiotolerans</i> K003, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . A análise foi realizada em comparação com o Banco de Dados de Genes de Resistência a Antibióticos (<i>ARDB</i>)	61

Tabela 2.9 Resultados da análise da presença de locus CRISPR na linhagem <i>M. radiotolerans</i> K003, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . A análise foi realizada através da ferramenta de reconhecimento CRISPRCasFinder	61
Tabela 2.10 Resultados da análise de <i>clusters</i> de genes biossintéticos para metabólitos secundários da linhagem <i>M. radiotolerans</i> K003, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . A análise foi realizada por meio do <i>software</i> AntiSMASH (v. 6.1.1) e apresentada de forma comparativa com a linhagem <i>M. radiotolerans</i> JCM 2831, acrônimo de CBMAI 824	62
Tabela 3.1 Propriedades fenotípicas que separam <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante <i>P. elgii</i> , das espécies relacionadas do gênero. Linhagens: 1, <i>Kitasatospora</i> K002; 2, <i>K. setalba</i> KM-6054; 3, <i>K. acidiphila</i> MMS16-C NU292T; 4, <i>K. azatica</i> KCTC 9699T; 5, <i>K. kifunensis</i> IFO 15206T; 6, <i>K. purpeofusca</i> NRRL B-1817T; 7, <i>K. nipponensis</i> HKI 0315T e 8, <i>K. humi</i> RB6PN24. +, Positivo, entre parênteses, crescimento ideal, -, sem dados disponíveis.....	85
Tabela 3.2 Resultados do teste de suscetibilidade a antibióticos de <i>Kitasatospora sp.</i> K002, obtida a partir do sobrenadante de <i>P. elgii</i> . O método adotado foi o de Kirby-Bauer (difusão em disco), onde são utilizados discos de papel filtro impregnados com uma concentração padrão de antibiótico e observada a formação de um halo de inibição. +, Suscetível; -, Resistente.....	86
Tabela 3.3 Composição de ácidos graxos (%) de <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> , em comparação a outras espécies do gênero. * <i>Summed Feature</i> ocorre quando há uma separação imperfeita do pico. <i>Summed feature</i> 3 representa C16:1 ω6c e/ou C16:1 ω7c, <i>summed feature</i> 8 representa C18:1 ω7c e/ou C18:1 ω6c, <i>summed feature</i> 9 representa iso-C17:1 w9c e /ou C16 0 10-methyl. Linhagens: 1, K002; 2, <i>K. setalba</i> KM-6054; 3, <i>K. acidiphila</i> MMS16-C NU292T; 4, <i>K. azatica</i> KCTC 9699T; 5, <i>K. kifunensis</i> IFO 15206T; 6, <i>K. purpeofusca</i> NRRL B-1817T and 7, <i>K. nipponensis</i> HKI 0315T	87
Tabela 3.4 Dados da montagem híbrida do genoma de <i>Kitasatospora</i> K002, obtida a partir do sobrenadante de <i>P. elgii</i> . A montagem foi realizada a partir dos sequenciamentos de leitura longa e curta através do programa Unicycler. A anotação foi realizada com o <i>software</i> Prokka e os parâmetros de qualidade foram avaliados nos <i>softwares</i> CheckM, Bowtie2 e Quast	89
Tabela 3.5 Identidade Média de Nucleotídeos (ANI) calculadas para <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . ANI _b : ANI com base no cálculo BLAST calculada na plataforma JSpeciesWS. ANI _m : ANI com base no cálculo MUMmer calculada na plataforma JSpeciesWS. ANI: calculada no <i>software</i> GTDB-Tk	92

Tabela 3.6 Anotação de genes potenciais responsáveis pela resistência a antibióticos de <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . A análise foi realizada em comparação com o Banco de Dados de Genes de Resistência a Antibióticos (ARDB)	97
Tabela 3.7 Resultados da análise da presença de locus CRISPR em <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . A análise foi realizada através da ferramenta de reconhecimento CRISPRCasFinder	98
Table 3.8 Resultados da análise de <i>clusters</i> de genes biossintéticos para metabólitos secundários de <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . A análise foi realizada por meio do <i>software</i> AntiSMASH (v. 6.1.1) e apresentada de forma comparativa com outras linhagens do gênero. Linhagens: 1, <i>Kitasatospora</i> K002; 2, <i>K. xanthocidica</i> JCM4862; 3, <i>K. acidiphila</i> MMS16-CNU292T; 4, <i>Kitasatospora sp.</i> MMS16-BH015	100
Tabela 3.9 Resultados da análise de genes envolvidos na biossíntese de peptídeos de <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> . A análise foi realizada por meio do servidor Bagel4	102
Tabela 3.10 Comparação entre as predições fenotípicas da montagem híbrida do genoma de <i>Kitasatospora sp.</i> K002, obtida a partir do uso do sobrenadante de <i>P. elgii</i> , e dos genomas de outras linhagens do gênero. Linhagens: 1, <i>Kitasatospora</i> K002; 2, <i>K. acidiphila</i> MMS16-CNU292T; 3, <i>K. azatica</i> KCTC 9699T; 4, <i>K. kifunensis</i> IFO 15206T; 5, <i>K. purpeofusca</i> NRRL B-1817T e 6, <i>K. nipponensis</i> HKI 0315T. +, Positivo; -, negativo, w, fraco	105

Glossário

°C – Grau Celsius

μl – Microlitro, unidade de medida de volume

μm – Micrômetro, unidade de medida de comprimento

AMP – *Antimicrobial peptides* (peptídeos antimicrobianos)

ANI – *Average Nucleotide Identity* (Identidade Média de Nucleotídeos)

ARDB – *Antibiotic Resistance Genes Database*

BGC – *Biosynthetic gene clusters* (Aglomerados de genes bissintéticos)

Cas – proteínas adjacentes associadas ao CRISPR

CBMAI – Coleção Brasileira de Micro-organismos de Ambiente e Indústria

CDS – *Coding Sequences* (Sequências Codificadoras)

cm – centímetro, unidade de medida de comprimento

COG – *Clusters of Orthologous Groups* (Grupos de Ortólogos)

CRISPR – matrizes agrupadas de repetições palindrômicas curtas regularmente interespaçadas

CRISPR-Cas – matrizes agrupadas de repetições palindrômicas curtas regularmente interespaçadas (CRISPR) e proteínas Cas adjacentes associadas ao CRISPR

dDDH – *Digital DNA-DNA hybridization* (Hibridização digital DNA-DNA)

DNA – *Deoxyribonucleic acid* (Ácido Desoxirribonucleico)

ESKAPE – acrônimo que compreende os nomes científicos de seis patógenos bacterianos altamente virulentos e resistentes a antibióticos

g – Grama, unidade de medida de massa

GC(%) – Conteúdo GC, proporção de guanina e citosina em sequências de DNA

GBDP – *Genome Blast Distance Phylogeny* (Distância filogenética entre genomas)

ha – hectare

JGI – *Joint Genome Institute*

KEGG – *Kyoto Encyclopedia of Genes and Genomes* (Enciclopédia de Kyoto de Genes e Genomas)

L – Litro, unidade de medida de volume

LPSN – *List of Prokaryotic names with Standing in Nomenclature* (Lista de Nomes Procarióticos com Base na Nomenclatura)

MEV – Microscopia Eletrônica de Varredura

MET – Microscopia Eletrônica de Transmissão

ml – Mililitro, unidade de medida de volume

mm – Milímetro, unidade de medida de comprimento

NCBI – *National Center for Biotechnology Information* (Centro Nacional de Informação Biotecnológica dos Estados Unidos)

NRPS – *Nonribosomal peptide synthase* (enzima sintase peptídica não ribossal)

GenBank – Base de dados do National Center for Biotechnology Information (Centro Nacional de Informação Biotecnológica dos Estados Unidos)

NGS Illumina – *Next Generation Sequencing* (Sequenciamento de nova geração em plataforma Illumina)

OGRI – *Overall Genome Related Index*

OMS – Organização Mundial de Saúde

ONT – *Oxford Nanopore Technologies*

p/v – Relação peso/volume

pb – Pares de bases

PCR – *Polymerase chain reaction* (Reação em Cadeia da Polimerase)

PKS – *Polyketide synthases* (enzima sintase policetílica)

RDP – Ribosomal Database Project da Michigan State University - USA

RIPP – *Ribosomally synthesized and post-translationally modified peptide* (peptídeos sintetizados ribossomalmente e modificados pós-traducionalmente)

RNA – *Ribonucleic Acid* (Ácido ribonucleico)

rRNA – *Ribosomal-RNA* (RNA ribossomal)

SSU RNA – *small subunit ribosomal*

tmRNA – *Transfer-messenger-RNA* (RNA transferência)

tRNA – *Transfer-RNA* (RNA transportador)

Apresentação

O presente trabalho visa a prospecção de linhagens bacterianas com potencial aplicação biotecnológica, com foco em compostos antibacterianos. O estudo faz uso de técnicas de microbiologia básica, biologia molecular, sequenciamento genético e análises de bioinformática para alcançar tal objetivo.

Este estudo está organizado em três capítulos. O primeiro capítulo apresenta uma revisão da literatura. O segundo descreve a metodologia para a prospecção de linhagens bacterianas de solo do Cerrado com potencial para a produção de compostos antibacterianos, a partir do sobrenadante de *Paenibacillus elgii* e a caracterização de uma das linhagens bacterianas selecionadas. O terceiro capítulo apresenta a caracterização taxonômica polifásica de uma das linhagens bacterianas selecionadas, identificada como uma nova espécie.

Em anexo seguem os *drafts* dos artigos que estão em fase de revisão para submissão, o artigo publicado em complemento à tese e demais atividades relevantes.

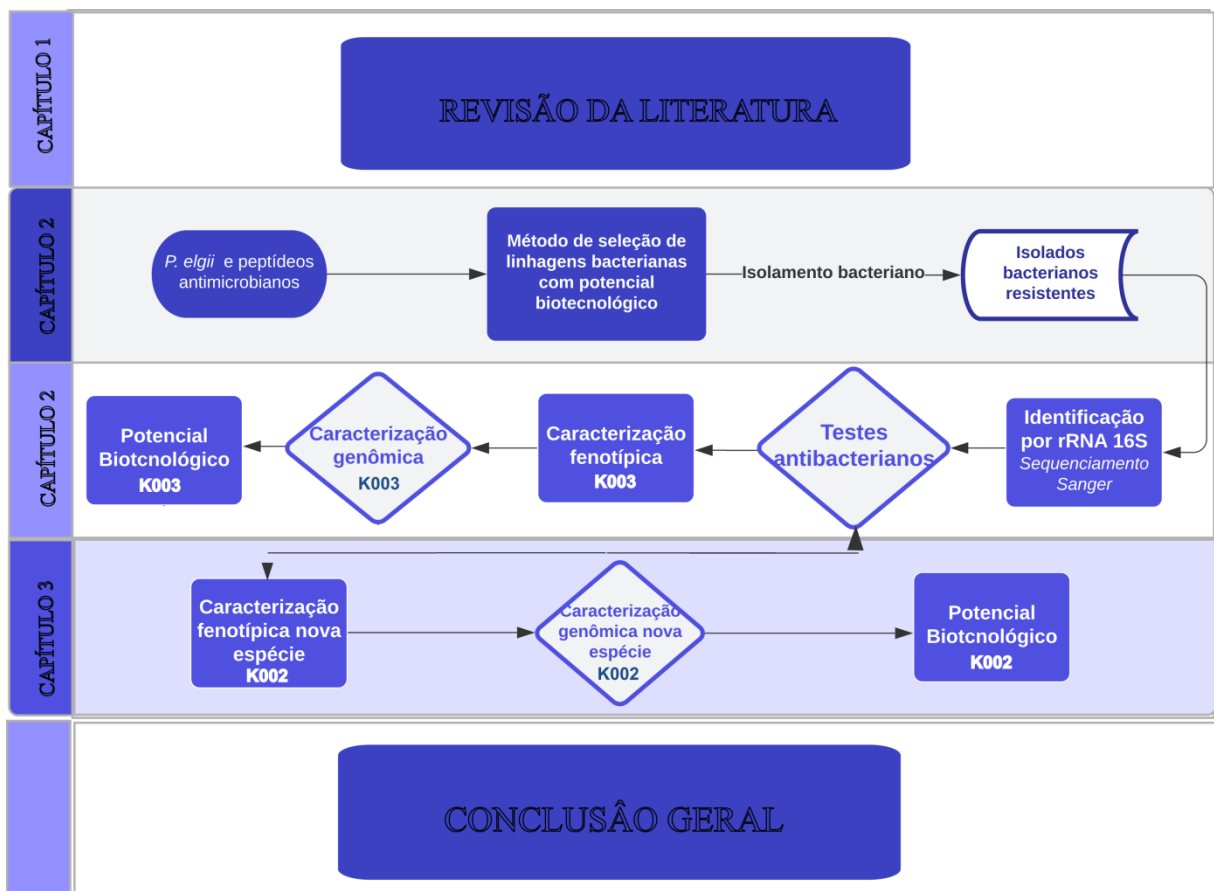


Figura 1.1 Fluxograma das principais etapas deste trabalho apresentado em capítulos.

SUMÁRIO

Agradecimentos	II
Resumo	III
Abstract	IV
Índice de figuras	V
Índice de tabelas	IX
Glossário	XII
Apresentação	XIV
CAPÍTULO 1 - FUNDAMENTAÇÃO TEÓRICA	18
1.1 INTRODUÇÃO GERAL	18
1.2 REVISÃO BIBLIOGRÁFICA	20
1.2.1 Micro-organismos e a bioprospecção.....	20
1.2.2 Importância dos antibióticos	21
1.2.3 Técnicas para obtenção de micro-organismos com potencial biotecnológico.....	22
1.2.4 <i>Paenibacillus elgii</i>	24
1.2.5 Peptídeos Antimicrobianos (AMP).....	25
1.2.6 Metabólitos Secundários.....	27
1.2.7 Tecnologias de sequenciamento genético e análise dos dados.....	28
1.2.8 Análises filogenéticas em procariotos.....	30
1.2.9 Anotação funcional de genomas	32
CAPÍTULO 2 - USO DO SOBRENANDANTE DE <i>PAENIBACILLUS ELGII</i> NA PROSPECÇÃO E SELEÇÃO DE LINHAGENS BACTERIANAS COM POTENCIAL PARA APLICAÇÃO BIOTECNOLÓGICA	34
2.1 INTRODUÇÃO	34
2.2 OBJETIVOS	36
2.2.1 Objetivo Geral	36
2.2.2 Objetivos Específicos	36
2.3 MATERIAL E MÉTODOS	37
2.3.1 Obtenção do sobrenadante de <i>Paenibacillus elgii</i>	37
2.3.2 Preparação dos meios de cultura enriquecidos com sobrenadante de <i>P. elgii</i>	37
2.3.3 Coleta e processamento de amostras de solo	37
2.3.4 Isolamento das linhagens bacterianas	38
2.3.5 Extração de DNA e amplificação do gene ribossomal rRNA 16S	38
2.3.6 Análise da atividade antibacteriana	39

2.3.7 Caracterização fenotípica <i>in vitro</i>	39
2.3.8 Sequenciamentos, montagem e análise do genoma	41
2.3.9 Análises filogenéticas	42
2.3.10 Predição de genes e anotação funcional	44
2.3.11 Caracterização fenotípica <i>in silico</i>	44
2.4 RESULTADOS	46
2.4.1 Uso do sobrenadante de <i>Paenibacillus elgii</i>	46
2.4.2 Identificação taxonômica das linhagens bacterianas	47
2.4.3 Atividade antibacteriana	49
2.4.4 Caracterização fenotípica <i>in vitro</i>	50
2.4.5 Sequenciamentos e análises filogenéticas	52
2.4.6 Características funcionais do genomas	59
2.4.7 Caracterização fenotípica <i>in silico</i>	63
2.5 DISCUSSÃO	66
2.6 CONCLUSÕES	72
CAPÍTULO 3 - CARACTERIZAÇÃO GENÔMICA E FISIOLÓGICA DA LINHAGEM <i>KITASATOSPORA SP.</i> OBTIDA A PARTIR DO USO DO SOBRENADANTE DE <i>PAENIBACILLUS ELGII</i> COMO MECANISMO DE SELEÇÃO.....	73
3.1 INTRODUÇÃO	73
3.2 OBJETIVOS	77
3.2.1 Objetivo Geral	77
3.2.2 Objetivos Específicos	77
3.3 MATERIAL E MÉTODOS	78
3.3.1 Caracterização fenotípica <i>in vitro</i>	78
3.3.2 Sequenciamentos, montagem e análise do genoma	79
3.3.3 Análises filogenéticas.....	81
3.3.4 Predição de genes e anotação funcional	82
3.3.5 Caracterização fenotípica <i>in silico</i>	83
3.4 RESULTADOS	84
3.4.1 Caracterização fenotípica <i>in vitro</i>	84
3.4.2 Sequenciamentos e análises filogenéticas	88
3.4.3 Características funcionais do genoma	95
3.4.4 Caracterização fenotípica <i>in silico</i>	102

3.5 DISCUSSÃO	106
3.6 CONCLUSÕES	112
4 CONCLUSÃO GERAL	113
5 REFERÊNCIAS BIBLIOGRÁFICAS	114
6 ANEXOS	135
6.1 <i>Draft</i> do artigo do Capítulo 2.....	135
6.2 <i>Draft</i> do artigo do Capítulo 3.....	178
6.3 Artigo completo publicado em periódico	219
6.4 Resumos publicados em anais de eventos	220
6.5 Participação em eventos	220
6.6 Capítulo de livro a ser publicado	220
6.7 Organização de evento	220
6.8 Atividades relevantes	221

CAPÍTULO 1 - FUNDAMENTAÇÃO TEÓRICA

1.1 INTRODUÇÃO GERAL

O Brasil conta com uma diversidade de ecossistemas que são ambientes favoráveis a prospecção de compostos bioativos. Entretanto, uma parcela significativa dos micro-organismos permanece inexplorada. Análises baseadas na avaliação do gene marcador *rRNA* 16S indicam a existência de novas espécies de bactérias nos ecossistemas, sugerindo o alto potencial biotecnológico a ser descoberto a partir dos metabólitos e enzimas desses micro-organismos (Gutleben *et al.*, 2018). Assim, há a necessidade de avanços na busca por novas tecnologias direcionadas à obtenção de micro-organismos produtores de antimicrobianos e outras atividades de interesse, visto que as técnicas clássicas de cultivo permitem o isolamento, na maioria das vezes, apenas de micro-organismos já conhecidos. (Overmann, 2013; Selvin *et al.*, 2012).

Nos últimos anos, o campo dos produtos naturais recebeu atenção com a descoberta, a partir da informação genômica, de que assim como a diversidade bacteriana ainda não foi totalmente descoberta, o repertório de compostos bioativos estabelecidos para os micro-organismos já conhecidos também não (Bachmann *et al.*, 2014). Assim, os potenciais antimicrobianos de diversos compostos produzidos por bactérias já descritas vêm sendo investigados, incluindo os peptídeos antimicrobianos, sejam eles de síntese ribossomal ou não. As bacteriocinas, por exemplo, representam uma alternativa potencial para atender à demanda por agentes antimicrobianos efetivos contra patógenos bacterianos. Elas têm sido utilizadas em várias aplicações, sendo úteis para prolongar a vida de prateleira de alimentos, controlar patógenos humanos e animais, além de modular processos fermentativos (Grady *et al.*, 2016).

O desenvolvimento de novos antimicrobianos exige uma série de etapas, onde a primeira é a seleção e caracterização de micro-organismos para detectar potenciais fontes de novos metabólitos bioativos. Devido ao um grande potencial econômico e científico em bioprodutos de origem bacteriana são necessários estudos de bioprospecção para que sejam fornecidas informações sobre compostos e suas linhagens produtoras. Esse conhecimento é necessário para as etapas seguintes na busca pelos antimicrobianos como caracterização e identificação dos compostos bioativos, otimização das condições de cultura, produção em larga escala e ensaios clínicos (Amin *et al.*, 2020).

As bactérias são organismos que se adaptam ao longo do tempo. Para isso, elas se ajustam ao ambiente e evoluem de forma a garantir sua existência (MacGowan e Macnaughton, 2017). Se algo interrompe sua capacidade de crescimento, como um

antibiótico, variações genéticas podem ser selecionadas e compartilhadas, gerando adaptações que permitem a sua sobrevivência (Munita e Arias, 2016). É o processo natural das bactérias desenvolverem resistência e se tornarem imunes aos medicamentos.

A resistência a antibióticos desenvolvida pelas bactérias tem sido uma ameaça crescente para o tratamento eficaz de infecções comuns causadas por esses micro-organismos (Dadgostar, 2019). Ela se agrava devido ao uso excessivo ou desnecessário de antibióticos, seja por diagnóstico inexato, prescrição inadequada ou automedicação (Chokshi *et al.*, 2019). Há também uma preocupação com o consumo mundial de fármacos antimicrobianos utilizados em animais para consumo humano. Os alimentos de origem animal, bem como a interação entre animais e seres humanos, têm sido considerados uma via importante de exposição dos seres humanos a bactérias resistentes, implicando em mais uma forma de propagação da resistência a antibióticos (Toutain *et al.*, 2016).

Segundo a Organização Mundial da Saúde (OMS, 2021), a resistência microbiana é um problema de saúde global. Os processos de resistência bacteriana juntamente com as informações relativas à toxicidade dos antibióticos já conhecidos ou em desenvolvimento levam a necessidade de busca de novos compostos que contornem esses problemas (Uddin *et al.*, 2021). Assim, uma das principais demandas biotecnológicas da indústria por compostos bioativos refere-se à produção de compostos antimicrobianos (Laws *et al.*, 2019).

Há a necessidade de descoberta de novos antibacterianos que substituam aqueles que se tornaram ineficazes e neutralizem microorganismos resistentes (Uddin *et al.*, 2021). Os micro-organismos, para sua sobrevivência, produzem pequenas moléculas conhecidas como metabólitos secundários ou produtos naturais que servem, entre outras funções, para suprimir organismos concorrentes. Essa capacidade de secretar compostos faz deles uma fonte potencial de antibacterianos (Challinor e Bode, 2015).

Diante do exposto, este trabalho visa a prospecção e caracterização de bactérias de solo do Cerrado com relevante potencial para aplicação biotecnológica utilizando o sobrenadante de *Paenibacillus elgii* como mecanismo de seleção. O presente estudo foi desenvolvido em continuidade ao trabalho de Carvalho (2013).

1.2 REVISÃO BIBLIOGRÁFICA

1.2.1 Micro-organismos e a bioprospecção

Devido à sua diversidade genômica e, portanto, metabólica e funcional, os micro-organismos servem como modelos ideais para estudos de biotecnologia. Muitos dos processos biotecnológicos complexos são possíveis graças a bactérias que desempenham tais tarefas de forma eficiente. Grande parte dos antimicrobianos descritos tem sua origem em procariotos que os produzem por competição no ambiente ou até mesmo comunicação com outros micro-organismos (Durand *et al.*, 2018).

Um dos principais obstáculos nos estudos de micro-organismos e na descoberta de novos compostos produzidos por eles é a incapacidade de se cultivar a maior parte da diversidade microbiana presente nos ecossistemas (Overmann *et al.*, 2017). Essa limitação de cultivo foi observada já em 1985, quando verificou-se divergência entre o número de células de bactérias que formavam colônias em placas e a contagem de células obtidas por observação microscópica. Tal constatação ficou conhecida como “grande anomalia da contagem em placa” (Staley e Konopka, 1985).

Após a década de 1970, graças às técnicas moleculares de identificação pelo gene marcador *rRNA* 16S de diferentes amostras ambientais, foi comprovada a presença de vários grupos de bactérias até então desconhecidas (Foster, 2017). Confirmou-se então, que a maior parte dos micro-organismos ainda não havia sido cultivada pelas metodologias clássicas de cultivo. Atualmente, graças à metagenômica, análise da diversidade microbiana através do *DNA* de uma amostra, e às técnicas de sequenciamento de alto rendimento, a diversidade bacteriana tem sido elucidada (Foster, 2017).

Estudos metagenômicos demonstraram que 99% das bactérias de solos ainda não foram cultivadas pelos métodos convencionais (Lloyd *et al.*, 2018). Desta forma, a abordagem metagenômica tem sido uma alternativa para garantir o acesso aos genomas desses ecossistemas (Handelsman, 2004; Foster, 2017). Esse campo tem sido um dos mais promissores para a descoberta de novas drogas, permitindo o acesso a *clusters* gênicos responsáveis pela produção de compostos bioativos que podem então ser introduzidos em micro-organismos de melhor cultivo em laboratório (Hua *et al.*, 2015).

A abordagem metagenômica tem sido uma alternativa para garantir o acesso aos genomas dos ecossistemas independentemente de cultivo (Handelsman, 2004; Foster, 2017). Porém o cultivo ainda assim é essencial, uma vez que previsões metagenômicas dependem amplamente da disponibilidade de genomas bem anotados de representantes cultivados. Logo,

o cultivo continuará sendo uma abordagem indispensável para a elucidação da microbiologia em geral (Choi *et al.*, 2017; Liu *et al.*, 2022; Ling *et al.*, 2015; Nowrotek *et al.*, 2019).

A recuperação de genomas a partir de metagenomas também pode disponibilizar dados para as previsões metagenômicas, bem como favorecer a descoberta de novos compostos bioativos como antibacterianos. Entretanto, trata-se de uma tarefa desafiadora em termos de bioinformática por ser influenciada por vários fatores como diversidade da comunidade microbiana, esforço de sequenciamento, comprimento de leitura de sequência e os algoritmos de montagem e binning usado (Meziti *et al.*, 2019; 2021).

1.2.2 Importância dos antibióticos

O primeiro antibiótico descoberto foi a penicilina, em 1928, acidentalmente. Alexander Fleming observou que algumas culturas de bactérias haviam sido contaminadas por um bolor e ao redor dele não havia mais bactérias. Descobriu-se que o bolor era um fungo do gênero *Penicillium spp.* que produzia uma substância com efeito bactericida, nomeada como penicilina. O fungo foi inicialmente identificado como *Penicillium notatum* e mais tarde foi chamado de *Penicillium chrysogenum*. A partir daí, milhares de compostos antimicrobianos foram descobertos, descritos e comercializados (Tan e Tatsumura, 2015).

Na década de 1940, os antibióticos eram definidos como substâncias produzidas por micro-organismos, capazes de inibir o crescimento de outros ou até destruí-los (da Cunha *et al.*, 2019). Desde então, muitos dos antibióticos utilizados atualmente são derivados de compostos produzidos por micro-organismos (Durand *et al.*, 2019).

A maioria dos antibióticos foi descoberta na década de 1950 como tetraciclina, cefalosporinas, aminoglicosídeos e macrolídeos (Nicolau e Rigol, 2018). Nas décadas posteriores, o gênero *Streptomyces* deu origem a maioria dos compostos antimicrobianos como actinomicina (derivada de *Streptomyces spp.*), neomicina (derivada de *Streptomyces fradiae*) e estreptomicina (derivada de *Streptomyces griseus*) que são utilizados até hoje (da Cunha *et al.*, 2019). Dentre as formas de atuação desses compostos sobre as bactérias, geralmente, eles são classificados em: interferência na síntese da parede celular, inibição da síntese de proteínas, inibição da síntese de ácidos nucleicos, inibição de vias metabólicas e rompimento da membrana. Além disso, os antibióticos podem ser subclassificados em bactericidas, quando matam o micro-organismo, e outros bacteriostáticos, quando interrompem o crescimento, este último podendo ser reversível após degradação (Kapoor *et al.*, 2018).

Em 2017, a Organização Mundial da Saúde declarou a ocorrência de resistência antimicrobiana em todo o mundo como um grande desafio para a saúde pública, o que causa a necessidade urgente de busca por novos antibióticos (Tillotson e Zinner, 2017). Nos primórdios do uso de antibióticos, também existiam bactérias resistentes, mas um fluxo constante de antibióticos experimentais oferecia terapias alternativas (Dodds *et al.*, 2017). No entanto, os antibióticos deixaram de chegar em maior número na década de 1980. A última vez que uma nova classe de antibióticos foi descoberta e lançada no mercado foi em 1987 e o último grupo de agentes de amplo espectro, as fluoroquinolonas, também foi descoberto na década de 1980 (Durand *et al.*, 2019). Desde então, tem havido uma escassez na área e atualmente existem apenas alguns novos grupos de antibióticos em desenvolvimento capazes de combater os níveis atuais de resistência antimicrobiana (Frieri *et al.*, 2017).

O surgimento da multirresistência entre as bactérias conhecidas como ESKAPE (acrônimo que compreende os nomes científicos de seis patógenos bacterianos altamente virulentos e resistentes a antibióticos) e, mais recentemente, em *Mycobacterium tuberculosis*, extremamente resistente, é uma grande ameaça à saúde pública mundial (Vajs *et al.*, 2017). Diante disso, explorar o potencial biossintético de espécies bacterianas pode ser uma alternativa para revelar novas estruturas com atividades biológicas eficazes contra organismos resistentes (Amin *et al.*, 2020).

1.2.3 Técnicas para obtenção de micro-organismos com potencial biotecnológico

A maioria das abordagens de cultivo tradicionais foi empírica, empregando diferentes meios e condições de incubação, monitorando o sucesso do cultivo e subcultivando espécies bacterianas de interesse. Isso é particularmente eficiente se as condições de cultivo atenderem aos requisitos das bactérias-alvo. Entretanto, cultivos direcionados a micro-organismos ainda não descritos, não cultivados e de interesse biotecnológico se fazem necessários para impulsionar o desenvolvimento da ecologia microbiana, biotecnologia e saúde pública (Overmann *et al.*, 2017).

Durante anos, a comunidade científica considerou os micro-organismos apenas como seres individuais autossuficientes, com auto-replicação descomplicada e incapazes de se comunicar ou se organizar em grupos (Njoroge e Sperandio, 2009). No entanto, sabe-se atualmente que bactérias e outros procariotos possuem um sistema de comunicação realizado por meio de diferentes processos que envolvem produção, liberação, detecção e respostas a compostos químicos chamados autoindutores. Essas pequenas moléculas semelhantes a

hormônios, permitem ao micro-organismo acessar a densidade celular de uma população local e coordenar a regulação gênica em uma população (Gray *et al.*, 2013).

Os sistemas de comunicação célula a célula denominados *quorum sensing* envolvem a formação dependente da densidade de moléculas autoindutoras de acil-homoserina lactona (AHL) ou peptídeos curtos e controla a produção de luz, expressão de fatores de virulência, formação de biofilme, agregação celular e competência genética. Interações metabólicas conhecidas também envolvem compostos participando do metabolismo central, como por exemplo, sideróforos (Papenfort e Bassler, 2016; Soares, 2022).

Em um cultivo tradicional, bactérias são normalmente cultivadas isoladamente com amplos nutrientes e recursos, o que contrasta fortemente com seu habitat natural complexo (Ho *et al.*, 2017). Matrizes complexas, como o solo, abrigam micro-organismos diversos que não apenas participam de importantes ciclos biogeoquímicos, mas também criam condições que possibilitam o crescimento de outros micro-organismos por compartilhamento metabólitos e substâncias essenciais para o crescimento desses organismos (Bahram *et al.*, 2018; Schink, 2002). Essas moléculas incluem aquelas que desempenham um papel na detecção de *quorum* e na comunidade (Jacoby *et al.*, 2017). A ausência de vizinhos na cultura pura e, conseqüentemente, a ausência dos metabólitos necessários, reduz a recuperação microbiana (Pande e Kost, 2017).

Assim, a maior parte dos micro-organismos do solo resiste ao cultivo em condições de laboratório e isso representa uma enorme lacuna de diversidade bioquímica. Para desbloquear esse potencial, precisamos simular melhor o contexto ecológico em que vivem as bactérias, pois isso pode permitir que acessemos essa fonte de produtos naturais ainda inexplorada (Kapinusova *et al.*, 2023).

Diante disso, há a necessidade do refinamento das estratégias existentes para o cultivo de micro-organismos, bem como o desenvolvimento de novas técnicas. Métodos que sejam mais eficazes na obtenção de micro-organismos até então não cultivados ou de interesse biotecnológico. Aprimoramentos em técnicas de cultivo podem incluir: uso de meios de cultura restritivos e uso de antimicrobianos na seleção de crescimento e assim, enriquecimentos específicos podem ser utilizados para construir coleções substancialmente maiores de micro-organismos de interesse (Bachmann *et al.*, 2014; Overmann, 2013). Outras possibilidades no melhoramento de técnicas de cultivo incluem o uso de co-culturas, câmara de difusão, sistema de membranas e biorreatores (Boilattabi *et al.*, 2021; Pascual *et al.*, 2017; van Dorst *et al.*, 2016; Chaudhary *et al.*, 2019).

O esforço necessário para enriquecer, isolar e manter novas linhagens de bactérias, muitas vezes é bastante exigente. Todavia, isso pode ser minimizado com informações sobre a fisiologia e bioquímica do organismo obtidas através de estudos ecológicos, genômicos, metagenômicos e fisiológicos independentes da cultura (Overmann *et al.*, 2017). Grandes avanços nas técnicas de sequenciamento permitem o acesso a informações sobre linhagens bacterianas por meio de suas sequências genômicas e ferramentas que prevêm as características de bactérias com base em seus genomas já estão disponíveis (Weimann *et al.*, 2016).

Espera-se que a diversidade bacteriana desconhecida ofereça novas alternativas no que diz respeito a prospecção de compostos bioativos de interesse farmacológico (Overmann *et al.*, 2017). O Cerrado, por sua grande extensão (2.000.000 km²), possui uma enorme biodiversidade em potencial (Ribeiro e Walter, 1998; Silva *et al.*, 2013). Diversos estudos já relataram o potencial biotecnológico proveniente do solo do Cerrado. Takahashi e colaboradores (2008) e Petite colaboradores (2009) identificaram fungos potenciais produtores de antibióticos, Peixoto e colaboradores (2016) descobriram bactérias degradadoras de polietileno, Da Cruz e colaboradores (2013) que detectaram potencial anti-inflamatório em actinobactérias e Da Costa e colaboradores (2022) caracterizaram uma linhagem de *Paenibacillus elgii* isolada de solo do Cerrado com atividade antimicrobiana.

1.2.4 *Paenibacillus elgii*

O gênero *Paenibacillus* faz parte da família *Paenibacillaceae*, ordem *Bacillales*, classe *Bacilli* e filo *Firmicutes*. O gênero era originalmente classificado como *Bacillus*, mas, graças a comparações de sequências de *rRNA* 16S, foi proposta a criação de um novo gênero, o *Paenibacillus* (Ash *et al.*, 1993).

Mais de 300 espécies para este gênero são listadas em *List of Prokaryotic names with Standing in Nomenclature* (<https://lpsn.dsmz.de/genus/kitasatospora> em 18/01/2023), encontradas nos mais variados ambientes como solo, leite, húmus e líquido cefalorraquidiano (Chowet *et al.*, 2012; Elo *et al.*, 2001; Scheldeman, 2004 e Wu *et al.*, 2011).

Vários trabalhos demonstram que espécies do gênero *Paenibacillus* possuem capacidade de produzir diferentes compostos de interesse biotecnológico, tais como fungicidas, enzimas hidrolíticas e larvicidas (Anandaraj *et al.*, 2009; Ding *et al.*, 2011; Chow *et al.*, 2012). Diversos antimicrobianos produzidos por linhagens de *Paenibacillus* foram

descobertos como antibióticos peptídicos incluindo polimixinas, jolipeptina, gavaserina, saltavalina, fusaricidina A–D e gatavalina (Li *et al.*, 2007; Wu *et al.*, 2011).

Bactérias formadoras de endosporos como *Paenibacillus* são bem conhecidas por sua capacidade de produzir antibióticos lipopeptídeos (Wang e Liu, 2008; Chen *et al.*, 2010). Os lipopeptídeos são peptídeos cíclicos ou lineares com um ácido graxo variável em sua região N-terminal produzidos pela enzima sintase não ribossomal (NRPS) (Cochrane e Vederas, 2016).

Dentre as bactérias pertencentes a esse gênero, destaca-se a *Paenibacillus elgii* que apresenta potentes antimicrobianos já explorados para o uso veterinário (Krüger *et al.*, 2013). Uma das características observadas em *P. elgii* foi a atividade antimicrobiana promovida por lipopeptídeos e por bacteriocinas denominados para essa espécie como pelgipeptinas (Da Costa *et al.*, 2022) e elgicinas (Teng *et al.*, 2012), respectivamente. Diferentes linhagens de *P. elgii* foram identificadas como produtoras de compostos que tem como características serem formados por peptídeos associados a diferentes grupos orgânicos como lipídeos (Teng *et al.*, 2012).

1.2.5 Peptídeos Antimicrobianos (AMP)

Os peptídeos antimicrobianos (AMPs) podem ser sintetizados por duas vias: a síntese ribossomal, como os lantibióticos e pediocinas, e síntese não ribossomal, como os lipopeptídeos não ribossomais, pela qual aminoácidos são incorporados independente de RNA mensageiro por sintases peptídicas não ribossomais (NRPS) (Grady *et al.*, 2016).

As bacteriocinas são peptídeos antimicrobianos sintetizados ribossomalmente e bem diversos em sua composição química, estrutural e funcional (Willey e Van Der Donk, 2007). Elas estão distribuídas entre espécies de bactérias e *Archaeas* (Morton *et al.*, 2015) e diferem em tamanho, sequência de aminoácidos e mecanismos de ação. Podem exibir atividade bactericida ou bacteriostática, formando poros em membranas de células sensíveis, inibindo a polimerização do peptidoglicano ou a síntese de proteínas e de ácidos nucleicos (Paiva *et al.*, 2011).

As estruturas das bacteriocinas revelaram muitos peptídeos que podem ser agrupados na classe de polipeptídeos modificados após a tradução. Esses tem origem como peptídeos precursores e posteriormente sofrem alterações que resultam em suas formas biologicamente ativas (Rea *et al.*, 2011). Normalmente, as bacteriocinas são sintetizadas como um pré-peptídeo biologicamente inativo contendo uma sequência líder na região N-terminal. A sequência líder auxilia o reconhecimento do pré-peptídeo pelas enzimas de maturação e

transporte e protege a estirpe produtora, mantendo a bacteriocina inativa até ser secretada no meio extracelular (van der Meer *et al.*, 1994, Perez *et al.*, 2014). Embora as bacteriocinas sejam estruturalmente diversas, os genes necessários para a biossíntese são geralmente conservados entre diferentes espécies de bactérias (Morton *et al.*, 2015).

Diferentes sistemas de classificação têm sido propostos para bacteriocinas ao longo dos anos. Esses sistemas foram modificados muitas vezes com base em novos desenvolvimentos em relação às estruturas e modos de ação das bacteriocinas (Lagha *et al.*, 2017). De acordo com a classificação delineada por Cotter e colaboradores (2013), que tem as vantagens de ser clara e simples e aplicar-se a bacteriocinas de bactérias Gram-positivas e Gram-negativas, as bacteriocinas foram categorizadas principalmente em duas classes fundamentais com base na presença ou não de modificações após a tradução, Classe I que reúne as bacteriocinas modificadas e classe II, as não modificadas. Além disso, esta classificação considerou apenas peptídeos antimicrobianos e não proteínas antimicrobianas maiores, como as colicinas de *Escherichia coli*. A Classe I é subdividida seguindo a nomenclatura RiPP (Arnison *et al.*, 2013). Inclui lanthipeptídeos, sactipeptídeos, peptídeos circulares e glicocinas de bactérias Gram-positivas (Cotter *et al.*, 2013), peptídeos lineares contendo azol(ina) (LAP) e peptídeos laço de bactérias Gram-positivas e Gram-negativas, e peptídeos nucleotídicos e peptídeos sideróforos de Gram-negativos (Duquesne *et al.*, 2007; Cotter *et al.*, 2013; Gabrielsen *et al.*, 2014). Além disso, linaridinas (Claesen e Bibb, 2010) e tiopeptídeos (Bagley *et al.*, 2005) produzidas essencialmente por Actinobactérias, e cianobactinas de diversas cianobactérias (Jaspars, 2014; Martins e Vasconcelos, 2015; Martins *et al.*, 2018), pertencem a esta classe. As bacteriocinas de classe II, por sua vez, são subdivididas em três classes: bacteriocinas semelhantes à pediocina, bacteriocinas não semelhantes à pediocina e bacteriocinas de dois peptídeos (Soltani *et al.*, 2021).

As elgicinas de *P. elgii* são bacteriocinas modificadas de Classe I, também chamadas de lantibióticos. Tratam-se de quatro novos peptídeos (elgicinas AI, AII, B e C) de pesos moleculares pequenos, com um amplo espectro inibitório. Essas bacteriocinas apresentaram atividade antibacteriana comprovada contra *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Escherichia coli* e *Proteus vulgaris* e parecem ser agentes antibacterianos promissores (Teng *et al.*, 2012).

Os lipopeptídeos de síntese não ribossomal atuam principalmente nas membranas citoplasmáticas e podem ser classificados como lineares catiônicos, cíclicos catiônicos ou não catiônicos (Grady *et al.*, 2016). Logo, variações ocorrem devido a síntese de cada família de lipopeptídeos ser realizada por diferentes NRPS, que é uma grande enzima modular, onde

cada módulo catalisa a adição de um único substrato específico no produto (Ding *et al.*, 2011; Finking e Marhiel, 2004; Qian *et al.*, 2012). Portanto, elas tendem a ser grandes complexos enzimáticos em consequência do tamanho e complexidade do produto (González *et al.*, 2016).

A grande diversidade estrutural apresentada pelos lipopeptídeos sugere diferentes funções naturais, podendo ser exclusivos para a fisiologia do micro-organismo. As principais funções descritas envolvem questões de antagonismo microbiano, motilidade, aderência a superfícies e sinalizadores para crescimento e diferenciação (Raaijmakers *et al.*, 2010).

Algumas investigações das finalidades fisiológicas dos lipopeptídeos para os micro-organismos produtores levantam hipóteses envolvendo sistemas regulatórios como o *quorum sensing*, que são metabólitos geralmente detectados na fase de crescimento exponencial ou estacionário tardio, sendo normalmente responsivos à densidade celular. Outro sistema regulatório são os transcricionais, regulando genes de biossíntese (Raaijmakers *et al.*, 2010)

As pelgipeptinas são lipopeptídeos descritos pela primeira vez em *P. elgii* B69. Elas são sintetizadas por sintases peptídicas não ribossomais (NRPSs) codificadas pelo *cluster* gênico *plp* identificado e caracterizado em *P. elgii* B69 (Qian *et al.*, 2012). Esses lipopeptídeos são conhecidos por sua atividade antimicrobiana de amplo espectro contra várias bactérias e fungos do solo, além de possuir ação larvicida (Ding *et al.*, 2011; Wu *et al.*, 2010; Da Costa *et al.*, 2023).

1.2.6 Metabólitos Secundários

Os metabólitos secundários de origem microbiana têm proporcionado uma grande quantidade de moléculas de interesse farmacêutico e outros campos na biotecnologia. São compostos de baixo peso molecular que geralmente exibem atividade biológica (Bérdy, 2012).

Há um debate contínuo sobre quais seriam melhores fontes de produtos naturais através da prospecção genômica de micro-organismos (Bachmann *et al.*, 2014). Considerando-se o número de genes das enzimas sintase policetílica tipo I (*PKS*) e sintase peptídica não ribossomal (*NRPS*) para medir o potencial de produzir metabólitos secundários, fica claro que micro-organismos com grandes genomas são fontes mais produtivas de metabólitos secundários, então os actinomicetos seriam os mais produtivos (Zhu *et al.*, 2011). Entretanto, sabe-se que muitos micro-organismos com genomas menores codificam peptídeos sintetizados ribossomalmente e modificados após a tradução (*RIPPs*) de grande interesse (Ju *et al.*, 2013). Além disso, os esforços mais tradicionais de descoberta de produtos naturais microbianos revelam continuamente uma nova diversidade molecular que

só é associável aos grupos de genes após sua constatação *in vitro*, devido à falta de precedência biossintética (Bachmann *et al.*, 2014).

Em bactérias, muitos estudos de identificação de peptídeos antimicrobianos (AMPs) baseiam-se em métodos dependentes de ensaios *in vitro* para avaliar a atividade antimicrobiana dos isolados potencialmente produtores desses compostos (Perez *et al.*, 2014). No entanto, a análise de dados genômicos através de ferramentas de bioinformática têm resultado na descoberta de novos peptídeos antimicrobianos produzidos por diferentes grupos de bactérias (Flaherty *et al.*, 2014).

1.2.7 Tecnologias de sequenciamento genético e análise dos dados

O sequenciamento genético tem evoluído e se tornado cada vez mais acessível. As novas tecnologias estão cada vez mais presentes nas rotinas de laboratório, pois permitem a análise completa de genomas a relativamente baixo custo e enriquecem as pesquisas científicas (Hu *et al.*, 2021).

Várias tecnologias de sequenciamento estão atualmente disponíveis, cada uma com suas vantagens e desvantagens exclusivas, permitindo gerar quantidades crescentes de dados a custos decrescentes (Reuter *et al.*, 2015). As sequências do gene marcador filogenético rRNA 16S são tradicionalmente geradas usando sequenciamento Sanger de “primeira geração”, oferecendo alta precisão de basecalling. No entanto, tal sequenciamento torna-se um processo difícil e trabalhoso quando se trata de identificação de comunidades bacterianas, pois esta tecnologia resultaria em sinais sobrepostos porque todas as bactérias são sequenciadas, exigindo, portanto, uma etapa intermediária de cultura para isolar primeiro as diferentes bactérias (Ley *et al.*, 2005).

Houve uma grande mudança de paradigma entre a chamada primeira tecnologia (Sanger) e as tecnologias de segunda ou nova geração (454, SOLiD, Illumina e Ion Torrent) referente ao tamanho das *reads* e ao volume de dados. As tecnologias de nova geração baseiam-se na geração de uma grande quantidade de dados, a um custo relativamente baixo, porém com *reads* em menor tamanho (150-300 bp), o que dificulta a montagem do genoma, especialmente para regiões altamente repetitivas, e a detecção de variações estruturais (Hu *et al.*, 2021).

As tecnologias de “segunda geração”, como a tecnologia Illumina, também são chamadas de tecnologias de sequenciamento de “leitura curta” porque os comprimentos de leitura alcançáveis são relativamente limitados, mas grandes quantidades de dados de sequenciamento precisos ou “leituras” podem ser geradas através de seu processamento

massivo. O instrumento Illumina é capaz de sequenciar no máximo 2×300 pb, o que é insuficiente para cobrir todo o gene 16S rRNA, entretanto é uma alternativa interessante para identificação bacteriana, especialmente para amostras mistas complexas, porque permitem a orientação rápida das amostras por não exigirem nenhuma etapa intermediária de cultura (Winand et al., 2019).

Nos últimos anos, a tendência reverteu-se e começaram a surgir as chamadas tecnologias de terceira geração, no momento concretizadas pelo *Pacific Biosciences (PacBio)* e pelo *Oxford Nanopore Technologies (ONT)*. O objetivo é a obtenção da sequência nucleotídica de uma molécula inteira de uma só vez (Garrido-Cardenas e Manzano-Agugliaro, 2017). No entanto, apesar das vantagens de obter *reads* maiores, as tecnologias de terceira geração ainda apresentam muitos erros, exigindo coberturas maiores e, portanto, aumentando o custo para obtenção de sequências com qualidade semelhante ou próxima à obtida por *Illumina*. Espera-se que nos próximos anos, com o refinamento das plataformas e protocolos, estas tecnologias se tornem cada vez mais vantajosas (Lin et al., 2021).

Os avanços tecnológicos nos sequenciamentos têm influenciado positivamente nos avanços na microbiologia moderna. Hoje em dia, até as áreas mais clássicas, como a taxonomia, necessitam do sequenciamento genômico para a descrição de novos táxons (Hu et al., 2021).

Com o grande volume de dados gerados nos sequenciamentos, é essencial que haja bancos de dados com genomas de referência para orientar os processos de montagem, anotação e análise dos genomas, além de métodos de computação capazes de processar e analisar todas essas informações. Avanços em metodologias de preparação das amostras também tem evoluído para melhorar a obtenção de dados (Pereira et al., 2020).

A quantidade de dados gerados no sequenciamento de um genoma bacteriano equivale a dezenas de vezes seu tamanho real. Então, alinhar *reads* a um genoma de referência para mapear um novo genoma tornou-se significativamente mais complexo e a quantidade de informação gerada impõe desafios ao processamento dos dados (Pereira et al., 2020).

Mesmo os algoritmos mais modernos não superam os problemas com o mapeamento e a montagem dos genomas. Analisar o volume de dados gerados pelos sequenciamentos é um processo complexo e computacionalmente intensivo. Para cada etapa da análise, pode ser necessário um conjunto de ferramentas ou *softwares* diferentes. Por exemplo, a montagem *de novo* é realizada por meio de uma combinação de vários alinhamentos, ferramentas de montagem e polimento, cada uma utilizando seus próprios parâmetros de entrada. Essas ferramentas são normalmente executadas a partir de uma linha de comando e requerem

extensos recursos computacionais, aumentando a complexidade do processo de análise e exigindo experiência dos analistas (De Koning *et al.*, 2020).

Por outro lado, algumas soluções baseadas na *web* também já são oferecidas sem que haja necessidade de programação por comando de linha. A plataforma *Galaxy* (disponível em <https://usegalaxy.org/>) oferece a análise dos dados com alto grau de flexibilidade, semelhante aos gerenciadores de fluxo de trabalho baseados em linha de comando, mas com uma interface *web* acessível (De Koning *et al.*, 2020).

Há também a plataforma KBase (disponível em <https://www.kbase.us/>) que permite que os usuários analisem, compartilhem e colaborem usando dados e ferramentas projetadas para ajudar a construir modelos cada vez mais realistas para a função biológica (Arkin *et al.*, 2018).

1.2.8 Análises filogenéticas em procariotos

Com o firmamento das técnicas moleculares, o estudo da filogenia dos procariotos através do sequenciamento da região do *DNA* que codifica genes marcadores permitiu analisar os passos evolutivos das espécies (Case *et al.*, 2007, Lane *et al.*, 1985, Wang *et al.*, 2007). São empregadas análises de vários genes para a classificação filogenética de micro-organismos, destacando-se o uso das sequências referentes à subunidade menor do *RNA* ribossomal (*SSU RNA – small subunit ribosomal*).

A subunidade *rRNA* 16S apresenta características similares ao gene da subunidade maior (23S *rRNA*), mas com a vantagem de ser menor e, por isso ser sequenciado mais rapidamente (Bricker, 2000, Woese e Fox 1977). Tal gene acumulou poucas modificações no decorrer do tempo, além de apresentar homologia funcional, possuir regiões variáveis que diferem em todos os níveis taxonômicos e ser pouco afetado por transferências gênicas horizontais, constituindo-se assim um adequado marcador filogenético (Nascimento, 2011). Por isso, a análise da sequência do gene *rRNA* 16S para a identificação molecular de micro-organismos tem sido uma estratégia muito utilizada em diferentes áreas de pesquisa (Srinivas *et al.* 2009, Whiteley *et al.*, 2012).

Para identificar um micro-organismo por meio da análise do *rRNA* 16S, é necessário realizar uma reação em cadeia da polimerase (*PCR – polymerase chain reaction*). Por meio desta reação, o *DNA* obtido de isolados bacterianos é amplificado com uso de oligonucleotídeos iniciadores e apenas a região específica do gene *rRNA* 16S do genoma é amplificada (Rappé e Giovannoni 2003). Com o gene *rRNA* 16S amplificado é possível fazer seu sequenciamento de forma simplificada e comparar os chamados *amplicons* com

sequências já depositadas em bancos de dados de domínio público (Wheeler *et al.*, 2007) tais como o *Ribosomal Data Base Project* (<http://rdp.cme.msu.edu/>) e *National Center for Biotechnology Information Blast Library* (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>). Estes bancos possuem ferramentas que permitem a rápida comparação de uma sequência do gene *rRNA* 16S com as sequências depositadas no banco de dados facilitando a identificação da linhagem.

Apesar do sucesso do sequenciamento do gene *rRNA* 16S para o entendimento filogenético dos procariotos, alguns problemas permanecem na circunscrição de táxons. As árvores filogenéticas de genes *rRNA* 16S, por exemplo, muitas vezes não possuem resolução para distinguir gêneros intimamente relacionados (Brenner *et al.*, 2005, Ludwig e Klenk, 2005).

Os avanços, especialmente na sistemática molecular, levaram à visão de que as classificações em todos os níveis da hierarquia taxonômica deveriam ser baseadas no uso integrado de dados genotípicos e fenotípicos (Wayne *et al.*, 1987; Stackebrandt, 1992). Essa abordagem, conhecida como taxonomia polifásica (Colwell, 1970), tornou-se a pedra angular da sistemática procariótica (Vandamme *et al.*, 1996; Gillis *et al.*, 2005; Kämpfer e Glaeser, 2012).

Para a classificação taxonômica de novas espécies, segue-se o *Overall Genome Related Index (OGRI)*, parâmetros estabelecidos para a classificação de bactérias de forma polifásica. O *OGRI*, além da análise das sequências do *rRNA* 16S, recomenda a caracterização fenotípica, a avaliação da identidade média de nucleotídeos (*ANI*), a hibridização do *DNA-DNA* digital (*dddH*), a avaliação do genoma (contaminação, cobertura e completude) e a construção de árvore filogenética baseada em múltiplos marcadores moleculares proteicos (Chun *et al.*, 2018).

Os parâmetros *OGRI* são utilizados para verificar se uma linhagem pertence a uma espécie conhecida, calculando a relação entre as sequências do genoma das linhagens e a linhagem tipo de uma espécie. A *ANI* e o *dddH* têm sido mais amplamente utilizados. Os limites de espécies propostos e geralmente aceitos para os valores de *ANI* e *dddH* são 95 ~ 96 e 70%, respectivamente (Meier-Kolthoff *et al.*, 2013). Uma combinação de similaridade entre o gene marcador *rRNA* 16S e os parâmetros *OGRI* é usada em um processo sistemático para identificar e reconhecer uma nova espécie. Nesta abordagem, a lista de espécies necessárias para comparação com a linhagem analisada é obtida através de uma pesquisa baseada no gene marcador *rRNA* 16S (Yoon *et al.*, 2017), onde apenas espécies que

apresentam 98,7% ou mais de similaridade são selecionadas para o cálculo dos parâmetros *OGRI* a partir de seus genomas (Meier-Kolthoff *et al.*, 2013).

Vários parâmetros estatísticos foram desenvolvidos e utilizados para descrever a montagem final de um genoma e garantir que ele seja adequado para fins taxonômicos. Entre eles: tamanho do genoma (soma do comprimento de todos os contigs), número de contigs (não pode ser um indicador da qualidade da sequência do genoma), N50 (definido como o comprimento do contig mais curto que mostra acumulativamente 50% ou mais do tamanho do genoma) e Profundidade de cobertura de sequenciamento (geralmente expresso como dobras, significa que cada base na montagem final foi lida em média x vezes, recomenda-se $\geq 50X$ para as plataformas NGS atualmente disponíveis) (Chun *et al.*, 2018).

Uma tarefa difícil é elucidar as relações filogenéticas entre os genomas. A maioria dos métodos filogenéticos amplamente utilizados foram desenvolvidos para inferir a filogenia de um gene, mas não de toda a sequência do genoma. Como muitos genes sofreram eventos de transferência horizontal, um processo geralmente aceito para inferir a filogenia do genoma completo é usar múltiplos genes que são considerados ortólogos ou que provavelmente sofreram eventos de transferência lateral. Genes funcionalmente importantes, como aqueles que codificam proteínas ribossômicas, são geralmente selecionados para tal árvore filogenômica (Chun *et al.*, 2018).

1.2.9 Anotação funcional de genomas

A anotação funcional tornou-se uma tarefa rotineira na análise de dados de sequenciamento do genoma bacteriano. Uma anotação completa do genoma é crucial para formar uma base estável para muitas análises posteriores, uma vez que tanto a precisão como a abrangência da anotação têm fortes impactos no resultado de estudos relacionados (Schwengers *et al.*, 2021).

Uma das bases utilizadas na microbiologia bacteriana é o *Clusters of Orthologous Groups of proteins (COGs)* (Tatusov *et al.*, 1997). Os bancos de dados COG tem sido uma ferramenta popular para genômica funcional e comparativa de bactérias. Ele permite a anotação das proteínas dos genomas em 26 categorias funcionais e a proporção de proteínas distribuídas em cada categoria tornou-se um dos parâmetros a serem analisados nos genomas sequenciados (Galperin *et al.*, 2021).

A anotação de matrizes CRISPR (repetições palindrômicas curtas agrupadas regularmente interespaçadas) também tem sido comum em análises genômicas. As matrizes CRISPR e suas proteínas associadas (Cas) exibem um sistema imunológico adaptativo

procariótico que memoriza infecções anteriores integrando sequências curtas de genomas invasores, os espaçadores, no locus CRISPR. Os espaçadores interespaçados com repetições são expressos como pequenos RNAs CRISPR guia (crRNAs) que são empregados pelas proteínas Cas para atingir a sequência invasora especificamente em uma infecção recorrente (Hille e Charpentier, 2016). O interesse no campo do CRISPR-Cas aumentou nos últimos anos. Estudos lançam luz sobre os processos genéticos e bioquímicos subjacentes ao sistema imunológico adaptativo, revelando o seu potencial na medicina moderna. Os sistemas CRISPR-Cas podem ser usados como ferramentas de engenharia genética para a edição do genoma programada por RNA e, assim contribuir para terapias celulares, descoberta de medicamentos, rastreamento de doenças, entre outros (Barrangou e Doudna, 2016).

A partir do genoma bacteriano, também pode-se detectar diferentes *clusters* gênicos biossintéticos (*BGC*), para domínios conservados típicos de enzimas que produzem moléculas com atividade biológica; como os policetídeos, peptídeos não ribossomais, terpenos, sideróforos, entre outras. Estes domínios indicam o grande potencial biotecnológico de microorganismos para obtenção de novos antibióticos, compostos antitumorais entre outras moléculas com aplicações biotecnológicas (Bérdy, 2005, 2012; Challis, 2014; Ikeda, 2003).

Ferramentas de bioinformática são indispensáveis na anotação de *clusters* gênicos e vias metabólicas de interesse biotecnológico a partir dos genomas. Elas também tornam possível a comparação entre os dados genômicos e metagenômicos gerados no laboratório e os diversos bancos de dados existentes. Muitas ferramentas online têm sido desenvolvidas nessa área. É o caso por exemplo do AntiSMASH, ferramenta online que permite a busca de genes e vias metabólicas ligadas a metabólitos secundários através da análise de *clusters* gênicos e do processamento simultâneo de múltiplos *contigs*. (Mardis, 2017, Blin *et al.* 2021).

O *Antimicrobial Peptide Database* (APD) é uma ferramenta bastante utilizada para análise de sequências relacionadas a produção de peptídeos antimicrobianos ou de metabólitos secundários em genomas microbianos (Wang e Wang, 2016). Alguns algoritmos destinados especificamente à busca por genes de bacteriocinas também estão disponíveis. O *software* BAGEL usa três bancos de dados diferentes contendo sequências de bacteriocinas e outros peptídeos, modificados ou não após a tradução (Van Heel *et al.*, 2013). Já o algoritmo BACTIBASE utiliza propriedades físico-químicas calculadas ou previstas de bacteriocinas produzidas por bactérias Gram-positivas e Gram-negativas para prever peptídeos antimicrobianos em genomas bacterianos (Hammami *et al.*, 2010).

CAPÍTULO 2 - USO DO SOBRENANDANTE DE *PAENIBACILLUS ELGII* NA PROSPECÇÃO E SELEÇÃO DE LINHAGENS BACTERIANAS COM POTENCIAL PARA APLICAÇÃO BIOTECNOLÓGICA.

2.1 INTRODUÇÃO

Após o surgimento de técnicas moleculares na década de 1970, como *PCR*, sequenciamento de nova geração e mais particularmente a metagenômica e a enorme quantidade de informações fornecidas por elas, o cultivo de bactérias pode ter sido negligenciado (Gutleben *et al.*, 2018). No entanto, as culturas bacterianas continuam sendo uma ferramenta fundamental para os microbiologistas visto que ambos os métodos, dependentes ou não de cultura, podem se complementar (Lagier *et al.* 2015).

O ritmo dos sequenciamentos bacterianos supera o isolamento de novas espécies e pode revelar capacidades fisiológicas interessantes nos micro-organismos e, assim, direcionar os esforços de cultivo (Wilson *et al.*, 2014). O cultivo de linhagens puras por outro lado, permite que seus genomas sejam bem anotados fornecendo dados para as previsões utilizadas nas análises genômicas e metagenômicas (Choi *et al.*, 2017; Liu *et al.*, 2022).

O cultivo associado à tecnologia e as informações fornecidas por abordagens independentes de cultura oferecem acesso a uma diversidade bacteriana inexplorada e novos esclarecimentos fundamentais sobre princípios biológicos (Overmann *et al.*, 2017). Muitas linhagens bacterianas pouco exploradas provavelmente apresentam novas vias biossintéticas e características bioquímicas desconhecidas, sugerindo alto potencial biotecnológico a ser descoberto a partir dos metabólitos e enzimas desses micro-organismos e possíveis aplicações para biotecnologia (Gutleben *et al.*, 2018).

Uma das principais demandas biotecnológicas da indústria por compostos bioativos refere-se à produção de compostos antimicrobianos. Há a necessidade de alternativas aos produtos já presentes no mercado aos quais muitas bactérias já se tornaram resistentes (Laws *et al.*, 2019). Muitos dos antimicrobianos descritos tem sua origem em bactérias que os produzem por competição no ambiente ou até mesmo comunicação com outros micro-organismos (Durand *et al.*, 2018). Entretanto, um dos principais obstáculos nos estudos de procariotos e na descoberta de novos compostos produzidos por eles é a incapacidade de cultivar em condições de laboratório a maior parte da diversidade ambiental (Overmann *et al.*, 2017).

Certos grupos de bactérias requerem condições específicas de crescimento exigindo o refinamento das estratégias existentes para o cultivo de micro-organismos, bem como o

desenvolvimento de novas técnicas. Métodos que sejam mais eficazes na obtenção de micro-organismos até então não cultivados ou de interesse biotecnológico (Kapinusova *et al.*, 2023).

Em seu ambiente natural, as bactérias vivem em comunidade e possuem sofisticadas redes de comunicação baseadas em alterações metabólicas que influenciam processos de comportamento basal como reprodução, esporulação, formação de biofilme e expressão de antimicrobianos (Papenfort e Bassler, 2016). Um tipo particular de interação química é a troca de compostos sinalizadores. Moléculas sinalizadoras podem ser essenciais para o crescimento de micro-organismos, já que alguns deles podem não crescer *in vitro*, pois não estão em suas comunidades nativas e, portanto, não recebem sinalização desta comunidade através desta comunicação intercelular denominada *quorum sensing* (Nichols *et al.*, 2008).

Dentre as moléculas sinalizadoras conhecidas, as acil-homoserina lactonas representam os compostos mais bem compreendidos e utilizados por muitas bactérias Gram-negativas (Papenfort e Bassler, 2016). Enquanto, peptídeos estão envolvidos na comunicação intercelular em bactérias Gram-positivas e podem ter uma função dupla, atuando como antimicrobiano e molécula sinalizadora (Verbeke *et al.*, 2017). Sideróforos, cuja biossíntese é controlada por um sistema de *quorum sensing* também podem estimular o crescimento bacteriano (Soares, 2022). *Paenibacillus elgii* é conhecido por produzir esses três tipos de moléculas sinal (homoserina lactonas, peptídeos e sideróforos) que podem estimular o crescimento bacteriano (da Costa *et al.*, 2022; Ortega *et al.*, 2018). Além disso, a espécie *P. elgii* apresenta potentes antimicrobianos já explorados para o uso veterinário (Krüger *et al.*, 2013; Teng *et al.*, 2012; da Costa *et al.*, 2022; Ortega *et al.*, 2018).

Espera-se que organismos produtores de antimicrobianos contenham mecanismos de auto-resistência contra suas próprias moléculas bioativas. Geralmente, bactérias produtoras de antibióticos possuem genes de resistência que as protegem de seus antimicrobianos (Mak *et al.*, 2014). Assim, espera-se que linhagens bacterianas resistentes a *P. elgii*, possuam resistência intrínseca por também produzirem compostos antimicrobianos. Este capítulo apresenta uma metodologia de cultivo bacteriano a partir do uso do sobrenadante de *P. elgii* como mecanismo de seleção de linhagens bacterianas de interesse biotecnológico. Uma das linhagens obtidas através desta metodologia foi caracterizada a fim de ilustrar uma potencial aplicação biotecnológica disponibilizada a partir do protocolo.

2.2 OBJETIVOS

2.2.1 Objetivo Geral

Este trabalho visa isolar linhagens bacterianas de solo do Cerrado com potencial biotecnológico a partir do uso do sobrenadante de *Paenibacillus elgii* como mecanismo de seleção.

2.2.2 Objetivos Específicos

I. Testar o uso de sobrenadante de *P. elgii* como agente de seleção de linhagens bacterianas com potencial biotecnológico.

II. Identificar as linhagens selecionadas através do gene marcador *rRNA* 16s.

III. Testar a atividade antibacteriana das linhagens contra *Bacillus subtilis*, *Escherichia coli* e *Pseudomonas aeruginosa*.

IV. Caracterizar fenotípica e genotipicamente uma linhagem obtida a partir do uso de sobrenadante de *P. elgii*.

2.3 MATERIAL E MÉTODOS

2.3.1 Obtenção do sobrenadante de *Paenibacillus elgii*

A linhagem *P. elgii* utilizada neste trabalho foi isolada do solo do Cerrado de uma área de reserva da fazenda Nova Índia em Cristalina (GO), uma propriedade rural privada. Para produção do sobrenadante da linhagem, fez-se um inóculo em meio de cultura LB Broth (NaCl, 5 g/L; Triptona, 10 g/L; Extrato de Levedura, 5 g/L) que cresceu *overnight* sob agitação de 200 rpm à 28°C. Após crescimento, o sobrenadante foi submetido a 4°C *overnight*, para induzir esporulação. Em seguida, a cultura foi centrifugada e o sobrenadante filtrado em seringa de 0,22µm de poro. O sobrenadante filtrado foi inoculado em placas com meio LB Broth/ágar e incubados por 48 h, em temperatura ambiente, para confirmar a ausência de formação de colônias de *P. elgii*.

2.3.2 Preparação dos meios de cultura enriquecidos com sobrenadante de *P. elgii*

As placas de cultivo foram preparadas contendo diferentes concentrações do sobrenadante: 300µl/L, 600µl/L, 1ml/L e 2ml/L de sobrenadante. Os meios de cultura utilizados para os ensaios foram o Middlebrook/ágar (Hidrolisado enzimático de caseína, 1g/L; Sulfato de amônio, 0.5 g/L, Ágar, 15 g/L; Fosfato monopotássico, 1.5 g/L; Fosfato dissódico, 1.5 g/L; Citrato de sódio, 0.4 g/L; Sulfato de magnésio, 0.05 g/L; ácido L-glutâmico, 0.5 g/L; Citrato férrico de amônio, 0.04 g/L; Piridoxina, 0.001 g/L; Biotina, 0.0005 g/L; Verde malaquita, 0.001 g/L; Ágar, 15 g/L) e o R2A/ágar (ágar, 15 g/L; hidrolisado ácido de caseína, 0.5 g/L; dextrose, 0.5 g/L; fosfato dipotássico, 0.3 g/L; sulfato de magnésio, 0.024 g/L; proteose peptona, 0.5 g/L; sódio piruvato, 0.3 g/L; amido solúvel, 0.5 g/L; extrato de levedura, 0.5 g/L). Esses meios de cultura foram vertidos em placas circulares de 115 mm de diâmetro.

2.3.3 Coleta e processamento de amostras de solo

As coletas de amostras foram feitas na Reserva Ecológica do IBGE (RECOR) nas proximidades do Córrego Pitoco. A RECOR é parte da Área de Proteção Ambiental (APA) Distrital Gama-Cabeça de Veado, localizado no Pediplano de Brasília e totaliza um total de 10.000 ha de área protegida. A formação geológica da região data da Era Terciária que juntamente com o Pediplano Contagem-Roncador constitui a parte do Planalto Central Goiano denominado Planalto do Distrito Federal. O clima da região é sazonal, característico do Brasil Central, com duas estações bem definidas, uma seca que abrange os meses de maio

a agosto, e uma chuvosa que abrange os meses de setembro a maio (fonte: <https://www.ibge.gov.br/>, consultado em 2022).

A coleta de amostras de solo foi feita com o uso de um cano de alumínio com indicação de profundidade. O cano foi introduzido com o auxílio de uma marreta de borracha no solo até atingir a marcação de 15 cm. Foram feitos 10 orifícios de coleta e todo o conteúdo amostrado foi peneirado em malha de 2 mm e misturado em um único saco plástico devidamente identificado.

As amostras ambientais de solo foram dissolvidas previamente em meio Middlebrook líquido na razão de 1g de solo em 50ml de meio, constituindo-se como inóculo inicial. O inóculo inicial foi diluído na proporção 1:20 e inoculado (150µL) nas placas enriquecidas com sobrenadante de *P. elgii*.

2.3.4 Isolamento das linhagens bacterianas

As colônias que cresceram nos meios de cultura com sobrenadante de *P. elgii* e apresentaram halo de inibição a colônias vizinhas foram transferidas para isolamento em novas placas a fim de serem identificadas e testadas quanto a atividade antimicrobiana.

2.3.5 Extração de DNA e amplificação do gene ribossomal *rRNA* 16S

As extrações de DNA das linhagens bacterianas obtidas a partir do uso do sobrenadante *P. elgii* foram realizadas com o kit comercial *Gen Elute Bacterial Genomic DNA Kit*, Sigma-Aldrich, de acordo com as instruções do fabricante. A eficiência da extração foi avaliada em gel de agarose (1% v/v). A quantificação do DNA foi realizada com o fluorômetro Qubit (Invitrogen, USA) através do kit *dsDNA BR Quant-it assay kit* (Invitrogen).

A reação em cadeia da (DNA) polimerase (PCR) para amplificação do *rRNA* 16S foi realizada com os oligonucleotídeos iniciadores 27F (5'- AGAGTTTGATCCTGGCTCAG-3') e 1492R (5'- GGTTACCTTGTTACGACTT-3') (Weisburg *et al.*, 1991), em termociclador, nas seguintes condições: desnaturação a 95°C durante 3 minutos, desnaturação a 95°C durante 30 segundos, anelamento a 51°C durante 30 segundos, extensão a 72°C durante 1 minuto e meio, repetição das etapas 2 a 4 por 30 vezes, extensão a 72°C durante 7 minutos. Para a purificação do produto de PCR foi utilizado o kit de purificação Thermo Scientific (Gene Jet PCR Purification Kit).

O sequenciamento de Sanger dos *amplicons* do gene 16S foi feito pela empresa Macrogen na Coreia do Sul. As amostras sequenciadas foram comparadas aos bancos de dados do NCBI (*National Center for Biotechnology Information*) através do BLAST.

2.3.6 Análise da atividade antibacteriana

Realizou-se um teste antimicrobiano por *overlay* para testar as linhagens bacterianas isoladas a partir do uso do sobrenadante de *P. elgii*. Os testes de atividade antibacteriana foram feitos em meios de cultura mínimo (R2A ou Middlebrook). As linhagens foram inoculadas no centro das placas e colocados em crescimento em temperatura ambiente (aproximadamente 25° graus) pelo período de 6 dias. No sétimo dia de crescimento, foi preparado um inóculo de soft-ágar (água ultra pura e ágar) com *Bacillus subtilis* que foi vertido sobre as placas de modo a formar uma camada fina e homogênea. Esta foi a cobertura ou *overlay*. Após a adição do soft-ágar era esperado o crescimento do *B. subtilis* por até 48 horas em temperatura ambiente. Algumas linhagens foram selecionadas para serem testadas também contra *E.coli* e *P. aeruginosa*. A seleção foi feita de modo a incluir pelo menos um representante de cada gênero encontrado.

Para comprovar a capacidade antibiótica de *P. elgii* e de seu sobrenadante, ambos foram também submetidos ao teste antimicrobiano por *overlay* contra *B. subtilis*.

2.3.7 Caracterização fenotípica *in vitro*

A linhagem K003 obtida a partir do uso do sobrenadante foi selecionada para ser caracterizada. A linhagem CBMAI 824 de *Methylobacterium radiotolerans* foi utilizada para a realização das análises fenotípicas e fisiológicas de forma comparativa com K003.

Após 18 h de crescimento em meio LB, os *pellets* de células das culturas de K003 e CBMAI 824 foram observados em microscopia óptica para a determinação das características de membrana, utilizando a coloração diferencial de Gram (Reddy *et al.*, 2007). A observação foi realizada em microscópio Leica DM750.

A caracterização fenotípica foi realizada com base nas metodologias adaptadas utilizadas por Kim *et al.* (2020) e Yun *et al.* (2020), onde a temperatura ótima de crescimento foi avaliada por uma semana em placas de *Nutrient Broth* (NB) ágar (Extrato de carne, 1 g/L; Extrato de Levedura, 2 g/L; Peptona, 5 g/L; Cloreto de Sódio, 5 g/L, Ágar 15 g/L) nas temperaturas de 5, 10, 15, 20, 30, 35, 40, 45, 50, 55 e 60 °C. A faixa de pH ideal para crescimento foi examinada entre 4.0 e 10.0 em intervalos de 1.0 após incubação à temperatura ideal verificada por 3 dias em meio NB líquido ajustado com HCl ou NaOH contendo tampão citrato/Na₂HPO₄ 50 mM (pH 4–7), tampão NaH₂PO₄ /Na₂HPO₄ 100 mM (pH 7–8) ou tampão NaHCO₃/Na₂CO₃ 100 mM (pH 9–10). A tolerância ao sal foi testada em NB líquido suplementado com NaCl (0-5% em intervalos de 1%) e incubação à temperatura ideal

verificada por 3 dias. Tanto salinidade quanto pH foram examinados medindo a OD₆₀₀ em SpectraMax M3 com *software* SoftMax Pro 6.0 (Zhu *et al.*, 2014).

A atividade da catalase foi testada pela adição de 3% de peróxido de hidrogênio à cultura cultivada em meio de cultura R2A a 30°C por 7 dias, e a efervescência foi monitorada (Tindall *et al.*, 2007). O teste de oxidase foi realizado através de *Oxidase Strips* 40560 (Millipore).

As células da linhagem K003 e sua referência, CBMAI 824, cultivadas em meio R2A sólido foram coletadas após 48h de crescimento, transferidas para tubos de vidro com tampa de teflon e então utilizadas para análises de ácidos graxos celulares, conforme descrito por Ramaprasad e colaboradores (2015). A saponificação dos ácidos graxos foi realizada com 1 mL de solução metanol/hidróxido de sódio (150 mL água deionizada, 150 mL metanol, 45 g NaOH) à 100°C por 30 minutos. Dois mililitros de metano 6M em HCL foram adicionados aos tubos, que ficaram incubados por 10 minutos à 80°C. A separação dos ésteres metílicos dos ácidos orgânicos foi realizada pela adição de 1,25 mL de hexano/éter terc-butílico (1: v/v), após agitação por 10 minutos. A fase mais pesada foi descartada e a fase superior contendo os ácidos graxos metil-ésteres foi recuperada e lavada com 0.25 mol/L NaOH e estocada até a análise (Sasser, 2001).

Os ésteres metílicos de ácidos graxos obtidos foram analisados em cromatógrafo gasoso (Agilent 7890^a) equipado com o *software* *Sherlock Microbial Identification System* (Sherlock, v. 6.2) usando o a biblioteca do método RTSBA6. Os picos obtidos foram então rotulados e os valores de comprimento de cadeia equivalente foram calculados pelo *software* Sherlock. O Sistema de Identificação Microbiana Sherlock é um sistema totalmente analítico automatizado de cromatografia gasosa para identificar bactérias com base em seus perfis únicos de ácidos graxos (Sasser, 2001).

A suscetibilidade a antibióticos foi avaliada em meio R2A sólido pelo método Kirby-Bauer (difusão com disco) utilizando discos de papel filtro impregnados com concentração padrão de antibiótico. A cultura foi espalhada com *swab* para a aplicação dos discos e observação da formação de halo de inibição foi realizada após incubação de 48h a 30° C. Os antibióticos testados foram (µg por disco): Amicacina (30), Amoxicilina/Ác. Clavulânico (20/10), Ampicilina (10), Aztreonam (30), Cefalexina (30), Cefazolina (30), Cefepima (30), Ceftazidima (30), Ceftriaxona (30), Ciprofloxacina (5), Clindamicina (2), Cloranfenicol (30), Doxiciclina (30), Eritromicina (15), Gentamicina (10), Imipenem (10), Meropenem (10), Norfloxacina (10), Ofloxacina (5), Oxacilina (1), Penicilina G (10), Piperacilina +

Tazobactam (110), Polimixina B (300), Sulfonamida (300), Tetraciclina (30) e Vancomicina (30).

2.3.8 Sequenciamentos, montagem e análise do genoma

Para a linhagem K003 caracterizada neste capítulo, foi feita a clonagem do gene marcador *rRNA* 16S, onde os fragmentos do *rRNA* 16S purificados foram ligados em vetor pGEM®-T Easy, Promega, de acordo com as instruções do fabricante. O objetivo foi ter uma sequência mais completa do gene marcador *rRNA* 16 para as análises filogenéticas.

A transformação bacteriana ocorreu em *Escherichia coli* DH10B por choque térmico. As células transformadas foram espalhadas em meio de cultura R2A sólido suplementado com Ampicilina (150 µg/mL), X-Gal (5-bromo-4-cloro-3-indoxil-β-D-galactopiranosídeo) (0.00625% p/v) e IPTG (isopropil-β-D-tiogalactopiranosídeo) (0.5 mM). As colônias positivas para a transformação com plasmídeo e inserto (brancas) foram inoculadas em meio líquido LB + Ampicilina (150 µg/mL).

O kit *QIAprep Spin Miniprep Kit*, Qiagen, foi utilizado para a extração e purificação do *DNA* plasmidial. A transformação foi confirmada através da verificação da presença dos insertos após realização de digestão com a enzima EcoR1 e avaliação do padrão de bandas em eletroforese em gel de agarose a 1% (p/v).

O *DNA* plasmidial dos clones foi sequenciado em plataforma Sanger na empresa MacroGen (Seoul, Coréia do Sul) utilizando os oligonucleotídeos iniciadores para os promotores T7 (5' GTAATACGACTCACTATAGGGC 3') e sp6 (5' ATTTAGGTGACACTATA 3').

A ferramenta Asparagin (disponível em <http://asparagin.cenargen.embrapa.br/phph/>) foi utilizada para avaliar a qualidade *Phred score* e remoção dos pares de bases provenientes do vetor (Ewing, 1998; Togawa, 2006). As sequências com *Phred score* maior que 30 foram mantidas para as análises seguintes. Em seguida, foi realizada a comparação por similaridade com sequências disponíveis no *GenBank* do *NCBI* (*National Center for Biotechnology Information*) usando a ferramenta BLAST N. Em adição, foi feita a comparação com as sequências do banco de dados Ezbiocloud. Nele, além das sequências de *rRNA* 16S de *amplicons* do *NCBI*, são consideradas também sequências 16S extraídas dos genomas disponíveis no *NCBI*, sequências compiladas do sequenciamento completo realizados com a tecnologia *Pacific Biosciences* (*PacBio*) de amostras de microbiomas, sequências extraídas da montagem de genomas do *JGI* (*Joint Genome Institute*) e sequências de genomoespécies que

são espécies provisoriamente novas suportadas por sequências genômicas inteiras (Yoon *et al.*, 2017).

O sequenciamento genômico da linhagem K003 foi realizado em plataforma *NGS Illumina* na empresa Macrogen (Seul, Coreia do Sul). A biblioteca de *DNA* foi preparada usando o *TruSeq Nano DNA Kit* e o sequenciamento foi realizado na plataforma NextSeq500, gerando leituras *paired-end* de 150 pares de bases.

A fim de melhorar a montagem do genoma, foi realizado o sequenciamento *NGS* utilizando a tecnologia *long reads* na plataforma *Oxford Nanopore Technologies (ONT)* (Oxford, Reino Unido). A biblioteca de *DNA* foi preparada com o *Rapid Sequencing Kit* (SQK-RAD004) e o sequenciamento foi realizado em *Flow cell FLO-MINSP6* no dispositivo MinION (MIN-101B).

Os dados brutos do sequenciamento *NGS Illumina* passaram por limpeza com a ferramenta *Trimmomatic* (Bolger *et al.*, 2014) para retirada de sequências dos adaptadores, eliminação de *reads forward* e *reverse* não-pareados e das sequências com *Phred score* menor do que 30 (Ewing *et al.*, 1998). Os dados brutos do sequenciamento *NGS ONT* passaram pelo processo de *basecalling* através do programa Guppy em modo de alta acurácia (Fraser *et al.*, 2020).

A montagem *de novo* de um rascunho do genoma da linhagem K003 foi realizada de forma híbrida com o programa Unicycler (V. 0.5.0) (Wicket *et al.*, 2016) para confirmação da identidade da linhagem. O *software* CheckM, por meio da plataforma Kbase, foi utilizado para determinação de sequências contaminantes e completude dos genomas (Parks *et al.*, 2015). A cobertura dos *reads* foi determinada com o *software* Bowtie2 (v. 2.4.4) (Langmead *et al.*, 2018) e os parâmetros L50 e N50 foram obtidos através do Quast (Mikheenko *et al.*, 2016). A anotação do genoma foi realizada usando o *software* Prokka (v. 1.14.6) (Seemann, 2014). O *software* Artemis foi utilizado para a visualização do genoma circular e suas regiões codificantes através do recurso *DNAPlotter* (Carver *et al.*, 2009).

2.3.9 Análises filogenéticas

A posição evolutiva da linhagem foi estimada pela construção de árvore filogenética pelo método *Maximum-likelihood* (ML) com 1000 repetições de amostragem (*Bootstrap*), utilizando vinte e três sequências representativas quanto *rRNA* 16S de espécies do gênero de linhagens válidas listadas em *List of Prokaryotic names with Standing in Nomenclature - LPSN* e depositadas no banco de dados GenBank/EMBL/DDBJ (Felsenstein, 1985, Kimura, 1980) e um *outgroup*. A ferramenta *clustal W* foi utilizada para fazer o alinhamento das

sequências (Larkin *et al.*, 2007). A construção da árvore filogenética baseada na sequência parcial do gene marcador molecular *rRNA* 16S foi realizada no programa Mega 11 (disponível para download em <https://www.megasoftware.net/>).

A classificação taxonômica da linhagem e a confirmação da identidade foram realizadas na plataforma TYGS – DSMZ, por hibridização digital de *DNA* (DDH) no servidor GGDC – DSMZ (Meier-Kolthoff *et al.*, 2022). Com base na sequência do *rRNA* 16S, são determinadas as dez linhagens filogeneticamente mais próximas do genoma a ser analisado. Essas onze linhagens seguem para comparação dos parâmetros genômicos de classificação taxonômica, como a distância entre genomas (*GBDP*).

A ANI (Identidade Média de Nucleotídeos) foi calculada na plataforma JSpeciesWS (disponível em: <https://jspecies.ribohost.com/>), onde ANI_b é a identidade nucleotídica média com base no cálculo BLAST e a ANI_m é a identidade nucleotídica média com base no cálculo MUMmer (Richter *et al.*, 2016). O software GTDB-Tk (v. 1.7.0), disponibilizado na plataforma Kbase também foi utilizado para o cálculo da ANI (Arkin *et al.*, 2018; Chaumeil *et al.*, 2019).

O alinhamento do genoma da linhagem K003 com o genoma de sua linhagem de referência, *M. radiotolerans* JCM 2831, acrônimo da linhagem *M. radiotolerans* CBMAI 824, utilizada nos ensaios *in vitro* como referência, foi realizado com o *pipeline* PlotNucmer (disponível em: https://github.com/fmalmeida/plot_nucmer), que é uma automação do MUMmer v3 .1 toolkit (Kurtz *et al.*, 2004) e Ggbio (Yin *et al.*, 2012), onde o primeiro alinha os genomas por comprimento e identidade, e o segundo gera uma visualização circular.

A árvore filogenética baseada em genes ortólogos centrais concatenados de sequências genômicas de 12 linhagens, incluindo K003, as linhagens estreitamente relacionadas de acordo com o TYGS-DSMZ e o *outgroup* *Enterovirga rhinocerotis* disponíveis no *NCBI* foi construída usando o M1CR0B1AL1Z3R servidor *web* (disponível em: <https://microbializer.tau.ac.il/>) (Avram *et al.*, 2019).

A análise comparativa de pangenoma foi realizada usando Roary versão 3.13.0 seguindo a anotação de Prokka, versão 1.14.6 (Seemann, 2014), disponível no Galaxy (Afgan *et al.*, 2016). Foram utilizadas a sequência do genoma de K003 e 10 sequências de suas linhagens mais próximas, segundo o TYGS-DSMZ disponíveis no *NCBI*. A identidade percentual mínima para BLASTp de 80% foi definida para agrupar os genes que codificam sequências completas de proteínas em genomas centrais (Corretto *et al.*, 2017).

Todas as árvores geradas neste trabalho foram editadas no Inkscape (disponível para download em: <https://inkscape.org/>).

2.3.10 Predição de genes e anotação funcional

As categorias funcionais baseadas na classificação *COG* de K003 foram obtidas pela versão on-line do eggNOG-mapper v.2 (Huerta-Cepas *et al.*, 2017), onde vias metabólicas e genes de interesse foram comparados ao banco de dados *Kyoto Encyclopedia of Genes and Genomes (KEGG)* (Kanehisa *et al.*, 2007).

O Banco de Dados de Genes de Resistência a Antibióticos (*ARDB*) foi utilizado para anotação de genes potenciais responsáveis pela resistência a antibióticos (Liu e Pop, 2009). A presença de *locus* CRISPR (agrupado regularmente interespaçado curto repetições palindrômicas) foram previstos usando a ferramenta de reconhecimento CRISPRCasFinder (Couvin *et al.*, 2018).

O *software* AntiSMASH (v. 6.1.1) (Blin *et al.* 2021) foi utilizado para a identificação de *clusters* gênicos relacionados ao metabolismo secundário (*BGCs*) dos micro-organismos. O BiG-SCAPE faz o agrupamento de similaridade de famílias de *BGCs* com base em suas semelhanças com sequências de proteínas de famílias Pfam (El-Gebali *et al.*, 2019) e com *BGCs* conhecidos disponíveis na versão mais recente do repositório MIBiG (v1.4) (Kautsar *et al.*, 2020). A fim de obter informações sobre a diversidade e similaridade de *BGCs* entre as linhagens K003 e sua linhagem de referência *M. radiotolerans* JCM 2831, as saídas do AntiSMASH foram agrupadas usando BiG-SCAPE em cortes de distância de 0.3 (Navarro-Munoz *et al.*, 2020).

O banco de dados CAMP_{R4} foi utilizado para predição de peptídios antimicrobianos (*AMPs*) (Gawde *et al.*, 2022). Trata-se de um banco com curadoria manual usando informações disponíveis de *AMPs* do *NCBI*, *PDB (RCSB Protein Data Bank - RCSB PDB)*, *PubMed* e *Lens (patent)*. Nele, prevê-se que a sequência seja antimicrobiana através de uma pontuação de probabilidade (0 a 1) usando métodos multivariados (Waghu *et al.*, 2014). Quanto maior a probabilidade, maior é a possibilidade de o peptídeo avaliado ser antimicrobiano.

2.3.11 Caracterização fenotípica *in silico*

A ferramenta de fenotipagem *in silico* Traitair (*Microbial Trait Analyzer*) foi utilizada para prever fenótipos com base nas sequências do genoma de K003 de forma comparativa com suas linhagens mais próximas de acordo com TYGS-DSMZ. Ele é capaz de prever 67 características fenotípicas utilizando o algoritmo Prodigal para predição de genes e família Pfam para anotação (Weimann *et al.*, 2016). O *software* usa dois modelos de predição, o modelo *phypat* (que prevê a presença/ausência de proteínas encontrado no fenótipo de 234

espécies bacterianas) e uma combinação de modelos *phypat+PGL* (que usa a informação de *phypat* combinada com a informação da aquisição ou perda de famílias de proteínas e fenótipos através da evolução), para determinar as características fenotípicas (Weimann *et al.*, 2016).

2.4 RESULTADOS

2.4.1 Uso do sobrenadante de *Paenibacillus elgii*

A espécie *P. elgii* apresentou atividade antibacteriana contra *B. subtilis* assim como o sobrenadante obtido de sua cultura líquida (Fig.2.1).

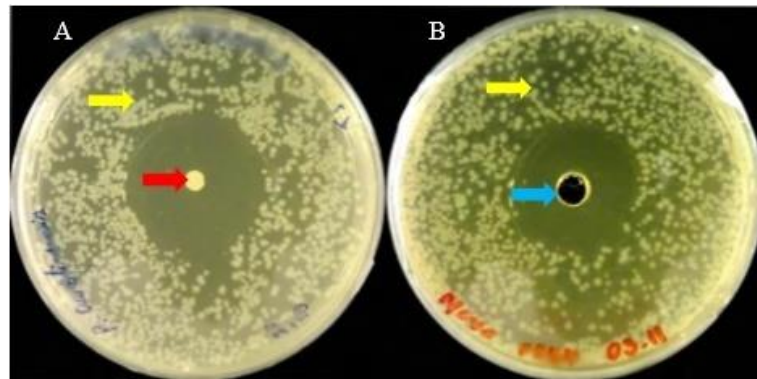


Figura 2.1 Atividade antibiótica de *P. elgii* e seu sobrenadante, demonstrando atividade da cultura indicada pela seta vermelha (A) e do sobrenadante (300µl/L) indicada pela seta azul (B). Os ensaios foram feitos contra *B. subtilis* (indicado por setas amarelas) (Carvalho, 2013).

A utilização do sobrenadante permitiu o crescimento de micro-organismos. A concentração do sobrenadante de *P. elgii* influenciou o crescimento de micro-organismos no meio de cultura após inóculo de amostras de solo. Os resultados demonstraram que há diferença de crescimento bacteriano nas diferentes concentrações do sobrenadante e houve maior crescimento na menor concentração, logo quanto maior a concentração do sobrenadante, mais seletivo é o meio enriquecido com ele (Fig. 2.2).

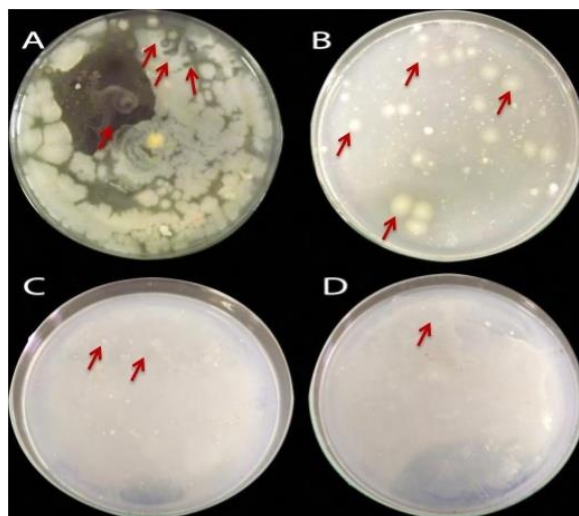


Figura 2.2 Crescimento de culturas a partir de amostras de solo em diferentes concentrações de sobrenadante de *P. elgii*. A placa A (300µl/L), a placa B (600 µl/L), a placa C (1ml/L) e a placa D (2ml/L). As linhagens de interesse coletadas por possuírem zonas de inibição são indicadas por setas vermelhas. (Carvalho, 2013).

2.4.2 Identificação taxonômica das linhagens bacterianas

Dentre as linhagens bacterianas crescidas a partir da cultura enriquecida com o sobrenadante de *P. elgii*, trinta e cinco (35) puderam ser cultivadas isoladamente e sem o sobrenadante. Os resultado das *PCRs* com os oligonucleotídeos 27F e 1492R para os 35 isolados foram bem-sucedidas e podem ser visualizadas na Figura 2.3.

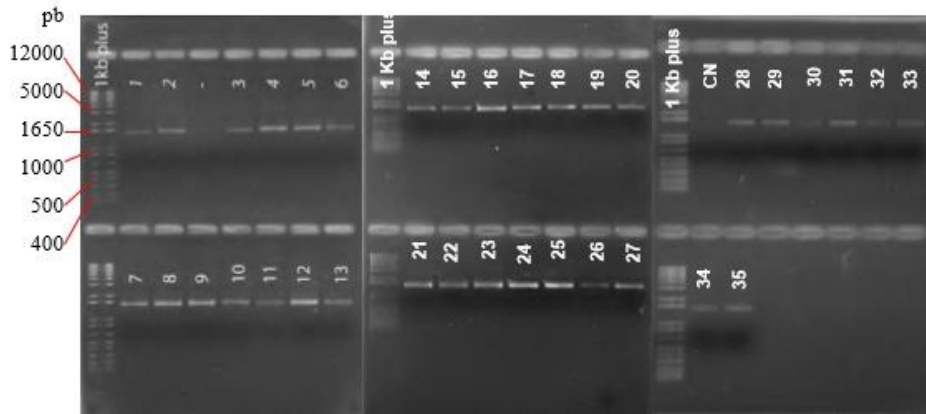


Figura 2.3 Gel (TBE/agarose 1% corado com brometo de etídeo) com resultados da *PCR* das linhagens. O marcador de referência de tamanho de bandas utilizado foi o 1kb plus *DNA Ladder* indicado a esquerda com algumas marcações de tamanho de bandas (Carvalho, 2013).

O resultado do sequenciamento do gene marcador *rRNA* 16S das linhagens obtidas são apresentados na Tabela 2.1. A identificação mostra presença dos gêneros *Paenibacillus*, *Pseudomonas*, *Enterobacter*, *Kitasatospora*, *Sphingomonas*, *Xanthobacter*, *Burkholderia* e *Methylobacterium*.

Tabela 2.1 Resultados da comparação dos *amplicons* das linhagens obtidas a partir da utilização do sobrenadante de *P. elgii* com os bancos de dados do *NCBI* através do *BLAST*.

Linhagem	Nº acesso	Descrição	Max Score	Total Score	Cobertura	Max Ident
A1	HQ236048.1	<i>P. elgii</i>	167	167	59%	92%
A2	JX174235.1	<i>Enterobacter sp.</i>	302	302	96%	95%
A3	AB739006.1	<i>P. ehimensis</i>	623	623	33%	92%
A4	AB739006.1	<i>P. ehimensis</i>	645	645	26%	98%
A5	HQ236085.1	<i>P. elgii</i>	1831	1831	98%	95%
A6	HM063032.1	<i>Paenibacillus sp.</i>	1943	1943	99%	97%
A7	AB739006.1	<i>P. ehimensis</i>	287	287	96%	93%
A8	AB739006.1	<i>P. ehimensis</i>	248	248	95%	90%
A9	HQ236085.1	<i>P. elgii</i>	474	474	98%	97%
A10	JQ735955.1	<i>P. elgii</i>	808	808	68%	98%
A11	HQ236048.1	<i>P. elgii</i>	1034	1034	99%	99%
A12	JX436405.1	<i>Pseudomonas sp.</i>	350	350	96%	99%
B1	JF701960.1	<i>Burkholderia sp.</i>	512	512	36%	92%
B2	GQ895737.1	<i>Sphingomonas sp.</i>	760	760	34%	95%
B3	JX090597.1	<i>P. aeruginosa</i>	22292	22292	99%	99%
B4	GU518381.1	Uncultured bacterium	115	115	83%	80%
B5	AF247494.1	<i>Burkholderia sp.</i>	642	642	97%	95%
B6	JX090597.1	<i>P. aeruginosa</i>	1000	1000	98%	98%
B7	EU139032.1	<i>Kitasatospora sp.</i>	263	263	94%	92%
B8	HQ651053.1	<i>Pseudomonas sp.</i>	1018	1018	98%	99%
B9	AF247494.1	<i>Burkholderia sp.</i>	388	388	48%	97%
B10	FR872408.1	<i>Burkholderia sp.</i>	329	329	90%	97%
B11	DQ095878.1	<i>P. aeruginosa</i>	176	176	64%	91%
B12	GU191167.1	<i>Burkholderia sp.</i>	294	294	88%	97%
C1	JX436403.1	<i>Pseudomonas sp.</i>	675	675	67%	96%
C2	JX436403.1	<i>Pseudomonas sp.</i>	497	675	63%	94%
C3	EU912440.1	<i>Methylobacterium sp.</i>	453	453	66%	88%
C4	HM771642.1	<i>P. aeruginosa</i>	990	990	99%	93%
C5	JX178938.1	<i>Xanthobacter sp.</i>	425	425	27%	95%
C6	EU912444.1	<i>Methylobacterium sp.</i>	193	193	73%	90%
C7	JX436403.1	<i>Pseudomonas sp.</i>	545	545	65%	94%
C8	JQ659918.1	<i>P. aeruginosa</i>	156	156	94%	81%
C10	JX436405.1	<i>Pseudomonas sp.</i>	279	279	98%	92%
C11	GU447238.1	<i>P. aeruginosa</i>	2017	2017	95%	96%
C12	AF247494.1	<i>Burkholderia sp.</i>	887	887	97%	95%

2.4.3 Atividade antibacteriana

As linhagens cultivadas de forma isolada foram submetidas ao ensaio *overlay* de atividade antibacteriana contra *B. subtilis*, onde todos apresentaram resultado positivo (Fig. 2.4 e Tab.2.2).

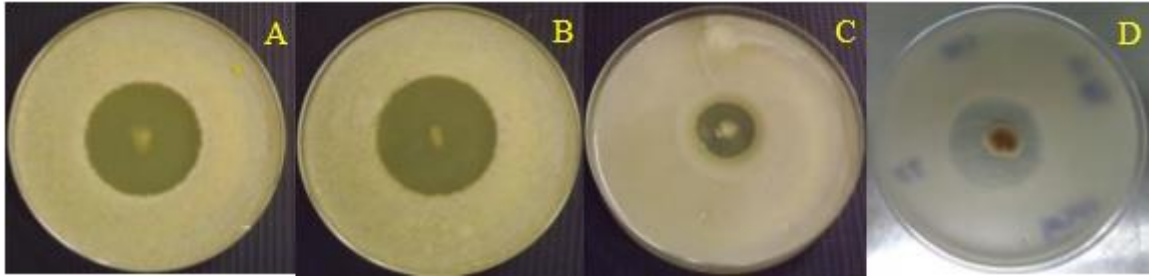


Figura 2.4 Ensaio *overlay* de atividade antibacteriana de algumas das linhagens obtidas a partir do uso de sobrenadante de *P. elgii*. Os halos de inibição são bastante visíveis. (A) Linhagem A3, (B) Linhagem A9, (C) Linhagem C5 e (D) Linhagem C12. Todos os ensaios foram realizados contra *B. subtilis* (Carvalho, 2013).

Tabela 2.2 Resultados de ensaios *overlay* de atividade antibacteriana das linhagens obtidas a partir do uso de sobrenadante de *P. elgii*. Os resultados foram observados após 48 h do inóculo de *B. subtilis*.

Linhagem	Halo de inibição (cm)	Linhagem	Halo de inibição (cm)
A1	2	B7	1
A2	3	B8	2
A3	2	B9	2
A4	3	B10	2
A5	3	B11	3
A6	2	B12	3
A7	2	C1	1
A8	3	C2	1
A9	3	C3	1
A10	3	C4	1
A11	3	C5	1
A12	3	C6	2
B1	2	C7	1
B2	2	C8	1
B3	3	C10	3
B4	3	C11	2
B5	1	C12	2
B6	1		

Dentre as trinta e cinco linhagens, cinco (B2, B4, B7, C3 e C6) foram selecionadas para serem testadas contra *E.coli* e *P. aeruginosa*. A seleção foi feita de modo a incluir ao menos um representante de cada gênero encontrado, exceto *Paenibacillus* e *Enterobacter*. A primeira por se tratar do gênero que originou o sobrenadante utilizado na obtenção das linhagens e a segunda por possuir potencial biotecnológico reduzido em relação as demais, conforme encontrado na bibliografia sobre o gênero. Todas as cinco linhagens testadas tiveram sua atividade antibacteriana confirmada sobre *E. coli* e apenas B7 e C6 apresentaram atividade antibacteriana sobre *P. aeruginosa*. Por esse motivo, B7 e C6, pertencentes aos gêneros *Kitasatospora* e *Methylobacterium*, respectivamente, foram escolhidas para serem caracterizadas genotípica e fenotipicamente neste trabalho e serão a partir de agora chamadas de K002 e K003, respectivamente. A caracterização da linhagem K002 será apresentada capítulo 3. A caracterização da linhagem K003 será apresentada a seguir.

2.4.4 Caracterização fenotípica *in vitro*

As linhagens K003 e *M. radiotolerans* CBMAI 824 apresentaram coloração Gram-negativa, cresceram aerobiamente em temperaturas de 15 a 40°C em meio NB. K003 cresceu em meio NB contendo 0-2% (p/v) de NaCl e em pH 5–9, enquanto CBMAI cresceu em 0-3% e em pH 4-9. O crescimento ótimo de ambas as linhagens ocorreu a 30 °C, pH 8 e com 0% (p/ v) de NaCl (Tab. 2.3). O crescimento a partir do pH 4 e em salinidade 3% distingue CBMAI 824 de K003. Ambas as linhagens foram positivas para catalase e oxidase.

Tabela 2.3 Propriedades fenotípicas comparativas entre a linhagem *Methylobacterium* K003, obtida a partir do sobrenadante de *P. elgii*, e a linhagem de referência *M. radiotolerans* CBMAI 824. +, Positivo, entre parênteses, crescimento ideal.

Faixa de crescimento (ideal):	K003	CBMAI
Temperatura (°C)	15-40 (30)	15-40 (30)
pH	5-9 (8)	4-9 (8)
NaCl (%)	0-2 (0)	0-3 (0)
Catalase	+	+
Oxidase	+	+

O perfil de ácidos graxos das linhagens K003 e CBMAI 824 é apresentado na Tabela 2.4. O ácido graxo mais abundante (>85%) em ambas as linhagens foi *Summed Feature 8* (C18:1 ω 7c e/ou C18:1 ω 6c), já os menos abundantes foram *Summed Feature 2* (possivelmente C12:0 aldeído) em K003 com 0,74% e C12:0 em CBMAI 824 com 0,67%. *Summed feature* ocorre quando há uma separação imperfeita do pico. Os perfis de ácidos graxos separaram claramente a linhagem K003 de CBMAI 824 pela ausência de C12:0 e 18:0 3OH.

Tabela 2.4 Composição de ácidos graxos (%) de forma comparativa entre a linhagem *Methylobacterium* K003, obtida a partir do uso do sobrenadante de *P. elgii*, e sua linhagem de referência *M. radiotolerans* CBMAI 824. O perfil de ácidos graxos foi obtido por cromatografia gasosa. *O *Summed Feature* ocorre quando há separação de pico imperfeita. O *Summed Feature 2* representa possivelmente aldeído 12:0, o *Summed Feature 3* representa C16:1 ω 6c e/ou C16:1 ω 7c e o *Summed Feature 8* representa C18: 1 ω 7c e/ou C18: 1 ω 6c. –, ausente.

Ác. graxo	Nome	Tipo	K003 (%)	CBMAI (%)
C12:0	ácido dodecanóico	Saturado de cadeia média	-	0.67
C16:0	ácido hexadecanóico	Saturado de Cadeia Longa	2.37	2.81
C18:0	ácido octadecanóico	Saturado de Cadeia Longa	5.75	5.59
18:0 3OH	ácido 3-hidroxiocetadecanóico	Saturado de cadeia longa com hidroxila	-	0.84
<i>Summed Feature 2*</i>	-	-	0.74	0.87
<i>Summed Feature 3*</i>	-	-	0.76	0.90
<i>Summed Feature 8*</i>	-	-	90.39	88.33

Dentre os vinte e seis antibióticos testados quanto a suscetibilidade das linhagens, K003 e CBMAI 824 foram resistentes à Amoxicilina/Ác. Clavulânico, Aztreonam, Cefazolina, Cefepima, Clindamicina, Meropenem, Oxacilina, Polimixina B e Sulfonamida. Além destes, K003 também foi resistente a Vancomicina. A tabela 2.5 apresenta os resultados do teste de suscetibilidade das duas linhagens aos antibióticos avaliados.

Tabela 2.5 Resultados do teste de suscetibilidade a antibióticos da linhagem *Methylobacterium* K003, obtida a partir do sobrenadante de *P. elgii*, e sua linhagem de referência *M. radiotolerans* CBMAI 824. O método adotado foi o de Kirby-Bauer (difusão em disco), onde são utilizados discos de papel filtro impregnados com uma concentração padrão de antibiótico e observada a formação de um halo de inibição. +, Suscetível; -, Resistente.

Antibiótico (µg per disco)	K003	CBMAI 824
Amicacina (30)	+	+
Amoxicilina/Ac. Clavulânico (20/10)	-	-
Ampicilina (10)	+	+
Aztreonam (30)	-	-
Cefalexina (30)	+	+
Cefazolina (30)	-	-
Cefepima (30)	-	-
Ceftazidima (30)	+	+
Ceftriaxona (30)	+	+
Ciprofloxacina (5)	+	+
Clindamicina (2)	-	-
Cloranfenicol (30)	+	+
Doxiciclina (30)	+	+
Eritromicina (15)	+	+
Gentamicina (10)	+	+
Imipenem (10)	+	+
Meropenem (10)	-	-
Norfloxacina (10)	+	+
Ofloxacina (5)	+	+
Oxacilina (1)	-	-
Penicilina G (10)	+	+
Piperacilina + Tazobactam (110)	+	+
Polimixina B (300)	-	-
Sulfonamida (300)	-	-
Tetraciclina (30)	+	+
Vancomicina (30)	-	+

2.4.5 Sequenciamentos e análises filogenéticas

O sequenciamento Sanger do *rRNA* 16S da linhagem K003 permitiu a identificação da sequência, onde o BLAST indicou maior similaridade (99,88%) com *Methylobacterium radiotolerans* JCM 2831. O banco de dados Ezbiocloud conferiu a mesma linhagem mais próxima com 99,35% de similaridade. A árvore filogenética baseada no gene marcador *rRNA* 16S de K003 está representada na Figura 2.5.

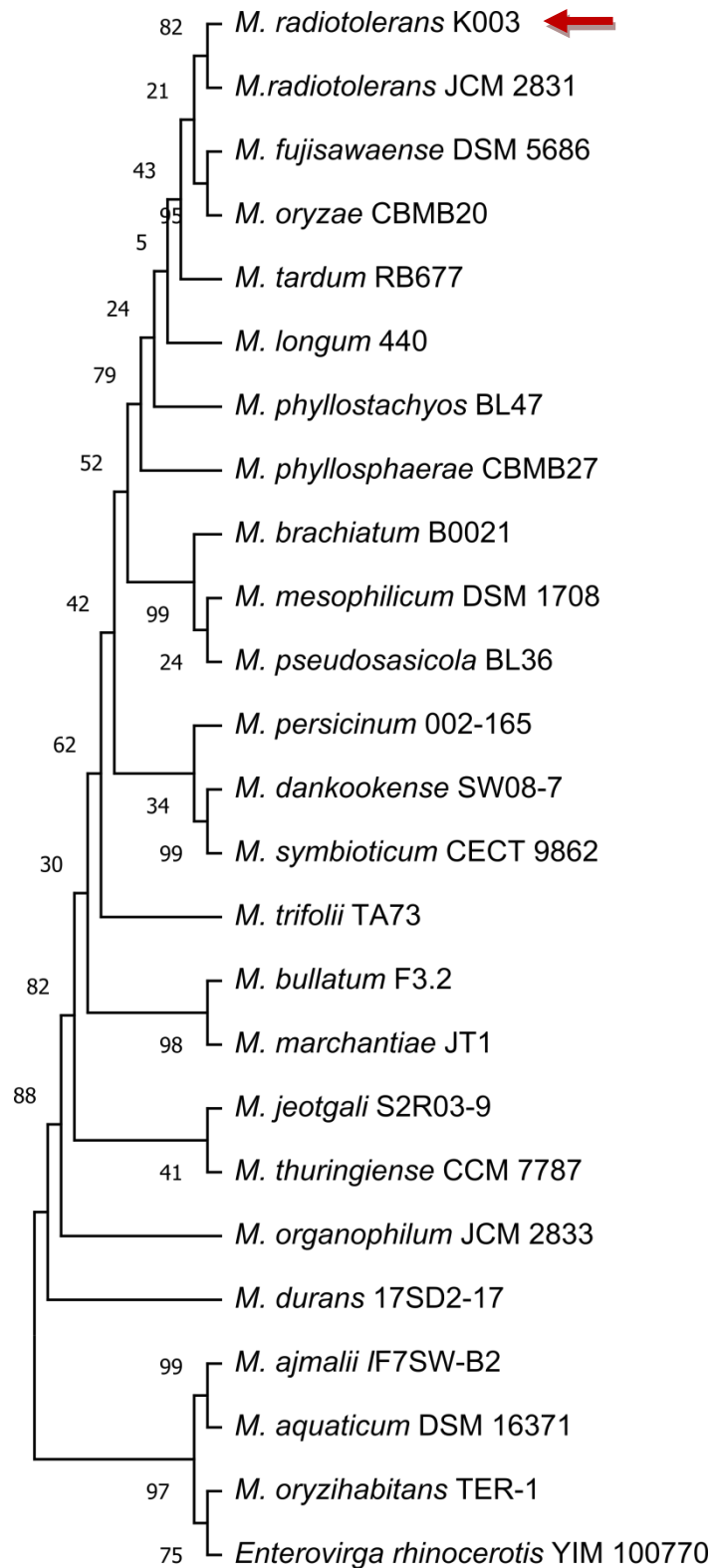


Figura 2.5 Árvore filogenética baseada na seqüência parcial do gene 16S ribossomal *rRNA* obtido por clonagem e sequenciamento de Sanger da linhagem *Methylobacterium* K003, obtido a partir da utilização do sobrenadante de *P. elgii*. A posição evolutiva da linhagem foi estimada pela construção de uma árvore filogenética usando o método *Maximum-likelihood (ML)* com 1000 repetições de amostragem (*Bootstrap*), usando seqüências representativas de 16S ribossomal *rRNA* de linhagens válidas do gênero listadas na *List of Prokaryotic names with Standing in Nomenclature - LPSN*.

O genoma da linhagem K003 foi montado usando uma combinação de abordagens de sequenciamento de leitura longa e curta, resultando em uma montagem 7.735.228 pb, distribuídos em 5 *contigs*, com Teor de GC de 70,84%. A anotação no *software* Prokka *Galaxy Version 1.14.5* mostrou que o genoma de K003 possui 9 *rRNA*, 73 *tRNA* (*RNA* transportador) e 7292 *CDS* (sequências codificadoras). (Fig. 2.6 e Tab. 2.6).

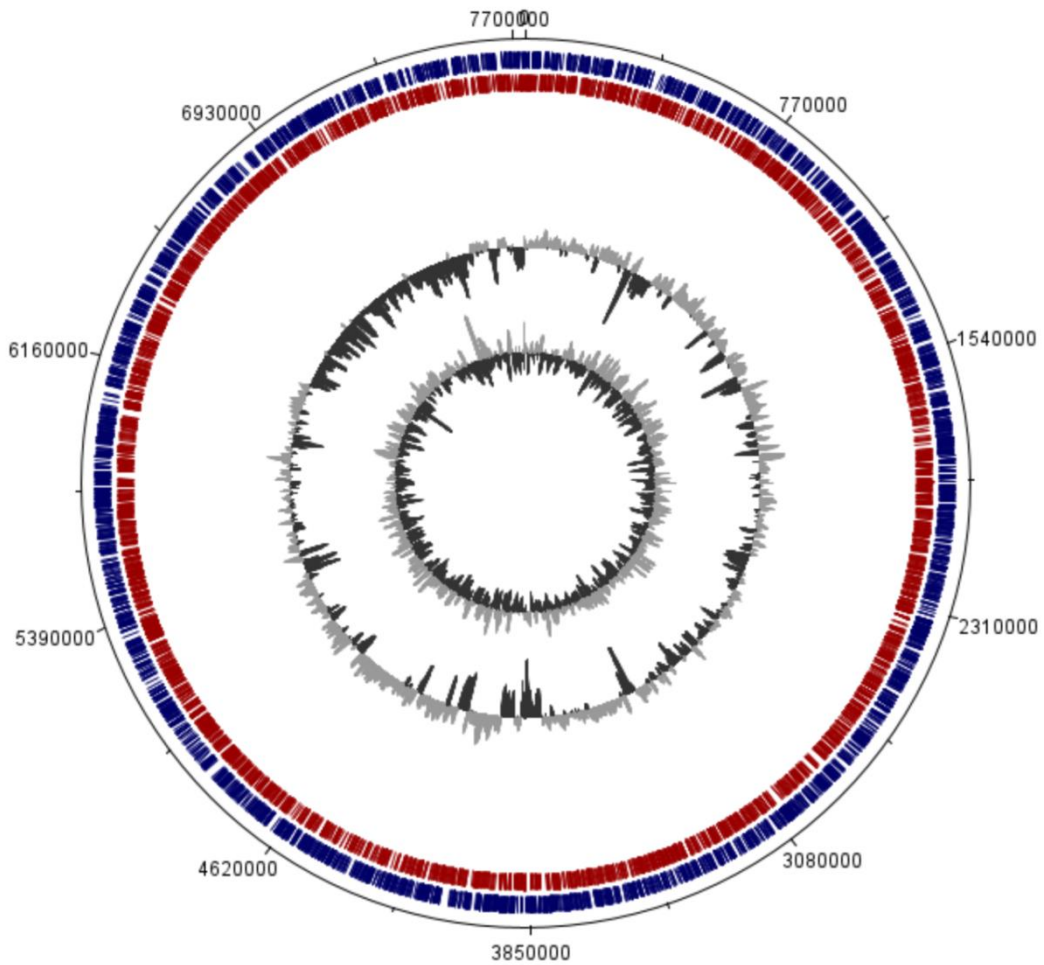


Figura 2.6 Mapa circular da montagem híbrida do genoma da linhagem *Methylobacterium* K003, obtida a partir do sobrenadante de *P. elgii*. A ilustração foi gerada pelo recurso DNAPlotter do *software* Artemis demonstrando as regiões codificantes e o conteúdo GG. De fora para o centro: *CDS* direto (azul escuro), *CDS* reverso (vermelho), conteúdo de GC e inclinação de GC (cinza escuro e cinza claro).

Tabela 2.6 Dados da montagem híbrida do genoma de *Methylobacterium* K003, obtida a partir do sobrenadante de *P. elgii*. A montagem foi realizada a partir dos sequenciamentos de leitura longa e curta através do programa Unicycler. A anotação foi realizada com o *software* Prokka e os parâmetros de qualidade foram avaliados nos *softwares* CheckM, Bowtie2 e Quast.

Dados da montagem	
<i>Contigs</i>	5
<i>CDS</i>	7292
<i>rRNA</i>	9
<i>Repeatregion</i>	0
<i>tRNA</i>	73
<i>TmRNA</i>	2
Teor de GC (%)	70.84
Contaminação (%)	2.46
Compleitude (%)	100
Cobertura	400x
L50	1
N50 (pb)	5241738
Tamanho Total (bp)	7735228
Maior <i>contig</i>	5241738

Na plataforma TYGS – DSMZ, a avaliação do genoma de K003 confirmou que a linhagem é uma *Methylobacterium radiotolerans* com *dDDH* de 90.3-94%. As árvores filogenéticas baseadas no gene marcador 16S *rRNA* e no genoma calculadas pelo TYGS estão apresentadas nas figuras 2.7 e 2.8, respectivamente.

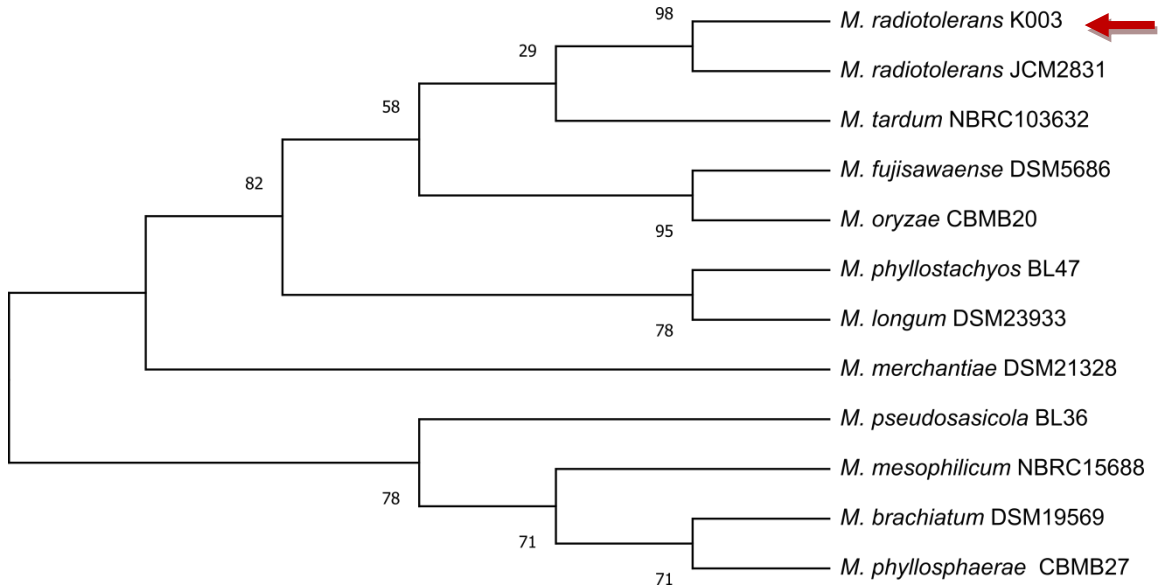


Figura 2.7 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de *Methylobacterium* K003, linhagem obtida a partir do sobrenadante de *P. elgii*, baseada na sequência do gene ribossomal *rRNA* 16S.

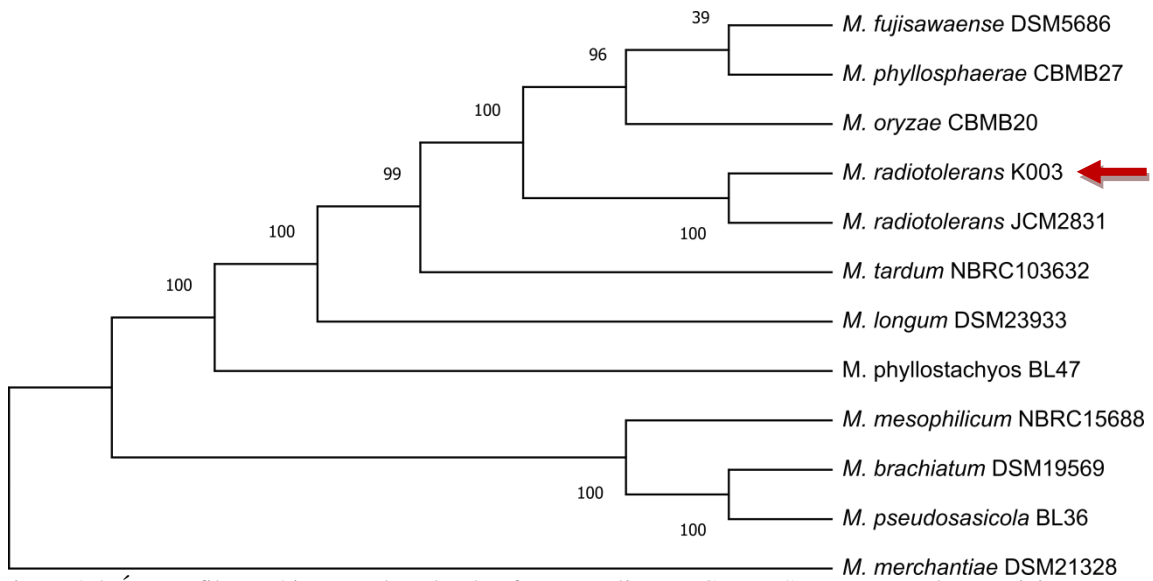


Figura 2.8 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de *Methylobacterium* K003, linhagem obtida a partir do sobrenadante de *P. elgii*, baseada em seu genoma total.

As ANIs calculadas para o genoma da linhagem K003 na plataforma JspeciesWS foram 97.51% (ANI_b) e 99.16% (ANI_m), ambas em relação a *Methylobacterium radiotolerans*. No software GTDB-tk, a ANI foi de 99.01% (Tab. 2.7). Esses valores de ANI > 95%, juntamente com o *dDDH* > 70%, citado anteriormente, confirmam a identidade da linhagem K003.

Tabela 2.7 Identidade Média de Nucleotídeos (ANI) calculadas para a linhagem *M. radiotolerans* K003, obtida a partir do uso do sobrenadante de *P. elgii*. *ANI_b*: ANI com base no cálculo BLAST calculada na plataforma JSpeciesWS. *ANI_m*: ANI com base no cálculo MUMmer calculada na plataforma JSpeciesWS. *ANI*: calculada nosoftware GTDB-Tk.

	JSpeciesWS		GTDB-tk
	<i>ANI_b</i> (%)	<i>ANI_m</i> (%)	<i>ANI</i> (%)
<i>M. radiotolerans</i> K003	97.51	99.16	99.01

O alinhamento dos genomas de K003 e sua linhagem mais próxima *M. radiotolerans* JCM 2831 mostra poucas regiões desalinhadas (16.95%) indicando que os genomas são semelhantes entre si (Fig. 2.9). Apenas os blocos de alinhamento do genoma com pelo menos 1.000 nucleotídeos e 90% de identidade foram considerados.

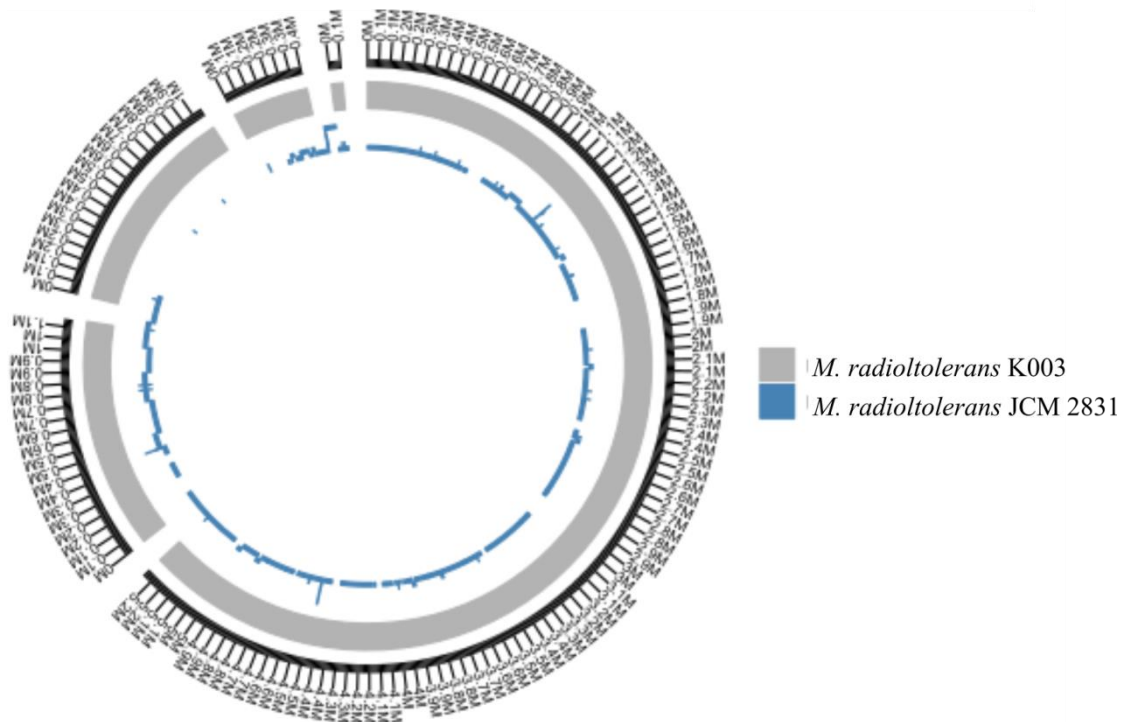


Figura 2.9 Visualização circular do alinhamento dos genomas da linhagem *M. radiotolerans* K003 (em cinza) e sua linhagem mais próxima de acordo com TYGS-DSMZ, *M. radiotolerans* JCM 2831 (em azul). O alinhamento e a ilustração foram gerados usando o Plot Nucmer. Apenas os blocos de alinhamento do genoma com pelo menos 1000 nucleotídeos e 90% de identidade são mostrados. A banda externa é a escala de tamanho dos *contigs*. Quando a banda azul tem duas camadas de profundidade, há duas sequências na linhagem mais próxima que se sobrepõem no mesmo local na linhagem *M. radiotolerans* K003.

Expandindo a análise de genômica comparativa para investigar as relações filogenéticas desse gênero, utilizamos as sequências do genoma de 12 linhagens (K003, 10 linhagens mais próximas e 1 outgroup) para extrair conjuntos de genes. Através do servidor *web* M1CR0B1AL1Z3R, um total de 233 núcleos de genes (subconjunto de genes compartilhados entre as 12 linhagens) foram recuperados e usados para construir uma árvore filogenética de máxima verossimilhança (Fig. 2.10).

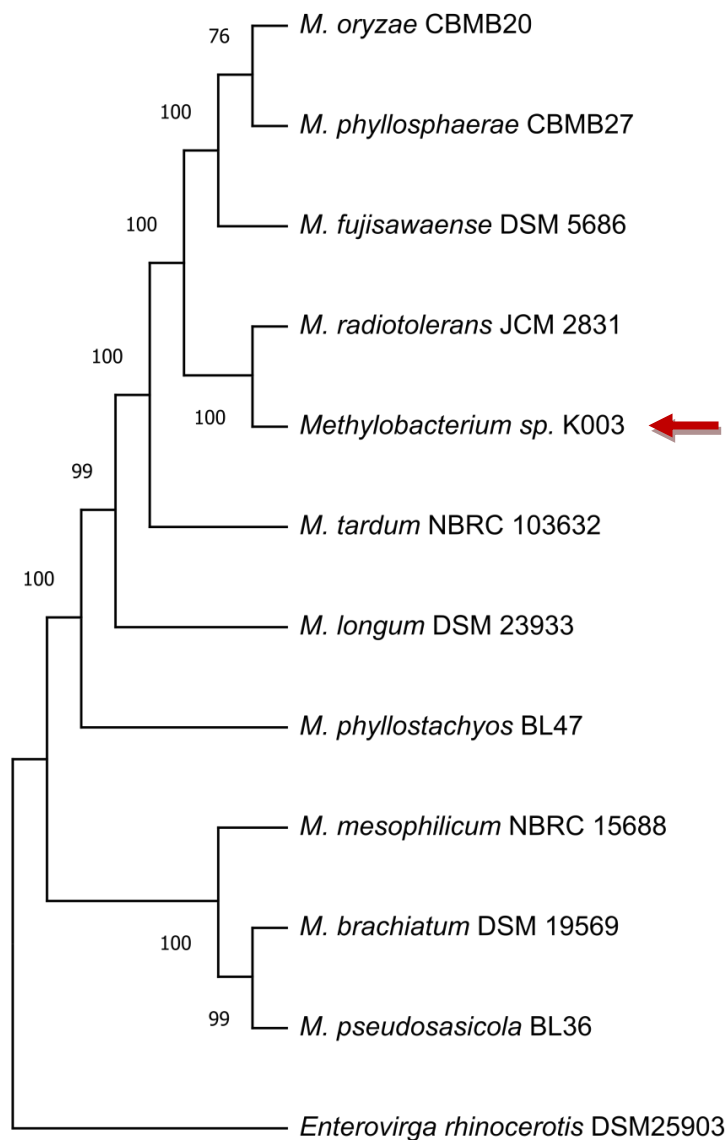


Figura 2.10 Árvore filogenética de máxima verossimilhança baseada em 233 genes ortólogos principais concatenados de sequências genômicas de 12 linhagens de *Methylobacterium*, incluindo *M. radiotolerans* K003, as linhagens intimamente relacionadas de acordo com o TYGS-DSMZ e o grupo externo *Enterovirga rhinocerotis*. O servidor *web* M1CR0B1AL1Z3R foi utilizado para concatenar os genes ortólogos e construir a árvore filogenética.

Foram possíveis identificar 21379 famílias gênicas compondo o pangenoma de K003 e suas linhagens mais próximas e, dessas, 2806 (13.12%) estavam presentes no núcleo do genoma e as restantes como acessórias. Os genes raros foram representados por 12387 famílias, ou seja, 57.7% do pangenoma (Fig. 2.11).

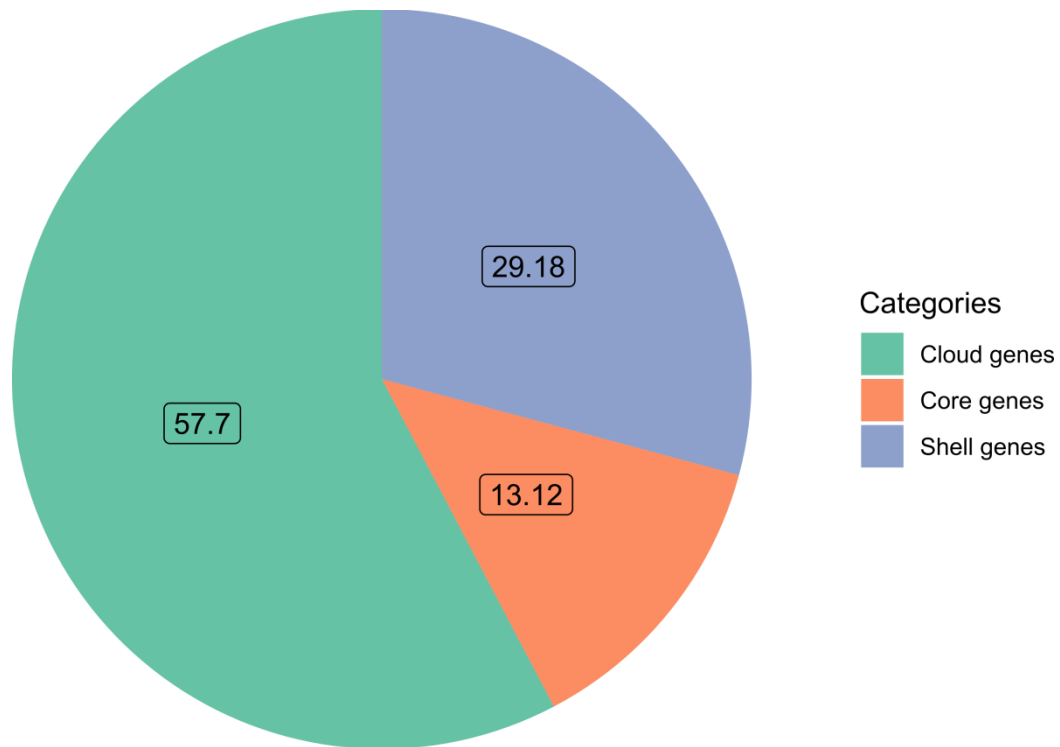


Figura 2.11 Análise comparativa do pangenoma realizada no Roary. O genoma de *M. radiotolerans* K003 e os 10 genomas mais próximos a ele, de acordo com TYGS-DSMZ, foram utilizadas. Uma identidade percentual mínima para BLASTp de 80% foi definida para agrupar os genes que codificam sequências de proteínas completas em genomas centrais.

Comparando apenas as linhagens K003 e JCM 2831, foram possíveis identificar 8455 famílias gênicas compondo o pangenoma e, dessas, 5332 (63.06%) estavam presentes no núcleo do genoma e as restantes como acessórias. Não foram encontrados genes específicos nas linhagens.

2.4.6 Características funcionais dos genomas

A figura 2.12 apresenta a distribuição das proteínas do genoma da linhagem K003 nas suas respectivas categorias *COGs*. A categoria S refere-se a proteínas de função desconhecida e representa 19% das proteínas do genoma de K003. A categoria *unknow*, representa 15%. Dentre as categorias bem anotadas, E (9%), K (6%) e C (6%), referentes ao metabolismo de aminoácidos, transcrição e produção e conversão de energia, respectivamente, estão entre as

mais abundantes no genoma. Entre as menos abundantes estão as categorias D (0.7%) e N (0.6%) responsáveis pelo ciclo celular e motilidade celular.

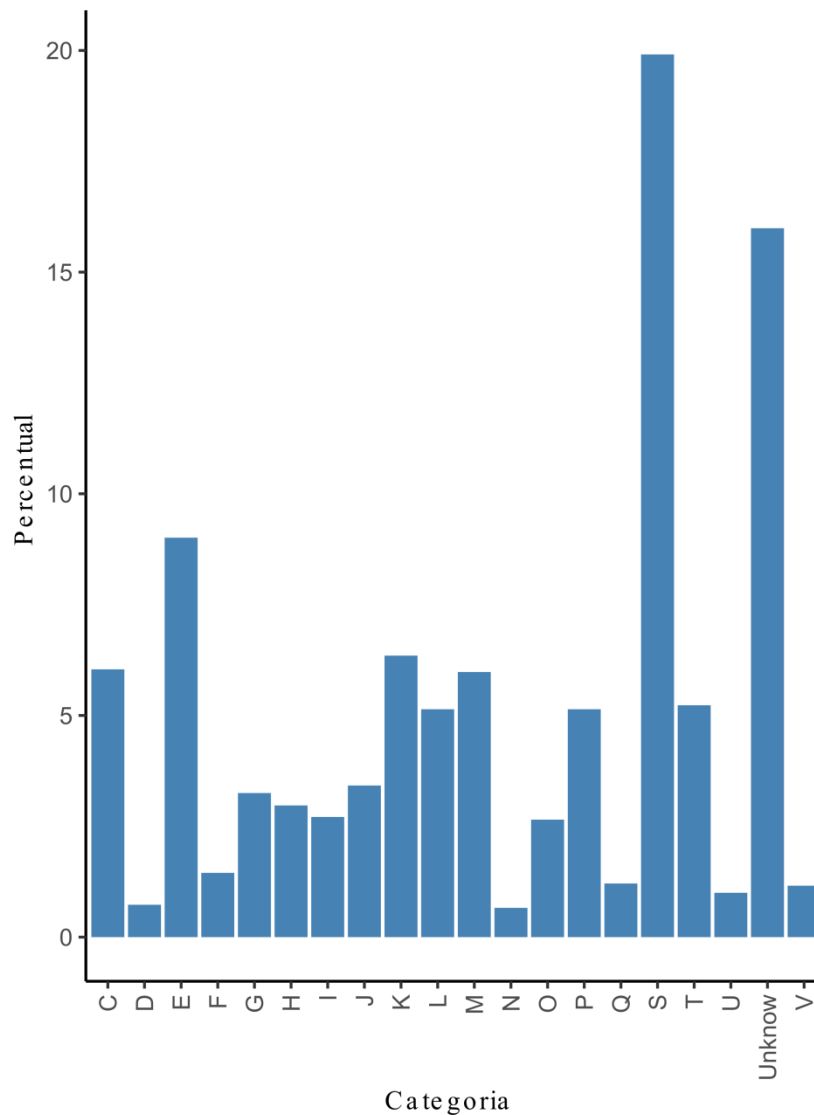


Figura 2.12 Proteínas preditas no genoma da linhagem *M. radiotolerans* K003, obtidas a partir do sobrenadante de *P. elgii*, classificadas de acordo com as categorias COGs. A classificação foi realizada usando o mapeador eggNOG contra o banco de dados da *Kyoto Encyclopedia of Genes and Genomes* (KEGG).

No que diz respeito a resistência antibióticos, foram encontradas quatro bombas de efluxo de antibiótico da família resistência-nodulação-divisão celular (RND) tanto em K003 quanto em JCM 2831, todas conferindo resistência aos antibióticos fluoroquinolona e tetraciclina. As similaridades variaram de 41 a 71% (Tab. 2.8).

Tabela 2.8 Anotação de genes potenciais responsáveis pela resistência a antibióticos da linhagem *M. radiotolerans* K003, obtida a partir do uso do sobrenadante de *P. elggi*. A análise foi realizada em comparação com o Banco de Dados de Genes de Resistência a Antibióticos (ARDB).

Gene	Família	Classe de droga	Mecanismo de resistência	Similaridade (%)
adeF	resistência-nodulação-divisão celular (RND)	fluoroquinolona e tetraciclina	Bomba de efluxo	44.59
adeF	resistência-nodulação-divisão celular (RND)	fluoroquinolona e tetraciclina	Bomba de efluxo	51.05
adeF	resistência-nodulação-divisão celular (RND)	fluoroquinolona e tetraciclina	Bomba de efluxo	71.52
adeF	resistência-nodulação-divisão celular (RND)	fluoroquinolona e tetraciclina	Bomba de efluxo	41.93

Os cinco *contigs* da montagem do genoma de *M. radiotolerans* K003 foram analisados quanto a presença de CRISPR. Foram encontrados CRISPR em dois *contigs* (1 e 3) e um caso Tipo IIIA sem matrizes CRISPR no *contig* 5 (Tab.2.9).

Tabela 2.9 Resultados da análise da presença de locus CRISPR na linhagem *M. radiotolerans* K003, obtida a partir do uso do sobrenadante de *P. elggi*. A análise foi realizada através da ferramenta de reconhecimento CRISPRCasFinder.

<i>Contig 1</i>			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CRISPR	23047	23136	1
CRISPR	65942	66033	1
CRISPR	379399	379492	1
CRISPR	417660	417754	1
CRISPR	3706204	3706583	6
CRISPR	5156346	5156465	1
<i>Contig 3</i>			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CRISPR	212125	212278	2
CRISPR	212489	212581	1
CRISPR	213350	213810	7
CRISPR	213951	214041	1
CRISPR	214554	214884	5
CRISPR	276689	276901	3
<i>Contig 5</i>			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CAS-Tipo IIIA	48893	49159	1

A anotação funcional dos genomas no Antismash mostrou um total de dez (10) agrupamentos gênicos para K003 (Tab. 2.10). A maioria desses agrupamentos gênicos estão relacionados à produção de compostos antibacterianos. A tabela 2.8 apresenta tais *clusters* em comparação com a linhagem JCM 2831 (12 clusters). As linhagens K003 e JCM 2831 são

bem similares, distinguidas apenas pela ausência de T3PKS em K003 e por essa possuir um *cluster* a menos de Homoserina lactona.

Tabela 2.10 Resultados da análise de *clusters* de genes biossintéticos para metabólitos secundários da linhagem *M. radiotolerans* K003, obtida a partir do uso do sobrenadante de *P. elgii*. A análise foi realizada por meio do *software* AntiSMASH (v. 6.1.1) e apresentada de forma comparativa com alinhagem *M. radiotolerans* JCM 2831, acrônimo de CBMAI 824.

Metabólitos secundários		
Tipo	Nº clusters	
	K003	JCM 2831
Homoserina lactona	2	3
NAPAA	1	1
NRPS	1	1
Redox-cofator	1	1
RIPP	1	1
T1PKS	1	1
T3PKS	0	1
Terpeno	3	3
Outros	0	0
TOTAL	10	12

De acordo com a análise no BiG-SCAPE, entre *BGCs* detectados para as linhagens de *M. radiotolerans* K003 (10) e JCM2831 (12), nove deles são semelhantes entre as ambas as linhagens. São eles: 3 terpenos, 1 *NRPS*, 1 homoserina lactona, 1 NPAA, 2 *RIPP* e 1 *T1PKS*. Entre os *BGCs* que distinguem as linhagens estão 1 homoserina lactona em K003 e 2 homoserina lactona e 1 *T3PKS* em JCM 2831 (Fig. 2.13).

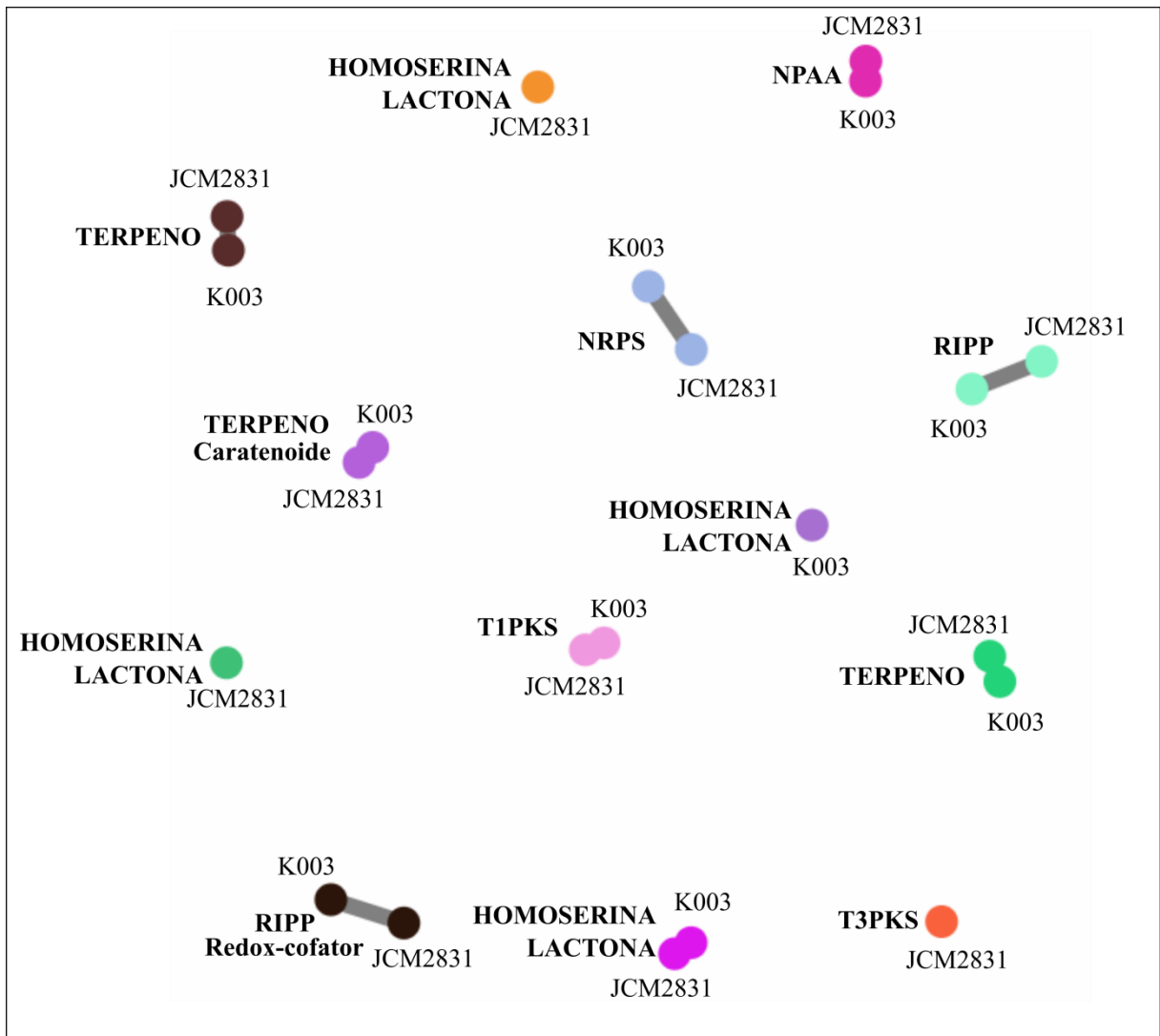


Figura 2.13 Análise de similaridade entre as famílias de *clusters* de genes usando BiG-SCAPE. Cada união representa *clusters* gênicos semelhantes entre as linhagens *M. radiotolerans* K003 e JCM2831. Os *clusters* específicos, ou seja, sem semelhanças entre as linhagens também são apresentados em forma isolada. Não há representação gráfica dos *clusters* gênicos semelhantes por se tratarem de *clusters* idênticos.

O banco de dados de dados CAMP_{R4} para predição de peptídios antimicrobianos (AMPs) mostrou 40 predições em K003. As predições apresentaram probabilidade superior a 0.68, com maior probabilidade de 0.75.

2.4.7 Caracterização fenotípica *in silico*

A figura 2.14 apresenta a comparação do fenótipo de K003 e suas linhagens mais próximas de acordo com DSMZ. A linhagem K003 se diferencia das demais por ser positiva para D-sorbitol. K003 compartilha várias características com *M. radiotolerans* JCM 2831, sua linhagem mais próxima segundo o TYGS-DSMZ e acrônimo a CBMAI, entretanto, K003 se

distingue por ser negativa para capnofílica e produção de coagulase e por ser positiva para D-sorbitol e L-arabinose.

De acordo com a predição fenotípica no Traitair, a linhagem mais próxima a K003 é *M. brachiatum* DSM 19569. A linhagem K003 é distinguida por ser negativa para crescimento a 42°C, capnofílica e produção de coagulase e por ser positiva para D-sorbitol.

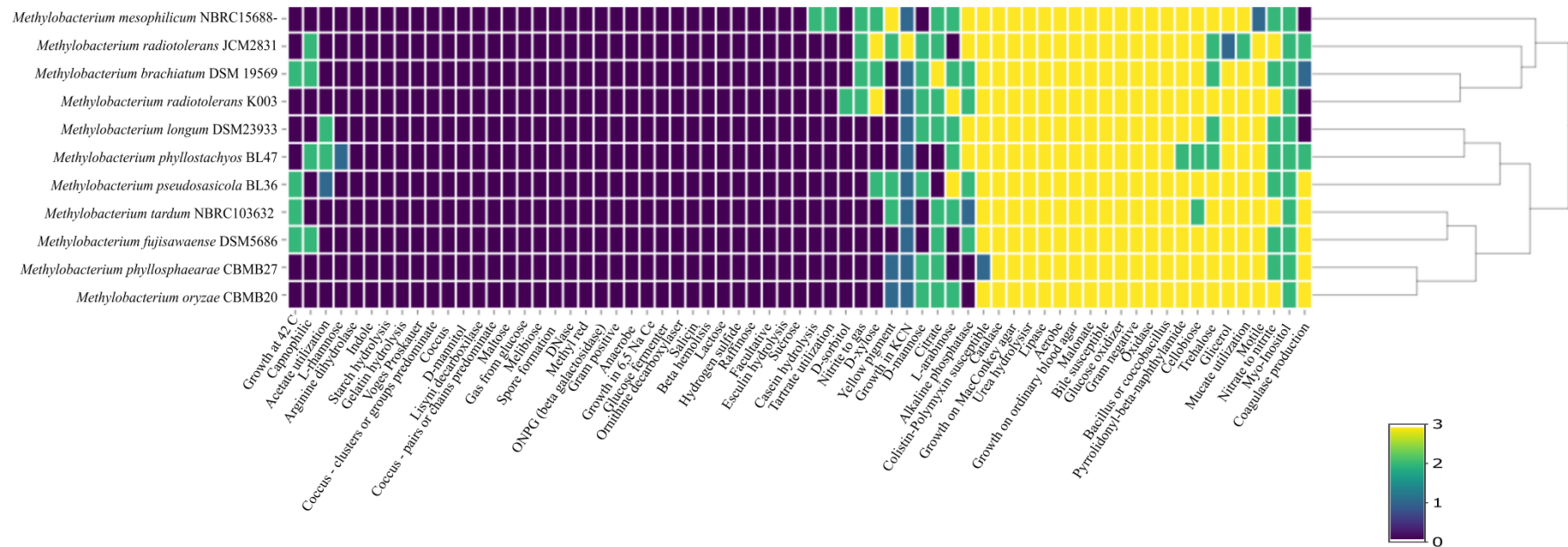


Figura 2.14 Comparação entre as predições fenotípicas da montagem híbrida do genoma da linhagem *M. radiotolerans* K003, obtida a partir do uso do sobrenadante de *P. elgii*, e dos genomas de suas linhagens mais próximas de acordo com o TYGS-DSMZ obtidas no banco de dados do *NCBI*. O valor zero significa que a linhagem é negativa para o fenótipo, os valores 1 e 2 correspondem ao positivo para o fenótipo nos preditores phypat e phypat+PGL, respectivamente. O valor 3 mostra que o fenótipo foi positivo nos dois preditores. A análise foi realizada pelo *software* Traitair.

2.5 DISCUSSÃO

Algumas bactérias só conseguem crescer em condições de laboratório quando em co-cultivo com outro membro da comunidade, também chamada de linhagem auxiliar, que pode ser uma bactéria filogeneticamente diferente ou até mesmo um organismo diferente (Boilattabi *et al.*, 2021). A co-cultura pode ser obtida por cultura direta da linhagem auxiliar ou usando seus sobrenadantes (Stewart, 2012). O uso do sobrenadante de *P. elgii* enriquece o cultivo com moléculas indutoras e compostos antimicrobianos que podem simular as relações químico-ecológicas que ocorrem naturalmente. Isso pode aumentar a recuperação microbiana de amostras ambientais e emergir como uma abordagem eficaz e direcionada para encontrar novos compostos bioativos com atividade contra os patógenos para os quais novos antibióticos são necessários (Turan *et al.*, 2017; van Bergeijk *et al.*, 2020).

Usando uma abordagem baseada no cultivo selecionado a partir do sobrenadante de *P. elgii*, foram isoladas e identificadas com sucesso trinta e cinco linhagens cultiváveis de crescimento rápido, a maioria de gêneros conhecidos por possuir potencial biotecnológico. São eles *Paenibacillus*, já exposto aqui por ser fonte produtora do sobrenadante, *Pseudomonas*, *Enterobacter*, *Kitasatospora*, *Sphingomonas*, *Xanthobacter*, *Burkholderia* que serão expostos a seguir e *Methylobacterium*, cuja a linhagem foi caracterizada neste capítulo para ilustrar o potencial da metodologia em selecionar micro-organismos com possível aplicação biotecnológica.

Espécies pertencentes ao gênero *Sphingomonas* possuem propriedades que vão desde a biorremediação de contaminações ambientais até a produção de fitohormônios altamente benéficos. Sua atividade antibacteriana foi comprovada por Akinsanya e colaboradores (2017) contra *S. Typhimurium*, *P. vulgaris*, *K. pneumoniae*, *E. coli*, *S. aureus*, *S. pyogenes*, *B. cereus*, e *C. albicans*. Estudos recentes têm mostrado o papel de espécies de *Sphingomonas* na degradação de compostos organometálicos (Asaf *et al.*, 2020).

As bactérias endofíticas do gênero *Xanthobacter* são conhecidas por serem promotoras do crescimento vegetal. Elas tem potencial interessante como biofertilizante, no controle biológico de ervas daninhas que competem com o cultivo de plantas e tem sido relatada como ferramenta biológica útil para a biorremediação de águas e solos poluídos por agentes químicos (Sánchez-Yañez, 2022).

O gênero *Kitasatospora* é caracterizado como produtor de metabólitos secundários de estruturas variadas, que são ativos inclusive contra organismos eucarióticos (Takahashi,

2017), além de apresentarem atividade contra patógenos Gram-positivos e Gram-negativos (Zhu *et al.*, 2014) e resistência a diversos metais pesados (Yun *et al.*, 2020).

O gênero *Burkholderia* é um grupo mais conhecido pelo seu papel como patógeno. No entanto, as diversas interações ambientais dessas bactérias estão agora apontando para múltiplas propriedades benéficas que se estendem desde suas capacidades na biorremediação e no controle biológico, até seu significativo potencial biotecnológico como produtores de antimicrobianos (Depoorter *et al.*, 2016; Wang *et al.* 2020).

Embora o gênero *Enterobacter* seja mais conhecido como fitopatógeno (Singh *et al.*, 2018), ele já foi mencionado como um bioinoculante eficaz para arroz sob estresse de salinidade alcalina (Sagar *et al.*, 2020) e por possuir capacidade de degradação de polietileno (Ren *et al.*, 2019).

No que diz respeito ao potencial biotecnológico de *Pseudomonas*, membros representativos da espécie foram identificados como promotores do crescimento de plantas (Glick, 2012), agentes de biorremediação (Dvořák *et al.*, 2017) e produtores de compostos naturais como ramnolipídios, terpenoides, policetídeos e peptídeos não ribossomais (Loeschcke e Thies, 2015), além de biopolímeros (Mozejko-Ciesielska *et al.*, 2019; Salvachúa *et al.*, 2020).

Com o resultado positivo dos testes antibacterianos feitos no sobrenadante de *P. elgii* e nas linhagens obtidas através do seu uso, comprova-se a capacidade do protocolo em selecionar linhagens com atividade semelhante ao sobrenadante. O sobrenadante de *P. elgii* é rico em sideróforos e lactonas autoindutoras, além de peptídeos de dupla propriedade, autoindutora e antibacteriana (Ortega *et al.* 2018), o que pode justificar o estímulo de crescimento nas diferentes concentrações do sobrenadantes de *P. elgii*, a diversidade de gêneros encontrada no cultivo e a seleção de crescimento direcionada a linhagens com potencial biotecnológico.

O resultado deste estudo incluiu a caracterização da linhagem K003 e um rascunho do seu genoma explorado através da mineração de genoma. Essa abordagem combinatória levou a resultados promissores para a produção de novos metabólitos antimicrobianos.

Methylobacterium, gênero ao qual pertence a linhagem K003, contém mais espécies do que qualquer outro gênero dentro da família *Methylobacteriaceae*, ordem Rhizobiales, classe Alphaproteobacteria (Kelly *et al.*, 2014). Tal gênero compreende 69 espécies com publicações válidas (<https://lpsn.dsmz.de/genus/methylobacterium> em 18/01/2023) e tem *Methylobacterium organophilum* como a espécie-tipo (Green, 2006).

As espécies de *Methylobacterium* são bacilos Gram-negativos aeróbios fastidiosos. O gênero abriga espécies de pigmentação rosa e metilotróficas facultativas, isto é, que têm a capacidade de crescer em metano ou metanol como fonte de carbono e energia, além de açúcares e ácidos orgânicos (Green e Bousfield, 1982).

Methylobacterium é um gênero de ampla distribuição em vários tipos de habitats incluindo ar (Park *et al.*, 2018), solo (Veyisoglu *et al.*, 2013), água doce (Gallego *et al.*, 2005) e no interior das plantas (Dourado *et al.*, 2012). Linhagens pertencentes à família *Methylobacteriaceae* foram isoladas de diferentes locais na Estação Espacial Internacional (ISS), em um experimento de rastreamento microbiano. Uma delas identificada como *Methylorubrum rhodesianum* e as outras classificadas como a nova espécie, *Methylobacterium ajmalii* (Bijlane, 2021).

Membros do gênero *Methylobacterium* podem ter uma associação benéfica com plantas e estudos mostram que exemplares do gênero fixam nitrogênio (Sy *et al.*, 2001), produzem celulase (Jayashree *et al.*, 2011) e interagem com outros patógenos das plantas induzindo resistência sistêmica (Ardanov *et al.*, 2012). Atividades contra patógenos também são relatadas para endófitos do gênero *Methylobacterium* (Wang, 2019). Assim, há um interesse contínuo em desenvolver drogas a partir de compostos endofíticos.

Extratos brutos e parcialmente purificados do gênero *Methylobacterium* demonstram possuir antimicrobiano, anticâncer e propriedades antioxidantes (Alamgir *et al.*, 2015, Balachandran *et al.*, 2012). Alguns relatos confirmaram o efeito antimicrobiano proveniente da produção de peptídeos antimicrobianos do tipo defensina (Tejesvi *et al.*, 2016). A atividade antimicrobiana (bactéria e fungos) e antioxidante foi comprovada para *M. radiotolerans* (Photolo *et al.*, 2020).

Linhagens de *Methylobacterium sp.* apresentaram resistência contra radiações. No entanto, as informações publicadas sobre essa resistência são relativamente limitadas (Kang e Srinivasan, 2018).

A linhagem K003 obtida neste estudo, apesar de ser tratar de uma espécie já descrita apresenta características próprias. Embora seja semelhante a sua linhagem de referência (CBMAI 824) quanto ao alinhamento de seus genomas, pangenoma e a diversidades *clusters* biossintéticos, K003 difere de CBMAI em crescimento em pH apenas partir de 5, ausência de crescimento em 3% de salinidade, ausência de certos ácidos graxos, não resistência a vancomicina, presença de enzimas cas Tipo IIIA, negativa para produção de coagulase e positiva para D-sorbitol e L-arabinose.

Comparando-se os resultados alcançados neste trabalho com o estudo de Knief e colaboradores (2011) com doze linhagens de *Methylobacterium*, o principal ácido graxo abundante é o mesmo, *Summed feature 8*. A linhagem K003 tolerara maior nível de salinidade (2%) em comparação as linhagens apresentadas nesse mesmo estudo, onde as linhagens suportaram até 1% de salinidade. No que diz respeito ao pH, todas as linhagens referidas cresceram em geral entre pH 4-5 e 8.

Alessa e colaboradores (2021) em seu estudo com sessenta e duas linhagens de *Methylobacterium* observou que o tamanho do genoma das linhagens varia de 4.4 a 8.8 Mb com *CDS* variando de 4171 a 8622 e GC% variando de 65.9 a 72.7%. Esses dados mostram que a linhagem K003 está dentro do esperado para o conteúdo GC (70.84%) e entre os maiores genomas do gênero (7.7 Mb).

Embora o teste *in vitro* tenha mostrado que K003 e CBMAI 824 sejam suscetíveis a tetraciclina, o resultado anotação de genes potenciais responsáveis pela resistência a antibióticos no *ARDB* detectou quatro bombas de efluxo de antibiótico da família resistência-nodulação-divisão celular que conferem potencial resistência a tetraciclina, além da fluoroquinolona.

O alto percentual (34%) das categorias S e Unknow em K003 reflete o atual nível de conhecimento sobre o proteoma desses micro-organismos (Galperin *et al.*, 2021) e apontam uma das limitações da anotação de genomas com referência a bancos de dados que se baseiam em prospecções apenas de estruturas já descritas.

Alessa e colaboradores (2021) verificaram que os genomas de suas linhagens eram altamente enriquecidos no processo de metabolismo o que condiz com o observado em K003, onde a categoria E relativa ao metabolismo de aminoácidos é a mais abundante dentre as categorias descritas.

A presença de sistema CRISPR em K003 representa um mecanismo de defesa de procariotos contra invasores como vírus e plasmídeos (Wimmer e Beisel 2020; Qiu *et al.*, 2016). Esse fator é interessante do ponto de vista biotecnológico devido ao seu poder de edição do *DNA* que favorece a engenharia genética (Sedeek *et al.*, 2019).

Muitas classes de produtos naturais vêm sendo identificadas com os terpenoides, os alcalóides, os policetídeos e os peptídeos não ribossomais (Arnison *et al.*, 2013). Alguns deles serão explorados a seguir por possuírem *clusters* responsáveis por sua produção presentes no genoma da linhagem obtida neste estudo.

O *cluster NRPS (non ribosomomal peptide synthase)*, presente em K003, é ligado a complexos de enzimas organizadas em módulos específicos de aminoácidos. A *NRPS* usa

aminoácidos não proteínogênicos, e não apenas os vinte possíveis em peptídeos ribossomais padrão, o que aumenta o número de modelos possíveis para várias centenas (Fischbach e Walsh, 2006; Ng *et al.*, 2009). Logo seus produtos naturais exibem diversas propriedades como atividade antiviral, antitumoral ou antibiótica contra patógenos humanos e vegetais (Qian *et al.*, 2011; Qian *et al.*, 2012) atraindo o interesse das indústrias biotecnológicas, farmacêuticas e agrícolas (Ng *et al.*, 2009; Cochrane e Vederas, 2016; Grady *et al.*, 2016; Biniarz *et al.*, 2017; Deng *et al.*, 2017).

As sintetases peptídicas não ribossomais (*NRPS*) produzem a maioria dos lipopeptídeos de *Paenibacillus*, gênero cujo sobrenadante foi utilizado para selecionar o crescimento da linhagem K003. O principal modo de ação desses lipopeptídeos geralmente envolve rompimento da membrana das células-alvo e, devido à dificuldade dessas células em reorganizar suas membranas, sugere-se que o desenvolvimento de resistência aos lipopeptídeos pode ser lento e limitado, o que é muito interessante do ponto de vista do desenvolvimento de novos fármacos (Finking e Marahiel, 2004, Cochrane e Vederas, 2016).

A presença de *clusters NRPS*, responsáveis pela produção das pelgipeptinas, peptídeos não ribossomais oriundas de *P.elgii* (Qian *et al.*, 2012), em K003, pode indicar a produção de compostos semelhantes entre as linhagens K003 e *P.elgii* e, assim, corroborar a premissa de que bactérias resistentes a certos antimicrobianos, podem produzir compostos semelhantes a eles.

Os *clusters PKS*, também encontrados em K003, juntamente com *NRPS*, representam dois grandes grupos de produtos naturais. Ambos possuem notável diversidade e atividades biológicas, como os antibióticos eritromicina, penicilina, daptomicina e vancomicina, os antifúngicos anfotericina e equinocandina, os agentes anticancerígenos epotilona e bleomicina, o agente redutor de colesterol lovastatina, os imunossupressores FK506 e ciclosporina e os antibióticos veterinários monensina e avermectina (Wang *et al.*, 2011).

As *PKS* são classificadas em Tipos I, II e III (Nivina *et al.*, 2019). Apenas Tipo I foi encontrado em K003. Dentro da *PKS* tipo I existe um grupo de produtos denominados poliéteres ionóforos. Eles causam a morte das células-alvo, sendo de grande importância por sua atividade antiparasitária, antifúngica, antiviral, herbicida, anti-inflamatória e imunorreguladora de amplo espectro (Wang *et al.*, 2011).

A homoserina lactona, presente em K003, está relacionada ao *quorum sensing*, uma vez que em bactérias Gram-negativas, lactonas homoserinas aciladas (acil-HSL) são moléculas sinais comuns que regulam diversas funções alvo e estão frequentemente envolvidos em interações com o hospedeiro. O *quorum sensing* desempenha um papel

fundamental no desenvolvimento de biofilmes microbianos para várias bactérias, incluindo patógenos humanos oportunistas. Assim, se os patógenos usam sinais de acil-HSL para estimular a expressão do gene de virulência, esses sistemas representam um alvo único para a descoberta de drogas (Kumar *et al.*, 2022). Vale ressaltar que em K003, foi encontrada uma homoserina lactona distinta de sua linhagem de referência JCM 2831, apesar de se tratar da mesma espécie, o que pode indicar produção de composto bioativo distinto por K003.

Poucos metabólitos terpenoides foram encontrados em procariotos em relação aos encontrados em fungos e plantas. As primeiras moléculas encontradas nas bactérias foram os terpenos odoríferos, principalmente a geosmina, um sesquiterpeno degradado, que fornece o odor característico de terra úmida (Cane e Ikeda, 2012) Em K003, estão presentes aglomerados terpenoides. Terpenos de origem microbiana têm sido usados como potentes agentes antimicrobianos, como o pentaleno ou a terapentecina (Yamada *et al.*, 2015).

Um dos *clusters* terpenoides de K003 foi identificado como caratenoide, importante componente para lidar com ambientes hostis devido a luz UV intensa, o que é esperado para *Methylobacterium*, visto que elas podem ser abundantes e adaptadas à filosfera, superfície de uma planta acima do solo (Yoshida *et al.*, 2017; Alessa *et al.*, 2021)

O NPAA é um *cluster* de poliaminoácidos semelhante ao Epsilon-poli-L-lisina (ϵ -PL) presente na linhagem K003. O ϵ -PL é um peptídeo antimicrobiano natural com um amplo espectro de atividade contra a deterioração de alimentos e bactérias de intoxicação alimentar *in vitro* (Wang *et al.*, 2021).

Outra grande classe de produtos naturais é formada por peptídeos sintetizados ribossomalmente e modificados após a tradução (*RIPP*). Trata-se de uma classe de produtos naturais com atividade antimicrobiana, cujo *cluster* também é encontrado em K003. As modificações pós-traducionais diversificam a estrutura destes produtos naturais, o que implica no melhor reconhecimento dos alvos e no aumento da estabilidade química e metabólica, melhorando sua funcionalidade química (Ongpipattanakul *et al.*, 2022).

Os resultados obtidos por meio do AntiSMASH, juntamente com os dados do CAMP_{R4} em K003 ilustram o potencial biotecnológico que pode ser encontrado nas linhagens obtidas a partir do uso do sobrenadante de *P. elgii*.

Dentre os *clusters* gênicos identificados em K003, exceto por um terpeno identificado como caratenoide, todos apresentaram baixa (<11%) ou nenhuma semelhança com compostos já conhecidos, o que sugere a potencial produção de novos compostos bioativos.

2.6 CONCLUSÕES

A metodologia descrita neste trabalho se mostrou eficiente para o isolamento de linhagens de gêneros distintos de bactérias de interesse biotecnológico. Foram eles: *Paenibacillus*, *Pseudomonas*, *Enterobacter*, *Kitasatospora*, *Sphingomonas*, *Xanthobacter*, *Burkholderia* e *Methylobacterium*.

Todas as linhagens obtidas a partir do uso do sobrenadante de *P. elgii* apresentaram atividade antibacteriana comprovada em ensaios *overlay*, atestando a eficiência do protocolo em obter linhagens com potencial para aplicação biotecnológica.

Através dos dados de sequenciamento genômico, confirmou-se que a linhagem K003 trata-se de uma *Methylobacterium radiotolerans*. Apesar de se tratar de uma espécie já conhecida, a linhagem K003 apresentou características próprias em relação a sua linhagem de referência quanto ao perfil de ácido graxos, resistência a vancomicina, *cluster* gênico biossintético distinto e características fenotípicas preditas.

O rascunho do genoma de *M. radiotolerans* K003 ilustrou uma amostra da diversidade biossintética que pode estar disponível nas linhagens obtidas a partir do uso do sobrenadante de *P. elgii* como mecanismo de seleção.

Os *clusters* gênicos encontrados na linhagem K003 podem estar relacionados à atividade antibacteriana constatada nos ensaios *overlay*. O *cluster NRPS* responsável pela produção de pelgipeptinas em *P. elgii* foi encontrado no genoma de K003. Esse *cluster*, pode comprovar a premissa de que bactérias resistentes a certos peptídeos antimicrobianos, provavelmente produzem compostos semelhantes a eles.

CAPÍTULO 3 - CARACTERIZAÇÃO GENÔMICA E FISIOLÓGICA DA LINHAGEM *KITASATOSPORA SP.* OBTIDA A PARTIR DO USO DO SOBRENADANTE DE *PAENIBACILLUS ELGII* COMO MECANISMO DE SELEÇÃO.

3.1 INTRODUÇÃO

Os produtos naturais têm grande potencial para a descoberta de novos medicamentos e as actinobactérias estão entre os produtores mais prolíficos de compostos naturais bioativos (van Bergeijk *et al.*, 2020). Pode-se dizer que os actinomicetos (Filo Actinobacteria), grupo ao qual pertence *Kitasatospora sp.*, são um grupo especial entre os procariotos por sua capacidade de produzir amplo espectro de produtos químicos essenciais e metabólitos secundários potencialmente úteis que são ativos inclusive contra organismos eucarióticos (Takahashi, 2017; Srivastava *et al.*, 2019).

A propensão a produzir moléculas bioativas e a riqueza dos genomas bacterianos em termos de diversidade de *clusters* gênicos biossintéticos foram correlacionadas com características do organismo como morfologia complexa e formação de esporos, características das actinobactérias. Além disso, devido a sua regulação metabólica complexa e capacidade de usar várias fontes de nutrientes que lhes permite adaptação a diversos ambientes e condições de crescimento, não é de se estranhar que as Actinobactérias sejam os mais importantes produtores de compostos naturais (van Bergeijk *et al.*, 2020).

Do ponto de vista comercial, *Streptomyces* são as Actinobactérias mais interessantes, pois são a fonte de cerca de metade dos produtos bacterianos totalmente naturais com aplicações biotecnológicas, especialmente na área médica (Hug *et al.*, 2018). Dentro desse filo, *Streptomycetaceae* é a família mais diversificada quanto ao potencial para aplicações biotecnológicas (van Bergeijk *et al.*, 2020). Tal família inclui os gêneros: *Allostreptomyces* (Huang *et al.*, 2017), *Embleya* (Nouioui *et al.*, 2018), *Streptacidiphilus* (Kim *et al.*, 2003), *Streptomyces* (Waksman e Henrici, 1943; Witt e Stackebrandt, 1990; Wellington *et al.*, 1992), *Yinghuangia* (Nouioui *et al.*, 2018) e *Kitasatospora* (Omura *et al.*, 1982; Zhang *et al.*, 1997).

O gênero *Kitasatospora* foi considerado controverso por muitos anos (Omura *et al.*, 1982; Wellington *et al.*, 1992; Zhang *et al.*, 1997). *Kitasatospora* foi proposto por Omura e colaboradores em 1982 para linhagens de actinomicetos que eram fenotipicamente semelhantes às linhagens de *Streptomyces*, mas continham grandes quantidades dos mesoisômeros do ácido diaminopimélico (DAP) e galactose em hidrolisados de células inteiras. Takahashi e colaboradores (1983) estudaram a distribuição dos dois isômeros de

DAP em células em diferentes estágios de diferenciação e crescimento. Descobriu-se que as espécies de *Streptomyces* continham apenas LL-DAP nos micélios aéreos e vegetativos, enquanto as espécies de *Kitasatospora* continham LL-DAP no micélio aéreo e meso-DAP no micélio vegetativo. Embora as quantidades relativas dos dois isômeros pareçam variar em diferentes experimentos e algumas espécies de *Streptomyces* também contenham grandes quantidades de meso-DAP, as espécies de *Kitasatospora* apresentaram consistentemente uma proporção muito maior de meso-DAP para LL-DAP do que as espécies de *Streptomyces* (Wellington *et al.*, 1992).

Wellington e colaboradores (1992) relataram que a sequência de *rRNA* 16S de *Kitasatospora setae* mostrou 91,6% de similaridade com a sequência de *rRNA* 16s de *Streptomyces baldaccii* e que uma sonda oligonucleotídica específica para *Streptomyces* poderia reconhecer todas as quatro espécies válidas de *Kitasatospora*. Com base nessas observações e propriedades fenotípicas compartilhadas pelas espécies de *Kitasatospora* e *Streptomyces*, esses autores propuseram que o nome *Kitasatospora* fosse reduzido a sinonímia com o nome *Streptomyces*. Ochi e Hiranuma (1994) posteriormente apoiaram essa proposta com base nos resultados de uma análise das sequências N-terminais da proteína ribossômica AT-L30. No entanto, a unificação desses dois gêneros foi equivocada (Nakagaito *et al.*, 1992; Kim *et al.*, 1996; Zhang *et al.*, 1997).

A classificação de *Kitasatospora* foi resolvida pela aplicação do sequenciamento do gene marcador *rRNA* 16S (Labeda *et al.*, 2012) e sequenciamento do genoma inteiro (Nouioui *et al.*, 2018; Li *et al.*, 2021). Nakagaito e colaboradores (1992), utilizando os resultados de reassociação *DNA-DNA* e estudos genéticos, constataram que espécies de *Kitasatospora* e espécies originais de *Streptomyces* formaram dois agrupamentos distintos. Kim e colaboradores (1996) conduziram uma análise filogenética baseada na sequência de *rRNA* 16S de um grande número de espécies de *Streptomyces* e observaram que três espécies de *Kitasatospora* formaram um clado distante e estável fora do clado que compreende as espécies de *Streptomyces*.

A fim de esclarecer melhor a relação filogenética entre os gêneros *Kitasatospora* e *Streptomyces*, Zhang e colaboradores (1997) determinaram as sequências de nucleotídeos de genes *rRNA* 16S e espaçadores 16S-23S *rRNA* de 12 linhagens de actinomicetos previamente classificadas como *Kitasatospora* ou *Streptomyces* e com grandes quantidades de meso-DAP em seus hidrolisados de célula inteira. Foi verificado que as espécies de *Kitasatospora* sempre formavam um clado monofilético estável. No entanto, o gênero *Kitasatospora* parece ser um táxon irmão do gênero *Streptomyces* ou uma linhagem que se originou dentro das espécies de

Streptomyces. Então, as espécies de *Streptomyces* e *Kitasatospora* foram consistentemente recuperadas como dois clados distintos. Assim, com base nas evidências filogenéticas, quimiotaxonômicas e fenotípicas, Zhang e colaboradores (1997) propuseram que o gênero *Kitasatospora* (Omura *et al.*, 1982) devesse ser revivido.

O gênero *Kitasatospora* é distinguível de *Streptomyces* com base na composição da parede celular, pois contém LL- e meso - DAP, glicina e galactose (Nouioui *et al.*, 2018) e pela resistência a fagos polivalentes de *Streptomyces* (Wellington *et al.*, 1992) e formação de esporos submersos em cultura líquida (Omura *et al.*, 1982, 1989; Takahashi, 2017), características raramente observadas em espécies de *Streptomyces*.

O gênero *Kitasatospora* é então um gênero legítimo. Trata-se de um gênero de actinobactérias aeróbicas, Gram-positivas, não álcool-ácido resistentes, que formam micélios intensamente ramificados e não fragmentados (Nouioui *et al.*, 2018). Atualmente, o gênero contém 41 espécies validamente descritas (<https://lpsn.dsmz.de/genus/kitasatospora> em 18/01/2023), cuja espécie-tipo é *Kitasatospora setae*.

Vários novos produtos químicos foram descobertos a partir de linhagens de *Kitasatospora*, incluindo compostos com estruturas únicas e atividades interessantes (kim *et al.*, 2020). Um estudo selecionou uma coleção de mais de 800 actinomicetos usando 40 condições de crescimento diferentes, que identificaram, entre outros, *Kitasatospora* sp. MBT66 como um produtor prolífico de antibióticos com atividade contra patógenos Gram-positivos e Gram-negativos (Zhu *et al.*, 2014).

Até agora, pelo menos 50 compostos bioativos foram descobertos a partir de linhagens de *Kitasatospora* (Takahashi, 2017), como as propioxatinas (inibidor da encefalinase B), a terpentecina (atividade antitumoral) (Takahashi, 2017), a setamicina (atividade nematocida e antifúngica) e a fosalacina (atividade herbicida) (Yun *et al.*, 2020). Antibióticos endofenásídeos também foram descobertos em *Kitasatospora* sp. MBT66 (Wu *et al.*, 2015).

Em uma linhagem de *Kitasatospora* também foi relatada uma alta resistência a diversos metais pesados por Yun e colaboradores (2020), atributo interessante em microorganismos que podem ser usados para biorremediação (Alvarez *et al.*, 2017). Além disso, bactérias resistentes a metais pesados isoladas de condições ambientais vêm sendo relatadas como fontes potenciais de novos compostos antimicrobianos (Tomova *et al.*, 2015).

O gênero *Kitasatospora* não é um grupo fácil de isolar como o gênero vizinho *Streptomyces*, e, portanto, seu potencial biossintético ainda não foi amplamente reconhecido (Yun *et al.*, 2020). Estudos sugerem que *Kitasatospora* é um recurso valioso para a descoberta de produtos naturais (Li *et al.*, 2021). Neste capítulo, descrevemos as

características genômicas e fisiológicas de *Kitasatospora sp.* K002 isolada a partir do uso do sobrenadante de *P. elgii* como mecanismo de seleção e propomos sua descrição como uma nova espécie do gênero. O presente trabalho deve ampliar o conhecimento geral sobre este táxon.

3.2 OBJETIVOS

3.2.1 Objetivo Geral

Caracterizar o genótipo, o fenótipo e o potencial biotecnológico de *Kitasatospora sp.* K002 selecionada a partir do uso do sobrenadante de *P. elgii*.

3.2.2 Objetivos específicos

- I. Efetuar caracterização fenotípica *in vitro* e *in silico* de *Kitasatospora sp.* K002.
- II. Sequenciar o genoma de *Kitasatospora sp.* K002 em *Illumina* e *Oxford Nanopore Technologies*.
- III. Efetuar a classificação taxonômica de *Kitasatospora sp.* K002 baseada em seus dados genômicos.
- IV. Prospectar as possíveis aplicações biotecnológicas de *Kitasatospora sp.* K002, como por exemplo os genes responsáveis pela atividade antibacteriana.

3.3 MATERIAL E MÉTODOS

3.3.1 Caracterização fenotípica *in vitro*

As células de *Kitasatospora sp.* K002 foram observadas em microscopia óptica para a determinação das características de membrana, utilizando a coloração diferencial de Gram (Reddy *et al.*, 2007). A observação foi realizada em microscópio Leica DM750.

Para a visualização em microscopia eletrônica de varredura, as células foram obtidas por centrifugação do cultivo em meio líquido NB e lavadas por 3 vezes em tampão salino de fosfato (PBS). A fixação foi em solução Karnovsky (2% (v/v) paraformaldeído, 2% (v/v) glutaraldeído, 3% (p/v) sacarose e 0.05 M de tampão cacodilato). A segunda fixação foi em solução de tetróxido de ósmio (1% p/v) seguida por adesão em lamínula com poli-L-lisina. As células foram então desidratadas em solução de acetona com concentração crescente (50/70/90/100%, v/v). A secagem em ponto crítico foi realizada com CO₂ e as células foram cobertas com uma fina camada de ouro (Souza, 2007). As imagens foram geradas em microscópio eletrônico de varredura JEOL JSM-7001F (JEOL Ltd., Tóquio, Japão).

Para microscopia eletrônica de transmissão, as células foram lavadas e fixadas como na microscopia de varredura e em seguida foi feita a contrastação *in block* em solução aquosa de acetato de uranila 0.5% *overnight*, desidratados em série de acetona (30/50/70/90/100/100%) e embebidos em resina Spurr's. Cortes ultrafinos foram contrastados com acetato de uranila e posteriormente examinados e fotografados em Microscópio Eletrônico de Transmissão Jeol 1011 (JEOL Ltd., Tóquio, Japão).

A temperatura ótima de crescimento foi avaliada por uma semana em placas de NB ágar nas temperaturas de 5, 10, 15, 20, 30, 35, 40, 45, 50, 55 e 60 °C. A faixa de pH ideal para crescimento foi examinada entre 4.0 e 10.0 em intervalos de 1.0 após incubação à temperatura ideal verificada por 3 dias em meio NB líquido ajustado com HCl ou NaOH contendo tampão citrato Na₂HPO₄ 50 mM (pH 4–7), tampão NaH₂PO₄/Na₂HPO₄ 100 mM (pH 7–8) ou tampão NaHCO₃/Na₂CO₃ 100 mM (pH 9–10). A tolerância ao sal foi testada em NB líquido suplementado com NaCl (0-5% em intervalos de 1%) e incubação à temperatura ideal verificada por 3 dias. Tanto salinidade quanto pH foram examinados medindo a OD₆₀₀ em SpectraMax M3 com *software* SoftMax Pro 6.0 (Zhu *et al.*, 2014).

A atividade da catalase foi testada pela adição de 3% de peróxido de hidrogênio à cultura cultivada em meio de cultura R2A a 30°C por 7 dias, e a efervescência foi monitorada (Tindall *et al.*, 2007).

As células de *Kitasatospora sp.* K002 cultivadas em meio R2A sólido foram coletadas após 48h de crescimento, transferidas para tubos de vidro com tampa de teflon e então

utilizadas para análises de ácidos graxos celulares, conforme descrito por Ramaprasad e colaboradores (2015). A saponificação dos ácidos graxos foi realizada com 1 mL de solução metanol/hidróxido de sódio (150 mL água deionizada, 150 mL metanol, 45 g NaOH) à 100°C por 30 minutos. Dois mililitros de metano 6M em HCL foram adicionados aos tubos, que ficaram incubados por 10 minutos à 80°C. A separação dos ésteres metílicos dos ácidos orgânicos foi realizada pela adição de 1.25 mL de hexano/éter terc-butílico (1:1 v/v), após agitação por 10 minutos. A fase mais pesada foi descartada e a fase superior contendo os ácidos graxos metil-ésteres foi recuperada e lavada com 0.25 mol/L NaOH e estocada até a análise (Sasser, 2001).

Os ésteres metílicos de ácidos graxos obtidos foram analisados em cromatógrafo gasoso (Agilent 7890^a) equipado com o *software Sherlock Microbial Identification System* (Sherlock, v. 6.2) usando o a biblioteca do método RTSBA6. Os picos obtidos foram então rotulados e os valores de comprimento de cadeia equivalente foram calculados pelo *software* Sherlock. O Sistema de Identificação Microbiana Sherlock é um sistema totalmente analítico automatizado de cromatografia gasosa para identificar bactérias com base em seus perfis únicos de ácidos graxos (Sasser, 2001).

A suscetibilidade a antibióticos foi avaliada em meio sólido pelo método Kirby-Bauer (difusão com disco) utilizando discos de papel filtro impregnados com concentração padrão de antibiótico e observação da formação de halo após aplicação do disco sobre cultura espalhada com *swab* e incubação de 48h a 30° C. Os antibióticos testados foram (µg per disco): Amicacina (30), Amoxicilina/Ác. Clavulânico (20/10), Ampicilina (10), Aztreonam (30), Cefalexina (30), Cefazolina (30), Cefepima (30), Ceftazidima (30), Ceftriaxona (30), Ciprofloxacina (5), Clindamicina (2), Cloranfenicol (30), Doxiciclina (30), Eritromicina (15), Gentamicina (10), Imipenem (10), Meropenem (10), Norfloxacina (10), Ofloxacina (5), Oxacilina (1), Penicilina G (10), Piperacilina + Tazobactam (110), Polimixina B (300), Sulfonamida (300), Tetraciclina (30) e Vancomicina (30).

3.3.2 Sequenciamentos, montagem e análise do genoma

Para a *Kitasatospora sp.* K002 caracterizada neste capítulo, foi feita a clonagem do gene marcados *rRNA* 16S, onde os fragmentos do *rRNA* 16S purificados foram ligados em vetor pGEM®-T Easy, Promega, de acordo com as instruções do fabricante. O objetivo foi ter uma sequência mais completa do gene marcador *rRNA* 16 para as análises filogenéticas.

A transformação bacteriana ocorreu em *Escherichia coli* DH10B por choque térmico. As células transformadas foram espalhadas em meio de cultura R2A sólido suplementado

com Ampicilina (150 µg/mL), X-Gal (0,00625 % p/v) e IPTG (0,5 mM). As colônias positivas para a transformação com plasmídeo e inserto (brancas) foram inoculadas em meio líquido LB + Ampicilina (150 µg/mL).

O kit *QIAprep Spin Miniprep Kit*, Qiagen foi utilizado para a extração e purificação do DNA plasmidial. A transformação foi confirmada através da verificação da presença dos insertos após realização de digestão com a enzima EcoR1 e avaliação do padrão de bandas em eletroforese em gel de agarose a 1% (p/v).

O DNA plasmidial dos clones foi sequenciado em plataforma Sanger na empresa MacroGen (Seoul, Coréia do Sul) utilizando os oligonucleotídeos iniciadores para os promotores T7 (5' GTAATACGACTCACTATAGGGC 3') e sp6 (5' ATTTAGGTGACACTATA 3').

A ferramenta Asparagin (disponível em <http://asparagin.cenargen.embrapa.br/phph/>) foi utilizada para avaliar a qualidade *Phred score* e remoção dos pares de bases provenientes do vetor (Ewing, 1998; Togawa, 2006). As sequências com *Phred score* maior que 30 foram mantidas para as análises seguintes. Em seguida, foi realizada a comparação por similaridade com sequências disponíveis no *GenBank* do *NCBI* (*National Center for Biotechnology Information*) usando a ferramenta BLAST N. Em adição, foi feita a comparação com as sequências do banco de dados Ezbiocloud. Nele, além das sequências de 16S de *amplicons* do *NCBI*, também são consideradas sequências 16S extraídas de genomas do *NCBI*, sequências compiladas do sequenciamento completo realizados com a tecnologia *Pacific Biosciences* (*PacBio*) de amostras de microbiomas, sequências extraídas da montagem de genomas do *JGI* (*Joint Genome Institute*) e sequências de genomoespécies que são espécies provisoriamente novas suportadas por sequências genômicas inteiras (Yoon *et al.*, 2017).

O sequenciamento genômico de *Kitasatospora sp.* K002 foi realizado em plataforma *NGS Illumina* na empresa MacroGen (Seul, Coreia do Sul). A biblioteca de DNA foi preparada usando o *TruSeq Nano DNA Kit* e o sequenciamento foi realizado na plataforma NextSeq500, gerando leituras *paired-end* de 150 pares de bases.

A fim de melhorar a montagem do genoma, o sequenciamento *NGS* foi realizado usando a tecnologia *long reads* na plataforma *Oxford Nanopore Technologies* (*ONT*) (Oxford, Reino Unido). A biblioteca de DNA foi preparada com o *Rapid Sequencing Kit* (SQK-RAD004) e o sequenciamento foi realizado em *Flow cell* FLO-MINSP6 no dispositivo MinION (MIN-101B).

Os dados brutos do sequenciamento *NGS Illumina* passaram por limpeza com a ferramenta *Trimmomatic* (Bolger *et al.*, 2014) para retirada de sequências dos adaptadores,

eliminação de *reads forward* e *reverse* não-pareados e das sequências com *Phred score* menor do que 30 (Ewing *et al.*, 1998). Os dados brutos do sequenciamento *NGS ONT* passaram pelo processo de *basecalling* através do programa Guppy em modo de alta acurácia (Fraser *et al.*, 2020).

A montagem *de novo* de um rascunho do genoma de *Kitasatospora sp.* K002 foi realizada de forma híbrida com o programa Unicycler (V. 0.5.0) para confirmação de sua identidade (Wick *et al.*, 2016). O *software* CheckM, por meio da plataforma Kbase, foi utilizado para determinação de sequências contaminantes e completude dos genomas (Parks *et al.*, 2015). A cobertura dos *reads* foi determinada com o *software* Bowtie2 (v. 2.4.4) (Langmead *et al.*, 2018) e os parâmetros L50 e N50 foram obtidos através do Quast (Mikheenko *et al.*, 2016). O *software* Artemis foi utilizado para a visualização do genoma circular e suas regiões codificantes através do recurso *DNAPlotter* (Carver *et al.*, 2009). A anotação do genoma foi realizada usando o *software* Prokka (v. 1.14.6) (SEEMANN, 2014).

3.3.3 Análises filogenéticas

A posição evolutiva de *Kitasatospora sp.* K002 foi estimada pela construção de árvore filogenética pelo método *Maximum-likelihood* (ML) com 1000 repetições de amostragem (*Bootstrap*), utilizando vinte e duas sequências representativas quanto *rRNA* 16S de espécies do gênero das linhagens válidas listadas em *List of Prokaryotic names with Standing in Nomenclature - LPSN* e depositadas no banco de dados GenBank/EMBL/DDBJ (Felsenstein, 1985; Kimura, 1980) e um *outgroup*. A ferramenta *clustal W* foi utilizada para fazer o alinhamento das sequências (Larkin *et al.*, 2007). A construção da árvore filogenética baseada na sequência parcial do gene marcador molecular *rRNA* 16S foi realizada no programa Mega 11 (disponível para download em <https://www.megasoftware.net/>).

A classificação taxonômica de *Kitasatospora sp.* K002 e a confirmação da identidade foram realizadas na plataforma TYGS – DSMZ, por hibridização digital de *DNA* (DDH) no servidor GGDC – DSMZ (Meier-Kolthoff *et al.*, 2022). Com base na sequência do *rRNA* 16S, são determinadas as dez linhagens filogeneticamente mais próximas do genoma a ser analisado. Essas onze linhagens seguem para comparação dos parâmetros genômicos de classificação taxonômica, como a distância entre genomas (*GBDP*).

A *ANI* (Identidade Média de Nucleotídeos) foi calculada na plataforma JSpeciesWS (<https://jspecies.ribohost.com/>), onde *ANIb* é a identidade nucleotídica média com base no cálculo BLAST e a *ANIm* é a a identidade nucleotídica média com base no cálculo MUMmer

(Richter *et al.*, 2016). O *software* GTDB-Tk (v. 1.7.0), disponibilizado na plataforma Kbase também foi utilizado (Arkin *et al.*, 2018, Chaumeil *et al.*, 2019).

O alinhamento do genoma de K002 e do genoma de sua linhagem mais próxima segundo o TYGS-DSMZ, *K. xanthocidica* JCM 4862, foi realizado com o *pipeline* PlotNucmer (disponível em: https://github.com/fmalmeida/plot_nucmer), que é uma automação do MUMmer v3 .1 toolkit (Kurtz *et al.*, 2004) e Ggbio (Yin *et al.*, 2012), onde o primeiro faz o alinhamento dos genomas por comprimento e identidade, e o segundo gera uma visualização circular.

A árvore filogenética baseada em genes ortólogos centrais concatenados de sequências genômicas de doze linhagens, incluindo K002, as linhagens estreitamente relacionadas de acordo com o TYGS-DSMZ e o *outgroup* *Streptomyces violaceorubidus* LMG 20319 disponíveis no NCBI foi construída usando o M1CR0B1AL1Z3R servidor *web* (<https://microbializer.tau.ac.il/>) (Avram *et al.*, 2019).

A análise comparativa de pangenoma foi realizada usando Roary versão 3.13.0 seguindo a anotação de Prokka, versão 1.14.6 (Seemann, 2014), disponível no Galaxy (Afgan *et al.*, 2016). Foram utilizadas a sequência do genoma de K002 e as sequências de suas dez sequências mais próximas, segundo o TYGS-DSMZ, disponíveis no NCBI. A identidade percentual mínima para BLASTp de 80% foi definida para agrupar os genes que codificam sequências completas de proteínas em genomas centrais (Corretto *et al.*, 2017).

Todas as árvores geradas neste trabalho foram editadas no Inkscape (disponível para download em <https://inkscape.org/>).

3.3.4 Predição de genes e anotação funcional

As categorias funcionais baseadas na classificação COG de K002 foram obtidas pela versão on-line do eggNOG-mapper v.2 (Huerta-Cepas *et al.*, 2017), onde vias metabólicas e genes de interesse foram comparados ao banco de dados *Kyoto Encyclopedia of Genes and Genomes* (KEGG) (Kanehisa *et al.*, 2007).

O Banco de Dados de Genes de Resistência a Antibióticos (ARDB) foi utilizado para anotação de genes potenciais responsáveis pela resistência a antibióticos (Liu e Pop, 2009). Os *locus* CRISPR (agrupado regularmente interespaçado curto repetições palindrômicas) foram previstos utilizando a ferramenta de reconhecimento CRISPRCasFinder (Couvin *et al.*, 2018).

O *software* AntiSMASH (v. 6.1.1) (Blin *et al.*, 2021) foi utilizado para a identificação de *clusters* gênicos relacionados ao metabolismo secundário (BCGs) dos micro-organismos.

O BiG-SCAPE faz o agrupamento de similaridade de famílias de *BGCs* com base em suas semelhanças com sequências de proteínas de famílias Pfam (El-Gebali *et al.*, 2019) e com *BGCs* conhecidos disponíveis na versão mais recente do repositório MIBiG (v1.4) (Kautsar *et al.*, 2020). A fim de obter informações sobre a diversidade e similaridade de *BCGs* entre de *Kitasatospora sp.* K002 e sua linhagem de referência, as saídas do antiSMASH foram agrupadas usando BiG-SCAPE em cortes de distância de 0.3 (Navarro-Munoz *et al.*, 2020).

O servidor *web* BAGEL4 foi utilizado para identificar e visualizar grupos de genes em *DNA* procariótico envolvidos na biossíntese de peptídeos sintetizados ribossomalmente e modificados após a tradução (*RIPPs*), bem como bacteriocinas (não modificadas) (Van Heel *et al.*, 2018).

O banco de dados de dados CAMP_{R4} foi utilizado para predição de peptídios antimicrobianos (*AMPs*) (Gawde *et al.*, 2022). Trata-se de um banco com curadoria manual usando informações disponíveis de *AMPs* do *NCBI* (*National Center for Biotechnology Information*), *PDB* (*RCSB Protein Data Bank - RCSB PDB*), *Pub Med* and *Lens* (*patent*). Nele, prevê-se que a sequência seja antimicrobiana através de uma pontuação de probabilidade (0 a 1) usando métodos multivariados (Waghu *et al.*, 2014). Quanto maior a probabilidade, maior é a possibilidade de o peptídeo avaliado ser antimicrobiano.

3.3.5 Caracterização fenotípica *in silico*

A ferramenta de fenotipagem *in silico* Traitair (*Microbial Trait Analyzer*) foi utilizada para prever fenótipos com base nas sequências do genoma de K002 de forma comparativa com suas linhagens mais próximas de acordo com TYGS-DSMZ. Este *software* é capaz de prever 67 características fenotípicas usando o algoritmo Prodigal para predição de genes e família Pfam para anotação (Weimann *et al.*, 2016). O *software* usa dois modelos de predição, o modelo *phypat* (que prevê a presença/ausência de proteínas encontrado no fenótipo de 234 espécies bacterianas) e uma combinação de modelos *phypat+PGL* (que usa a informação de *phypat* combinada com a informação da aquisição ou perda de famílias de proteínas e fenótipos através da evolução), para determinar as características fenotípicas (Weimann *et al.*, 2016).

3.4 RESULTADOS

3.4.1 Caracterização fenotípica *in vitro*

As colônias de *Kitasatospora sp.* K002 eram Gram-positivas e em meio sólido R2A, apresentaram coloração branco acinzentada, forma irregular e aspecto coriáceo. Em microscopia eletrônica, foram observados micélios vegetativos e esporos submersos a partir de sua cultura em meio líquido. Os esporos submersos brotam e se alongam formando os micélios vegetativos sem nenhuma fragmentação. Nenhum micélio aéreo e cadeia de esporos foram observados (Fig. 3.1).

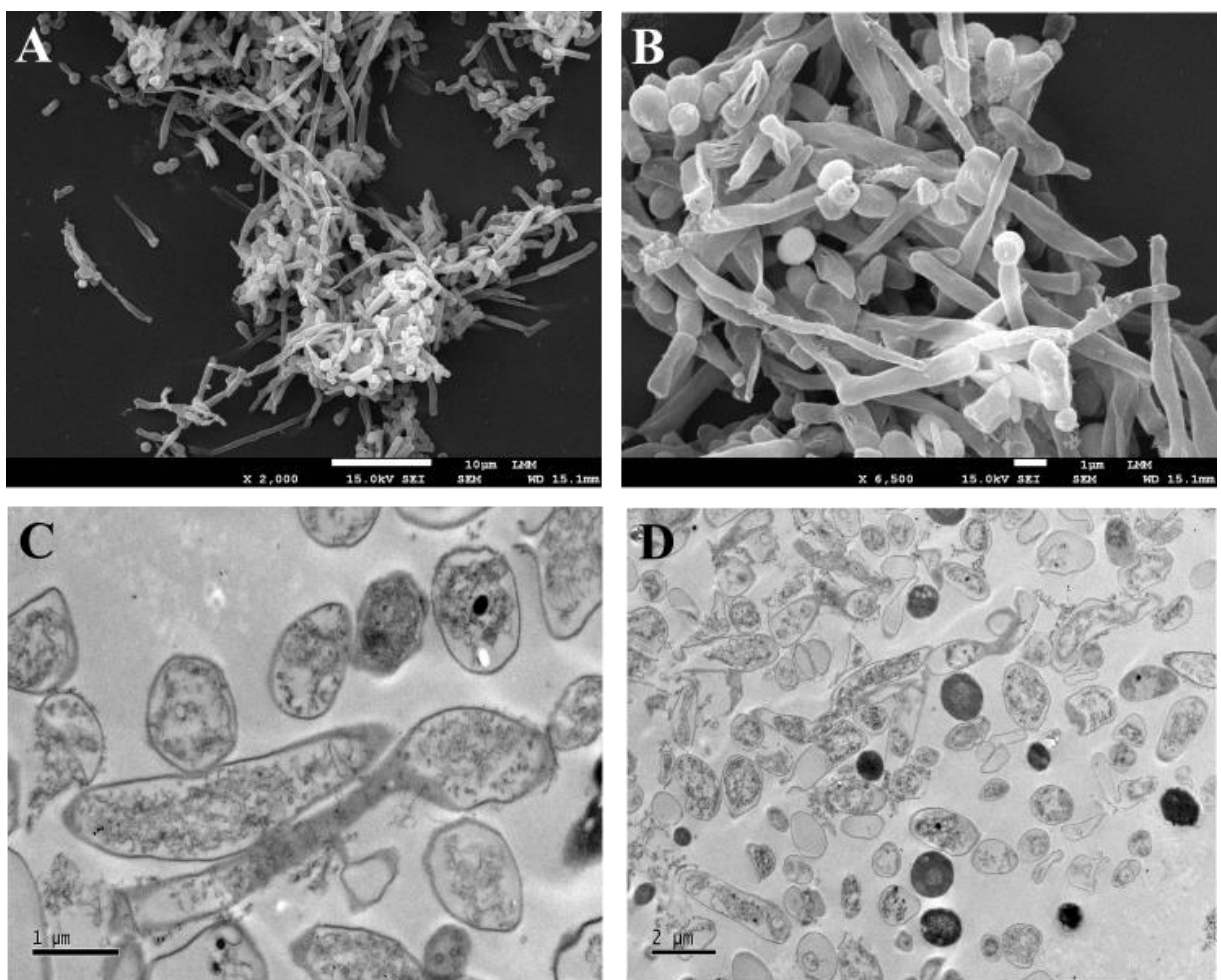


Figura 3.1 Micrografias eletrônicas de varredura (A e B) e de transmissão (C e D) de *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante de *P. elgii*. Podem ser observados micélios vegetativos (forma alongada) e esporos submersos (forma arredondada).

O crescimento de *Kitasatospora sp.* K002 ocorreu aerobiamente entre as temperaturas de 15–40 °C em meio NB contendo 0-3% (p/v) de NaCl em pH 7–9,0, e o crescimento ótimo ocorreu a 30 °C, pH 7.0 e com 0 % (p/v) de NaCl. As características fenotípicas diferenciais entre K002 e outras espécies do gênero (Õmura *et al.*, 1982; Kim *et al.*, 2020; Klaysubun *et*

al., 2022) estão resumidas na Tabela 3.1. O crescimento ideal em pH 8 distingue K002 das demais.

Tabela 3.1 Propriedades fenotípicas que separam *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante *P. elgii*, das espécies relacionadas do gênero. Linhagens: 1, *Kitasatospora* K002; 2, *K. setalba* KM-6054; 3, *K. acidiphila* MMS16-C NU292T; 4, *K. azatica* KCTC 9699T; 5, *K. kifunensis* IFO 15206T; 6, *K. purpeofusca* NRRL B-1817T; 7, *K. nipponensis* HKI 0315T e 8, *K. humi* RB6PN24. +, Positivo, entre parênteses, crescimento ideal, -, sem dados disponíveis.

Faixa de crescimento (ideal):	1	2	3	4	5	6	7	8
Temperatura (°C)	15-35 (30)	22-37 (30)	15-37 (30)	10-30 (30)	10-30 (25)	10-30 (25)	10-30 (25)	15-39 (30)
pH	7-9(8)	-	4-9(5)	5-8 (6)	5-9 (7)	5-9 (7)	5-9 (7)	4-8 (7)
NaCl (%)	0-3(0)	0-2(0)	0-3(0)	0-2 (0.5)	0-1 (0)	0-2 (0.5)	0-2 (0.5)	0-2 (0)

Dentre os vinte e seis antibióticos testados, K002 foi resistente a sete. São eles: Amoxicilina/Ác. Clavulânico, Aztreonam, Cefepima, Ceftazidima, Meropenem, Oxacilina e Sulfonamida. A tabela 3.2 apresenta os resultados do teste de suscetibilidade de *Kitasatospora sp.* K002 aos antibióticos avaliados.

Tabela 3.2 Resultados do teste de suscetibilidade a antibióticos de *Kitasatospora sp.* K002, obtida a partir do sobrenadante de *P. elgii*. O método adotado foi o de Kirby-Bauer (difusão em disco), onde são utilizados discos de papel filtro impregnados com uma concentração padrão de antibiótico e observada a formação de um halo de inibição. +, Suscetível; -, Resistente.

Antibiótico (μg per disco)	Resultado	Halo de inibição (cm)
Amicacina (30)	+	1.5
Amoxicilina/Ác. Clavulânico (20/10)	-	-
Ampicilina (10)	+	2
Aztreonam (30)	-	-
Cefalexina (30)	+	2.5
Cefazolina (30)	+	1.5
Cefepima (30)	-	-
Ceftazidima (30)	-	-
Ceftriaxona (30)	+	2
Ciprofloxacina (5)	+	3
Clindamicina (2)	+	1
Cloranfenicol (30)	+	1.5
Doxiciclina (30)	+	3.5
Eritromicina (15)	+	4
Gentamicina (10)	+	1.5
Imipenem (10)	+	5
Meropenem (10)	-	-
Norfloxacina (10)	+	1.5
Ofloxacina (5)	+	2.5
Oxacilina (1)	-	-
Penicilina G (10)	+	1.3
Piperacilina + Tazobactam (110)	+	2.5
Polimixina B (300)	+	0.8
Sulfonamida (300)	-	-
Tetraciclina (30)	+	2
Vancomicina (30)	+	2

O perfil de ácidos graxos de *Kitasatospora sp.* K002 e de outras espécies do gênero (Kim *et al.*, 2020) é apresentado na Tabela 3.3. Dentre os ácidos graxos mais abundantes (>15%) estão anteiso-C15:0 (27.07%) e iso-C16:0 (18.44%), já os menos abundantes foram C18:0 (0.52%) e *Summed Feature 8* (0.57%). Os perfis de ácidos graxos separaram claramente a linhagem K002 do restante devido aos componentes iso-C18:0, C18:0 e *Summed Feature 8*.

Tabela 3.3 Composição de ácidos graxos (%) de *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante de *P. elgii*, em comparação a outras espécies do gênero. **Summed Feature* ocorre quando há uma separação imperfeita do pico. *Summed feature 3* representa C16:1 ω6c e/ou C16:1 ω7c, *summed feature 8* representa C18:1 ω7c e/ou C18:1 ω6c, *summed feature 9* representa iso-C17 : 1 w9c e /ou C16:0 10-methyl. Linhagens: 1, *Kitasatospora* K002; 2, *K. setalba* KM-6054; 3, *K. acidiphila* MMS16-C NU292T; 4, *K. azatica* KCTC 9699T; 5, *K. kifunensis* IFO 15206T; 6, *K. purpeofusca* NRRL B-1817T and 7, *K. nipponensis* HKI 0315T.

Ácido graxo	Nome	Tipo	Percentual					
			1	2	3	4	5	6
iso-C14:0	Ácido 12-metil tridecanóico	Saturado de Cadeia Longa	2.68	2.0	2.5	3.6	1.1	2.8
C14:0	ácido tetradecanóico	Saturado de Cadeia Longa	1.09	0	0	2.5	2.0	0
iso-C15:0	ácido 13-metiltetradecanóico	Cadeia longa saturada	13.11	17.4	12.9	14.7	13.8	17.7
anteiso-C15:0	ácido 12-metiltetradecanóico	Ácido graxo ramificado, ramificado metílico, ácido graxo saturado e derivados.	27.07	29.0	17.6	25.3	18.8	21.9
iso-C16:1 H	-	cadeia longa saturada	0.72	0	2.5	0	0	1.6
iso-C16:0	ácido 14-metilpentadecanóico	Saturado de Cadeia Longa	18.44	14.8	17.9	6.8	17.0	6.2
C16:0	ácido hexadecanóico	Saturado de Cadeia Longa	10.07	6.6	16.4	16.5	23.8	9.2
anteiso-C17:1 w9c	Ácido (7Z)-13-Metil-7-Hexadecenóico	Saturado de Cadeia Longa	1.54	2.9	2.7	1.2	1.0	2.0
iso-C17:0	ácido heptadecanóico	Saturado de Cadeia Longa	3.26	4.0	2.5	1.2	3.8	3.0
anteiso-C17:0	Ácido 14-metil-hexadecanóico	Metil Ramificado Saturado	11.98	11.9	5.6	3.1	7.0	5.76
cyclo-C17:0	Ácido cis-9,10-metileno-hexadecanóico	Cadeia longa saturada com um anel de ciclopropano em C9 e C10	4.51	4.6	3.4	14.2	4.9	1.1
iso-C18:0	ácido 16-metil-heptadecanóico	Metil ramificado	1.12	0	0	0	0	0
C18:0	ácido octadecanóico	Saturado de Cadeia Longa	0.52	0	0	0	0	0
<i>Summed Feature 3*</i>	-	-	1.49	0	9.3	6.1	10.8	8.6
<i>Summed Feature 8*</i>	-	-	0.57	0	0	0	0	0
<i>Summed Feature 9*</i>	-	-	1.83	3.1	4.3	2.2	3.07	6.7

3.4.2 Sequenciamentos e análises filogenéticas

O sequenciamento Sanger do *rRNA* 16S de *Kitasatospora sp.* K002 permitiu a confirmação da identidade da sequência, onde o BLAST indicou maior similaridade (99.24%) com *Kitasatospora kifunensis* JCM 9081. Já o banco de dados Ezbicloud apontou maior similaridade (99.44%) com *Kitasatospora xanthocidica* NBRC13469. A árvore filogenética baseada no *rRNA* 16S de K002 está representada na Figura 3.2.

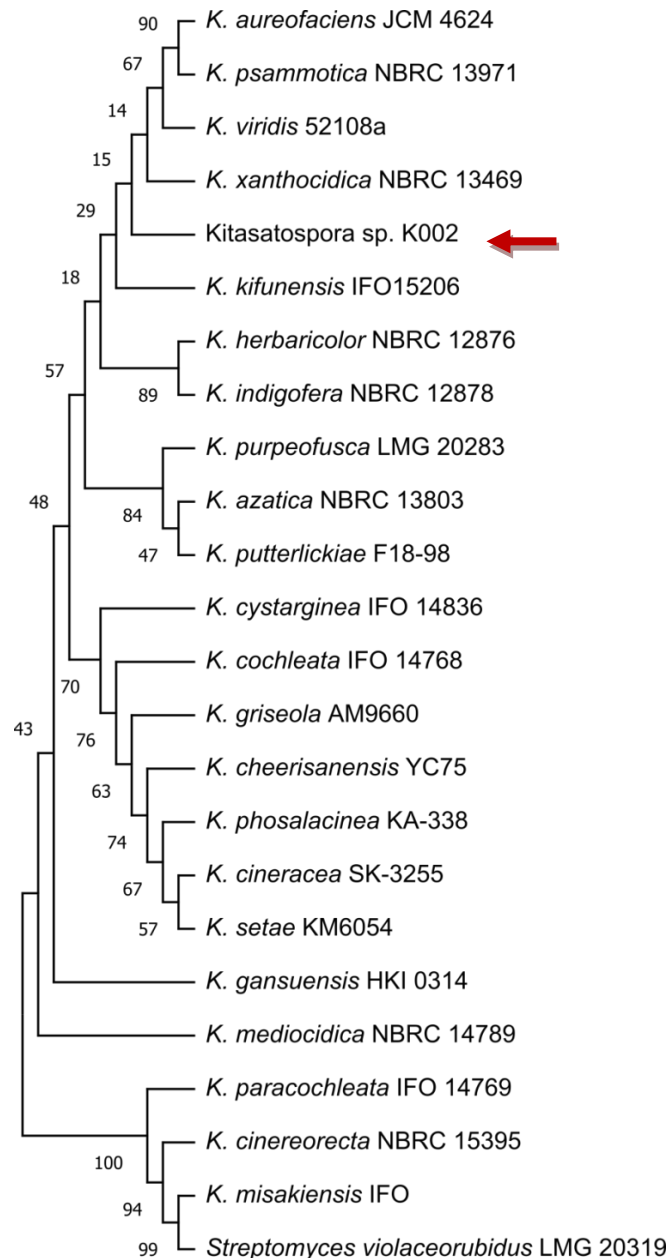


Figura 3.2 Árvore filogenética baseada na sequência parcial do gene 16S ribossomal *rRNA* obtido por clonagem e sequenciamento de Sanger de *Kitasatospora sp.* K002, obtida a partir da utilização do sobrenadante de *P. elgii*. A posição evolutiva da linhagem foi estimada pela construção de uma árvore filogenética usando o método *Maximum-likelihood (ML)* com 1000 repetições de amostragem (*Bootstrap*), usando sequências representativas de 16S ribossomal *rRNA* de linhagens válidas do gênero listadas na *List of Prokaryotic names with Standing in Nomenclature - LPSN*.

O genoma de *Kitasatospora sp.* K002 foi montado usando uma combinação de abordagens de sequenciamento de leitura longa e curta, resultando em uma montagem totalizando 7.951.960 pb, distribuídos em 14 *contigs*, com Teor de GC de 72.8%. A anotação no *software* Prokka *Galaxy Version1.14.5* mostrou que no genoma de K002 há cerca de 7.951.960 bp, 19 *rRNA*, 90 *tRNA* e 6708 *CDS* (sequências codificadoras). (Tab.3.4 e Fig. 3.3).

Tabela 3.4 Dados da montagem híbrida do genoma de *Kitasatospora* K002, obtida a partir do sobrenadante de *P. elgii*. A montagem foi realizada a partir dos sequenciamentos de leitura longa e curta através do programa Unicycler. A anotação foi realizada com o *software* Prokka e os parâmetros de qualidade foram avaliados nos *softwares* CheckM, Bowtie2 e Quast.

Dados da montagem	
<i>Contigs</i>	14
<i>CDS</i>	6708
<i>RRNA</i>	19
<i>Repeatregion</i>	11
<i>TRNA</i>	90
<i>TmRNA</i>	1
Teor de GC (%)	72.8
Contaminação (%)	1.32
Compleitude (%)	99.47
Cobertura	300x
L50	4
N50 (pb)	761556
Tamanho Total (bp)	7951960
Maior <i>contig</i>	1713370

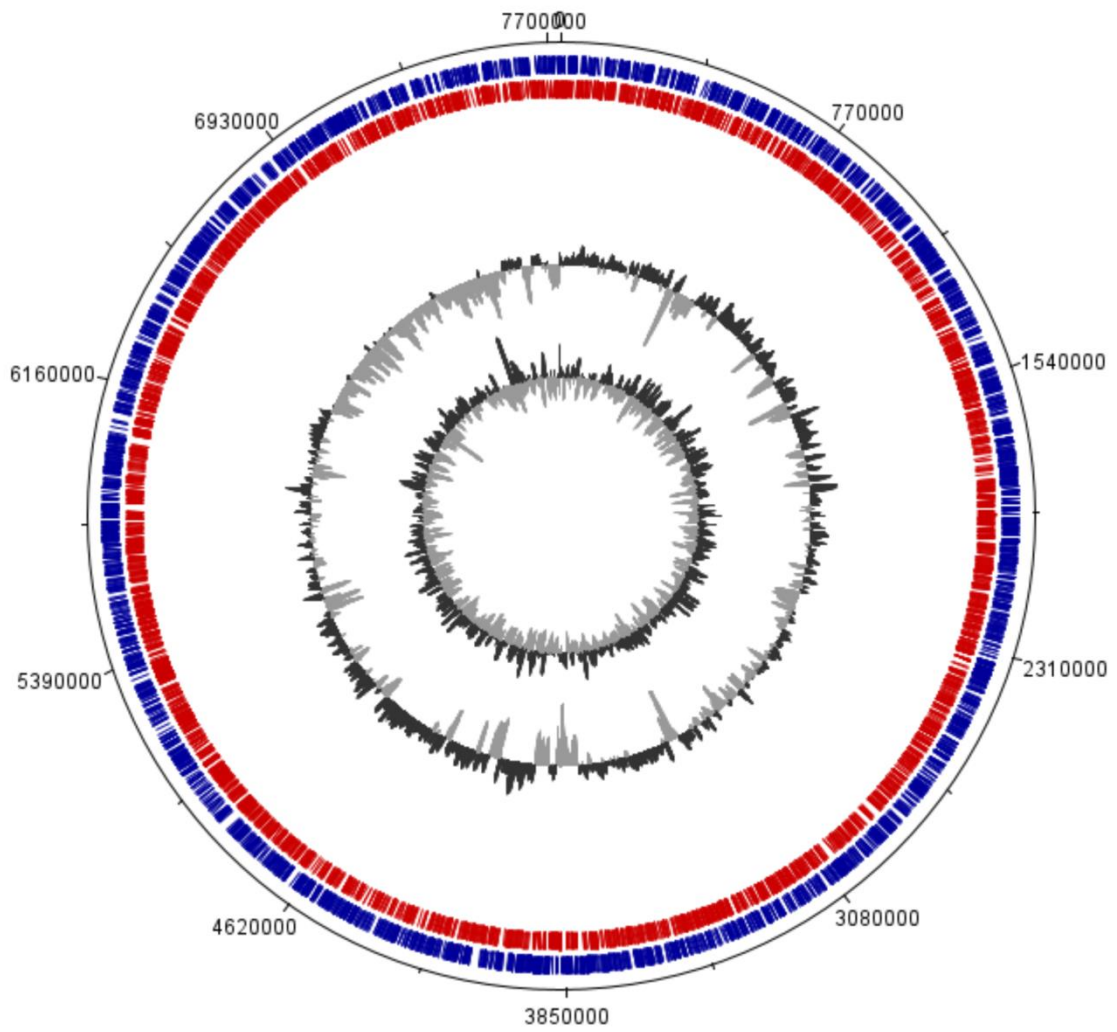


Figura 3.3 Mapa circular da montagem híbrida do genoma de *Kitasatospora sp. K002*, obtida a partir do sobrenadante de *P. elgii*. A ilustração foi gerada pelo recurso DNAPlotter do software Artemis demonstrando as regiões codificantes e o conteúdo GG. De fora para o centro: CDS direto (azul escuro), CDS reverso (vermelho), conteúdo de GC e inclinação de GC (cinza escuro e cinza claro claro).

A avaliação da montagem do do genoma de *Kitasatospora sp.* K002 realizada na plataforma TYGS – DSMZ confirmou que a linhagem mais próxima a sua é uma *Kitasatospora xanthocidica* com *dddH* 32.8-37.8%. A conclusão do protocolo automatizado da plataforma indica que K002 é potencialmente uma nova espécie. As árvores filogenéticas baseadas no *rRNA* 16S e no genoma propostas pela plataforma estão apresentadas nas figuras 3.4 e 3.5, respectivamente.

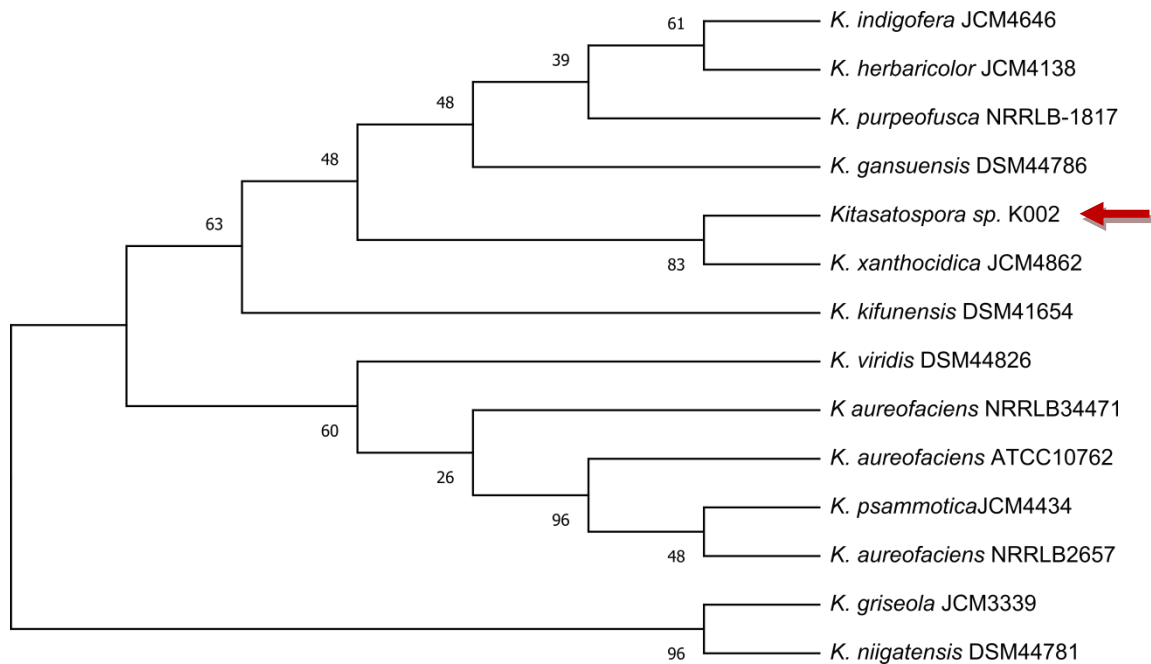


Figura 3.4 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de *Kitasatospora* K002, linhagem obtida a partir do sobrenadante de *P. elgii*, baseada na sequência do gene ribossomal *rRNA* 16S.

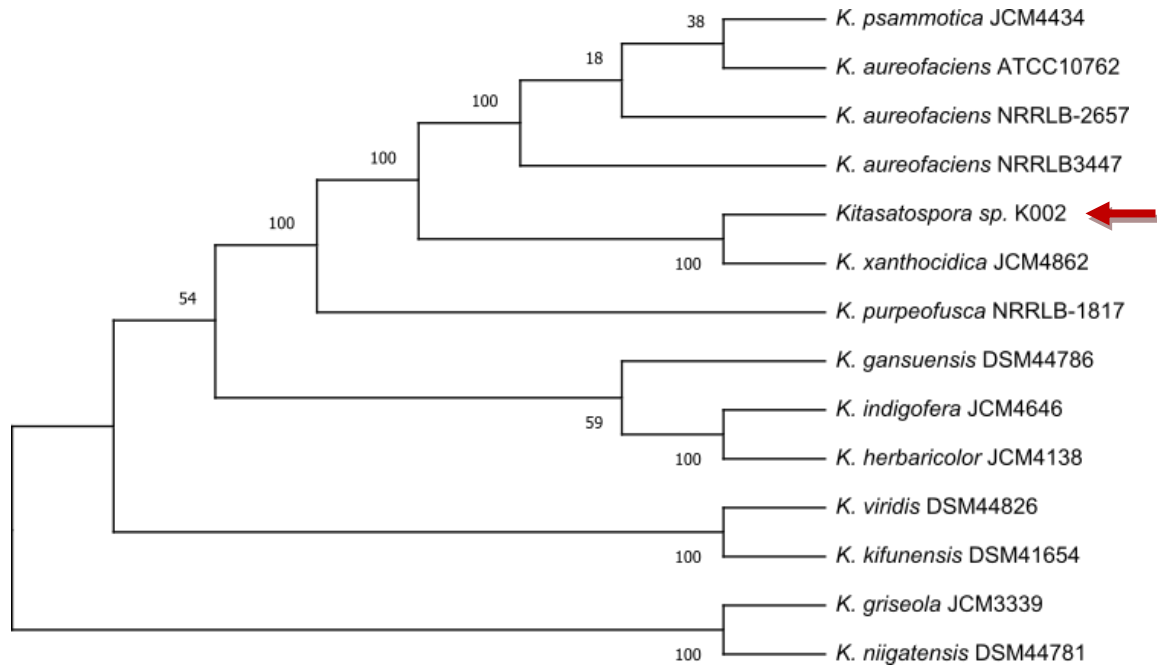


Figura 3.5 Árvore filogenética gerada pela plataforma on-line TYGZ – DSMZ mostrando o posicionamento taxonômico de *Kitasatospora* K002, linhagem obtida a partir do sobrenadante de *P. elgii*, baseada em seu genoma total.

As ANIs (*Average Nucleotide Identity*) calculadas para o genoma de *Kitasatospora* sp. K002 na plataforma JspeciesWS foram 86.86% (*ANI_b*: Identidade média de nucleotídeos com base no BLAST+) e 90.09% (*ANI_m*: Identidade média de nucleotídeos com base em MUMmer), ambas em relação a *Kitasatospora xanthocidica*. No software GTDB-tk, a ANI foi de 93.41% em relação a *Kitasatospora* sp. (Tab. 3.5).

Tabela 3.5 Identidade Média de Nucleotídeos (*ANI*) calculadas para *Kitasatospora* sp. K002, obtida a partir do uso do sobrenadante de *P. elgii*. *ANI_b*: ANI com base no cálculo BLAST calculada na plataforma JSpeciesWS. *ANI_m*: ANI com base no cálculo MUMmer calculada na plataforma JSpeciesWS. *ANI*: calculada no software GTDB-Tk.

	JspeciesWS		GTDB-tk
	<i>ANI_b</i> (%)	<i>ANI_m</i> (%)	<i>ANI</i> (%)
<i>Kitasatospora</i> sp. K002	86.86	90.09	93.41

O alinhamento dos genomas de K002 e sua linhagem mais próxima *K. xanthocidica* JCM 4862 mostra um número considerável de regiões desalinhadas (53.45%) indicando que os genomas são divergentes entre si (Fig. 3.6). Apenas os blocos de alinhamento do genoma com pelo menos 1.000 nucleotídeos e 90% de identidade foram considerados.

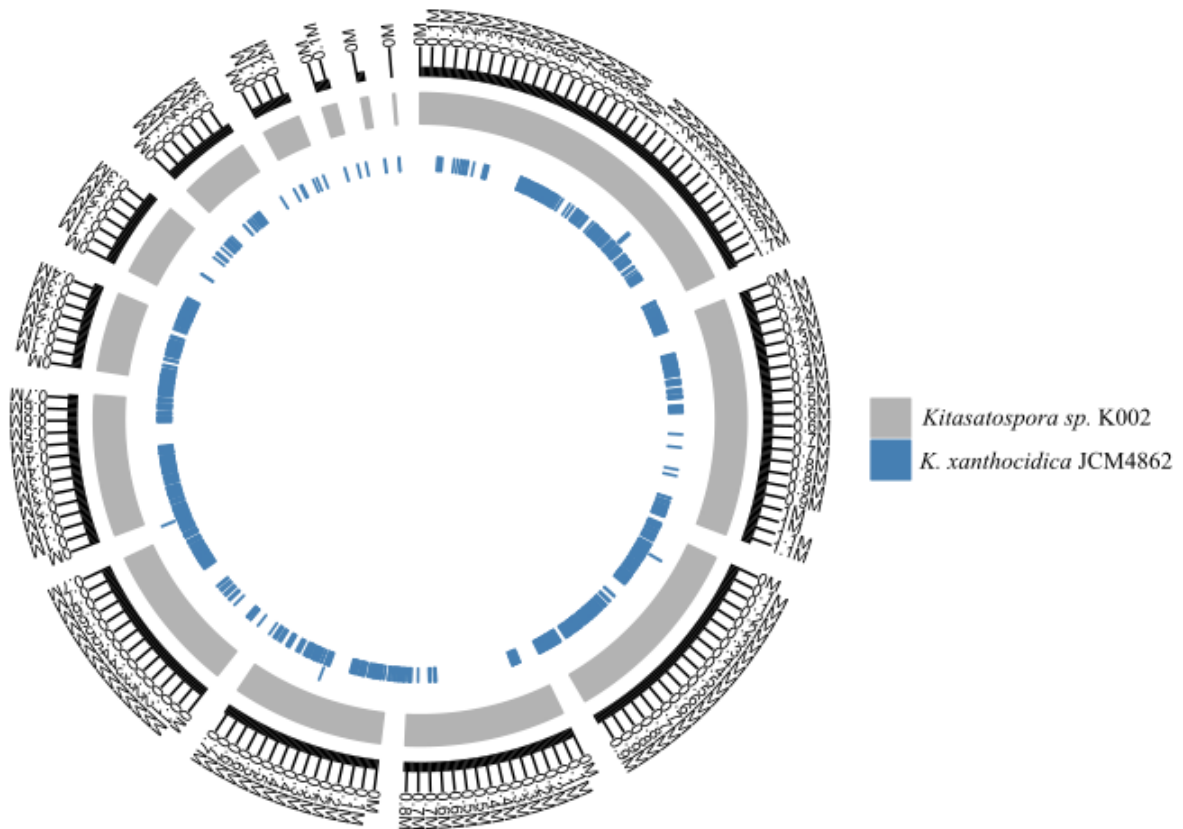


Figura 3.6 Visualização circular do alinhamento dos genomas de *Kitasatospora sp.* K002 (em cinza) e sua linhagem mais próxima de acordo com TYGS-DSMZ, *K. xanthocidica* JCM 4862(em azul). O alinhamento e a ilustração foram gerados usando o PlotNucmer. Apenas os blocos de alinhamento do genoma com pelo menos 1000 nucleotídeos e 90% de identidade são mostrados. A banda externa é a escala de tamanho dos *contigs*. Quando a banda azul tem duas camadas de profundidade, há duas sequências na linhagem mais próxima que se sobrepõem no mesmo local em *Kitasatospora sp.* K002.

Expandindo a análise de genômica comparativa para investigar as relações filogenéticas desse gênero, utilizamos as sequências do genoma de 12 linhagens (K002, 10 espécies mais próximas segundo o TYGS-DSMZ e 1 *outgroup*) para extrair conjuntos de genes ortólogos. Usando o servidor *web* M1CR0B1AL1Z3R, um total de 520 núcleos de

genes (subconjunto de genes compartilhados entre as 12 linhagens) foram recuperados e usados para construir uma árvore filogenética de máxima verossimilhança (Fig. 3.7).

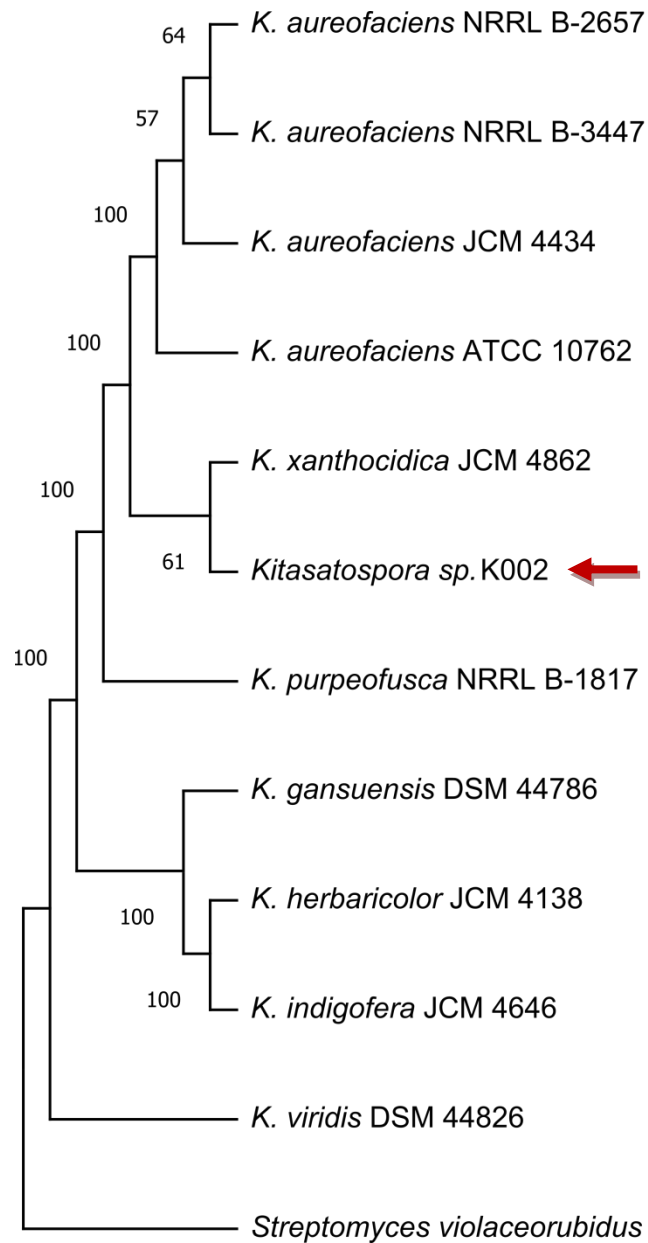


Figura 3.7 Árvore filogenética de máxima verossimilhança baseada em 520 genes ortólogos principais concatenados de seqüências genômicas de 12 linhagens de *Kitasatospora*, incluindo *Kitasatospora sp. K002*, as linhagens intimamente relacionadas de acordo com o TYGS-DSMZ e o grupo externo *Streptomyces violaceorubidus*. O servidor web MICR0BIAL1Z3R foi utilizado para concatenar os genes ortólogos e construir a árvore filogenética.

Foram identificadas 36568 famílias gênicas compondo o pangenoma de K002 e suas linhagens mais próximas e, dessas, 1383 (3.7%) estavam presentes no núcleo do genoma e as restantes como acessórias. Os genes raros foram representados por 23308 famílias, ou seja, 63.73% do pangenoma (Fig. 3.8).

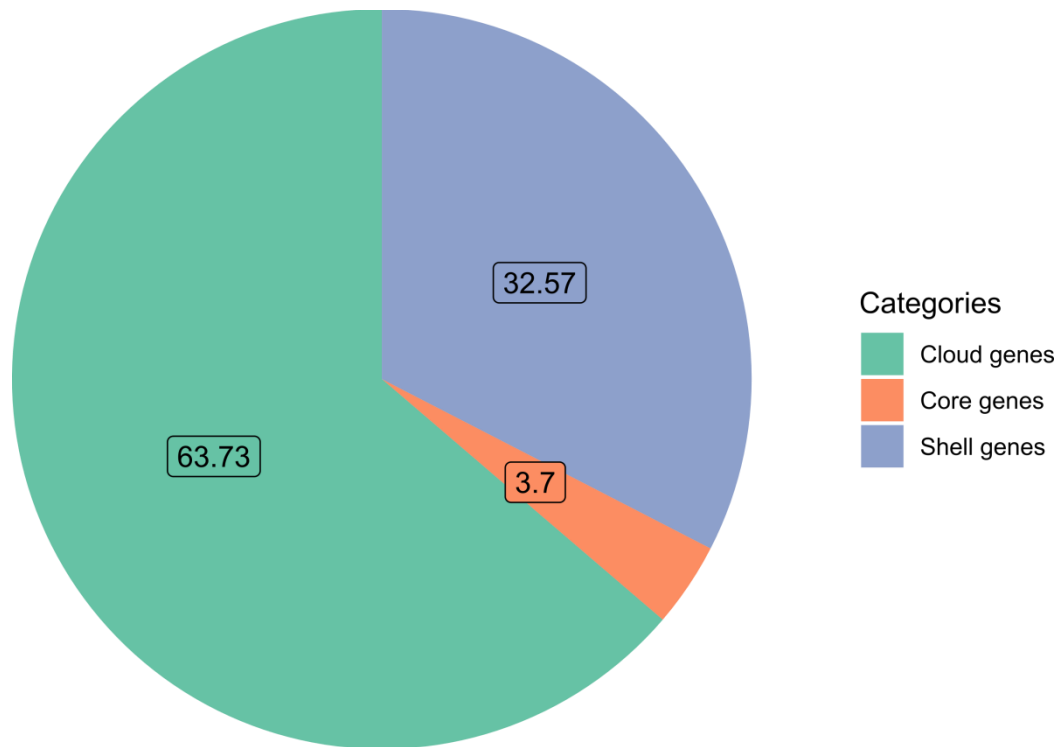


Figura 3.8 Análise comparativa de pangenoma realizada no Roary. O genoma de *Kitasatospora sp.* K002e os 10 genomas mais próximos a ele, de acordo com TYGS-DSMZ, foram utilizadas. Uma identidade percentual mínima para BLASTp de 80% foi definida para agrupar os genes que codificam sequências de proteínas completas em genomas centrais.

3.4.3 Características funcionais do genoma

A figura 3.9 apresenta a distribuição das proteínas do genoma do isolado K002 nas suas respectivas categorias *COGs*. A categoria S refere-se a proteínas de função desconhecida. Aproximadamente 17% das proteínas do genoma de K002 foram atribuídas a esse grupo. A categoria *unknow* (13%) representa as proteínas que não foram anotadas em nenhuma das categorias *COG*. Entre as categorias *COG* bem anotadas no genoma de K002, as categorias K (11%), E (7%) e Q (5%), referentes a transcrição, ao metabolismo de aminoácidos e biossíntese/catabolismo de metabólitos secundários, respectivamente, estão entre as mais abundantes no genoma.

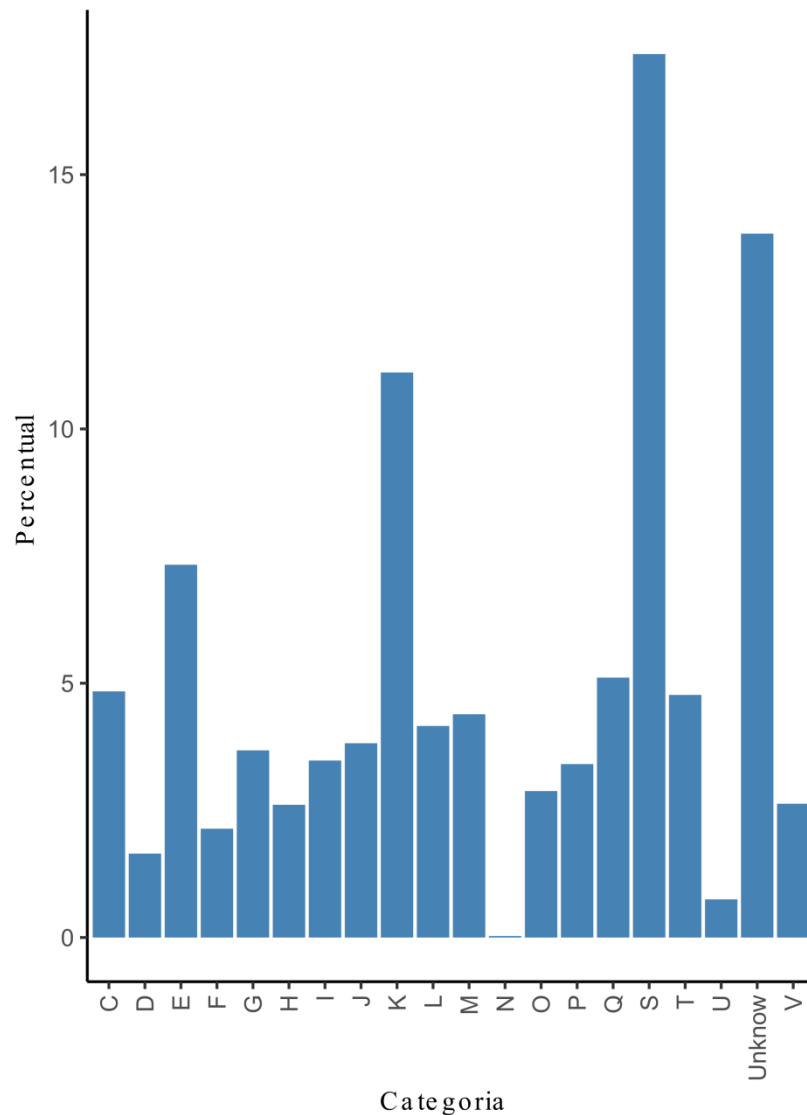


Figura 3.9 Proteínas preditas no genoma de *Kitasatospora sp.* K002, obtidas a partir do sobrenadante de *P. elgii*, classificadas de acordo com as categorias COGs. A classificação foi realizada usando o mapeador eggNOG contra o banco de dados da *Kyoto Encyclopedia of Genes and Genomes (KEGG)*.

No que diz respeito a resistência antibióticos, foram encontradas uma RNA polimerase do tipo helicase (81.32% de similaridade) e uma Rifampicina glicosiltransferase (77.7% de similaridade) em K002, ambos conferindo resistência a rifamicina. Os mecanismos de resistência foram proteção e inativação do antibiótico, respectivamente. Uma bomba de efluxo de antibiótico também foi detectada com 44.23% de similaridade (Tab. 3.6).

Tabela 3.6 Anotação de genes potenciais responsáveis pela resistência a antibióticos de *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante de *P. elggi*. A análise foi realizada em comparação com o Banco de Dados de Genes de Resistência a Antibióticos (ARDB).

Gene	Família	Classe de droga	Mecanismo de resistência	Similaridade (%)
HelR	RNA polimerase RNA tipo helicase	Rifamicina	Proteção de alvo antibiótico	81.32
rgt1438	Rifampicina glicosiltransferase	Rifamicina	Inativação de antibiótico	77.7
qacJ	Bomba de de pequena resistência a múltiplas drogas (SMR)	Agentes desinfetantes e antissépticos	Bomba de efluxo	44.23

Os catorze *contigs* de *Kitasatospora sp.* K002 foram analisados quanto a presença de CRISPR. Foram encontrados CRISPR em seis *contigs* (1, 2, 3, 7, 8 e 9) e no *contig* 8 também foram encontradas proteínas cas Tipo IE (Tab. 3.7).

Tabela 3.7 Resultados da análise da presença de locus CRISPR em *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante de *P. elgii*. A análise foi realizada através da ferramenta de reconhecimento CRISPRCasFinder.

Contig 1			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CRISPR	342676	342949	3
CRISPR	606097	606188	1
CRISPR	901808	901904	1
CRISPR	1515073	1516077	4
Contig 2			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CRISPR	226307	226383	1
CRISPR	484355	484448	1
Contig 3			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CRISPR	875621	875706	1
Contig 7			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CRISPR	298093	299098	16
Contig 8			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CAS-TypeIE	38089	46348	6
CRISPR	175343	175614	4
CRISPR	178426	178699	4
CRISPR	200917	201982	17
CRISPR	228377	228956	9
CRISPR	230591	231351	12
CRISPR	232885	233584	11
CAS-TypeIE	298659	306442	5
CRISPR	306661	306994	5
CRISPR	310217	310794	9
Contig 9			
Elemento	Início (pb)	Fim (pb)	Nº de espaçadores
CRISPR	241749	241865	1

Um total de 60 *clusters* gênicos biossintéticos de 22 tipos diferentes foram detectados através do AntiSMASH em *Kitasatospora sp.* K002. Dentre eles destaca-se *NRPS*, o mais abundante (23), lantipeptídeos de 4 classes diferentes, *PKS* dos 3 tipos, entre outros.

A tabela 3.8 apresenta os *clusters* de K002 em comparação aos presentes em *K. xanthocidica* JCM 4862, onde 76 *clusters* gênicos de 26 tipos diferentes foram detectados, em *K. acidiphila* MMS16-CNU292T (Kim *et al.*, 2020), onde 43 *clusters* gênicos de 11 tipos diferentes foram detectados, e em *Kitasatospora sp.* MMS16-BH015 (Yun *et al.*, 2020), onde 38 *clusters* gênicos de 14 tipos diferentes foram detectados. Ao todo 217 *clusters* gênicos

foram identificados nos 4 genomas de *Kitasatospora*. Embora existam clusters específicos, todos possuem *NRPS*, terpeno e pelo menos um tipo de lantipeptídeo e *PKS*. A presença dos *clusters* Betalactona e Betalactâmico distingue de *Kitasatospora sp.* K002 das demais.

Tabela 3.8 Resultados da análise de *clusters* de genes biossintéticos para metabólitos secundários de *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante de *P. elgii*. A análise foi realizada por meio do *software* AntiSMASH (v. 6.1.1) e apresentada de forma comparativa com outras linhagens do gênero. Linhagens: 1, *Kitasatospora* K002; 2, *K. xanthocidica* JCM 4862; 3, *K. acidiphila* MMS16-CNU292T; 4, *Kitasatospora sp.* MMS16-BH015.

	Metabólitos secundários			
	Nº clusters			
	1	2	3	4
Aminoglicosídeo/Aminociclitol	1	0	0	1
Bacteriocina	0	0	0	1
Betalactona	1	0	0	0
Betalactâmico	1	0	0	0
Butirolactona	2	4	0	2
CDPS	0	1	0	0
Ectoína	0	1	0	0
Furano	1	1	0	0
Guanidinotitas	0	1	0	0
Homoserina lactona	0	0	0	1
Indol	0	0	1	0
Ladderano	0	1	1	0
Lantipeptídeo classe I	1	2	6	4
Lantipeptídeo classe II	1	2	0	0
Lantipeptídeo classe III	1	2	2	0
Lantipeptídeo classe IV	1	4	0	0
Lantipeptídeo classe V	0	2	0	0
<i>LAP</i>	1	2	0	0
Lassopectídeo	0	1	1	0
Melanina	0	1	0	0
<i>NPAA</i>	1	2	0	0
<i>NRPS</i>	23	20	12	8
Nucleosídeo	0	0	0	1
Outro	1	2	0	3
<i>PKS</i>	1	0	2	0
Rantipeptídeo	0	1	0	0
<i>RIPP</i>	2	2	0	0
Sideróforo	1	2	2	1
<i>T1PKS</i>	7	9	7	5
<i>T2PKS</i>	1	2	0	2
<i>T3PKS</i>	2	1	4	1
Terpeno	7	8	5	5
Tiopeptídeo	1	1	0	3
TransAT- <i>PKS</i>	2	1	0	0
TOTAL	60	76	43	38

De acordo com a análise no BiG-SCAPE, foram detectados 81 famílias de BCGs para as linhagens de *Kitasatospora* K002 e *K. xanthocidica* JCM 4862, onde apenas nove são semelhantes entre as ambas as linhagens. São elas: 4 terpenos, 2 NRPS, 1 T3PKS, 1 sideróforo e 1 butirolactona (Fig. 2.10).

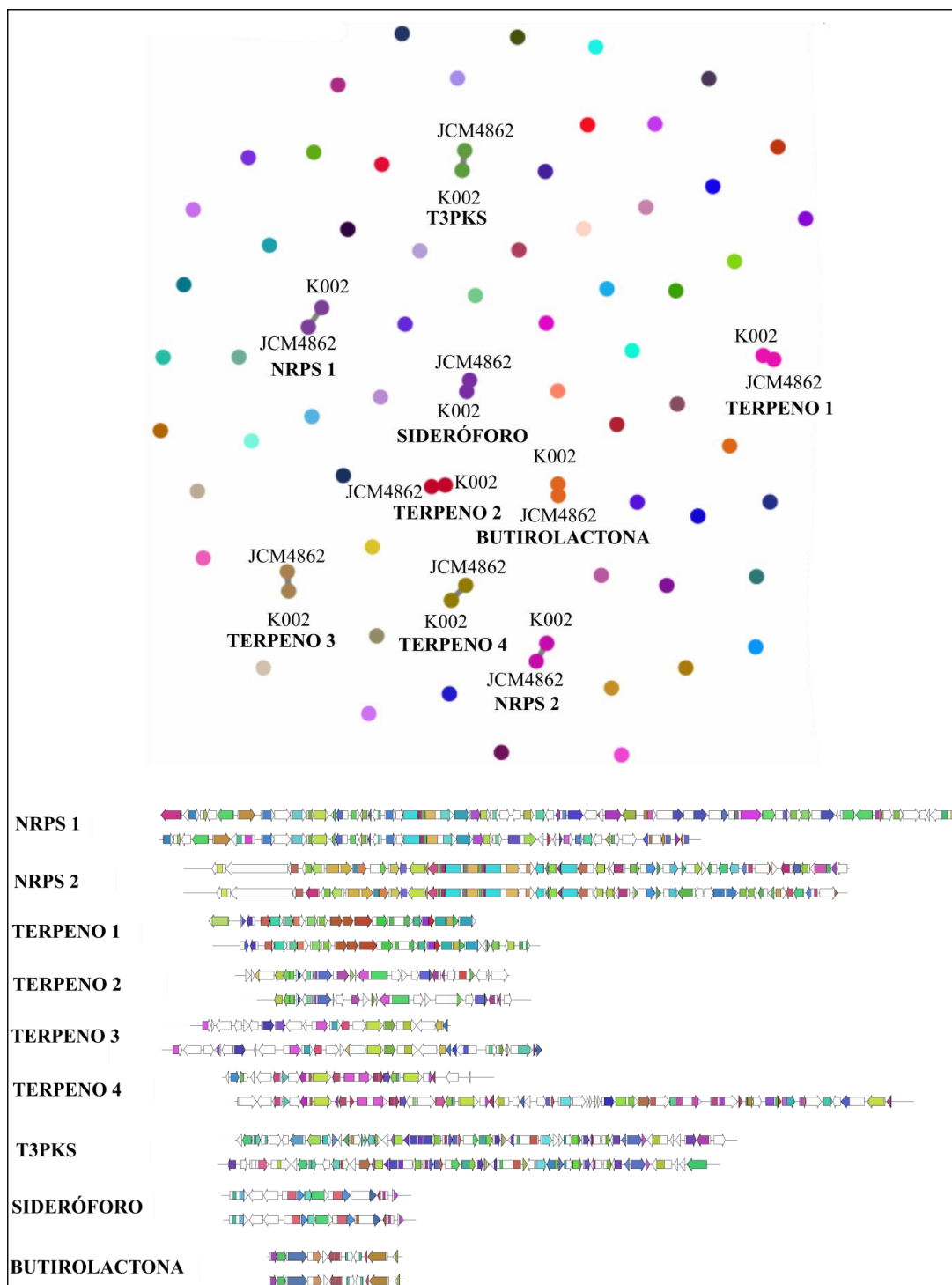


Figura 3.10 Análise de similaridade entre as famílias de *clusters* de genes usando BiG-SCAPE. Cada união representa *clusters* gênicos semelhantes entre as linhagens *Kitasatospora* K002 e *K. xanthocidica* JCM 4862. Os *clusters* específicos, ou seja, sem semelhanças entre as linhagens também são apresentados em forma isolada. A representação gráfica dos *clusters* gênicos semelhantes também é apresentada.

O banco de dados de dados CAMP_{R4} para predição de peptídios antimicrobianos (AMPs) mostrou 59 predições, em K002. Nele, prevê-se que a sequência seja antimicrobiana através de uma pontuação de probabilidade (0 a 1) (Waghu *et al.*, 2014). Quanto maior a probabilidade, maior é a possibilidade do peptídeo ser antimicrobiano. As predições de K002 apresentaram probabilidade superior a 0.68, com maior probabilidade de 0.78.

No genoma de K002, oito áreas potenciais ligadas a produção de bacteriocinas sintetizadas ribossalmente foram encontradas através do BAGEL4, são elas: um *cluster* gênico ligado a produção de zoocina A (bacteriocina classe III), um *cluster* gênico ligado a produção de peptídeos lineares contendo azol (LAPs) (bacteriocinas de classe I), um *cluster* gênico putativo relacionado à biossíntese de bacteriocinas de classe não identificada, três *clusters* gênicos relacionados com a biossíntese de lantibióticos (bacteriocina classe I) e dois *clusters* gênicos associados a produção de sactipeptídeos (bacteriocina classe I) (Tab. 3.9).

Tabela 3.9 Resultados da análise de genes envolvidos na biossíntese de peptídeos de *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante de *P. elgii*. A análise foi realizada por meio do servidor Bagel4.

Início (bp)	Fim (bp)	Classe
141602	161902	<i>Zoocin A</i>
362141	382141	<i>LAPs</i>
770112	790871	<i>PutativeBacteriocinfamilyprotein</i>
62228	82228	<i>Lanthipeptideclass IV</i>
647639	667639	<i>Lanthipeptideclass II</i>
43982	63982	<i>Lanthipeptideclass I</i>
131183	151183	<i>Sactipeptides</i>
239900	259900	<i>Sactipeptides</i>

3.4.4 Caracterização fenotípica *in silico*

A figura 3.11 apresenta a comparação do fenótipo de K002 e suas linhagens mais próximas de acordo com DSMZ. A *Kitasatospora sp.* K002 se diferencia das demais por ser negativa para fermentação de glicose. K002 compartilha várias características com *Kitasatospora xanthocidica* JCM 4862, sua linhagem mais próxima segundo o DSMZ, entretanto, K002 difere por ser negativa para pigmento amarelo, para crescimento em DNase, tartarato, fermentação de glicose, celobiose, l -arabinose, l -ramnose, melibiose, o -nitrofenil-β- D -galactopiranosídeo (β-galactosidase), rafinose e salicina, e positiva para arginina desidrolase.

De acordo com a predição fenotípica no TraitAr, as linhagens mais próximas a K002 são *Kitasatospora aureofaciens* JCM4434 e NRRLB3447. A *Kitasatospora* sp. K002 difere das demais por ser positiva para oxidase, d -xilose, d -Manitol e indol, e por ser negativa para pigmento amarelo, crescimento em tartarato e fermentação de glicose. K002 difere apenas de NRRLB3447 por não crescer em *DNAse* e por crescer em mio-inositol. E difere apenas de JCM4434 por oxidar glicose.

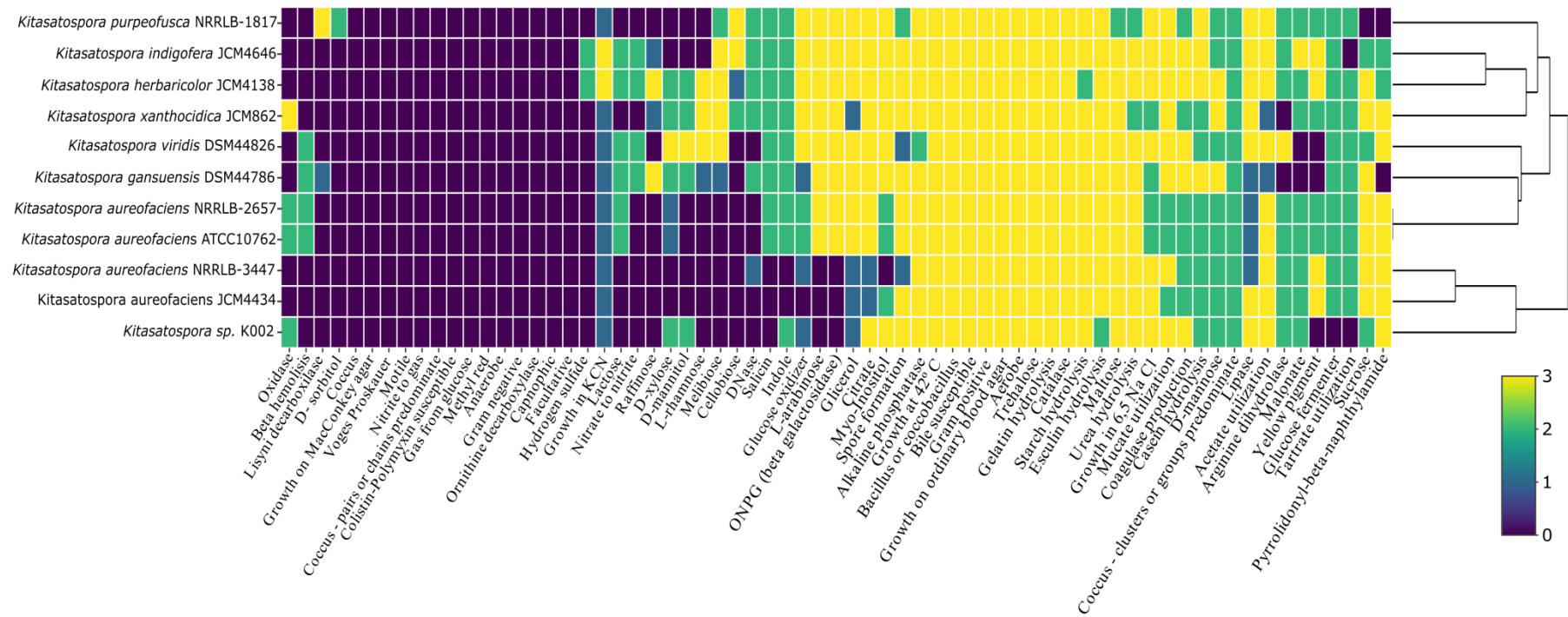


Figura 3.11 Comparação entre as predições fenotípicas da montagem híbrida do genoma de *Kitasatospora* sp. K002, obtida a partir do uso do sobrenadante de *P. elgii*, e dos genomas de suas linhagens mais próximas de acordo com o TYGS-DSMZ obtidas no banco de dados do NCBI. O valor zero significa que a linhagem é negativa para o fenótipo, os valores 1 e 2 correspondem ao positivo para o fenótipo nos preditores *phypat* e *phypat+PGL*, respectivamente. O valor 3 mostra que o fenótipo foi positivo nos dois preditores. A análise foi realizada pelo *software*

Traitair.

A tabela 3.10 mostra a predição fenotípica de K002 no Traitair em comparação com as características de outras linhagens do gênero (Kim *et al.*, 2020). *Kitasatospora sp.* K002 difere das demais por ser positiva para Arginina dehidrolase, produção de indol e utilização de citrato e maltose.

Tabela 3.10 Comparação entre as predições fenotípicas da montagem híbrida do genoma de *Kitasatospora sp.* K002, obtida a partir do uso do sobrenadante de *P. elgii*, e dos genomas de outras linhagens do gênero. Linhagens: 1, *Kitasatospora* K002; 2, *K. acidiphila* MMS16-C NU292T; 3, *K. azatica* KCTC 9699T; 4, *K. kifunensis* IFO 15206T; 5, *K. purpeofusca* NRRL B-1817T e 6, *K. nipponensis* HKI 0315T.+, Positivo; -, negativo, w, fraco.

Carcacterísticas	1	2	3	4	5	6
Hidrólise:						
Caseína	+	+	-	-	-	+
DNA	-	-	-	+	+	-
Amido	+	+	+	+	+	+
Enzima:						
Produção de ácido a partir da glicose	-	-	-	-	+	-
Arginina dehidrolase	+	-	-	-	-	-
β -Galactosidase	-	-	+	+	+	+
Liquefação de gelatina	+	w	-	+	+	+
β -Glucosidase (hidrólise de esculina)	+	-	+	-	+	+
Produção de indol	+	-	-	-	-	-
Redução de nitrato	-	-	-	-	-	-
Urease	+	-	-	-	-	+
Fosfato alcalino	+	+	+	+	+	+
Lipase	+	-	-	-	+	-
Crescimento em fontes únicas de carbono:						
Arabinose	-	-	+	-	-	+
Citrato	+	-	-	-	-	-
Glicose	+	W	+	-	+	+
Maltose	+	-	-	-	-	-
Manitol	+	-	-	-	-	+
Manose	+	-	-	-	-	+

3.5 DISCUSSÃO

A família *Streptomycetaceae* é caracterizada por um intrincado ciclo de vida micelial e características marcantes de multicelularidade (Flardh e Buttner, 2009; Claessen *et al.*, 2014). E ao contrário da maioria dos Streptomycetos, o gênero *Kitasatospora* produz esporos submersos com micélios em cultura líquida verificado nas micrografias de K002. A morfologia variada em micrografias de *Kitasatospora sp.* como observada também em K002 já foi questionada gerando dúvida se havia contaminação ou não, mas por fim foi confirmado a variedade de tipos de células encontradas em culturas líquidas deste gênero (Takahashi, 2017).

A análise de ácidos graxos por cromatografia gasosa tem sido rotineiramente usada na identificação de bactérias com base em seus perfis únicos de ácidos graxos. Mais de 300 ácidos graxos e compostos relacionados foram encontrados em bactérias analisadas no *MIDI Research and Development Laboratory*. A riqueza de informações contidas nesses compostos pode ser estimada considerando não apenas a presença ou ausência de cada ácido, mas também usando os dados de forma quantitativa (Sasser, 2001). O perfil de ácidos graxos de K002 concordam com o relatado por Kämpfer (2012) e verificado em Kim e colaboradores (2020) em *K. acidiphilae* por Klaysubun e colaboradores (2022) em *K. humi*, onde ácidos graxos iso e anteiso predominam em linhagens do gênero *Kitasatospora*.

Quanto a atividade antibacteriana, os resultados alcançados nesse trabalho concordam com os achados de Srivastava e colaboradores (2019), onde extratos de *Kitastospora* apresentaram atividade antibacteriana contra *B. subtilis* e *E. coli*.

Embora a análise do gene *rRNA* 16S de K002 através do BLAST tenha indicado grande similaridade (99,24%) com *K. kifunensis*, a análise genômica indicou maior similaridade com *K. xanthocidica* concordando com o banco de dados Ezbiocloud para *rRNA* 16S. Os valores de *ANI* < 95% e o *dddH* < 70% calculados para K002, de acordo com os parâmetros genômicos *OGRI* (*Overall Genome Related Index*) para o delineamento de novas espécies, indicam que se trata de uma espécie distinta do gênero *Kitasatospora* (Chun *et al.*, 2018).

Por se tratar de uma nova espécie, é possível que o enriquecimento do cultivo com moléculas indutoras a partir do uso do sobrenadante de *P. elgii* tenha favorecido a recuperação microbiana, já que algumas bactérias só conseguem crescer em laboratório quando em co-cultivo com outro membro da comunidade (Boilatabi *et al.*, 2021). Assim, o sobrenadante funcionou como uma co-cultura (Stewart, 2012), simulando algumas das relações químico-ecológicas que ocorreriam naturalmente (Turan *et al.*, 2017).

De acordo com o estudo de quarenta e cinco linhagens de *Kitasatospora* feito por Li e colaboradores (2021), o tamanho dos genomas de *Kitasatospora* variam em tamanho de 6.3 a 12.36 Mbp, com tamanho médio de 8.7 ± 1.0 Mbp. Isso mostra que K002 não está entre os maiores genomas do gênero (7.9 Mbp). O estudo também mostra que o conteúdo de GC do gênero varia entre 71.2 e 74.5% com média de 73.1 % (Li *et al.*, 2021), logo K002 corresponde a média com 72.8%. O mesmo trabalho mostra número de sequências codificantes por genoma de *Kitasatospora* varia entre 5482 e 9863 e K002 apresenta 6708.

No estudo pangenômico feito também por Li e colaboradores (2021), o genoma central desse gênero foi estimado em 1.476 famílias de genes e a semelhança com os valores de K002 (1383) sugere que K002 retém os componentes essenciais do genoma central de *Kitasatospora*. Li e colaboradores (2021) em sua análise verificaram que o genoma central era formado 2.7% das famílias, enquanto K002 possui 3.7% das famílias em seu genoma central e que os genes raros representavam 70.5% do pangenoma, enquanto em K002 representa 63.73%. Esses resultados indicam que o conteúdo gênico do gênero é altamente divergente.

Assim como em *Kitasatospora sp.* MMS16-BH015, *K. griseola*, *K. mediocidica*, *K. setae* e *Kitasatospora sp.* OK780, onde foram relatados mais de 20% de proteínas desconhecidas (Yun *et al.*, 2020) classificadas nas categorias S e Unknow, em K002 ambas frações das categorias somam 30% e apontam uma das limitações da anotação de genomas com referência em bancos de dados, já que eles baseiam suas prospecções apenas em estruturas já descritas.

No geral, todos os genomas de *Kitasatospora* exibem perfil COG semelhante para as proteínas conhecidas, embora algumas diferenças sutis tenham sido observadas. Em comparação com o genoma de *Kitasatospora sp.* MMS16-BH015, as categorias K e E também são relatadas como mais abundantes assim como em K002, entretanto a terceira categoria mais abundante em *Kitasatospora sp.* MMS16-BH015 é a G relativa a transporte e metabolismo de carboidratos. A abundância da categoria Q (terceira mais abundante em K002) concorda com o observado em genomas de *Streptomyces* (Bentley *et al.*, 2002; Ikeda *et al.*, 2003; Wang *et al.*, 2013; Zaburannyi *et al.*, 2014), gênero intimamente relacionado a *Kitasatospora*, onde o metabolismo secundário representa uma porção significativa do genoma. Vale ressaltar que categoria T, a quinta mais abundante em K002 (4,76%) relativa a transdução de sinal é conhecida por seus papéis reguladores na biossíntese de antibióticos em actinomicetos (Rodríguez *et al.*, 2013).

Um locus CRISPR-Cas consiste em matrizes agrupadas de repetições palindrômicas curtas regularmente interespaçadas (CRISPR) e proteínas Cas adjacentes associadas ao

CRISPR (Koonin e Makarova, 2019). A linhagem *Kitasatospora sp.* K002 codificou um *locus* CRISPR-Cas completo. As células procarióticas usam esses *locus* como um sistema imunológico adaptativo para se proteger de vírus e outros elementos genéticos móveis potencialmente prejudiciais (Wimmer e Beisel 2020; Qiu *et al.*, 2016). A presença de uma provável matriz CRISPR sugere capacidade de aceitar *DNA* estranho. Esse fator gera entusiasmo em relação ao sistema CRISPR-Cas quanto ao seu poder de edição do *DNA* que favorece a engenharia genética (Sedeek *et al.*, 2019). Isso pode oferecer uma abordagem revolucionária para a criação de novos tratamentos para uma ampla gama de doenças genéticas (Huang *et al.*, 2022).

Embora *Streptomyces* seja o modelo da fisiologia bacteriana e um importante recurso para a descoberta de antibióticos, seus parentes próximos como *Kitasatospora* ainda não foram amplamente estudados (Li *et al.*, 2021). Entretanto, vários novos produtos químicos já foram descobertos a partir de linhagens de *Kitasatospora*, incluindo compostos com estruturas únicas e atividades interessantes (Kim *et al.*, 2020). *Kitasatospora sp.* MBT66 foi identificada como um produtor prolífico de antibióticos com atividade contra patógenos Gram-positivos e Gram-negativos (Zhu *et al.*, 2014). Espécies de *Kitasatospora* também foram caracterizadas como produtoras de novos metabólitos secundários de estruturas variadas, que são principalmente ativos contra organismos eucarióticos (Takahashi, 2017).

Dentre os sessenta *clusters* gênicos identificados em K002, vinte e dois *clusters* apresentaram baixa semelhança (<10%) com compostos já conhecidos, o que sugere a potencial produção de novos compostos bioativos. São eles: T1PKS (1), T3PKS (1), Butirolactona (2), NRPS (7), LAP (1), Tiopeptídeo (1), RIPP (2), Sideróforo (1), NPAA (1), Terpeno (3), Lantipeptídeo de classe IV (1) e Betalactâmico (1). A média para as sequências de referência de *Kitasatospora* é cinquenta e quatro *clusters* (Zin *et al.*, 2021) e ilustra como K002 está acima da média, mesmo não estando entre os maiores genomas do gênero. O baixo número de famílias *clusters* gênicos similares entre *Kitasatospora sp.* K002 e sua linhagem de referência, reforça o fato de K002 se tratar de uma espécie distinta.

De acordo com Li e seus colaboradores (2021), *clusters* biossintéticos de *Kitasatospora* podem ser agrupados em oito classes biossintéticas principais: RIPPs, NRPS, terpeno, PKS-NRP híbridos, outras PKS, PKS tipo I, sacarídeo e de não classificáveis. Exceto por sacarídeos, todos os *clusters* estão presentes em K002.

A presença de vias biossintéticas da sintase peptídica não ribossomal (NRPS) e da sintase de policetídeo (PKS) em uma Actinobactéria está altamente relacionada com o seu potencial biossintético (Amim *et al.*, 2020). Os *clusters* PKS e NRPS, ambos encontrados em

K002, representam dois grandes grupos de produtos naturais. Eles possuem notável diversidade e atividades biológicas como os antibióticos eritromicina, penicilina, daptomicina e vancomicina, os antifúngicos anfotericina e equinocandina, os agentes anticancerígenos epotilona e bleomicina, o agente redutor de colesterol lovastatina, os imunossuppressores FK506 e ciclosporina e os antibióticos veterinários monensina e avermectina (Wang *et al.*, 2011).

As *PKS* são classificadas em Tipos I, II e III, todos encontrados dentre os *clusters* de K002 (Hopwood, 1997). Dentro das *PKS* tipo I encontra-se um grupo de produtos chamados poliéteres ionóforos. Eles causam a morte das células alvo, sendo de alta importância por sua atividade antiparasitária, antifúngica, antiviral, herbicida, anti-inflamatória e imunorreguladora (Wang *et al.*, 2011). No gênero *Streptomyces*, pertencente a mesma família que *Kitasatospora*, identifica-se uma via biossintética de *PKS* dos tipos II e III para a produção de um pigmento característico dos esporos (Funa *et al.*, 2005).

Enquanto as vias *NRPS* e *PKS* são bem documentadas em actinomicetos, existem muitos aglomerados de genes biossintéticos, que precisam ser desvendados para a produção de metabólitos potencialmente novos (Becerril *et al.*, 2018). Os esforços de sequenciamento do genoma da primeira década do século 21 revelaram que uma grande classe de produtos naturais é formada por peptídeos sintetizados ribossomalmente e modificados após a tradução (*RIPP*). Trata-se de uma classe de produtos naturais com atividade antimicrobiana, cujo *cluster* é encontrado em K002. As modificações pós-traducionais diversificam a estrutura destes produtos naturais, o que implica no melhor reconhecimento dos alvos e no aumento da estabilidade química e metabólica, melhorando sua funcionalidade química. Eles são categorizados em cerca de vinte classes diferentes como *LAPs*, lantipeptídeos e tiopeptídeos, cujos *clusters* específicos também estão presentes em K002 (Arnison *et al.*, 2013). *LAPs* constituem uma grande classe de *RIPPs* e até o momento, são conhecidos dezessete *LAPs*, a maioria dos quais apresenta atividade citolítica (Ex. estreptolisina S, clostridiolisina S, listeriolisina S) ou atividades antibacterianas (Travin *et al.*, 2020).

O *cluster* Betalactona, presente em K002, refere-se a um inibidor de protease contendo β -lactona. As β -lactonas são estruturas encontradas em muitos produtos naturais. Elas geralmente têm propriedades anticancerígenas ou antibióticas. Eles são uma fonte rica para pesquisa de novas drogas, mas costumam ser ofuscadas pelos β -lactâmicos, também presentes em K002, assim seu potencial terapêutico permanece pouco explorado em relação aos β -lactâmicos, que compreendem mais de 50% dos antibióticos no mercado hoje (Robinson *et al.*, 2018).

O *cluster* Furano, encontrado em K002, refere-se ao anel furano, substância biologicamente ativa com altas propriedades terapêuticas, dentre elas: antibacterianos, antivirais, antiinflamatórios, antifúngicos, antitumorais, anti-hiperglicêmicos, analgésicos, anticonvulsivantes, entre outros (Koca *et al.*, 2005; Kirilmis *et al.*, 2008).

Dentre outras propriedades, os sideróforos, cujo *cluster* também está presente em K002 têm sido relacionados a mecanismos de virulência em microorganismos. Além disso, eles têm aplicações biotecnológicas, sendo usados como fertilizantes de solo e agentes de biocontrole, tratamento de anemias e malária, além de ser usado na administração seletiva de antibióticos em bactérias resistentes. Essa é uma aplicação potencialmente poderosa que usa as habilidades de transporte de ferro dos sideróforos para transportar drogas para as células por meio da preparação de conjugados entre sideróforos e agentes antimicrobianos (Saha *et al.*, 2016).

O *cluster* Butirolactona relacionado a produção de butirolactonas pode se relacionar com atividade antibiótica de K002. Moléculas sinalizadoras chamadas γ -butirolactonas são produzidas por espécies de *Streptomyces*, gênero irmão a *Kitasatospora*, nas quais regulam a produção de antibióticos (Takano, 2006). Estas moléculas são produzidas pela maioria de *Streptomyces* e outras actinobactérias e estão envolvidas na ativação do metabolismo secundário (Bibb, 2005).

Poucos metabólitos terpenoides têm sido detectados em procariotos em relação aos encontrados em fungos e plantas. As primeiras moléculas encontradas em bactérias foram os terpenos odoríferos, principalmente, a geosmina, um sesquiterpeno degradado, que proporciona o odor característico de terra húmida. Em K002, estão presentes *clusters* terpenoides. Terpenos de origem microbiana têm sido usados como potentes agentes antimicrobianos como o pentaleneno ou a terapentecina (Cane e Ikeda, 2012; Ikeda *et al.*, 2003; Yamada *et al.*, 2015).

Aminoácidos não proteínogênicos (NPAA) são aqueles que não ocorrem naturalmente no código genético de um organismo, porém possuem importância biológica, pois muitos são utilizados como intermediários metabólicos ou são encontrados em proteínas com modificações pós-traducionais. O NPAA é um *cluster* de poliaminoácidos semelhantes a Epsilon-poli-L-lisina (ϵ -PL) presente no isolado K002. A ϵ -PL é um peptídeo antimicrobiano natural com um amplo espectro de atividade contra a deterioração de alimentos e bactérias de intoxicação alimentar *in vitro* (Chang *et al.*, 2010; Ongpipattanakul *et al.*, 2022).

A busca por antibióticos valiosos começa com a triagem de bactérias para detectar a melhor fonte de novos metabólitos bioativos. Dessa forma, a análise funcional de genes biossintéticos feitas neste trabalho confirma o gênero *Kitasatospora* como uma importante fonte de compostos com potencial para aplicação biotecnológica. O uso do sbrenandante de *P. elgii* na seleção de crescimento de *Kitasatospora sp.* K002 mostrou-se uma abordagem eficaz e direcionada para encontrar novos potenciais produtores de compostos bioativos com atividade contra os patógenos para os quais novos antibióticos são necessários (van Bergeijk *et al.*, 2020).

3.6 CONCLUSÕES

Os resultados das análises filogenéticas e das características morfofisiológicas permitem afirmar que *Kitasatospora sp.* K002 pertence ao gênero *Kitasatospora*. No entanto, que K002 representa uma nova espécie do gênero, para a qual o nome *Kitasatospora brasiliensis* sp. nov. é proposto.

Os resultados desse trabalho indicam que o genoma de K002 abriga vários genes responsáveis pela produção biossintética de metabólitos secundários confirmando o potencial do gênero em ser uma valiosa fonte de compostos bioativos, o que é consistente com seu fenótipo de atividade antibacteriana constatada nos ensaios *overlay* do capítulo anterior.

A obtenção de K002 a partir do uso do sobrenadante de *P. elgii* atesta a eficiência do protocolo em obter linhagens com potencial biotecnológico. Além disso, pode indicar que a metodologia favorece o isolamento de novas espécies bacterianas, visto que se trata de uma espécie ainda não descrita.

O genoma de K002 apresenta diversos *clusters* gênicos que podem estar relacionados à resistência ao sobrenadante de *P. elgii*. O *cluster NRPS* responsável pela produção de pelgipeptinas em *P. elgii* e *clusters* de lantipeptídeos, responsáveis pela produção de elgicinas em *P. elgii* podem comprovar a premissa do primeiro capítulo de que bactérias resistentes a certos peptídeos antimicrobianos, provavelmente produzem compostos semelhantes a eles.

Este trabalho amplia o entendimento geral e reforça o interesse biotecnológico sobre este táxon.

4 CONCLUSÃO GERAL

O enriquecimento de meios de cultura com o sobrenadante de *P. elgii* como mecanismo de seleção para obtenção de linhagens com potencial para aplicação biotecnológica se mostrou eficiente. Linhagens de gêneros distintos de bactérias foram isoladas, onde todos já são conhecidos por seu interesse biotecnológico em diversas áreas. Uma destas, que foi o foco deste trabalho, foi a atividade antibacteriana comprovada em todos os isolados através dos ensaios *overlay*.

A metodologia utilizada neste trabalho poderá ser utilizada com diferentes meios, concentrações e amostras a fim de refinar e otimizar o protocolo, ampliar as coleções de linhagens com potencial biotecnológico e até permitir o isolamento de linhagens não descritas.

A caracterização taxonômica polifásica de *Kitasatospora sp.* K002 permitiu sua identificação como uma nova espécie do gênero *Kitasatospora*. A caracterização da linhagem K003, apesar de se tratar da espécie já conhecida *M. radiotolerans*, também foi importante, pois permitiu ampliar o conhecimento sobre a espécie no que diz respeito ao seu potencial para produção de compostos bioativos.

A anotação funcional de K002 e K003, onde vários genes responsáveis pela produção biossintética de metabólitos secundários foram encontrados, permitiu a comprovação do potencial para produção de compostos de interesse biotecnológico das linhagens obtidas a partir do uso do sobrenadante de *P. elgii*.

O isolamento das culturas puras permitirá maior exploração do potencial dessas linhagens. Os compostos antimicrobianos poderão ser caracterizados por espectrometria de massa e purificados para futura caracterização de suas estruturas, bem como sua comparação com compostos já caracterizados para *P. elgii*, espécie utilizada na seleção da linhagem.

A identificação e isolamento de metabólitos poderá ser guiado pela percepção genômica das origens biossintéticas de seus metabólitos potenciais obtidos neste estudo. Ensaios *in vivo* e *in vitro* poderão ser realizados para avaliar o potencial de aplicação dos antimicrobianos obtidos como produtos viáveis para a indústria farmacêutica.

Assim, este estudo mostrou a eficiência do uso do sobrenadante de *P. elgii* em selecionar micro-organismos com potencial para aplicação biotecnológica e em favorecer o isolamento de novas espécies bacterianas. A maioria dos *clusters* de genes biossintéticos das linhagens obtidas não possuíam similaridade com *clusters* publicados e, portanto, representam uma nova fonte de produtos farmacêuticamente relevantes.

5 REFERÊNCIAS BIBLIOGRÁFICAS

Afgan E Baker D VAN Den Beek M Blankenberg D Bouvier D et al (2016) The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2016 Update. *Nucleic Acids Res* 44:W3-W10. <https://doi.org/10.1093/nar/gkw343>

Alamgir KM Masuda S Fujitani Y Fukuda F Tani A (2015) Production of ergothioneine by *Methylobacterium* species. *Front Microbiol* 27:6-1185. <https://doi.org/10.3389/fmicb.2015.01185>

Alessa O Ogura Y Fujitani Y Takami H Hayashi T Sahin N Tani A (2021) Comprehensive Comparative Genomics and Phenotyping of *Methylobacterium* Species. *Front Microbiol* 6:12-740610. <https://doi.org/10.3389/fmicb.2021.740610>

Amin DH Abdallah NA Abolmaaty A et al. (2020) Microbiological and molecular insights on rare Actinobacteria harboring bioactive prospective. *Bull Natl Res Cent* 44:5. <https://doi.org/10.1186/s42269-019-0266-8>

Amin MA Hoque MN Siddiki AZ Saha S Kamal MM (2020) Antimicrobial resistance situation in animal health of Bangladesh. *Vet World* 13(12):2713-2727. <https://doi.org/10.14202/vetworld.2020.2713-2727>

Anandaraj B VellaichamyA Kachman M Selvamanikandan A Pegu S Murugan V (2009) Co-production of two new peptide antibiotics by a bacterial isolate *Paenibacillus alvei* NP75. *BiochemBiophys Res Commun* 6;379(2):179-85. <https://doi.org/10.1016/j.bbrc.2008.12.007>

Ardanov P SessitschA Häggman H Kozyrovska N Pirttilä AM (2012) *Methylobacterium*-induced endophyte community changes correspond with protection of plants against pathogen attack. *PLoS One* 7(10):e46802. <https://doi.org/10.1371/journal.pone.0046802>

Arnison PG Bibb MJ Bierbaum G Bowers AA Bugni TS Bulaj G et al. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* 30(1):108-60. <https://doi.org/10.1039/c2np20085f>. PMID: 23165928

Arkin AP Cottingham RW Henry CS Harris N, Stevens RL Maslov S (2018) KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* 6;36(7):566-569. <https://doi.org/10.1038/nbt.4163>

Asaf S Numan M Khan AL Al-Harrasi A (2020) *Sphingomonas*: from diversity and genomics to functional role in environmental remediation and plant growth. *Crit Rev Biotechnol* 40(2):138-152. <https://doi.org/10.1080/07388551.2019>

Ash C Priest FG Collins MD (1993) Molecular identification of *rRNA* group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Proposal for the creation of a new genus *Paenibacillus*. *Antonie Van Leeuwenhoek*. 1993-1994;64(3-4):253-60. <https://doi.org/10.1007/BF00873085>

Avram O Rapoport D Portugez S Pupko T (2019) M1CR0B1AL1Z3R-a user-friendly web server for the analysis of large-scale microbial genomics data. *Nucl Acids Res* 47:W88–W92. <https://doi.org/10.1093/nar/gkz423>

Bachmann BO Van Lanem SG Baltz RH (2014) Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J Ind Microbiol Biotechnol* 41(2):175-84. <https://doi.org/10.1007/s10295-013-1389-9>.

Babenko LM Kosakivska IV Romanenko KO (2022) Molecular mechanisms of N-acetyl homoserine lactone signals perception by plants. *Cell Biol Int* 46(4):523-534. <https://doi.org/10.1002/cbin.11749>

Bagley MC Dale JW Merritt EA Xiong X (2005) Thiopeptide antibiotics. *Chem Rev* 105(2):685-714. <https://doi.org/10.1021/cr0300441>

Bahram M Hildebrand F Forslund SK Anderson JL Soudzilovskaia NA et al. (2018). Structure and function of the global topsoil microbiome. *Nature* 560(7717):233-237. <https://doi.org/10.1038/s41586-018-0386-6>

Balachandran C Duraipandiyar V Ignacimuthu S (2012) Cytotoxic (A549) and antimicrobial effects of *Methylobacterium* sp. isolate (ERI-135) from Nilgiris forest soil, India. *AsianPac J Trop Biomed* 2(9):712-6. [https://doi.org/10.1016/S2221-1691\(12\)60215-9](https://doi.org/10.1016/S2221-1691(12)60215-9)

Barrangou R Doudna JA (2016) Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* 34(9):933-941. <https://doi.org/10.1038/nbt.3659>

Becerril A Alvarez S Brana AF Rico S Diaz M et al. (2018) Uncovering production of specialized metabolites by *Streptomyces argillaceus*: activation of cryptic biosynthesis gene clusters using nutritional and genetic approaches. *PLoS One* 13(5):e0198145. <https://doi.org/10.1371/journal.pone.0198145>

Bentley SD Chater KF Cerdeño-Tárraga AM Challis GL Thomson NR et al. (2002). Complete Genome Sequence of the Model Actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417: 141–147. <https://doi.org/10.1038/417141a>

Bérdy J (2005) Bioactive microbial metabolites. *J Antibiot (Tokyo)* 58(1):1-26. <https://doi.org/10.1038/ja.2005.1>.

Bérdy J (2012) Thoughts and facts about antibiotics: where we are now and where we are heading. *J Antibiot (Tokyo)* 65(8):385-95. <https://doi.org/10.1038/ja.2012.27>

Bibb MJ (2005) Regulation of secondary metabolism in *streptomycetes*. *Curr Opin Microbiol* 8(2):208-15. <https://doi.org/10.1016/j.mib.2005.02.016>

Bijlani S Singh NK Eedara VV Podile AR Mason CE Wang CCC Venkateswaran K (2021) *Methylobacterium ajmalii* sp. nov., Isolated From the International Space Station. *Front Microbiol* 12:639396. doi: 10.3389/fmicb.2021.639396

Biniarz P Łukaszewicz M Janek T (2017) Screening concepts, characterization and structural analysis of microbial-derived bioactive lipopeptides: a review. *Crit Rev Biotechnol* 37(3):393-410. <https://doi.org/10.3109/07388551.2016.1163324>

Blin K Shaw S Kloosterman AM Charlop-Powers Z van Wezel GP et al. (2021) AntiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49(W1):W29-W35. <https://doi.org/10.1093/nar/gkab335>

Boilattabi N Barrassi L Bouanane-Darenfed A La Scola B (2021). Isolation and identification of *Legionella spp.* from hot spring water in Algeria by culture and molecular methods. *J Appl Microbiol* 130:1394–1400. <https://doi.org/10.1111/jam.14871>

Brenner DJ Krieg NR Staley JT Garrity GM (2005) *Bergey's Manual of Systematic Bacteriology*. New York, NY: Springer. 2005.

Bricker BJ (2000) Characterization of the three ribosomal RNA operons *rRNA*, *rrnB*, and *rrnC*, from *Brucella melitensis*. *Gene* 5;255(1):117-26. [https://doi.org/10.1016/s0378-1119\(00\)00273-0](https://doi.org/10.1016/s0378-1119(00)00273-0)

Carvalho LS (2013) Utilização de Elgicinas para Obtenção de Cepas Bacterianas de Interesse Biotecnológico. Dissertação, Universidade de Brasília.

Cain CC Henry AT Waldo RH Casida LJ Jr Falkinham JO (2000) Identification and characteristics of a novel *Burkholderia* strain with broad-spectrum antimicrobial activity. *Appl Environ Microbiol* 66(9):4139-41. <https://doi.org/10.1128/AEM.66.9.4139-4141.2000>

Cane DE Ikeda H (2012) Exploration and mining of the bacterial terpenome. *Acc Chem Res* 45(3):463-72. <https://doi.org/10.1021/ar200198d>

Carver T Thomson N Bleasby A Berriman M Parkhill J (2009) DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics.* 25(1):119-20. <https://doi.org/10.1093/bioinformatics/btn578>

Case RJ Boucher Y Dahllöf I Holmström C Doolittle WF Kjelleberg S (2007) Use of 16S *rRNA* and *rpoB* genes as molecular markers for microbial ecology studies. *Appl Environ Microbiol* 73(1):278-88. <https://doi.org/10.1128/AEM.01177-06>

Challinor VL Bode HB (2015) Bioactive natural products from novel microbial sources. *Ann N Y Acad Sci* 1354:82-97. <https://doi.org/10.1111/nyas.12954>

Challis GL (2014) Exploitation of the *Streptomyces coelicolor* A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. *J Ind Microbiol Biotechnol* 41(2):219-32. <https://doi.org/10.1007/s10295-013-1383-2>

Chang SS Lu WY Park SH Kang DH (2010) Control of foodborne pathogens on ready-to-eat roast beef slurry by epsilon-polylysine. *Int J Food Microbiol* 141(3):236-41. <https://doi.org/10.1016/j.ijfoodmicro.2010.05.021>

Chaudhary DK Khulan A Kim J (2019). Development of a novel cultivation technique for uncultured soil bacteria. *Sci. Rep.* 9:6666. <https://doi.org/10.1038/s41598-019-43182-x>

Chaumeil PA Mussig AJ Hugenholtz P Parks DH (2022) GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38(23):5315-5316. <https://doi.org/10.1093/bioinformatics/btac672>

- Chen L Wang N Wang X Hu J Wang S (2010) Characterization of two anti-fungal lipopeptides produced by *Bacillus amyloliquefaciens* SH-B10. *Bioresour Technol* 101(22):8822-7. <https://doi.org/10.1016/j.biortech.2010.06.054>
- Choi J Yang F Stepanauskas R Cardenas E Garoutte A et al. (2017). Strategiesto improve referencedatabases for soilmicrobiomes. *ISME J.* 11:829–34.<https://doi.org/10.1038/ismej.2016.168>
- Chokshi A Sifri Z Cennimo D Horng H (2019) Global ContributorstoAntibioticResistance. *J Glob Infect Dis* 11(1):36-42. https://doi.org/10.4103/jgid.jgid_110_18
- Chow V Nong G St John FJ Rice JD Dickstein E Chertkov O (2012) Complete genome sequence of *Paenibacillus* sp. strain JDR-2. *Stand Genomic Sci* 19;6(1):1-10. <https://doi.org/10.4056/sigs.2374349>
- Chun J Oren A Ventosa A Christensen H Arahall DR da Costa MS (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst EvolMicrobiol.* 2018 Jan;68(1):461-466. <https://doi.org/10.1099/ijsem.0.002516>
- Chun J Rainey FA (2014) Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *Int J Syst EvolMicrobiol* 64:316–24. <https://doi.org/10.1099/ijse.0.054171-0>
- Claesen J Bibb M (2010) Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified peptides. *Proc Natl Acad Sci U S A* 107(37):16297-302. <https://doi.org/10.1073/pnas.1008608107>
- Claessen D Rozen DE Kuipers OP Søggaard-Andersen L van Wezel GP (2014) Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat Rev Microbiol* 12(2):115-24. <https://doi.org/10.1038/nrmicro3178>.
- Cochrane SA Vederas JC (2016) Lipopeptides from *Bacillus* and *Paenibacillus* spp.: A Gold Mine of Antibiotic Candidates. *Med Res Rev* 36(1):4-31. <https://doi.org/10.1002/med.21321>
- Colwell RR (1970) Polyphasic taxonomy of the genus vibrio: numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. *J Bacteriol* 104(1):410-33. <https://doi.org/10.1128/jb.104.1.410-433.1970>
- Cotter PD Ross RP Hill (2013) Bacteriocins - a viable alternative to antibiotics? *Nat Rev Microbiol* 11(2):95-105. <https://doi.org/10.1038/nrmicro2937>
- Corretto E Antonielli L SessitschACompant S Höfer C et al. (2017) Complete Genome Sequence of the Heavy Metal Resistant Bacterium *Agromyces aureus* AR33T and Comparison with Related Actinobacteria. *Stand. Genomic Sci.* 12:2. <https://doi.org/10.1186/s40793-016-0217-z>
- Couvin D Bernheim AToffano-Nioche C Touchon M Michalik J et al. (2018) CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* 46(W1):W246-W251. <https://doi.org/10.1093/nar/gky425>

da Costa RA Andrade IEPC Pinto OHB de Souza BBP Fulgêncio DLA Mendonça ML (2022) A novel family of non-secreted tridecaptin lipopeptide produced by *Paenibacilluselgii*. *Amino Acids* 54(11):1477-1489. <https://doi.org/10.1007/s00726-022-03187-9>

Da Cruz RB Galdino PM Penna KG Hoffmann K Costa EA Bataus LA (2013) Acetone extract from *Streptovorticillium* sp., a bacterium isolated from Brazilian Cerrado soil, induces anti-inflammatory activity in mice. *An Acad Bras Cienc* 85(2):595-603. <https://doi.org/10.1590/S0001-37652013000200009>

da Cunha BR Fonseca LP Calado CRC (2019) Antibiotic Discovery: Where Have We Come from, Where Do We Go? *Antibiotics* (Basel) 24;8(2):45. <https://doi.org/10.3390/antibiotics8020045>

Dadgostar P (2019) Antimicrobial Resistance: Implications and Costs. *Infect Drug Resist* 20;12:3903-3910. <https://doi.org/10.2147/IDR.S234610>

de Koning W Miladi M Hiltemann S Heikema A Hays JP Flemming S (2020) NanoGalaxy: Nanopore long-read sequencing data analysis in Galaxy. *Gigascience* 17;9(10):giaa105. <https://doi.org/10.1093/gigascience/giaa105>

Depoorter E Bull MJ Peeters C Coenye T Vandamme P Mahenthiralingam E (2016) *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers. *Appl Microbiol Biotechnol* 100(12):5215-29. <https://doi.org/10.1007/s00253-016-7520-x>

Ding R Li Y Qian C Wu X (2011) Draft genome sequence of *Paenibacilluselgii* B69, a strain with broad antimicrobial activity. *J Bacteriol* 193(17):4537. <https://doi.org/10.1128/JB.00406-11>

Dodds WK Zeglin LH Ramos RJ Platt TG Augusto FB (2020) Connections and feedback: aquatic, plant, and soil microbiomes in heterogeneous and changing environments. *BioScience* 70(7):548–62. <https://doi.org/10.1093/biosci/biaa046>

Dourado MN Andreote FD Dini-Andreote F Conti R Araújo JM Araújo WL (2012) Analysis of 16S *rRNA* and *mxoF* genes revealing insights into *Methylobacterium* niche-specific plant association. *Genet Mol Biol* 35(1):142-8. <https://doi.org/10.1590/s1415-47572012005000017>

Duquesne S Destoumieux-Garzón D Peduzzi J Rebuffat S (2007) Microcins, gene-encoded antibacterial peptides from enterobacteria. *Nat Prod Rep* 24(4):708-34. <https://doi.org/10.1039/b516237h>

Durand GA Raoult D Dubourg G (2019) Antibiotic discovery: history, methods and perspectives. *Int J Antimicrob Agents* 53:371–82, <http://dx.doi.org/10.1016/j.ijantimicag.2018.11.010>

Elo S Suominen I Kämpfer P Juhanoja J Salkinoja-Salonen M Haahtela K (2001) *Paenibacillus borealis* sp. nov., a nitrogen-fixing species isolated from spruce forest humus in Finland. *Int J Syst Evol Microbiol* 51(Pt 2):535-545. <https://doi.org/10.1099/00207713-51-2-535>

El-Gebali S Mistry J Bateman A Eddy SR Luciani A *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* 47(D1):D427-D432. doi: 10.1093/nar/gky995

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–91. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>

Finking R Marahiel MA (2004) Biosynthesis of nonribosomal peptides1. *Annu Rev Microbiol* 58:453-88. <https://doi.org/10.1146/annurev.micro.58.030603.123615>

Flaherty RA Freed SD Lee SW (2014) The wide world of ribosomally encoded bacterial peptides. *PLoS Pathog* 10(7):e1004221. <https://doi.org/10.1371/journal.ppat.1004221>

Flärdh K Buttner MJ (2009) *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat Rev Microbiol* 7(1):36-49. <https://doi.org/10.1038/nrmicro1968>.

Foster TJ (2017) Antibiotic resistance in *Staphylococcus aureus*. Current status and future prospects. *FEMS Microbiol Rev* 41:430–49, <http://dx.doi.org/10.1093/femsre/fux007>

Fraser BA Whiting JR Paris JR Weadick CJ Parsons PJ Charlesworth D et al. (2020) Improved Reference Genome Uncovers Novel Sex-Linked Regions in the Guppy (*Poecilia reticulata*). *Genome Biology and Evolution* 12:1789–805. <https://doi.org/10.1093/gbe/evaa187>

Frieri M Kumar K Boutin A (2017) Antibiotic resistance. *J Infect Public Health* 10:369–78, <http://dx.doi.org/10.1016/j.jiph.2016.08.007>.

Funa N Funabashi M Ohnishi Y Horinouchi S (2005) Biosynthesis of hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in *Streptomyces griseus*. *J Bacteriol* 187(23):8149-55. <https://doi.org/10.1128/JB.187.23.8149-8155.2005>

Gabrielsen C Brede DA Nes IF Diep DB (2014) Circular bacteriocins: biosynthesis and mode of action. *Appl Environ Microbiol* 80(22):6854-62. <https://doi.org/10.1128/AEM.02284-14>

Gallego V García MT Ventosa A (2005) *Methylobacterium variabile* sp. nov., a methylotrophic bacterium isolated from an aquatic environment. *Int J Syst Evol Microbiol* ;55(Pt 4):1429-1433. <https://doi.org/10.1099/ijs.0.63597-0>

Galperin MY Wolf YI Makarova KS Vera Alvarez R Landsman D Koonin EV (2021) COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res* 49(D1):D274-D281. <https://doi.org/10.1093/nar/gkaa1018>

Garrido-Cardenas JA Manzano-Agugliaro F (2017) The metagenomics worldwide research. *Curr Genet* 63(5):819-829. <https://doi.org/10.1007/s00294-017-0693-8>

Gawde U Chakraborty S Waghugh FH Barai RS Khanderkar A et al. (2023) CAMPR4: a database of natural and synthetic antimicrobial peptides. *Nucleic Acids Research*.

Gillis P Vandamme P De Vos P Swings J Kersters K (2005) *Polyphasic taxonomy*. In: Bonne DR Castenholz RW Garrity GM (eds) *Bergey's Manual of Systematic Bacteriology*. New York, NY: Springer 43–48. 2005.

González GDT Sigristi R PAULO BS (2016) Avanços Recentes na Manipulação Genética de Organismos para a Produção de Peptídeos não Ribossomais. *Revista Virtual de Química* 6:1998–2025. Disponível em: <https://rvq.s bq.org.br/>

Grady EN MacDonald J Liu L Richman A Yuan ZC (2016) Current knowledge and perspectives of *Paenibacillus*: a review. *Microb Cell Fact* 15(1):203. <https://doi.org/10.1186/s12934-016-0603-7>

Gray B Hall P Gresham H (2013) Targeting agr- and agr-Like quorum sensing systems for development of common therapeutics to treat multiple gram-positive bacterial infections. *Sensors (Basel)* 13(4):5130-66. <https://doi.org/10.3390/s130405130>

Green PN (2006) *Methylobacterium*. In: Dworkin M Falkow S Rosenberg E Schleifer KH Stackebrandt E (eds) *The Prokaryotes*, 3rd edn, 257–265. New York: Springer. 2006.

Green PN Bousfield IJ (1982) *A taxonomic study of some Gram-negative facultatively methylotrophic bacteria*. *Microbiol* 128:623–638. <https://doi.org/10.1099/00221287-128-3-623>

Grossi CEM Fantino E Serral F Zawoznik MS Fernandez Do Porto DA Ulloa RM (2020) *Methylobacterium* sp. 2A Is a Plant Growth-Promoting Rhizobacteria That Has the Potential to Improve Potato Crop Yield Under Adverse Conditions. *Front Plant Sci* 11:71. <https://doi.org/10.3389/fpls.2020.00071>

Gurevich A Saveliev V Vyahhi N Tesler G (2013) QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29(8):1072-5. <https://doi.org/10.1093/bioinformatics/btt086>

Gutleben J Loureiro C Ramírez Romero L A Shetty S Wijffels R H Smidt H et al. (2020). Cultivation of bacteria from *Aplysinaaerophoba*: effects of oxygen and nutrient gradients. *Front Microbiol* 11:175. <https://doi.org/10.3389/fmicb.2020.00175>

Hammami R Zouhir A Le Lay C Ben Hamida J Fliss I (2010) BACTIBASE second release: a database and tool platform for bacteriocin characterization. *BMC Microbiol* 10:22. <https://doi.org/10.1186/1471-2180-10-22>

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* ;68(4):669-85. <https://doi.org/10.1128/MMBR.68.4.669-685.2004>

Hille F Charpentier E (2016) CRISPR-Cas: biology, mechanisms and relevance. *Philos Trans R Soc Lond B Biol Sci* 371(1707):20150496. <https://doi.org/10.1098/rstb.2015.0496>

Ho A Di Lonardo DP Bodelier PL (2017) Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiol Ecol* 93(3). <https://doi.org/10.1093/femsec/fix006>

Hu T Chitnis N Monos D Dinh A (2021) Next-generation sequencing technologies: An overview. *Hum Immunol* 82(11):801-811. <https://doi.org/10.1016/j.humimm.2021.02.012>.

Hua ZS Han YJ Chen LX Liu J Hu M Li SJ Kuang JL Chain PS Huang LN Shu WS (2015) Ecological roles of dominant and rare prokaryotes in acid mine drainage revealed by

metagenomics and metatranscriptomics. ISME J 9(6):1280-94. <https://doi.org/10.1038/ismej.2014.212>

Huang J Zhou Y Li J Lu A Liang C (2022) CRISPR/Cas systems: Delivery and application in gene therapy. Front BioengBiotechnol 22:10-942325. <https://doi.org/10.3389/fbioe.2022.942325>.

Huang S Liu Y Liu WQ Neubauer P Li J (2021) The Nonribosomal Peptide Valinomycin: From Discovery to Bioactivity and Biosynthesis. Microorganisms 9(4):780. <https://doi.org/10.3390/microorganisms9040780>

Huerta-Cepas J Forslund K Coelho LP Szklarczyk D Jensen LJ et al. (2017) Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. Mol Biol Evol 34(8):2115-2122. <https://doi.org/10.1093/molbev/msx148>

Hug JJ Bader CD Remškar M Cirnski K Müller R (2018) Concepts and Methods to Access Novel Antibiotics from Actinomycetes. Antibiotics (Basel) 7(2):44. doi: 10.3390/antibiotics7020044

Ichikawa N Oguchi A Ikeda H Ishikawa J Kitani S et al. (2010) Genome Sequence of *Kitasatospora setae* NBRC 14216T: an Evolutionary Snapshot of the Family Streptomycetaceae. DNA Res 17:393–406. <https://doi.org/10.1093/dnares/dsq026>

Ikeda H Ishikawa J Hanamoto A Shinose M Kikuchi H Shiba et al. (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat Biotechnol 21(5):526-31. <https://doi.org/10.1038/nbt820>

Jacoby R Peukert M SuccurroAKoprivova A Kopriva S (2017) The Role of Soil Microorganisms in Plant Mineral Nutrition-Current Knowledge and Future Directions. Front Plant Sci 8:1617. <https://doi.org/10.3389/fpls.2017.01617>

Jaspars M (2014) The origins of cyanobactin chemistry and biology. Chem Commun (Camb 50(71):10174-6. <https://doi.org/10.1039/c3cc49252d>

Jayashree S Lalitha R Vadivukkarasi P Kato Y Seshadri S (2011) Cellulase production by pink pigmented facultative methylotrophic strains (PPFMs). Appl BiochemBiotechnol 164(5):666-80. <https://doi.org/10.1007/s12010-011-9166-6>

Ju KS Doroghazi JR Metcalf WW (2014) Genomics-enabled discovery of phosphonate natural products and their biosynthetic pathways. J Ind MicrobiolBiotechnol ;41(2):345-56. <https://doi.org/10.1007/s10295-013-1375-2>

Kämpfer P. (2012) Genus *Incertaesedis* I. *Kitasatospora*. Bergey's Manual of Systematic Bacteriology 5:1768–1777.

Kämpfer P Glaeser SP (2012) Prokaryotic taxonomy in the sequencing era--the polyphasic approach revisited. Environ Microbiol 14(2):291-317. <https://doi.org/10.1111/j.1462-2920.2011.02615.x>

Kanehisa M Araki M Goto S Hattori M Hirakawa M et al. (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36(Database issue):D480-4. <https://doi.org/10.1093/nar/gkm882>

Kang MS Srinivasan S (2018) Complete genome sequence of *Methylobacterium* sp. 17Sr1-43, a radiation-resistant bacterium. *Mol Cell Toxicol* 14:453–457 (2018). <https://doi.org/10.1007/s13273-018-0050-0>

Kapoor G Saigal S Elongavan A (2017) Action and resistance mechanisms of antibiotics: A guide for clinicians. *J Anaesthesiol Clin Pharmacol* 33(3):300-305. https://doi.org/10.4103/joacp.JOACP_349_15

Kapinusova G Lopez Marin MA Uhlik O (2023) Reaching unreachables: Obstacles and successes of microbial cultivation and their reasons. *Front Microbiol* 14:1089630. <https://doi.org/10.3389/fmicb.2023.1089630>

Kautsar SA Blin K Shaw S Navarro-Muñoz JC Terlouw BR *et al.* (2020) MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res* 48(D1):D454-D458

Kelly DP McDonald IR Wood AP (2014) *The family Methylobacteriaceae* In: Rosenberg E DeLong EF Lory S Stackebrandt E Thompson F (eds) *The Prokaryotes- Alphaproteobacteria and Betaproteobacteria*. Berlin: Springer 313–340.

Kim MJ Roh SG Kim M K Park C Kim S Kim SB (2020). *Kitasatosporaacidiphila* sp. nov., isolated from pine grove soil, exhibiting antimicrobial potential. *Int J Syst Evol Microbiol* 70:5567–5575. <https://doi.org/10.1099/ijsem.0.004450>

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16:111–20. <https://doi.org/10.1007/BF01731581>

Kirilmis C Ahmedzade M Servi S Koca M Kizirgil A Kazaz C (2008) Synthesis and antimicrobial activity of some novel derivatives of benzofuran: part 2. The synthesis and antimicrobial activity of some novel 1-(1-benzofuran-2-yl)-2-mesitylethanone derivatives. *Eur J Med Chem* 43(2):300-8. <https://doi.org/10.1016/j.ejmech.2007.03.023>

Klaysubun C Srisuk N Duangmal K (2022) *Kitasatosporahumi* sp. nov., isolated from a tropical peat swamp forest soil, and proposal for the reclassification of *Kitasatosporapsammitica* as a later heterotypic synonym of *Kitasatospora aureofaciens*. *Int J Syst Evol Microbiol* 72(5). <https://doi.org/10.1099/ijsem.0.005356>.

Knief C Dengler V Bodelier PL Vorholt JA (2012) Characterization of *Methylobacterium* strains isolated from the phyllosphere and description of *Methylobacterium longum* sp. nov. *Antonie Van Leeuwenhoek* 101(1):169-83. <https://doi.org/10.1007/s10482-011-9650-6>

Koca M Servi S Kirilmis C Ahmedzade M Kazaz C Ozbek B Otük G (2005) Synthesis and antimicrobial activity of some novel derivatives of benzofuran: part 1. Synthesis and antimicrobial activity of (benzofuran-2-yl)(3-phenyl-3-methylcyclobutyl) ketoxime derivatives. *Eur J Med Chem* 40(12):1351-8. <https://doi.org/10.1016/j.ejmech.2005.07.004>

Koonin E Makarova K (2019) Origins and evolution of CRISPR-Cas systems. *Philos Trans R Soc Lond B Biol Sci* 374-20180087. <https://doi.org/10.1098/rstb.2018.0087>

Krüger RH Henrique CH Araujo JFKurok A (2013) Peptídeos com atividade antimicrobiana, composições medicamentosas para o tratamento de animais, composições para o tratamento e profilaxia de plantas, usos dos ditos peptídeos, e uso do extrato de *Paenibacillus elgiouofinensis*. Patente: Privilégio de Inovação. Número do registro: JP2014-504122, Instituição de registro: Japanese Patent Office.

Kumar L Patel SKS Kharga K Kumar R Kumar P Pandohee J et al. (2022) Molecular Mechanisms and Applications of N-Acyl Homoserine Lactone-Mediated Quorum Sensing in Bacteria. *Molecules* 27(21):7584. <https://doi.org/10.3390/molecules27217584>

Kurtz S Phillippy A Delcher AL Smoot M Shumway M et al. (2004) Versatile and open software for comparing large genomes. *Genome Biol.* 5(2):R12. <https://doi.org/10.1186/gb-2004-5-2-r12>

Labeda DP Goodfellow M Brown R Ward AC Lanoot B et al. (2012) Phylogenetic Study of the Species within the Family Streptomycetaceae. *Antonie Van Leeuwenhoek* 101:73–104. <https://doi.org/10.1007/s10482-011-9656-0>

Lagha B A Haas B Gottschalk M Grenier D (2017) Antimicrobial potential of bacteriocins in poultry and swine production. *Vet Res* 48(1):22. <https://doi.org/10.1186/s13567-017-0425-6>

Lagier JC Hugon P Khelaifia S Fournier PE La Scola B Raoult D (2015) The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev* 28(1):237-64. <https://doi.org/10.1128/CMR.00014-14>

Lane DJ Pace B Olsen GJ Stahl DA Sogin ML Pace NR (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci USA* 82(20):6955-9. <https://doi.org/10.1073/pnas.82.20.6955>

Langmead B Wilks C Antonescu V Charles R (2019) Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics.* 35(3):421-432. <https://doi.org/10.1093/bioinformatics/bty648>

Larkin MA Blackshields G Brown NP Chenna R McGettigan PA (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-8. <https://doi.org/10.1093/bioinformatics/btm404>

Laws M Shaaban A Rahman KM (2019) Antibiotic resistance breakers: current approaches and future directions. *FEMS Microbiol Rev* 43(5):490-516. <https://doi.org/10.1093/femsre/fuz014>

Ley RE Bäckhed F Turnbaugh P Lozupone CA Knight RD Gordon JI (2005) Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* 102(31):11070-5. <https://doi.org/10.1073/pnas.0504978102>

Li J Beatty PK Shah S Jensen SE (2007) Use of PCR-targeted mutagenesis to disrupt production of fusaricidin-type antifungal antibiotics in *Paenibacilluspolymyxa*. *Appl Environ Microbiol* 73(11):3480-9. [https://doi.org/ 10.1128/AEM.02662-06](https://doi.org/10.1128/AEM.02662-06)

Li Y Wang M Sun ZZ Xie BB (2021) Comparative genomic insights into the taxonomic classification, diversity, and secondary metabolic potentials of *Kitasatospora*, a genus closely related to *Streptomyces*. *Front Microbiol* 12:683814. [https:// doi. org/ 10. 3389/ fmicb. 2021. 683814](https://doi.org/10.3389/fmicb.2021.683814)

Lin B Hui J Mao H (2021) Nanopore Technology and Its Applications in Gene Sequencing. *Biosensors (Basel)* 11(7):214. [https://doi.org/ 10.3390/bios11070214](https://doi.org/10.3390/bios11070214)

Ling LL Schneider T Peoples AJ Spoering AL Engels I Conlon BP (2015) A new antibiotic kills pathogens without detectable resistance. *Nature* 520(7547):388. [https://doi.org/ 10.1038/nature14303](https://doi.org/10.1038/nature14303)

Liu B Pop M (2009) ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res.* 37(Database issue):D443-7. <https://doi.org/10.1093/nar/gkn656>

Liu S Moon, CD Zheng N Huws S Zhao S Wang J (2022). Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. *Microbiome* 10:76. [https://doi.org/ 10.1186/s40168-022-01272-5](https://doi.org/10.1186/s40168-022-01272-5)

Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J., Crosby, L., and Neufeld, J. D. (2018). Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* 3:e00055–e00018. [https://doi.org/ 10.1128/mSystems.00055-18](https://doi.org/10.1128/mSystems.00055-18)

Loeschcke A Thies S (2015) *Pseudomonas putida*-a versatile host for the production of natural products. *Appl MicrobiolBiotechnol* 99(15):6197-214. [https://doi.org/ 10.1007/s00253-015-6745-4](https://doi.org/10.1007/s00253-015-6745-4)

Ludwig W Klenk HP (2005) *Overview: a phylogenetic backbone and taxonomic framework for procaryotic systematics*. In: Brenner DJ Krieg NR Staley JT Garrity GM (eds) *Bergey's Manual of Systematic Bacteriology*. New York, NY: Springer.

MacGowan A and Macnaughton E (2017) Antibiotic resistance. *Medicine*, 45(10):622-628. [https://doi.org/ 10.1016/j.mpmed.2017.07.006](https://doi.org/10.1016/j.mpmed.2017.07.006)

Mak S Xu Y Nodwell JR (2014) The expression of antibiotic resistance genes in antibiotic-producing bacteria. *Mol Microbiol* 93(3):391-402. [https://doi.org/ 10.1111/mmi.12689](https://doi.org/10.1111/mmi.12689)

Mardis ER (2017) DNA sequencing technologies: 2006-2016. *Nat Protoc* 12(2):213-218. [https://doi.org/ 10.1038/nprot.2016.182](https://doi.org/10.1038/nprot.2016.182)

Martins J Leikoski N Wahlsten M Azevedo J Antunes J et al. (2018) Sphaerocyclamide, a prenylated cyanobactin from the cyanobacterium *Sphaerospermopsis* sp. LEGE 00249. *Sci Rep* 8(1):14537. [https://doi.org/ 10.1038/s41598-018-32618-5](https://doi.org/10.1038/s41598-018-32618-5)

Martins J Vasconcelos V (2015) Cyanobactins from Cyanobacteria: Current Genetic and Chemical State of Knowledge. *Mar Drugs* 13(11):6910-46. [https://doi.org/ 10.3390/md13116910](https://doi.org/10.3390/md13116910)

Meier-Kolthoff JP Carbasse JS Peinado-Olarte RL Göker M (2022) TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 50(D1):D801-D807. <https://doi.org/10.1093/nar/gkab902>

Mikheenko A Prjibelski A Saveliev V Antipov D Gurevich A (2018) Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34(13):i142-i150. <https://doi.org/10.1093/bioinformatics/bty266>

Meziti A Rodriguez-R LM Hatt JK Peña-Gonzalez A Levy K Konstantinidis KT (2021) The Reliability of Metagenome-Assembled Genomes (MAGs) in Representing Natural Populations: Insights from Comparing MAGs against Isolate Genomes Derived from the Same Fecal Sample. *Appl Environ Microbiol* 87(6):e02593-20. <https://doi.org/10.1128/AEM.02593-20>

Meziti A Tsementzi D Rodriguez-R LM Hatt JK Karayanni H Kormas KA Konstantinidis KT (2019) Quantifying the changes in genetic diversity within sequence-discrete bacterial populations across a spatial and temporal riverine gradient. *The ISME Journal* 13(3):767-779. <https://doi.org/10.1038/s41396-018-0307-6>

Morton JT Freed SD Lee SW Friedberg I (2015) A large scale prediction of bacteriocin gene blocks suggests a wide functional spectrum for bacteriocins. *BMC Bioinformatics* 16:381. <https://doi.org/10.1186/s12859-015-0792-9>.

Mozejko-Ciesielska J Szacherska K Marciniak P (2019) *Pseudomonas* Species as Producers of Eco-friendly Polyhydroxyalkanoates. *J Polym Environ* 27:1151–1166 <https://doi.org/10.1007/s10924-019-01422-1>

Munita JM Arias CA (2016) Mechanisms of Antibiotic Resistance. *MicrobiolSpectr* 4(2):10. <https://doi.org/10.1128/microbiolspec.VMBF-0016-2015>

Nakagaito Y Yokota A (1992) Three new species of the genus *Streptomyces*: *Streptomyces cochleatus* sp. nov., *Streptomyces paracochleatus* sp. nov., and *Streptomyces azaticus* sp. nov. *J Gen App Microbiol* 38(2):105-120. <https://doi.org/10.2323/jgam.38.105>

Nascimento AMA (2011) *Use of the rRNA operon and genomic repetitive sequences for the identification of bacteria*. In: Brui JNJ Willey J (eds) *Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches* 800p.

Navarro-Muñoz JC Selem-Mojica N Mullaney MW Kautsar SA Tryon *et al.* (2020) A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol* 16(1):60-68. doi: 10.1038/s41589-019-0400-9

Nicolaou KC Rigol S (2018) A brief history of antibiotics and select advances in their synthesis. *J Antibiot (Tokyo)* 71(2):153-184. <https://doi.org/10.1038/ja.2017.62>

Nichols D Lewis K Orjala J Mo S Ortenberg R O'Connor P *et al.* (2008) Short peptide induces an "uncultivable" microorganism to grow in vitro. *Appl Environ Microbiol* 74(15):4889-97. <https://doi.org/10.1128/AEM.00393-08>

Nivina A Yuet KP Hsu J Khosla C (2019) Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chem Rev* 119(24):12524-12547. <https://doi.org/10.1021/acs.chemrev.9b00525>

Njoroge J Sperandio V (2009) Jamming bacterial communication: new approaches for the treatment of infectious diseases. *EMBO Mol Med* 1(4):201-10. <https://doi.org/10.1002/emmm.200900032>

Nouioui I Carro L García-López M Meier- Kolthoff JP Woyke T et al. (2018) Genome-based taxonomic classification of the phylum Actinobacteria. *Front Microbiol* 9:2007. <https://doi.org/10.3389/fmicb.2018.02007>

Nowrotek M Jałowicki Ł Harnisz M Płaza GA (2019) Culturomics and metagenomics: in understanding of environmental resistome. *Front Environ Sci En.* 13:40. <https://doi.org/10.1007/s11783-019-1121-8>

Ojewumi ME Okeniyi JO Ikotun JO Okeniyi ET Ejemen VA Popoola API (2018) Bioremediation: Data on *Pseudomonas aeruginosa* effects on the bioremediation of crude oil polluted soil. *Data Brief* 3;19:101-113. <https://doi.org/10.1016/j.dib.2018.04.102>

Omura S Takahashi Y Iwai Y Tanaka H (1982) *Kitasatosporia*, a new genus of the order Actinomycetales. *J Antibiot* 35:1013-1019. <https://doi.org/10.7164/antibiotics.35.1013>

Ongpipattanakul C Desormeaux EK DiCaprio A van der Donk WA Mitchell DA Nair SK (2022) Mechanism of Action of Ribosomally Synthesized and Post-Translationally Modified Peptides. *ChemRev* 122(18):14722-14814. <https://doi.org/10.1021/acs.chemrev.2c00210>

Ortega DB Costa RA Pires AS Araújo TF Araújo JF et al. (2018) Draft Genome Sequence of the Antimicrobial-Producing Strain *Paenibacillus elgii* AC13. *Genome Announc* 6(26):e00573-18. <https://doi.org/10.1128/genomeA.00573-18>

Overmann J (2013) Principles of enrichment, isolation, cultivation, and preservation of bacteria. In E Rosenberg, DeLong EF, Stackebrandt E, Lory S, Thompson F (ed) *The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*, 4th edn. New York: Springer, pp 149-207

Overmann J Abt B Sikorski J (2017) Present and Future of Culturing Bacteria. *Annu Rev Microbiol* 71:711-730. <https://doi.org/10.1146/annurev-micro-090816-093449>.

Paiva AD Breukink E Mantovani HC (2011) Role of lipid II and membrane thickness in the mechanism of action of the lantibiotic bovicin HC5. *Antimicrob Agents Chemother* 55(11):5284-93. <https://doi.org/10.1128/AAC.00638-11>

Pande S Kost C (2017) Bacterial Unculturability and the Formation of Intercellular Metabolic Networks. *Trends Microbiol* 25(5):349-361. <https://doi.org/10.1016/j.tim.2017.02.015>

Papenfort K Bassler BL (2016) Quorum sensing signal-response systems in Gram-negative bacteria. *Nat Rev Microbiol* 14(9):576-88. <https://doi.org/10.1038/nrmicro.2016.89>

Park C Lee YS Park SY Park W (2018) *Methylobacterium currus* sp. nov., isolated from a car air conditioning system. *Int J Syst Evol Microbiol* 68(11):3621-3626. <https://doi.org/10.1099/ijsem.0.003045>

Parks DH Imelfort M Skennerton CT Hugenholtz P Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25(7):1043-55. <https://doi.org/10.1101/gr.186072.114>

Pascual J García-López M Carmona C Sousa Tda S de Pedro N et al. (2014) *Pseudomonas soli* sp. nov., a novel producer of xantholysin congeners. *Syst Appl Microbiol* 37:412–416. <https://doi.org/10.1016/j.syapm.2014.07.003>

Peixoto J Silva LP Krüger RH (2017) Brazilian Cerrado soil reveals an untapped microbial potential for unpretreated polyethylene biodegradation. *J Hazard Mater* 324(Pt B):634-644. doi: 10.1016/j.jhazmat.2016.11.037

Pereira R, Oliveira J, Sousa M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *J Clin Med*. 2020 Jan 3;9(1):132. <https://doi.org/10.3390/jcm9010132>

Petit P Lucas EMF Abreu LM Pfenning LH Takahashi JA (2009) Novel antimicrobial secondary metabolites from a *Penicillium* sp. isolated from Brazilian cerrado soil. *Electron. J Biotechnol* 12(4):8-9. <http://dx.doi.org/10.4067/S0717-34582009000400008>

Perez RH Zendo T Sonomoto K (2014) Novel bacteriocins from lactic acid bacteria (LAB): various structures and applications. *Microb Cell Fact* 13 Suppl 1(Suppl 1):S3. <https://doi.org/10.1186/1475-2859-13-S1-S3>

Photolo MM Mavumengwana V Sitole L Tlou MG (2020) Antimicrobial and Antioxidant Properties of a Bacterial Endophyte, *Methylobacterium radiotolerans* MAMP 4754, Isolated from *Combretum erythrophyllum* Seeds. *Int J Microbiol*. 18:2020-9483670. <https://doi.org/10.1155/2020/9483670>

Pop M (2009) Genome assembly reborn: recent computational challenges. *Brief Bioinform* 10(4):354-66. <https://doi.org/10.1093/bib/bbp026>

Qian CD Liu TZ Zhou SL Ding R Zhao WP Li O Wu XC (2012) Identification and functional analysis of gene cluster involvement in biosynthesis of the cyclic lipopeptide antibiotic pelgipeptin produced by *Paenibacillus elgii*. *BMC Microbiol* 8:12-197. <https://doi.org/10.1186/1471-2180-12-197>

Qiu Y Wang S Chen Z Guo Y Song Y (2016) An active type I-E CRISPR-Cas system identified in *Streptomyces avermitilis*. *PLoS ONE* 11:e0149533. <https://doi.org/10.1371/journal.pone.0149533>

Raaijmakers JM De Bruijn I Nybroe O Ongena M (2010) Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol Rev* 34(6):1037-62. <https://doi.org/10.1111/j.1574-6976.2010.00221.x>

Ramaprasad EVV Mahidhara G Sasikala C Ramana CV (2018) *Rhodococcus electrodiphilus* sp. nov., a marine electro active actinobacterium isolated from coral reef. *Int J Syst Evol Microbiol* 68(8):2644-2649. <https://doi.org/10.1099/ijsem.0.002895>

Raney F Oren A (2011) Taxonomy of Prokaryotes. ISBN

Rappé MS Giovannoni SJ (2003) The uncultured microbial majority. *Annu Rev Microbiol* ;57:369-94. [https://doi.org/ 10.1146/annurev.micro.57.030502.090759](https://doi.org/10.1146/annurev.micro.57.030502.090759)

Ravin DY Watson ZL Meteleev M Ward FR Osterman IA et al. (2019) Structure of ribosome-bound azole-modified peptide phazolicin rationalizes its species-specific mode of bacterial translation inhibition. *Nat Commun* 8:10(1)-4563. <https://doi.org/10.1038/s41467-019-12589-5>

Reddy CA Beveridge TJ Breznak JA Marzluf GA Schmidt TM Snyder LR (2007) *Methods for General and Molecular Microbiology*. American Society of Microbiology. <https://doi.org/10.1128/9781555817497>

Ren L Men L Zhang Z Guan F Tian J Wang B Wang J Zhang Y Zhang W (2019) Biodegradation of Polyethylene by *Enterobacter* sp. D1 from the Gut of Wax Moth *Galleria mellonella*. *Int J Environ Res Public Health* 16(11):1941. <https://doi.org/10.3390/ijerph16111941>

Reuter JA Spacek DV Snyder MP (2015) High-throughput sequencing technologies. *Mol Cell* 58(4):586-97. [https://doi.org/ 10.1016/j.molcel.2015.05.004](https://doi.org/10.1016/j.molcel.2015.05.004)

Ribeiro JF Walter BMT (1998) Fitofisionomias do bioma Cerrado. In: Sano SM Almeida SP (eds) *Cerrado: ambiente e flora*. 1st edn. Embrapa CPAC.

Richter M Rosselló-Móra R Oliver Glöckner F Peplies (2016) JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32(6):929-31. <https://doi.org/10.1093/bioinformatics/btv681>

Robinson SL Christenson JK Wackett LP (2019) Biosynthesis and chemical diversity of β -lactone natural products. *Nat Prod Rep* 36(3):458-475. <https://doi.org/10.1039/c8np00052b>

Rodríguez H Rico S Díaz M Santamaría RI (2013) Two-component systems in *Streptomyces*: key regulators of antibiotic complex pathways. *Microb Cell Fact* 12:127. [https://doi.org/ 10.1186/1475-2859-12-127](https://doi.org/10.1186/1475-2859-12-127)

Sagar A Sayyed RZ Ramteke PW Sharma S Marraiki N Elgorban AM Syed A (2020) ACC deaminase and antioxidant enzymes producing halophilic *Enterobacter* sp. PR14 promotes the growth of rice and millets under salinity stress. *Physiol Mol Biol Plants* 26(9):1847-1854. [https://doi.org/ 10.1007/s12298-020-00852-9](https://doi.org/10.1007/s12298-020-00852-9)

Saha M Sarkar S Sarkar B Sharma BK Bhattacharjee S Tribedi P (2016) Microbial siderophores and their potential applications: a review. *Environ Sci Pollut Res Int* 23(5):3984-99. <https://doi.org/10.1007/s11356-015-4294-0>

Sah S Krishnani S Singh R (2021) *Pseudomonas* mediated nutritional and growth promotional activities for sustainable food security. *Curr Res Microb Sci* 2:100084. <https://doi.org/10.1016/j.crmicr.2021.100084>

Salvachúa D Rydzak T Auwae R De Capite A Black BA et al. (2020) Metabolic engineering of *Pseudomonas putida* for increased polyhydroxyalkanoate production from lignin. *Microb Biotechnol* 13(1):290-298. <https://doi.org/10.1111/1751-7915.13481>

Sánchez-Yañez JM (2022) *Xanthobacter autotrophicus* an Endophytic Beneficial Bacterium for Wheat and Other Plants: A Short Review. *Current Trends in Wheat Research* pp 73. <https://doi.org/10.5772/intechopen.102066>.

Sasser M (2001) Technical Note # 101 Identification of Bacteria by Gas Chromatography of Cellular Fatty Acids.

Scheldeman P Goossens K Rodriguez-Diaz M Pila Goris J Herman L et al. (2004) *Paenibacillus lactis* sp. nov., isolated from raw and heat-treated milk. *Int J Syst Evol Microbiol* 54(Pt 3):885-891. <https://doi.org/10.1099/ijs.0.02822-0>

Schink B (2002) Synergistic interactions in the microbial world. *Antonie Van Leeuwenhoek* 81(1-4):257-61. <https://doi.org/10.1023/a:1020579004534>

Schwengers O Jelonek L Dieckmann MA Beyvers S Blom J Goesmann A (2021) Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* 7(11):000685. <https://doi.org/10.1099/mgen.0.000685>

Sedeek KEM Mahas A Mahfouz M (2019) Plant Genome Engineering for Targeted Improvement of Crop Traits. *Front Plant Sci* 10:114. <https://doi.org/10.3389/fpls.2019.00114>

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14):2068-9. <https://doi.org/10.1093/bioinformatics/btu153>

Selvin J Kennedy J Lejon DP Kiran GS Dobson AD (2012) Isolation identification and biochemical characterization of a novel halo-tolerant lipase from the metagenome of the marine sponge *Haliclona simulans*. *Microb Cell Fact* 11:11-72. <https://doi.org/10.1186/1475-2859-11-72>

Singh RP Mishra S Jha P Raghuvanshi S Jha PN (2018) Effect of inoculation of zinc-resistant bacterium *Enterobacter ludwigii* CDP-14 on growth, biochemical parameters and zinc uptake in wheat (*Triticum aestivum* L.) plant. *Ecological Engineering* 116:163-173. <https://doi.org/10.1016/j.ecoleng.2017.12.033>.

Soares EV (2022) Perspective on the biotechnological production of bacterial siderophores and their use. *Appl Microbiol Biotechnol* 106(11):3985-4004. <https://doi.org/10.1007/s00253-022-11995-y>

Soltani S Hammami R Cotter PD Rebuffat , Said LB et al. (2021) Bacteriocins as a new generation of antimicrobials: toxicity aspects and regulations. *FEMS Microbiol Rev* 45(1):fuaa039. <https://doi.org/10.1093/femsre/fuaa039>

Souza W (2007) Técnicas de microscopia eletrônica aplicadas às ciências biológicas. Sociedade Brasileira de Microscopia. Rio de Janeiro

Srinivas TN Nageswara Rao SS Vishnu VardhanReddy P Pratibha MS Sailaja B Kavya B et al. Bacterial diversity and bioprospecting for cold-active lipases, amylases and proteases, from culturable bacteria of kongsfjorden and Ny-alesund, Svalbard, Arctic. *CurrMicrobiol* 59(5):537-47. [https://doi.org/ 10.1007/s00284-009-9473-0](https://doi.org/10.1007/s00284-009-9473-0).

Srivastava N Nandi I IbeyaimaA Gupta S Sarethy IP (2019) Microbial diversity of a Himalayan forest and characterization of rare actinomycetes for antimicrobial compounds. *3 Biotech*. 9(1):27. [https://doi.org/ 10.1007/s13205-018-1556-9](https://doi.org/10.1007/s13205-018-1556-9)

Stackebrandt E (1988) Phylogenetic relationships vs. phenotypic diversity: how to achieve a phylogenetic classification system of the eubacteria. *Can J Microbiol* 34(4):552-6. [https://doi.org/ 10.1139/m88-094](https://doi.org/10.1139/m88-094)

Staley JT Konopka A (1985) Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol* 39:321-46. [https://doi.org/ 10.1146/annurev.mi.39.100185.001541](https://doi.org/10.1146/annurev.mi.39.100185.001541)

Stewart EJ (2012) Growing unculturable bacteria. *J Bacteriol* 194(16):4151-60. [https://doi.org/ 10.1128/JB.00345-12](https://doi.org/10.1128/JB.00345-12)

Sy A Giraud E Jourand P Garcia N Willems A de Lajudie P Prin Y Neyra M Gillis M Boivin-Masson C Dreyfus B (2001) Methylophilic Methylobacterium bacteria nodulate and fix nitrogen in symbiosis with legumes. *J Bacteriol* 183(1):214-20. [https://doi.org/ 10.1128/JB.183.1.214-220.2001](https://doi.org/10.1128/JB.183.1.214-220.2001)

Takahash Y (2017) Genus *Kitasatospora*, taxonomic features and diversity of secondary metabolites. *J Antibiot* 70:506–513. <https://doi.org/10.1038/ja.2017.8>

Takahashi A Castro MCM Souza GG Lucas EMF Bracarense AAP Abreu LM Marriel IE Oliveira MSFloreano MB Oliveira TS (2008) Isolation and screening of fungal species isolated from Brazilian cerrado soil for antibacterial activity against *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhimurium*, *Streptococcus pyogenes* and *Listeria monocytogenes*. *J MycolMédicale* 18(4):198-204. <https://doi.org/10.1016/j.mycmed.2008.08.001>.

Takano E (2006) Gamma-butyrolactones: *Streptomyces* signalling molecules regulating antibiotic production and differentiation. *CurrOpinMicrobiol* 9(3):287-94. <https://doi.org/10.1016/j.mib.2006.04.003>

Tatusov RL Koonin E V Lipman DJ (1997) A Genomic Perspective on Protein Families. *Science* 278:631– 7. <https://doi.org/10.1126/science.278.5338.631>

Tejesvi MV Andersen B Antcheva N Brinch K Koskimäki J Kristensen H Tossi A Pirttilä A (2016) MB1533 is a defensin-like antimicrobial peptide from the intracellular meristem endophyte of Scots pine *Methylobacterium extorquens* DSM13060. *J Microbial Biochem Technology* 8(1):1–5. doi: 10.4172/1948-5948.1000252

Teng Y Zhao W Qian C Li O Zhu L Wu X (2012) Gene cluster analysis for the biosynthesis of elgicins, novel lantibiotics produced by *Paenibacillus elgii* B69. *BMC Microbiol* 12:45. [https://doi.org/ 10.1186/1471-2180-12-45](https://doi.org/10.1186/1471-2180-12-45).

Tillotson GS Zinner SH (2017) Burden of antimicrobial resistance in an era of decreasing susceptibility. *Expert Rev Anti Infect Ther* 15(7):663-676. <https://doi.org/10.1080/14787210.2017.1337508>.

Tindall B Sikorski J Smibert R Krieg N (2007) *Phenotypic characterization and the principles of comparative systematics*. In: Reddy C Berridge T Breznak J Marzluf G Schmidr T Snyder L (eds) *Methods for General and Molecular Microbiology*. Washington, DC: ASM Press, 330–393.

Toutain PL Ferran AA Bousquet-Melou A Pelligand L Lees P (2016) Veterinary Medicine Needs New Green Antimicrobial Drugs. *Front Microbiol* 3:7-1196. <https://doi.org/10.3389/fmicb.2016.01196>

Turan NB Chormey DS Büyükpınar C Engin GO Bakirdere S (2017) Quorum sensing: Little talks for an effective bacterial coordination, *TrAC Trends in Analytical Chemistry* 91:1-11. <https://doi.org/10.1016/j.trac.2017.03.007>

Uddin TM Chakraborty AJ Khusro A Zidan BRM Mitra S *et al.* (2021) Antibiotic resistance in microbes: History, mechanisms, therapeutic strategies and future prospects. *J Infect Public Health* 14(12):1750-1766. <https://doi.org/10.1016/j.jiph.2021.10.020>

Vajs J Proud C Brozovic A Gazvoda M Lloyd A Roper DI *et al.* (2017). Diaryltriazenes as antibacterial agents against methicillin resistant *Staphylococcus aureus* (MRSA) and *Mycobacterium smegmatis*. *Eur J Med Chem* 127:223-234. <https://doi.org/10.1016/j.ejmech.2016.12.060>.

van Bergeijk DA Terlouw BR Medema MH van Wezel GP (2020) Ecology and genomics of Actinobacteria: new concepts for natural product discovery. *Nat Rev Microbiol* 18(10):546-558. <https://doi.org/10.1038/s41579-020-0379-y>

Vandamme P Pot B Gillis M de Vos P Kersters K Swings J (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* 60(2):407-38. <https://doi.org/10.1128/mr.60.2.407-438.1996>

van der Meer JR Rollema HS Siezen RJ Beerthuyzen MM Kuipers OP de Vos WM (1994) Influence of amino acid substitutions in the nisin leader peptide on biosynthesis and secretion of nisin by *Lactococcus lactis*. *J Biol Chem* 269(5):3555-62.

van Dorst J M Hince G Snape I Ferrari B C (2016). Novel culturing techniques select for heterotrophs and hydrocarbon degraders in a subantarctic soil. *Sci Rep* 6:36724. <https://doi.org/10.1038/srep36724>

van Heel AJ de Jong A Song C Viel JH Kok J Kuipers OP (2018) BAGEL4: a user-friendly web server to thoroughly mine *RIPPs* and bacteriocins. *Nucleic Acids Res* 46(W1):W278-W281. <https://doi.org/10.1093/nar/gky383>.

Verbeke F De Craemer S Debonne N Janssens Y Wynendaele E *et al.* Peptides as Quorum Sensing Molecules: Measurement Techniques and Obtained Levels In Vitro and In Vivo. *Front Neurosci* 11:183. <https://doi.org/10.3389/fnins.2017.00183>

Veyisoglu A Camas M Tatar D Guven K Sazak A Sahin N (2013) *Methylobacterium tarhaniae* sp. nov., isolated from arid soil. *Int J Syst Evol Microbiol* 63(Pt 8):2823-2828. <https://doi.org/10.1099/ijs.0.049551-0>

Waghu FH Barai RS Gurung P Idicula-Thomas S (2016) CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res* 44:D1094–D1097. <https://doi.org/10.1093/nar/gkv1051>

Wang G Li X Wang Z (2016) APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res* 44(D1):D1087-93. <https://doi.org/10.1093/nar/gkv1278>

Wang H Liu N Xi L Rong X Ruan J Huang Y (2011) Genetic screening strategy for rapid access to polyether ionophore producers and products in actinomycetes. *Appl Environ Microbiol* 77(10):3433-42. <https://doi.org/10.1128/AEM.02915-10>

Wang J Gao C Chen X Liu L (2021) Expanding the lysine industry: biotechnological production of l-lysine and its derivatives. *Adv Appl Microbiol* 115:1-33. <https://doi.org/10.1016/bs.aambs.2021.02.001>

Wang LT Lee FL Tai CJ Kasai H (2007) Comparison of *gyrB* gene sequences, 16S *rRNA* gene sequences and DNA-DNA hybridization in the *Bacillus subtilis* group. *Int J Syst Evol Microbiol* 57(Pt 8):1846-1850. <https://doi.org/10.1099/ijs.0.64685-0>

Wang SS Liu JM Sun J Sun YF Liu JN Jia N Fan B Dai XF (2019) Diversity of culture-independent bacteria and antimicrobial activity of culturable endophytic bacteria isolated from different *Dendrobium* stems. *Sci Rep* 9:10389. <https://doi.org/10.1038/s41598-019-46863-9>

Wang X Zhao J Liu C Wang J Shen Y et al. (2013). *Nonomuraea solani* sp. nov., an actinomycete isolated from eggplant root (*Solanum melongena* L.). *Int J Syst Evol Microbiol* 63:2418–2423. <https://doi.org/10.1099/ijs.0.045617-0>

Wang Y Hoffmann JP Chou CW Höner Zu Bentrup K et al. (2020) *Burkholderia thailandensis* outer membrane vesicles exert antimicrobial activity against drug-resistant and competitor microbial species. *J Microbiol* 58(7):550-562. <https://doi.org/10.1007/s12275-020-0028-1>

Wang ZW Liu XL (2008) Medium optimization for antifungal active substances production from a newly isolated *Paenibacillus* sp. using response surface methodology. *Bioresour Technol* 99(17):8245-51. <https://doi.org/10.1016/j.biortech.2008.03.039>

Waksman SA Henrici AT (1943) The Nomenclature and Classification of the Actinomycetes. *J Bacteriol* 46(4):337-41. doi: 10.1128/jb.46.4.337-341.1943

Wayne LG Brenner DJ Colwell RR Grimont PAD Kandler O Krichevsky MI et al. (1987) *Report of the ad hoc committee on reconciliation of approaches to bacterial systematics*. *Int J Syst Bacteriol* 37:463–464. [https://doi.org/10.1016/s0176-6724\(88\)80120-2](https://doi.org/10.1016/s0176-6724(88)80120-2).

Webster G Mullins AJ Cunningham-Oakes E Renganathan A Aswathanarayan JB et al. (2020) Culturable diversity of bacterial endophytes associated with medicinal plants of the

Western Ghats, India. FEMS Microbiol Ecol 96(9):f1aa147. <https://doi.org/10.1093/femsec/f1aa147>

Weimann A Mooren K Frank J Pope PB Bremges A McHardy AC (2016) From genomes to phenotypes: TraitAr, the microbial trait analyzer. mSystem. <https://doi.org/10.1101/043315>

Weisburg WG Barns SM Pelletier DA Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173(2):697-703. <https://doi.org/10.1128/jb.173.2.697-703.1991>

Wellington EMH Stackebrandt E Sanders D Wolstrup J Jorgensen NO G (1992) Taxonomic Status of *Kitasatosporia*, and Proposed Unification with *Streptomyces* on the Basis of Phenotypic and 16S *rRNA* Analysis and Emendation of *Streptomyces* Waksman and Henrici 1943, 339AL. Int J Syst Bacteriol 42:156–160. <https://doi.org/10.1099/00207713-42-1-156>

Wheeler DL Barrett T Benson DA Bryant SH Canese K Chetvernin V et al. (2007) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 35(Database issue):D5-12. <https://doi.org/10.1093/nar/gkl1031>

Whiteley AS Jenkins S Waite I Kresoje N Payne H Mullan B Allcock R O'Donnell A (2012) Microbial 16S *rRNA* Ion Tag and community metagenome sequencing using the Ion Torrent (PGM) Platform J Microbiol Methods 91(1):80-8. <https://doi.org/10.1016/j.mimet.2012.07.008>

Wick RR Judd LM Gorrie CL Holt KE (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 8:13(6)-e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>

Willey JM van der Donk WA (2007) Lantibiotics: peptides of diverse structure and function. Annu Rev Microbiol 61:477-501. <https://doi.org/10.1146/annurev.micro.61.080706.093501>

Wilson L Iqbal KM Simmons-Ehrhardt T Bertino MF Shah MR et al. (2019) Customizable 3D printed diffusion chambers for studies of bacterial pathogen phenotypes in complex environments. J Microbiol Methods 162:8-15. <https://doi.org/10.1016/j.mimet.2019.05.002>

Wimmer F Beisel CL (2020) CRISPR-Cas Systems and the Paradox of Self-Targeting Spacers. Front Microbiol 22:10:3078. <https://doi.org/10.3389/fmicb.2019.03078>

Winand R Bogaerts B Hoffman S Lefevre L Delvoeye M et al. (2019) Targeting the 16s *rRNA* gene for bacterial identification in complex mixed samples: Comparative evaluation of second (illumina) and third (oxford nanopore technologies) generation sequencing technologies. Int J Mol Sci 21(1):298. <https://doi.org/10.3390/ijms21010298>

Witt D Stackebrandt E (1990) Unification of the Genera *Streptoverticillum* and *Streptomyces*, and Amendation of *Streptomyces* Waksman and Henrici 1943. Syst App Microbiol 13(4):361-371. [https://doi.org/10.1016/S0723-2020\(11\)80234-1](https://doi.org/10.1016/S0723-2020(11)80234-1)

Woese CR Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 74(11):5088-90. <https://doi.org/10.1073/pnas.74.11.5088>

Wu C van Wezel GP Hae Choi Y (2015) Identification of novel endophenaside antibiotics produced by *Kitasatospora* sp. MBT66. *J Antibiot (Tokyo)* 68:445–452. <https://doi.org/10.1038/ja.2015.14>

Wu X Fang H Qian C Wen Y Shen X Li O Gao H (2011) *Paenibacillustianmuensis* sp. nov., isolated from soil. *Int J Syst Evol Microbiol* 61(Pt 5):1133–1137. <https://doi.org/10.1099/ijs.0.024109-0>

Yamada Y Kuzuyama T Komatsu M Shin-Ya K Omura S et al. (2015) Terpene synthases are widely distributed in bacteria. *Proc Natl Acad Sci USA* 112(3):857–62. <https://doi.org/10.1073/pnas.1422108112>

Yin T Cook D Lawrence M (2012) Ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biol* 13(8):R77. <https://doi.org/10.1186/gb-2012-13-8-r77>

Yoon SH Ha SM Kwon S Lim J Kim Y Seo H Chun J (2017) Introducing EzBioCloud: a taxonomically united database of 16S *rRNA* gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 67(5):1613–1617. <https://doi.org/10.1099/ijsem.0.001755>

Yoshida S Hiradate S Koitabashi M Kamo T Tsushima S (2017) Phyllosphere *Methylobacterium* bacteria contain UVA-absorbing compounds. *J Photochem Photobiol B* 167:168–175. <https://doi.org/10.1016/j.jphotobiol.2016.12.019>

Yun BR Malik A Kim SB (2020) Genome based characterization of *Kitasatospora* sp. MMS16-BH015, a multiple heavy metal resistant soil actinobacterium with high antimicrobial potential. *Gene* 5:733–144379. <https://doi.org/10.1016/j.gene.2020.144379>

Zaburannyi N Rabyk M Ostash B Fedorenko V Luzhetskyy A (2014) Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics* 5:15–97. <https://doi.org/10.1186/1471-2164-15-97>

Zhang Z Wang Y Ruan J (1997) A proposal to revive the genus *Kitasatospora* (Omura, Takahashi, Iwai, and Tanaka 1982). *Int J Syst Bacteriol* 47(4):1048–54. <https://doi.org/10.1099/00207713-47-4-1048>

Zhu D Xie C Huang Y Sun J Zhang W (2014) Description of *Comamonasserinivorans* sp. nov., isolated from wheat straw compost. *Int J Syst Evol Microbiol*. 64(Pt 12):4141–4146. <https://doi.org/10.1099/ijs.0.066688-0>

Zhu F Qin C Tao L Liu X Shi Z Ma X Jia J Tan Y Cui C Lin J Tan C Jiang Y Chen Y (2011) Clustered patterns of species origins of nature-derived drugs and clues for future bioprospecting. *Proc Natl Acad Sci USA* 108(31):12943–8. <https://doi.org/10.1073/pnas.1107336108>

Zin NM Ismail A Mark DR Westrop G Schniete JK Herron PR (2021) Adaptation to Endophytic Lifestyle Through Genome Reduction by *Kitasatospora* sp. SUK42. *Front Bioeng Biotechnol* 12:9-740722. <https://doi.org/10.3389/fbioe.2021.740722>

1 **6 ANEXOS**

2 6.1 *Draft* do artigo do Capítulo 2

3 **Use of *Paenibacilluselgii* supernatant in prospecting and selection of bacterial strains** 4 **with potential for biotechnological application**

5 Cunha-Ferreira, I.C.^{1*}, Vizzotto, C.S.^{2*}, Frederico, T.D.¹, Quirino, B.F.^{3,4,5}, Tótola, M.R.⁶ and
6 Krüger, R.H.¹

7
8 ¹Laboratory of Enzymology, Department of Cellular Biology, University of Brasília (UNB),
9 Brasília, Brazil.

10 ²Laboratory of Environmental Sanitation, Department of Civil and Environmental
11 Engineering, University of Brasília (UNB), Brasília, Brazil.

12 ³Genetics and Biotechnology Laboratory, Embrapa-Agroenergia, Brasília, Brazil

13 ⁴Genomic Sciences and Biotechnology Graduate Program, Catholic University of Brasília,
14 Brasília, Brazil

15 ⁵Microbial Biology Graduate Program, University of Brasília, Brasília, Brazil

16 ⁶Laboratory of Biotechnology and Biodiversity for the Environment, Department of
17 Microbiology, Federal University of Viçosa (UFV), Viçosa, Brazil.

18
19 *Corresponding authors: kruger@unb.br

21 **Abstract**

22 One of the main biotechnological demands of the industry for bioactive compounds refers to
23 the production of antimicrobial compounds due to the need for alternatives to products
24 already present on the market which a variety of bacteria have already become resistant.
25 Classical cultivation techniques for prospecting microorganisms make it difficult to discover
26 new antimicrobial producers. Thus, this study describes a protocol for the selection of
27 bacterial strains with biotechnological potential from *Paenibacilluselgii* supernatant. Culture
28 media were enriched with the supernatant which is rich in signal molecules and antimicrobial
29 peptides. Soil samples were inoculated in these media in order to select only bacteria resistant
30 to that supernatant and, therefore, with potential for the production of similar compounds.
31 Bacterial strains obtained through this protocol had their antibacterial capacity demonstrated.
32 Furthermore the functional annotation of the genome of one of the strains obtained through the
33 protocol, gene clusters of biotechnological interest were detected for the strain isolated in this

34 study. This study attested to the efficiency of using the supernatant in the selection of growth
35 of microorganisms with potential for biotechnological application.

36 **Keywords:** prospecting, supernatant, antibacterial activity

37

38 **Introduction**

39 After the emergence of molecular techniques in the 1970s, such as PCR, next
40 generation sequencing and more particularly metagenomics and the enormous amount of
41 information provided by them, the cultivation of bacteria may have been neglected (Gutleben
42 et al., 2018). However, culture media remain a fundamental tool for microbiologists since
43 both methods, culture-dependent or not, can complement each other (Lagier et al. 2015).

44 The pace of bacterial DNA sequencing surpasses the isolation of new species and can
45 reveal interesting physiological capacities in microorganisms and thus direct cultivation
46 efforts (Wilson et al. 2019). Cultivation, on the other hand, provides data through its well-
47 annotated genomes for predictions used in genomic and metagenomic analyzes (Choi et al.
48 2017; Liu et al. 2022).

49 Technology associated cultivation in combination with information provided by
50 culture independent experimental approaches offers the chance to access unexplored bacterial
51 diversity that offers fundamental new insights into biological principles (Overmann et al.
52 2017). Many of the underexplored bacterial strains probably present new biosynthetic
53 pathways and unknown biochemical characteristics, suggesting the high biotechnological
54 potential to be discovered from the metabolites and enzymes of these microorganisms and
55 potential applications for biotechnology (Gutleben et al. 2020).

56 One of the main biotechnological demands of the industry for bioactive compounds
57 refers to the production of antimicrobial compounds, due to the need for alternatives to
58 products already present on the market which a variety of bacteria have already become
59 resistant (Laws et al. 2019). Many of the antimicrobials described have their origin in bacteria
60 that produce them by competition in the environment or even communication with other
61 microorganisms (Durand et al. 2018). However, one of the main obstacles in studying
62 prokaryotes and discovering new compounds produced by them is the inability to cultivate
63 most of the environmental diversity under laboratory conditions (Overmann et al. 2017).

64 Certain groups of bacteria require specific growth conditions, so there is a need for
65 refinement of existing strategies for growing microorganisms, as well as the development of
66 new techniques and methods that are more effective in obtaining previously uncultivated
67 microorganisms and/or of biotechnological interest. Improvements in culture techniques may

68 include: use of restrictive culture media and use of antimicrobials in growth selection, for
69 example (Overmann 2013). Other possibilities for improving cultivation techniques include
70 the use of co-cultures, diffusion chambers, membrane systems and bioreactors (Boilattabi et
71 al. 2021; Pascual et al. 2017; van Dorst et al. 2016; Chaudhary et al. 2017). Thus, specific
72 enrichments can be used to build substantially larger collections of microorganisms of interest
73 (Bachmann et al. 2014).

74 In their natural environment, bacteria live in communities and have sophisticated
75 communication networks based on metabolic changes that influence basal behavior processes
76 such as reproduction, sporulation, biofilm formation and antimicrobial expression (Xiao et al.
77 2022). A particular type of chemical interaction is the exchange of signaling compounds.
78 Signaling molecules may be essential for the growth of microorganisms, since some
79 microorganisms may not grow *in vitro*, as they are not in their native communities and,
80 therefore, do not receive signaling from this community through this intercellular
81 communication called quorum sensing (Nichols et al. 2008).

82 Among known signaling molecules, acylhomoserine lactones represent the best
83 understood compounds used by many Gram-negative bacteria (Babenko et al. 2022; Pappenfort
84 and Bassler 2016). Whereas, autoinductive peptides are involved in intercellular
85 communication in Gram-positive bacteria and may have a dual function, acting as an
86 antimicrobial and signaling molecule (Verbeke et al. 2017). Siderophores, whose biosynthesis
87 is controlled by a quorum sensing system, can also stimulate bacterial growth (Soares 2022).
88 *Paenibacillus elgii* is known to produce peptides, self-inducing lactones and siderophores,
89 signal molecules that can stimulate bacterial growth (da Costa et al. 2022; Ortega, et al. 2018).

90 Antimicrobial producing organisms are expected to contain self resistance
91 mechanisms against their own bioactive molecules. Generally, antibiotic producing bacteria
92 encode resistance genes that protect them from their antimicrobials (Mak et al. 2014). It is
93 expected that bacterial strains resistant to *P. elgii* have intrinsic resistance because they also
94 produce antimicrobial compounds. Therefore, in this work, we used a bacterial culture
95 methodology based on the use of *P. elgii* supernatant as a selection mechanism for the growth
96 of bacterial strains of biotechnological interest. Furthermore, a strain obtained through this
97 methodology was characterized in order to illustrate the potential for biotechnological
98 application of the obtained strains.

99

100 **Material and methods**

101 **Obtaining the Supernatant**

102 To produce the supernatant, an inoculum of *P. elgii* was made in LB Broth (NaCl,
103 5 g/L; Tryptone, 10 g/L; Yeast Extract, 5 g/L) culture medium. Growth was overnight under
104 200 rpm agitation at 28°C. After growth, the supernatant was subjected to 4°C overnight to
105 induce sporulation. Then, the culture was centrifuged and filtered through a 0.22µm syringe.
106 The filtered supernatant was inoculated onto plates with LB Broth medium and incubated for
107 48 h to confirm the absence of *P. elgii* colonies.

108

109 Preparation of culture media with *Paenibacillus elgii* supernatant

110 Plates containing different concentrations of the supernatant were prepared: 300µl/L,
111 600µl/L, 1ml/L and 2ml/L of supernatant. The culture media used for the tests were
112 Middlebrook agar (Casein enzymatic hydrolysate, 1(g/L); Ammonium sulfate, 0.5(g/L);
113 Monopotassium phosphate, 1.5(g/L); Disodium phosphate, 1.5(g/L); Sodium citrate, 0.4(g/L);
114 Magnesium sulfate, 0.05(g/L); L-Glutamic acid, 0.5(g/L); Ferric ammonium citrate,
115 0.04(g/L); Pyridoxine, 0.001(g/L); Biotin, 0.0005(g/L); Malachite green, 0.001(g/L); Agar,
116 15.0(g/L) and R2A agar (agar, 1.9 g/L; 15 g/L; casein acid hydrolysate, 0.5 g/L; dextrose,
117 0.5 g/L; dipotassium phosphate, 0.3 g/L; magnesium sulfate, 0.024 g/L; proteose peptone,
118 0.5 g/L; sodium pyruvate, 0.3 g/L; starch soluble, 0.5 g/L; yeast extract, 0.5 g/L).

119

120 Processing of environmental samples in the laboratory

121 Soil samples were collected using an aluminum pipe with a 15cm depth indication in a
122 native area of the Cerrado biome, Brasília, Federal District, Brazil. The pipe was introduced
123 with the aid of a rubber mallet into the ground until reaching the 15 cm mark. Ten collection
124 holes were made and the entire sampled content was sieved through a 2 mm mesh and mixed
125 in a single properly identified plastic bag.

126 Environmental soil samples were previously dissolved in liquid Middlebrook medium
127 at the rate of 1g of soil in 50 ml of medium (this was considered undiluted). Soon after, they
128 were diluted in 1:20 and inoculated in culture medium plates enriched with *P. elgii*
129 supernatant.

130

131 DNA extraction and 16S ribosomal *rRNA* analysis of the obtained strains

132 The DNA extractions of the bacterial strains obtained from the use of the *P. elgii*
133 supernatant were performed with the commercial kit 'GenElute Bacterial Genomic DNA Kit',
134 Sigma-Aldrich, according to the manufacturer's instructions both for the analysis of the *rRNA*
135 marker 16S and for genomic sequencing. The extraction efficiency was evaluated in agarose

136 gel (1% v/v). DNA quantification was performed with a Qubit fluorometer (Invitrogen, USA)
137 using the dsDNA BR Quant-it assay kit (Invitrogen).

138 Polymerase chain reaction (PCR) (DNA) for amplification of 16S *rRNA* was
139 performed with primers 27F and 1492R. Purification of the PCR product was performed with
140 the Thermo Scientific purification kit (Gene Jet PCR Purification Kit). Bidirectional
141 sequencing was performed on Macrogen's Sanger platform (Seoul, South Korea). Comparison
142 by similarity with sequences available in the GenBank of the *NCBI* (National Center for
143 Biotechnology Information) was performed using the BLAST N tool.

144

145 Antimicrobial Activity Analysis Protocol.

146 An overlay test was performed to confirm the antibiotic capacity of *P. elgii* and its
147 supernatant. Antibacterial activity tests were performed in minimal culture media (R2A or
148 Middlebrook). The *P. elgii* strain and its supernatant were inoculated in the center of the
149 plates and placed in growth at room temperature for a period of 6 days. On the seventh day of
150 growth, a soft-agar inoculum with *Bacillus subtilis* was prepared and poured on to the plate,
151 in order to form a thin and homogeneous layer. This was the cover or overlay. After the
152 addition of soft-agar, growth of *B. subtilis* was expected for up to 48 hours at room
153 temperature.

154 Bacterial strains isolated from the use of *P. elgii* supernatant were also submitted to
155 the overlay assay against *B. subtilis*, and some against *Escherichia coli* and *Pseudomonas*
156 *aeruginosa*.

157

158 *In vitro* phenotypic characterization

159 The CBMAI 824 strain of *Methylobacterium radiotolerans* (CBMAI) will be used to
160 perform the phenotypic and physiological analyzes in a comparative way with the K003 strain
161 of *Methylobacterium*.

162 Cells were observed under light microscopy to determine membrane characteristics
163 using differential Gram staining (Reddy et al. 2007). Observation was performed using a
164 Leica DM750 microscope. The optimum growth temperature was evaluated for one week on
165 NB agar (Beef extract, 1 g/L; Yeast Extract, 2 g/L; Peptone, 5 g/L; Sodium Chloride, 5 g/L;
166 agar, 15 g/L) plates at temperatures of 5, 10, 15, 20, 30, 35, 40., 45, 50, 55 and 60°C. The
167 optimum pH range for growth was examined between 4.0 and 10.0 in intervals of 1.0 after
168 incubation in NB (Beef extract, 1 g/L; Yeast Extract, 2 g/L; Peptone, 5 g/L; Sodium Chloride,
169 5 g/L) at 30°C for 3 days in liquid medium adjusted with HCl or NaOH containing 50 mM

170 citrate/Na₂HPO₄ buffer (pH 4–7), 100 mM NaH₂PO₄ Buffer 4 /Na₂HPO₄ (pH 7–8) or 100
171 mM NaHCO₃ /Na₂CO₃ buffer (pH 9–10). Salt tolerance was tested in NB in liquid
172 medium supplemented with NaCl (0-5% in 1% intervals) and incubation at 30°C for 3 days.
173 Both salinity and pH were examined by measuring the OD₆₀₀ in SpectraMax M3 with
174 SoftMax Pro 6.0 software (Zhu et al. 2014).

175

176 Physiological characterization

177 Catalase activity was observed by the formation of bubbles after application of a 3%
178 (v/v) hydrogen peroxide solution (Raney and Oren 2011). Oxidase activity was checked using
179 Oxidase Strips 40560 (Millipore).

180 K003 and CBMAI cells were cultured in solid R2A medium and collected after 48h of
181 growth, transferred to glass tubes with teflon lids and then used for cellular fatty acid analysis,
182 as described by Ramaprasad et al. 2015. Esters Fatty acid methyl concentrations (FAME)
183 obtained were analyzed on a gas chromatograph (Agilent 7890^a) equipped with the Sherlock
184 Microbial Identification System software (Sherlock, v. 6.2) using the RTSBA6 method library
185 (Sasser 2001).

186 The susceptibility to antibiotics was evaluated in solid medium by the Kirby-Bauer
187 method (diffusion with disc) using filter paper discs impregnated with a standard
188 concentration of antibiotic and observation of halo formation after application of the disc on a
189 culture spread with swab and inoculated for 48h at 30° C. The antibiotics tested were (µg per
190 disc): Amikacin (30), Amoxicillin/Clavulanic acid (20/10), Ampicillin (10), Aztreonam (30),
191 Cephalexin (30), Cefazolin (30), Cefepime (30), Ceftazidime (30), Ceftriaxone (30),
192 Ciprofloxacin (5), Clindamycin (2), Chloramphenicol (30), Doxycycline (30), Erythromycin
193 (15), Gentamicin (10), Imipenem (10), Meropenem (10), Norfloxacin (10), Ofloxacin (5),
194 Oxacillin (1), Penicillin G (10), Piperacillin + Tazobactam (110), Polymyxin B (300),
195 Sulfonamide (300), Tetracycline (30) and Vancomycin (30).

196

197 Genome sequencing and assembly

198 Whole sequencing of the K003 NGS Illumina genome was performed at Macrogen
199 (Seoul, South Korea). The DNA library was prepared using the TruSeq Nano DNA Kit and
200 sequencing was performed on the NextSeq500 platform, generating paired-end reads of 150
201 base pairs.

202 In order to improve genome assembly, NGS sequencing was performed using long
203 reads technology on the Oxford Nanopore Technologies (ONT) platform (Oxford, UK). The

204 DNA library was prepared with the Rapid Sequencing Kit (SQK-RAD004) and the
205 sequencing was performed in Flow cell FLO-MINSP6 in the MinION device (MIN-101B).

206 The de novo assembly of a draft genome of the strain obtained from the use of the
207 supernatant was performed in a hybrid way with the Unicycler program (V. 0.5.0) (Wick et al.
208 2016). The CheckM software, through the Kbase platform, was used to determine
209 contaminating sequences and genome completeness (Parks et al. 2015). The coverage of the
210 Illumina reads was determined using the Bowtie2 software (v. 2.4.4) (Langmead et al. 2018)
211 and the L50 and N50 parameters were obtained using Quast (Gurevich et al. 2013). Artemis
212 software was used to visualize the circular genome and its coding regions through the
213 DNAPlotter resource (Carver et al. 2009). Genome annotation was performed using Prokka
214 software (v. 1.14.6) (Seemann 2014).

215

216 Phylogenetic analyzes

217 The taxonomic classification of the strain and identity confirmation were performed
218 on the TYGS – DSMZ platform, by digital DNA hybridization (DDH) on the GGDC – DSMZ
219 server (Meier-Kolthoff et al. 2022). The ANI (Average Nucleotide Identity) was calculated on
220 the JSspecieWS platform (Richter et al. 2016).

221 Alignment of the genome of K003 and the genome of its reference strain, *M.*
222 *radiotolerans* JCM 2831, was performed with the Plot Nucmer (available in:
223 https://github.com/fmalmeida/plot_nucmer), which is an automation of MUMmer v3.1 toolkit
224 (Kurtz et al. 2004) and Ggbio (Yin et al. 2012), where the first aligns filtering by length and
225 identity, and the second generates a circular visualization. It is noteworthy that *M.*
226 *.radiotolerans* JCM 2831 is an acronym for *M. radiotolerans* CBMAI 824.

227 A phylogenetic tree based on concatenated central orthologous genes from genomic
228 sequences of 12 strains, including the K003 strain the closely related strains according to the
229 TYGS-DSMZ and the *Enterovirgarhinocerotis* outgroup available at the *NCBI* was
230 constructed using the MICR0B1AL1Z3R web server (<https://microbializer.tau.ac.il/>) (Avram
231 et al. 2019).

232 Comparative pan-genome analysis was performed using Roary version 3.13.0
233 following Prokka annotation, version 1.14.6 (Seemann, 2014), available from Galaxy (Afgan
234 et al. 2016). Complete genome sequences of K003 and its 10 closest sequences were used,
235 according to TYGS-DSMZ. The minimum percentage identity for blastp of 80% (Corretto et
236 al. 2017) was defined to group the genes that encode complete sequences of proteins in
237 central genomes

238 Gene prediction and functional annotation

239 Functional categories based on the COG (Clusters of Orthologous Groups)
240 classification were obtained using the online version of eggNOG-mapper v.2 (Huerta-Cepas
241 et al. 2017), where metabolic pathways and genes of interest were compared to the Kyoto
242 Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa et al. 2007).

243 The Antibiotic Resistance Gene Database (ARDB) was used for annotation of
244 potential genes responsible for antibiotic resistance (Liu and Pop, 2009). The CRISPR
245 (Clustered Regularly Interspaced Short Palindromic Repeats) loci were predicted using the
246 CRISPRCasFinder recognition tool (Couvin et al. 2018).

247 Two online software were used in the *in silico* analysis of the biotechnological
248 potential: AntiSMASH (v. 6.1.1) (Blin et al. 2021) to identify genes and gene clusters related
249 to secondary metabolism and CAMPR4 to predict antimicrobial peptides (AMPs) (Gawde et
250 al. 2022), where the sequence is predicted to be antimicrobial through a probability score (0 to
251 1). The higher the probability, the greater the possibility that the peptide is antimicrobial
252 (Waghu et al. 2014).

253

254 *In silico* phenotypic characterization

255 The *in silico* phenotyping tool Traitar (Microbial Trait Analyzer) (Weimann et al.
256 2016) was used to predict phenotypes based on K003 genome sequences in a comparative
257 way with its closest strains according to TYGS-DSMZ. This software is able to predict sixty
258 seven phenotypic traits using Prodigal for gene prediction and Pfam family for annotation
259 (Weimann et al. 2016). The software uses two prediction models, the phyPat model (which
260 predicts the presence/absence of proteins found in the phenotype of 234 bacterial species) and
261 a combination of phyPat+PGL models (which uses the phyPat information combined with the
262 acquisition or loss of protein families and phenotypes through evolution), to determine
263 phenotypic traits (Weimann et al. 2016).

264

265 **Results**

266 Efficiency of using *Paenibacillus elgii* supernatant

267 The species *P. elgii* showed antibacterial activity against *B. subtilis*, as well as the
268 supernatant obtained from its liquid culture (Supplementary Figure 1). The use of the
269 supernatant was efficient in stimulating the growth of microorganisms. The concentration of
270 the *P. elgii* supernatant influenced the growth of microorganisms in the culture medium after

271 inoculum of soil samples. The results demonstrated that there is a difference in bacterial
272 growth in the different concentrations of the supernatant (Supplementary Figure 2).

273 Taxonomic identification of strains

274 Among the bacterial strains grown from the supernatant rich culture of *P. elgii*, thirty-
275 five (35) could be grown alone and without the supernatant. The results of the 16S gene
276 sequencing of the strains obtained are presented in Table 1. Identification shows the presence
277 of the genus *Paenibacillus*, *Pseudomonas*, *Enterobacter*, *Kitasatospora*, *Sphingomonas*,
278 *Xanthobacter*, *Burkholderia* and *Methylobacterium*.

279

280 In vitro antibacterial activity of the prospected strains

281 The cultivated strains were submitted to the overlay assay of antibacterial activity
282 against *B. subtilis*, where all showed a positive result (Supplementary Figure 3 and
283 Supplementary Table 1).

284 Among the thirty-five strains, five (B2, B4, B7, C3 and K003) were selected to be
285 tested against *E.coli* and *P. aeruginosa*. The selection was made in order to include at least
286 one strain of each genus found, except for *Paenibacillus* and *Enterobacter*. The first because
287 it is the genus that originated the supernatant used to obtain the strains and the second because
288 its has lower biotechnological potential than the others strains. All had their antibacterial
289 activity confirmed against *E. coli*. And B7 and K003 also showed antibacterial activity on *P.*
290 *aeruginosa*.

291 The K003 strain, belonging to the genus *Methylobacterium*, was selected for genomic
292 analysis. This genus is known to promote plant growth (Grossi et al. 2020), show resistance
293 against radiation (kim et al. 2020) and have antimicrobial (bacteria and fungi), antioxidant
294 (Photolo et al. 2020) and against bacterial endophytes (Wang, 2019).

295

296 In vitro phenotypic characterization

297 Strain K003 and its reference strain were grown aerobically at temperatures from 15 to
298 40°C in NB medium. K003 grew in NB medium containing 0-2% (w/v) NaCl, while CBMAI
299 grew in 0-3%. K003 grew at pH 5–9 and CBMAI at pH 4-9. Optimum growth of both strains
300 occurred at 30 °C, pH 8 and 0 % (w/v) NaCl (Tab. 2). Growth from pH 4 distinguishes
301 CBMAI from K003. Both strains were positive for catalase and oxidase.

302 The fatty acid profile of strain K003 compared to CBMAI is shown in Table 3. The
303 most abundant fatty acid (>85%) in both strains was Summed Feature 8, while the least
304 abundant were Summed Feature 2 (0.74%) at K003 and C12:0 (0.67%) on CBMAI. The fatty

305 acid profiles clearly separated the CBMAI and K003 strain by the absence of C12:0 and 18:0
306 3OH.

307 Among the twenty-six antibiotics tested for susceptibility of the strains, K003 and
308 CBMAI were resistant to amoxicillin/Ac. Clavulanic, Aztreonam, cefazolin, cefepime,
309 clindamycin, meropenem, oxacillin, polymyxin B and sulfonamide. In addition to these,
310 CBMAI was also resistant to vancomycin. Table 4 shows the list of antibiotics tested and their
311 susceptibility.

312

313 Sequencing and phylogenetic analysis

314 Sanger sequencing of the 16S *rRNA* of K003 strain allowed the identification of the
315 sequence, where the blast indicated greater similarity (99.88%) with
316 *Methylobacteriumradiotolerans*. The phylogenetic tree based on the 16S *rRNA* of K003 and
317 the *Enterovirgarhinoceotis* outgroup is represented in Figure 1.

318 The genome sequence of the K003 strain was assembled using a combination of long
319 and short reading sequencing approaches, resulting in an assembly of 7735228 bp, distributed
320 in 5 *contigs*, with a GC Content of 70.84% (Tab. 5 and Fig. 2).

321 On the TYGS – DSMZ platform, the genome evaluation of K003 confirmed that the
322 strain is a *Methylobacteriumradiotolerans* with dDDH of 90.3-94%. The ANI (Average
323 Nucleotide Identity) calculated for the genome on the JspeciesWS platform was 97.51%.

324 The alignment of the genomes of K003 and its closest strain *M. radiotolerans*
325 JCM2831 shows few misaligned regions (16.95%) indicating that the genomes are similar to
326 each other (Fig. 3). Only genome alignment blocks with at least 1000 nucleotides and 90%
327 identity are considered.

328 Expanding comparative genomic analysis to investigate the phylogenetic relationships
329 of this genus, we used the genome sequences of 12 strains (strain K003, 10 closest species
330 and the *Enterovirgarhinoceotis* outgroup) to extract sets of orthologous genes. Using the
331 MICR0B1AL1Z3R web server, a total of 233 gene cores (subset of genes shared among the
332 12 strains) were retrieved and used to construct a maximum likelihood phylogenetic tree (Fig.
333 4).

334 It was possible to identify 21379 gene families composing the pan-genome of K003
335 and its closest strains and, of these, 2806 (13.12%) were present in the nucleus of the genome
336 and the rest as accessories. Rare genes were represented by 123387 families, that is, 57.7% of
337 the pan-genome (Fig. 5).

338 Comparing only the K003 and JCM 2831 strains, it was possible to identify 8455 gene
339 families composing the pan-genome and, of these, 5332 (63.06%) were present in the nucleus
340 of the genome and the rest as accessories. No rare genes were found among them.

341

342 Functional features of genomes

343 Figure 6 shows the distribution of proteins in the genome of the K003 strain in their
344 respective COG categories. The S category refers to proteins of unknown function and
345 represents 19% of the proteins in the K003 genome and the unknown category, 15%. Among
346 the well-noted categories, E (9%), K (6%) and C (6%), referring to amino acid metabolism,
347 transcription and energy production and conversion, respectively, are among the most
348 abundant in the genome. Among the least abundant are categories D (0.7%) and N (0.6%)
349 responsible for the cell cycle and cell motility.

350 With regard to antibiotic resistance, four resistance nodulation cell division (RND)
351 antibiotic efflux pumps were found in both K003 and JCM 2831, all conferring resistance to
352 fluoroquinolone and tetracycline antibiotics. Similarities ranged from 41 to 71%.

353 Two CRISPR matrices with 11 and 19 spacers, respectively, and a Type IIIA case
354 without CRISPR matrices were found at K003. No complete CRISPR-cas system was
355 detected. While in CBMAI, two CRISPR matrices with 1 and 6 spacers, respectively, were
356 also found, but no cas was found.

357 The functional annotation of genomes in Antismash showed a total of ten (10) gene
358 clusters, for K003 (Tab. 6). Most of these gene clusters are related to the production of
359 antibacterial compounds. Table 6 presents such clusters in comparison with the JCM 2831
360 strain (12 clusters), acronym to K003, and the ME94 strain (8 clusters) isolated from
361 medicinal plant leaves in India (Webster et al. 2020). K003 and JCM 2831 are very similar,
362 distinguished only by the absence of T3PKS in K003 and the fact that it has one less cluster of
363 Hserlactone. ME94 presents only Hserlactone, T1PKS and Terpene among the clusters in
364 common with K003 and JCM 2831, in addition it presents two unidentified clusters.

365 The CAMPR4 database for predicting antimicrobial peptides (AMPs) showed 40
366 predictions in K003. The predictions had a probability greater than 0.68, with a greater
367 probability of 0.75.

368

369 In silico phenotypic characterization

370 Figure 7 shows the comparison of the phenotype of K003 and its closest strains
371 according to DSMZ. The K003 strain differs from the others by being positive for d-sorbitol.

372 K003 shares several characteristics with *M. radiotolerans* JCM 2831, its closest strain
373 according to the DSMZ and acronym to CBMAI, however, K003 is distinguished by being
374 negative for capnophila and coagulase production and positive for d-sorbitol and l-
375 arabinose.

376 According to the phenotypic prediction in Traitar, the closest strain to K003 is *M.*
377 *brachiatum* DSM 19569. The K003 strain is distinguished by being negative for growth at
378 42°C, capnophilic and coagulase production and by being positive for D-sorbitol.

379

380 **Discussion**

381 For years, the scientific community considered microorganisms only as self-sufficient
382 individual beings and incapable of communicating or organizing themselves in groups
383 (Njoroge and Sperandio, 2009). However, it is now known that bacteria and other prokaryotes
384 have a communication system carried out through different processes that involve production,
385 release, detection and responses to chemical compounds called autoinducers, small molecules
386 similar to hormones that allow the microorganism to access the density cell of a local
387 population and coordinate gene regulation in a population (Gray et al. 2013).

388 Cell-to-cell communication systems called quorum sensing involve the density-
389 dependent formation of autoinducing acylhomoserine lactone (AHL) molecules or short
390 peptides and control light production, expression of virulence factors, swarming, biofilm
391 formation, cell aggregation and genetic competence (Papenfort and Bassler, 2016). Known
392 metabolic interactions also involve compounds participating in central metabolism, for
393 example, siderophores (Soares, 2022).

394 Most soil microorganisms resist cultivation under laboratory conditions and therefore
395 represent an enormous gap of biochemical diversity. To unlock this potential, it is necessary
396 that simulate the ecological context in which bacteria live, as this may allow us to access this
397 still untapped source of natural products (Kapinusova et al. 2023).

398 In a traditional culture, bacteria are normally grown in isolation with ample nutrients
399 and resources, which is in stark contrast to their complex natural habitat (Ho et al. 2017).
400 Complex environments, such as soil, shelter diverse microorganisms (Bahram et al. 2018) that
401 create conditions that allow the growth of other microorganisms by sharing metabolites and
402 essential substances for growth (Schink, 2002). These molecules include those that play a role
403 in quorum sensing and the community (Jacoby et al. 2017). The absence of neighbors in pure
404 culture reduces microbial recovery (Pande and Kost, 2017).

405 Some bacteria can only grow under laboratory conditions when in co-culture with
406 another member of the community, which may be another bacterium or even another
407 microorganism (Boilattabi et al. 2021). Coculture can be obtained by directly culturing the
408 microorganisms together or using supernatants (Stewart, 2012). The use of *P. elgii*
409 supernatant enriches the culture with inducer molecules and antimicrobial compounds that can
410 mimic naturally occurring chemical-ecological relationships. This may enhance microbial
411 recovery from environmental samples (Turan et al. 2017) and emerge as an effective and
412 targeted approach to finding new bioactive compounds with activity against the pathogens for
413 which new antibiotics are needed (van Bergeijk et al. 2020).

414 Using a culture-selected approach from *P. elgii* supernatant, 35 fast-growing cultivable
415 strains were successfully isolated and identified, most from genera known to have
416 biotechnological potential. They are *Paenibacillus*, already exposed here for being the base of
417 the supernatant, *Pseudomonas*, *Enterobacter Kitasatospora*, *Sphingomonas*, *Xanthobacter*,
418 *Burkholderia* that will be exposed below and *Methylobacterium*, whose strain was
419 characterized in this work to illustrate the potential of the methodology in selecting
420 microorganisms with possible biotechnological application.

421 Species belonging to the genus *Sphingomonas* have properties ranging from the
422 bioremediation of environmental contamination to the production of highly beneficial
423 phytohormones. Its antibacterial activity was proven by Akinsanya et al. (2017) against *S.*
424 *Typhimurium*, *P. vulgaris*, *K. pneumoniae*, *E. coli*, *S. aureus*, *S. pyogenes*, *B. cereus*, and *C.*
425 *albicans*. Recent studies have shown an intriguing role of *Sphingomonas* species in the
426 degradation of organometallic compounds (Asaf et al. 2020).

427 Endophytic bacteria of the genus *Xanthobacter* are known to promote plant growth. It
428 has interesting potential as a biofertilizer, in the biological control of weeds that compete with
429 plant cultivation, and has been reported as a useful biological tool for the bioremediation of
430 water and soil polluted by chemical agents (Sánchez-Yañez, 2022).

431 The genus *Kitasatospora* is characterized as a producer of new secondary metabolites
432 of varied structures, which are active even against eukaryotic organisms (Takahashi 2017), in
433 addition to showing activity against Gram-positive and Gram-negative pathogens (Zhu et al.
434 2014) and resistance to several heavy metals (Yun et al. 2020).

435 The genus *Burkholderia* is a more congruent group known for its role as a pathogen.
436 However, the diverse environmental interactions of these bacteria are now pointing to
437 multiple beneficial properties that extend from their capabilities in bioremediation and

438 biological control, to their significant biotechnological potential as antimicrobial producers
439 (Depoorter et al. 2016; Cain et al. ., 2000; Wang et al. 2020).

440 Although the genus *Enterobacter* is known as a phytopathogen (Singh et al. 2018), it
441 has already been mentioned as an effective bioinoculant for rice under alkaline salinity stress
442 (Sagar et al. 2020) and for having the ability to degrade polyethylene (Ren et al. 2019).

443 With regard to the biotechnological potential of *Pseudomonas*, representative
444 members of the species have been identified as plant growth promoters (Sah et al 2021),
445 bioremediation agents (Ojewumi et al. 2018) and producers of natural compounds such as
446 rhamnolipids, terpenoids, polyketides and peptides non-ribosomal (Loeschcke and Thies,
447 2015) and biopolymers (Mozejko-Ciesielska et al. 2019 ; Salvachúa et al. 2020).

448 The supernatant of *P. elggi* is rich in siderophores and self-inducing lactones, in
449 addition to peptides with dual properties, self-inducing and antibacterial (Ortega, *et al.*
450 2018), which may explain the growth stimulus at different concentrations of *P. elgii*
451 supernatants, the diversity of genus found in cultivation and selection of growth directed to
452 strains with biotechnological potential.

453 With the positive result of the antibacterial tests performed on the supernatant of *P.*
454 *elgii* and on the strains obtained through its use, the protocol's ability to select strains with
455 activity similar to the supernatant is demonstrated.

456 The result of this study included the characterization of the K003 strain and a draft of
457 its genome explored through genome mining. This combinatorial approach has led to
458 promising results of new antimicrobial metabolites. The K003 strain obtained in this study,
459 despite being a previously described species, has its own characteristics. Although it is similar
460 to its reference strain (CBMAI) regarding the alignment of its genomes, pangenome and the
461 diversity of biosynthetic clusters, K003 differs from CBMAI in growth in pH only from 5,
462 absence of certain fatty acids, non-resistance to vancomycin, presence from a Type IIIA case,
463 negative for coagulase production and positive for d-sorbitol and l-arabinose.

464 Comparing our results with the study by Knief et al. (2012) with 12 strains of
465 *Methylobacterium*, the main fatty acid summed feature 8 agrees with what they observed. Our
466 strains tolerated a higher salinity level (2% - K003 and 3% - CBMAI) compared to the same
467 study, where the strains only tolerated 1% salinity. With regard to pH, all strains mentioned
468 grew in general between pH 4-5 and 8.

469 Alessa et al. (2021) in their study with sixty-two strains of *Methylobacterium* and
470 *Methylorubrum* observed that the genome size of the strains ranges from 4.4 to 8.8 Mb with
471 CDS ranging from 4171 to 8622 and GC% ranging from 65.9 to 72.7 %. These data show that

472 the K003 strain is within the expected range for GC content (70.84%) and among the largest
473 genomes described (7.7 Mb).

474 Although the in vitro test showed that K003 and CBMAI are susceptible to
475 tetracycline, the result annotation of genes potentially responsible for antibiotic resistance in
476 the ARDB detected four resistance-nodulation-cell division (RND) antibiotic efflux pump that
477 confer resistance to tetracycline in addition to fluoroquinolone.

478 The high percentage (34%) of categories S and Unknown in K003 reflects the current
479 level of knowledge about the proteome of these microorganisms (Galperin et al. 2021) and
480 points out one of the limitations of annotation of genomes with reference in databases that are
481 based on prospecting only the structures already described. Alessa et al. (2021) verified that
482 the genomes of their sixty-two strains analyzed were highly enriched in the metabolism
483 process, which is consistent with what was observed in K003, where category E related to
484 amino acid metabolism is the most abundant among the described categories.

485 The presence of the CRISPR system in K003, which represents a defense mechanism
486 of prokaryotes against invaders such as viruses and plasmids (Wimmer and Beisel 2020) is
487 interesting from a biotechnological point of view due to its DNA editing power that favors
488 genetic engineering (Sedeek et al. 2019).

489 Many classes of natural products have been identified, including terpenoids, alkaloids,
490 polyketides and non-ribosomal peptides (Arnison et al. 2013). Some of them will be explored
491 below because they have clusters responsible for their production are present in the genome of
492 the strain obtained in this study.

493 The NRPS (nonribosomal peptide synthase) cluster, present at K003, is linked to
494 enzyme complexes organized into specific amino acid modules. NRPS uses non-
495 proteinogenic amino acids, not just the twenty possible in standard ribosomal peptides, which
496 increases the number of possible models to several hundred (Ding et al. 2020). Soon its
497 natural products exhibit several properties such as antiviral, antitumor or antibiotic activity
498 against human and plant pathogens (Huang et al. 2021) and attracting the interest of
499 biotechnological, pharmaceutical and agricultural industries (Biniarz et al. 2017).

500 Non-ribosomal peptide synthetases (NRPS) produce the majority of lipopeptides from
501 *Paenibacillus*, the genus whose supernatant was used to select for growth of the K003 strain.
502 The main mode of action of these lipopeptides usually involves disruption of the target cell
503 membrane and due to the difficulty of these cells to reorganize their membranes, it is
504 suggested that the development of resistance to lipopeptides may be slow and limited, which

505 is very interesting from the point of view point of view of the development of new drugs
506 (Cochrane and Vederas, 2016).

507 PKS clusters, together with NRPS, also found in K003, represent two large groups of
508 natural products with remarkable diversity and biological activities, such as the antibiotics
509 erythromycin, penicillin, daptomycin and vancomycin, the antifungals amphotericin and
510 echinocandin, the anticancer agents epothilone and bleomycin, the cholesterol-lowering agent
511 lovastatin, the immunosuppressants FK506 and cyclosporine, and the veterinary antibiotics
512 monensin and avermectin (Wang et al. 2011).

513 PKS are classified into Types I, II and III. Only Type I was found in K003. Within
514 type I PKS there is a group of products called polyethers ionophores that cause the death of
515 target cells, being of great importance for their antiparasitic activity and antifungal, antiviral,
516 herbicide, anti-inflammatory and broad-spectrum immunoregulatory action (Nivina et al.
517 2019).

518 The Hserlactone cluster, present at K003, refers to a homoserine lactone and is related
519 to quorum sensing, since in Gram-negative bacteria, acylated homoserine lactones (acyl-HSL)
520 are common signal molecules that regulate several target functions and are often involved in
521 interactions with the host. Quorum sensing plays a key role in the development of microbial
522 biofilms for various bacteria, including opportunistic human pathogens. Thus, if pathogens
523 use acyl-HSL signals to stimulate virulence gene expression, these systems represent a unique
524 target for drug discovery (Kumar et al. 2022).

525 Fewer terpenoid metabolites have been found in prokaryotes compared to those found
526 in fungi and plants. The first molecules found in bacteria were odoriferous terpenes, mainly
527 geosmin, a degraded sesquiterpene, which provides the characteristic odor of damp earth
528 (Cane and Ikeda, 2012) In K003, terpenoid clusters are present. Terpenes of microbial origin
529 have been used as potent antimicrobial agents, such as pentalene or therapeutecin (Yamada et
530 al. 2015).

531 One of the K003 terpenoid clusters was identified as a carotenoid, an important
532 component to deal with hostile environments due to intense UV light, which is expected for
533 *Methylobacterium*, since they can be abundant and adapted to the phyllosphere (Yoshida et al.
534 2017 and Alessa et al. 2021).

535 NPAA is a cluster of polyamino acids similar to Epsilon-poly-L-lysine (ϵ -PL) present
536 in K003. ϵ -PL is a natural antimicrobial peptide with a broad spectrum of activity against food
537 spoilage and food poisoning bacteria in vitro (Wang et al. 2021).

538 Another large class of natural products is formed by ribosomally synthesized and post-
539 translationally modified peptides (*RIPP*). This is a class of natural products with antimicrobial
540 activity, whose cluster is also found in K003. Post-translational modifications diversify the
541 structure of these natural products, which implies better recognition of targets and increased
542 chemical and metabolic stability, improving their chemical functionality (Ongpipattanakul et
543 al. 2022).

544 The results obtained using AntiSMASH, together with data from CAMPR4 in K003
545 illustrate the biotechnological potential that can be found in the strains obtained from the use
546 of *P. elgii* supernatant.

547 The presence of NRPS clusters, responsible for the production of pelgipeptins, non-
548 ribosomal peptides originating from *P.elgii* (Qian, et al. 2012), in K003, may indicate the
549 production of similar compounds between the K003 and *P.elgii* strains and, like this,
550 corroborate the premise that bacteria resistant to certain antimicrobials can produce
551 compounds similar to them.

552 Among the gene clusters identified in K003, except for a terpene identified as a
553 carotenoid, all showed little (<11%) or no similarity with known compounds, which suggests
554 the potential production of new bioactive compounds.

555

556 **Conclusions**

557 The methodology tested in this work proved to be efficient for the isolation of
558 different genus of bacteria of biotechnological interest, such as: *Paenibacillus*, *Pseudomonas*,
559 *Enterobacter*, *Kitasatospora*, *Sphingomonas*, *Xanthobacter*, *Burkholderia* and
560 *Methylobacterium*.

561 This methodology can be used with different media, concentrations and samples in
562 order to refine and optimize the protocol, expand collections of strains with biotechnological
563 potential and even allow the isolation of strains that have not been cultivated so far.

564 All strains obtained from the use of *P. elgii* supernatant demonstrate antibacterial
565 activity in overlay assays, attesting to the efficiency of the protocol in obtaining strains with
566 biotechnological potential. The isolation of pure cultures will allow future exploration of the
567 potential of these strains.

568 Genomic sequencing data confirmed that the K003 strain is a
569 *Methylobacteriumradiotolerans*. The draft genome of K003 allowed a sample about the
570 biosynthetic diversity that may be available in the strains obtained in this work. Future work
571 in identifying and isolating K003 metabolites can be guided by genomic insight into the

572 biosynthetic origins of its potential metabolites. Most of the biosynthetic gene clusters did not
573 have homology with published clusters and therefore represent a new source of
574 pharmaceutically relevant products.

575

576 **Acknowledgments**

577 We thank the Laboratories of the Department of Microbiology (Federal University of
578 Viçosa, Brazil) for the infrastructure to carry out part of the experiments in this work.

579

580 **Funding**

581 This study was funded by the National Council for Scientific and Technological
582 Development (CNPq), the Coordination for the Improvement of Higher Education Personnel
583 (CAPES), which also provided financial support to the first author, and the Federal District
584 Research Support Foundation (FAP -DF).

585

586 **Protologue**

587 The type strain is K003, isolated from Cerrado soil, Brasília, Brazil.

588

589 **Declaration of interests**

590 The authors declare that they have no known competing financial interests or personal
591 relationships that could have appeared to influence the work reported in this paper.

592

593 **Author Contributions**

594 Cunha-Ferreira, Vizzotto and Krüger designed the study. The collection of sample was
595 performed by Krüger. The tests were carried out by Cunha-Ferreira, Vizzotto and Tótola.
596 Data analyzes were carried out by Cunha-Ferreira, Vizzotto and Frederico. The manuscript
597 was prepared by Cunha-Ferreira and Vizzotto and revised by all authors. Quirino critically
598 revised and substantially edited the manuscript. All authors read and approved the manuscript.
599 Cunha-Ferreira and Vizzotto contributed equally on this paper.

600

601 **References**

602 Afgan E Baker D VAN Den Beek M Blankenberg D Bouvier D et al (2016) The Galaxy
603 Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2016 Update.
604 Nucleic Acids Res 44:W3-W10. <https://doi.org/10.1093/nar/gkw343>

- 605 Alessa O Ogura Y Fujitani Y Takami H Hayashi T Sahin N Tani A (2021) Comprehensive
606 Comparative Genomics and Phenotyping of *Methylobacterium* Species. Front Microbiol.
607 6:12-740610. [https://doi.org/ 10.3389/fmicb.2021.740610](https://doi.org/10.3389/fmicb.2021.740610)
- 608 Arnison PG Bibb MJ Bierbaum G Bowers AA Bugni TS Bulaj G et al. (2013) Ribosomally
609 synthesized and post-translationally modified peptide natural products: overview and
610 recommendations for a universal nomenclature. Nat Prod Rep 30(1):108-60. [https://doi.org/](https://doi.org/10.1039/c2np20085f)
611 10.1039/c2np20085f. PMID: 23165928
- 612 Asaf S Numan M Khan AL Al-Harrasi A (2020) *Sphingomonas*: from diversity and
613 genomics to functional role in environmental remediation and plant growth. Crit Rev
614 Biotechnol 40(2):138-152. [https://doi.org/ 10.1080/07388551.2019](https://doi.org/10.1080/07388551.2019)
- 615 Avram O Rapoport D Portugez S Pupko T (2019) M1CR0B1AL1Z3R-a user-friendly web
616 server for the analysis of large-scale microbial genomics data. Nucl Acids Res 47:W88–W92.
617 <https://doi.org/10.1093/nar/gkz423>
- 618 Babenko LM Kosakivska IV Romanenko KO (2022) Molecular mechanisms of N-
619 acyl homoserine lactone signals perception by plants. Cell Biol Int 46(4):523-534.
620 [https://doi.org/ 10.1002/cbin.11749](https://doi.org/10.1002/cbin.11749)
- 621 BachmannBO Van Lanem SG Baltz RH (2014) Microbial genome mining for accelerated
622 natural products discovery: is a renaissance in the making? J Ind MicrobiolBiotechnol
623 41(2):175-84. [https://doi.org/ 10.1007/s10295-013-1389-9](https://doi.org/10.1007/s10295-013-1389-9).
- 624 Bahram M Hildebrand F Forslund SK Anderson JL Soudzilovskaia NA et al. (2018).
625 Structure and function of the global topsoil microbiome. Nature 560(7717):233-237.
626 [https://doi.org/ 10.1038/s41586-018-0386-6](https://doi.org/10.1038/s41586-018-0386-6)
- 627 Biniarz P Łukaszewicz M Janek T (2017) Screening concepts, characterization and structural
628 analysis of microbial-derived bioactive lipopeptides: a review. Crit Rev Biotechnol 37(3):393-
629 410. [https://doi.org/ 10.3109/07388551.2016.1163324](https://doi.org/10.3109/07388551.2016.1163324)
- 630 Blin K Shaw S Kloosterman AM Charlop-Powers Z van Wezel GP et al. (2021) AntiSMASH
631 6.0: improving cluster detectionandcomparisoncapabilities. Nucleic Acids Res. 49(W1):W29-
632 W35. <https://doi.org/10.1093/nar/gkab335>
- 633 Boilattabi N Barrassi L Bouanane-Darenfed A La Scola B (2021). Isolation and identification
634 of *Legionella spp.* from hot spring water in Algeria by culture and molecular methods. J Appl
635 Microbiol 130:1394–1400. [https://doi.org/ 10.1111/jam.14871](https://doi.org/10.1111/jam.14871)
- 636 Cain CC Henry AT Waldo RH Casida LJ Jr Falkinham JO (2000) Identification and
637 characteristics of a novel *Burkholderia* strain with broad-spectrum antimicrobial activity.
638 Appl Environ Microbiol 66(9):4139-41. [https://doi.org/ 10.1128/AEM.66.9.4139-4141.2000](https://doi.org/10.1128/AEM.66.9.4139-4141.2000)

- 639 Cane DE Ikeda H (2012) Exploration and mining of the bacterial terpenome. *Acc Chem Res*
640 45(3):463-72. <https://doi.org/10.1021/ar200198d>
- 641 Carver T Thomson N Bleasby A Berriman M Parkhill J (2009) DNAPlotter: circular and
642 linear interactive genome visualization. *Bioinformatics*. 25(1):119-20.
643 <https://doi.org/10.1093/bioinformatics/btn578>
- 644 Chaudhary DK Khulan A Kim J (2019). Development of a novel cultivation technique for
645 uncultured soil bacteria. *Sci. Rep.* 9:6666. <https://doi.org/10.1038/s41598-019-43182-x>
- 646 Choi J Yang F Stepanauskas R Cardenas E Garoutte A et al. (2017). Strategies to improve
647 reference databases for soil microbiomes. *ISME J.* 11:829–34.[https://doi.org/](https://doi.org/10.1038/ismej.2016.168)
648 [10.1038/ismej.2016.168](https://doi.org/10.1038/ismej.2016.168)
- 649 Cochrane SA Vederas JC (2016) Lipopeptides from *Bacillus* and *Paenibacillus spp.*: A Gold
650 Mine of Antibiotic Candidates. *Med Res Rev* 36(1):4-31. <https://doi.org/10.1002/med.21321>
- 651 Corretto E Antonielli L Sessitsch A Compant S Höfer C et al. (2017) Complete Genome
652 Sequence of the Heavy Metal Resistant Bacterium *Agromyces aureus* AR33T and
653 Comparison with Related Actinobacteria. *Stand. Genomic Sci.* 12:2.
654 <https://doi.org/10.1186/s40793-016-0217-z>
- 655 Couvin D Bernheim A Toffano-Nioche C Touchon M Michalik J et al. (2018)
656 CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced
657 performance and integrates search for Cas proteins. *Nucleic Acids Res.* 46(W1):W246-W251.
658 <https://doi.org/10.1093/nar/gky425>
- 659 Depoorter E Bull MJ Peeters C Coenye T Vandamme P Mahenthiralingam E (2016)
660 *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers.
661 *Appl Microbiol Biotechnol* 100(12):5215-29. <https://doi.org/10.1007/s00253-016-7520-x>
- 662 Ding Y Ting JP Liu J Al-Azzam S Pandya P Afshar S (2020) Impact of non-proteinogenic
663 amino acids in the discovery and development of peptide therapeutics. *Amino*
664 *Acids* 52(9):1207-1226. <https://doi.org/10.1007/s00726-020-02890-9>
- 665 Durand GA Raoult D Dubourg G (2019) Antibiotic discovery: history, methods and
666 perspectives. *Int J Antimicrob Agents* 53:371–82, [http://dx.doi.org/10.](http://dx.doi.org/10.1016/j.ijantimicag.2018.11.010)
667 [1016/j.ijantimicag.2018.11.010](http://dx.doi.org/10.1016/j.ijantimicag.2018.11.010).
- 668 Galperin MY Wolf YI Makarova KS Vera Alvarez R Landsman D Koonin EV (2021) COG
669 database update: focus on microbial diversity, model organisms, and widespread pathogens.
670 *Nucleic Acids Res* 49(D1):D274-D281. <https://doi.org/10.1093/nar/gkaa1018>
- 671 Gawde U Chakraborty S Waghu FH Barai RS Khanderkar A et al. (2023) CAMPR4: a
672 database of natural and synthetic antimicrobial peptides. *Nucleic Acids Research.*

- 673 Gray B Hall P Gresham H (2013) Targeting agr- and agr-Like quorum sensing systems for
674 development of common therapeutics to treat multiple gram-positive bacterial infections.
675 *Sensors (Basel)* 13(4):5130-66. [https://doi.org/ 10.3390/s130405130](https://doi.org/10.3390/s130405130)
- 676 Grossi CEM Fantino E Serral F Zawoznik MS Fernandez Do Porto DA Ulloa RM
677 (2020) *Methylobacterium* sp. 2A Is a Plant Growth-Promoting Rhizobacteria That Has the
678 Potential to Improve Potato Crop Yield Under Adverse Conditions. *Front Plant Sci* 11:71.
679 [https://doi.org/ 10.3389/fpls.2020.00071](https://doi.org/10.3389/fpls.2020.00071)
- 680 Gurevich A Saveliev V Vyahhi N Tesler G (2013) QUASt: quality assessment tool for
681 genome assemblies. *Bioinformatics* 29(8):1072-5.
682 <https://doi.org/10.1093/bioinformatics/btt086>
- 683 Gutleben J Loureiro C Ramírez Romero L A Shetty S Wijffels R H Smidt H et al. (2020).
684 Cultivation of bacteria from *Aplysina* aerophoba: effects of oxygen and nutrient gradients.
685 *Front Microbiol* 11:175. [https://doi.org/ 10.3389/fmicb.2020.00175](https://doi.org/10.3389/fmicb.2020.00175)
- 686 Ho A Di Lonardo DP Bodelier PL (2017) Revisiting life strategy concepts in environmental
687 microbial ecology. *FEMS Microbiol Ecol* 93(3). [https://doi.org/ 10.1093/femsec/fix006](https://doi.org/10.1093/femsec/fix006)
- 688 Huang S Liu Y Liu WQ Neubauer P Li J (2021) The Nonribosomal Peptide Valinomycin:
689 From Discovery to Bioactivity and Biosynthesis. *Microorganisms* 9(4):780. [https://doi.org/](https://doi.org/10.3390/microorganisms9040780)
690 [10.3390/microorganisms9040780](https://doi.org/10.3390/microorganisms9040780)
- 691 Huerta-Cepas J Forslund K Coelho LP Szklarczyk D Jensen LJ et al. (2017) Fast Genome-
692 Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol Biol*
693 *Evol* 34(8):2115-2122. <https://doi.org/10.1093/molbev/msx148>
- 694 Jacoby R Peukert M Succurro AKoprivova A Kopriva S (2017) The Role of Soil
695 Microorganisms in Plant Mineral Nutrition-Current Knowledge and Future Directions. *Front*
696 *Plant Sci* 8:1617. [https://doi.org/ 10.3389/fpls.2017.01617](https://doi.org/10.3389/fpls.2017.01617)
- 697 Kanehisa M Araki M Goto S Hattori M Hirakawa M et al. (2008) KEGG for linking genomes
698 to life and the environment. *Nucleic Acids Res.* 36(Database issue):D480-4.
699 <https://doi.org/10.1093/nar/gkm882>
- 700 Kapinusova G Lopez Marin MA Uhlik O (2023) Reaching unreachables: Obstacles and
701 successes of microbial cultivation and their reasons. *Front Microbiol* 14:1089630.
702 [https://doi.org/ 10.3389/fmicb.2023.1089630](https://doi.org/10.3389/fmicb.2023.1089630)
- 703 Kim J Chhetri G Kim I Kim MK Seo T (2020) *Methylobacterium durans* sp. nov., a radiation-
704 resistant bacterium isolated from gamma ray-irradiated soil. *Antonie Van Leeuwenhoek*
705 113(2):211-220. [https://doi.org/ 10.1007/s10482-019-01331-2](https://doi.org/10.1007/s10482-019-01331-2)

- 706 Knief C Dengler V Bodelier PL Vorholt JA (2012) Characterization of *Methylobacterium*
707 strains isolated from the phyllosphere and description of *Methylobacterium longum* sp. nov.
708 Antonie Van Leeuwenhoek 101(1):169-83. [https://doi.org/ 10.1007/s10482-011-9650-6](https://doi.org/10.1007/s10482-011-9650-6)
- 709 Kumar L Patel SKS Kharga K Kumar R Kumar P Pandohee J et al. (2022) Molecular
710 Mechanisms and Applications of N-Acyl Homoserine Lactone-Mediated Quorum Sensing in
711 Bacteria. *Molecules* 27(21):7584. [https://doi.org/ 10.3390/molecules27217584](https://doi.org/10.3390/molecules27217584)
- 712 Kurtz S Phillippy A Delcher AL Smoot M Shumway M et al. (2004) Versatile and open
713 software for comparing large genomes. *Genome Biol.* 5(2):R12. [https://doi.org/10.1186/gb-](https://doi.org/10.1186/gb-2004-5-2-r12)
714 [2004-5-2-r12](https://doi.org/10.1186/gb-2004-5-2-r12)
- 715 Lagier JC Hugon P Khelaifia S Fournier PE La Scola B Raoult D (2015) The rebirth of
716 culture in microbiology through the example of culturomics to study human gut microbiota.
717 *Clin Microbiol Rev* 28(1):237-64. [https://doi.org/ 10.1128/CMR.00014-14](https://doi.org/10.1128/CMR.00014-14)
- 718 Langmead B Wilks C Antonescu V Charles R (2019) Scaling read aligners to hundreds of
719 threads on general-purpose processors. *Bioinformatics.* 35(3):421-432.
720 <https://doi.org/10.1093/bioinformatics/bty648>
- 721 Laws M, Shaaban A, Rahman KM. Antibiotic resistance breakers: current approaches and
722 future directions. *FEMS Microbiol Rev.* 2019 Sep 1;43(5):490-516. [https://doi.org/](https://doi.org/10.1093/femsre/fuz014)
723 [10.1093/femsre/fuz014](https://doi.org/10.1093/femsre/fuz014).
- 724 Liu B Pop M (2009) ARDB--AntibioticResistance Genes Database. *NucleicAcids Res.*
725 37(Databaseissue):D443-7. <https://doi.org/10.1093/nar/gkn656>
- 726 Liu S Moon, CD Zheng N Huws S Zhao S Wang J (2022). Opportunities and challenges of
727 using metagenomic data to bring uncultured microbes into cultivation. *Microbiome* 10:76.
728 [https://doi.org/ 10.1186/s40168-022-01272-5](https://doi.org/10.1186/s40168-022-01272-5)
- 729 Loeschcke A Thies S (2015) *Pseudomonas putida*-a versatile host for the production of
730 natural products. *Appl MicrobiolBiotechnol* 99(15):6197-214. [https://doi.org/](https://doi.org/10.1007/s00253-015-6745-4)
731 [10.1007/s00253-015-6745-4](https://doi.org/10.1007/s00253-015-6745-4)
- 732 Mak S Xu Y Nodwell JR (2014) The expression of antibiotic resistance genes in antibiotic-
733 producing bacteria. *Mol Microbiol* 93(3):391-402. [https://doi.org/ 10.1111/mmi.12689](https://doi.org/10.1111/mmi.12689)
- 734 Meier-Kolthoff JP Carbasse JS Peinado-Olarte RL Göker M (2022) TYGS and LPSN: a
735 database tandem for fast and reliable genome-based classification and nomenclature of
736 prokaryotes. *Nucleic Acids Res* 50(D1):D801-D807. <https://doi.org/10.1093/nar/gkab902>
- 737 Mozejko-Ciesielska J Szacherska K Marciniak P (2019) *Pseudomonas* Species as Producers
738 of Eco-friendly Polyhydroxyalkanoates. *J Polym Environ* 27:1151–1166
739 <https://doi.org/10.1007/s10924-019-01422-1>

- 740 Nichols D Lewis K Orjala J Mo S Ortenberg R O'Connor P et al. (2008) Short peptide induces
741 an "uncultivable" microorganism to grow in vitro. *Appl Environ Microbiol* 74(15):4889-97.
742 [https://doi.org/ 10.1128/AEM.00393-08](https://doi.org/10.1128/AEM.00393-08)
- 743 Nivina A Yuet KP Hsu J Khosla C (2019) Evolution and Diversity of Assembly-Line
744 Polyketide Synthases. *Chem Rev* 119(24):12524-12547.
745 <https://doi.org/10.1021/acs.chemrev.9b00525>
- 746 Njoroge J Sperandio V (2009) Jamming bacterial communication: new approaches for the
747 treatment of infectious diseases. *EMBO Mol Med* 1(4):201-10. [https://doi.org/
748 10.1002/emmm.200900032](https://doi.org/10.1002/emmm.200900032)
- 749 Ojewumi ME Okeniyi JO Ikotun JO Okeniyi ET Ejemen VA Popoola API (2018)
750 Bioremediation: Data on *Pseudomonas aeruginosa* effects on the bioremediation of crude oil
751 polluted soil. *Data Brief* 3;19:101-113. [https://doi.org/ 10.1016/j.dib.2018.04.102](https://doi.org/10.1016/j.dib.2018.04.102)
- 752 Ongpipattanakul C Desormeaux EK DiCaprio A van der Donk WA Mitchell DA Nair SK
753 (2022) Mechanism of Action of Ribosomally Synthesized and Post-Translationally Modified
754 Peptides. *ChemRev* 122(18):14722-14814. <https://doi.org/10.1021/acs.chemrev.2c00210>
- 755 Ortega DB Costa RA Pires AS Araújo TF Araújo JF et al. (2018) Draft Genome Sequence of
756 the Antimicrobial-Producing Strain *Paenibacilluselgii* AC13. *Genome Announc*
757 6(26):e00573-18. [https://doi.org/ 10.1128/genomeA.00573-18](https://doi.org/10.1128/genomeA.00573-18)
- 758 Overmann J (2013) Principles of enrichment, isolation, cultivation, and preservation of
759 bacteria. In E Rosenberg, DeLong EF, StackebrandtE, Lory S, Thompson F (ed) *The*
760 *Prokaryotes: Prokaryotic Biology and Symbiotic Associations*, 4th edn. New York: Springer,
761 pp 149-207
- 762 Overmann J Abt B Sikorski J (2017) Present and Future of Culturing Bacteria. *Annu Rev*
763 *Microbiol* 71:711-730. [https://doi.org/ 10.1146/annurev-micro-090816-093449](https://doi.org/10.1146/annurev-micro-090816-093449).
- 764 Pande S Kost C (2017) Bacterial Unculturability and the Formation of Intercellular Metabolic
765 Networks. *Trends Microbiol* 25(5):349-361. [https://doi.org/ 10.1016/j.tim.2017.02.015](https://doi.org/10.1016/j.tim.2017.02.015)
- 766 Papenfort K Bassler BL (2016) Quorum sensing signal-response systems in Gram-negative
767 bacteria. *Nat Rev Microbiol* 14(9):576-88. [https://doi.org/ 10.1038/nrmicro.2016.89](https://doi.org/10.1038/nrmicro.2016.89)
- 768 Parks DH Imelfort M Skennerton CT Hugenholtz P Tyson GW (2015) CheckM: assessing the
769 quality of microbial genomes recovered from isolates, single cells, and metagenomes.
770 *Genome Res* 25(7):1043-55. <https://doi.org/10.1101/gr.186072.114>
- 771 Pascual J García-López M Carmona C Sousa Tda S de Pedro N et al. (2014) *Pseudomonas*
772 *solii* sp. nov., a novel producer of xantholysin congeners. *Syst Appl Microbiol* 37:412–416.
773 [https://doi.org/ 10.1016/j.syapm.2014.07.003](https://doi.org/10.1016/j.syapm.2014.07.003)

- 774 Photolo MM Mavumengwana V Sitole L Tlou MG (2020) Antimicrobial and Antioxidant
775 Properties of a Bacterial Endophyte, *Methylobacteriumradiotolerans* MAMP 4754, Isolated
776 from *Combretum erythrophyllum* Seeds. Int J Microbiol 18:2020-9483670. doi:
777 10.1155/2020/9483670
- 778 Ramaprasad EVV Mahidhara G Sasikala C Ramana CV (2018) *Rhodococcuselectrodiphilus*
779 sp. nov., a marine electro active actinobacterium isolated from coral reef. Int J Syst
780 EvolMicrobiol 68(8):2644-2649. <https://doi.org/10.1099/ijsem.0.002895>
- 781 Raney F Oren A (2011) Taxonomy of Prokaryotes. ISBN
- 782 Reddy CA Beveridge TJ Breznak JA Marzluf GA Schmidt TM Snyder LR (2007) Methods
783 for General and Molecular Microbiology. American Society of Microbiology.
784 <https://doi.org/10.1128/9781555817497>
- 785 Ren L Men L Zhang Z Guan F Tian J Wang B Wang J Zhang Y Zhang W (2019)
786 Biodegradation of Polyethylene by *Enterobacter* sp. D1 fromtheGutsofWaxMoth *Galleria*
787 *mellonella*. Int J Environ Res Public Health 16(11):1941. [https://doi.org/](https://doi.org/10.3390/ijerph16111941)
788 10.3390/ijerph16111941
- 789 Richter M Rosselló-Móra R Oliver Glöckner F Peplies (2016) JSpeciesWS: a web server for
790 prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics
791 32(6):929-31. <https://doi.org/10.1093/bioinformatics/btv681>
- 792 Sagar A Sayyed RZ Ramteke PW Sharma S Marraiki N Elgorban AM Syed A (2020) ACC
793 deaminase and antioxidant enzymes producing halophilic *Enterobacter* sp. PR14 promotes
794 the growth of rice and millets under salinity stress. Physiol Mol Biol Plants 26(9):1847-1854.
795 [https://doi.org/ 10.1007/s12298-020-00852-9](https://doi.org/10.1007/s12298-020-00852-9)
- 796 Sah S Krishnani S Singh R (2021) *Pseudomonas* mediated nutritional and growth promotional
797 activities for sustainable food security. Curr Res Microb Sci 2:100084. [https://doi.org/](https://doi.org/10.1016/j.crmicr.2021.100084)
798 10.1016/j.crmicr.2021.100084
- 799 Salvachúa D Rydzak T Auwae R De CapiteA Black BA et al. (2020) Metabolic engineering
800 of *Pseudomonas putida* for increased polyhydroxyalkanoate production from lignin.
801 MicrobBiotechnol 13(1):290-298. [https://doi.org/ 10.1111/1751-7915.13481](https://doi.org/10.1111/1751-7915.13481)
- 802 Sánchez-Yañez JM (2022) *Xanthobacter autotrophicus*an Endophytic Beneficial Bacterium
803 for Wheat and Other Plants: A Short Review. Current Trends in Wheat Research pp 73.
804 [https://doi.org/ 10.5772/intechopen.102066](https://doi.org/10.5772/intechopen.102066)
- 805 Sasser M (2001) Technical Note # 101 Identification of Bacteria by Gas Chromatography of
806 Cellular Fatty Acids.

- 807 Schink B (2002) Synergistic interactions in the microbial world. *Antonie Van Leeuwenhoek*
808 81(1-4):257-61. [https://doi.org/ 10.1023/a:1020579004534](https://doi.org/10.1023/a:1020579004534)
- 809 Sedeek KEM Mahas A Mahfouz M (2019) Plant Genome Engineering for Targeted
810 Improvement of Crop Traits. *Front Plant Sci* 10:114. [https://doi.org/ 10.3389/fpls.2019.00114](https://doi.org/10.3389/fpls.2019.00114)
- 811 Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*
812 30(14):2068-9. <https://doi.org/10.1093/bioinformatics/btu153>
- 813 Singh RP Mishra S Jha P Raghuvanshi S Jha PN (2018) Effect of inoculation of zinc-resistant
814 bacterium *Enterobacter ludwigii* CDP-14 on growth, biochemical parameters and zinc uptake
815 in wheat (*Triticum aestivum* L.) plant. *Ecological Engineering* 116:163-173.
816 <https://doi.org/10.1016/j.ecoleng.2017.12.033>.
- 817 Soares EV (2022) Perspective on the biotechnological production of bacterial siderophores
818 and their use. *Appl Microbiol Biotechnol* 106(11):3985-4004. [https://doi.org/ 10.1007/s00253-](https://doi.org/10.1007/s00253-022-11995-y)
819 [022-11995-y](https://doi.org/10.1007/s00253-022-11995-y)
- 820 Stewart EJ (2012) Growing unculturable bacteria. *J Bacteriol* 194(16):4151-60.
821 [https://doi.org/ 10.1128/JB.00345-12](https://doi.org/10.1128/JB.00345-12)
- 822 Turan NB Chormey DS Büyükpınar C Engin GO Bakirdere S (2017) Quorum sensing: Little
823 talks for an effective bacterial coordination, *TrAC Trends in Analytical Chemistry* 91:1-11.
824 <https://doi.org/10.1016/j.trac.2017.03.007>
- 825 van Bergeijk DA Terlouw BR Medema MH van Wezel GP (2020) Ecology and genomics of
826 Actinobacteria: new concepts for natural product discovery. *Nat Rev Microbiol* 18(10):546-
827 558. [https://doi.org/ 10.1038/s41579-020-0379-y](https://doi.org/10.1038/s41579-020-0379-y)
- 828 van Dorst J M Hince G Snape I Ferrari B C (2016). Novel culturing techniques select for
829 heterotrophs and hydrocarbon degraders in a subantarctic soil. *Sci Rep* 6:36724.
830 [https://doi.org/ 10.1038/srep36724](https://doi.org/10.1038/srep36724)
- 831 Verbeke F De Craemer S Debunne N Janssens Y Wynendaele E et al. Peptides as Quorum
832 Sensing Molecules: Measurement Techniques and Obtained Levels In Vitro and In
833 Vivo. *Front Neurosci* 11:183. [https://doi.org/ 10.3389/fnins.2017.00183](https://doi.org/10.3389/fnins.2017.00183)
- 834 Waghu FH Barai RS Gurung P Idicula-Thomas S (2016) CAMPR3: a database on sequences,
835 structures and signatures of antimicrobial peptides. *Nucleic Acids Res* 44:D1094-
836 D1097. <https://doi.org/10.1093/nar/gkv1051>
- 837 Wang H Liu N Xi L Rong X Ruan J Huang Y (2011) Genetic screening strategy for rapid
838 access to polyether ionophore producers and products in actinomycetes. *Appl Environ*
839 *Microbiol* 77(10):3433-42. <https://doi.org/10.1128/AEM.02915-10>

- 840 Wang J Gao C Chen X Liu L (2021) Expanding the lysine industry: biotechnological
841 production of l-lysine and its derivatives. *Adv Appl Microbiol* 115:1-33. [https://doi.org/](https://doi.org/10.1016/bs.aambs.2021.02.001)
842 10.1016/bs.aambs.2021.02.001
- 843 Wang SS Liu JM Sun J Sun YF Liu JN Jia N Fan B Dai XF (2019) Diversity of culture-
844 independent bacteria and antimicrobial activity of culturable endophytic bacteria isolated
845 from different *Dendrobium* stems. *Sci Rep* 9:10389. [https://doi.org/10.1038/s41598-019-](https://doi.org/10.1038/s41598-019-46863-9)
846 46863-9
- 847 Wang Y Hoffmann JP Chou CW Höner Zu Bentrup K et al. (2020) *Burkholderiathailandensis*
848 outer membrane vesicles exert antimicrobial activity against drug-resistant and competitor
849 microbial species. *J Microbiol* 58(7):550-562. doi: 10.1007/s12275-020-0028-1
- 850 Webster G Mullins AJ Cunningham-Oakes E Renganathan A Aswathanarayan JB et al.
851 (2020) Culturable diversity of bacterial endophytes associated with medicinal plants of the
852 Western Ghats, India. *FEMS Microbiol Ecol* 96(9):fiae147. [https://doi.org/](https://doi.org/10.1093/femsec/fiae147)
853 10.1093/femsec/fiae147
- 854 Weimann A Mooren K Frank J Pope PB Bremges A McHardy AC (2016) From genomes to
855 phenotypes: Traitair, the microbial trait analyzer. *mSystem*. <https://doi.org/10.1101/043315>
- 856 Wick RR Judd LM Gorrie CL Holt KE (2017) Unicycler: Resolving bacterial genome
857 assemblies from short and long sequencing reads. *PLoS Comput Biol* 8:13(6)-e1005595.
858 <https://doi.org/10.1371/journal.pcbi.1005595>
- 859 Wilson L Iqbal KM Simmons-Ehrhardt T Bertino MF Shah MR et al. (2019) Customizable
860 3D printed diffusion chambers for studies of bacterial pathogen phenotypes in complex
861 environments. *J Microbiol Methods* 162:8-15. doi: 10.1016/j.mimet.2019.05.002
- 862 Wimmer F Beisel CL (2020) CRISPR-Cas Systems and the Paradox of Self-Targeting
863 Spacers. *Front Microbiol* 22:10:3078. <https://doi.org/10.3389/fmicb.2019.03078>
- 864 Yamada Y Kuzuyama T Komatsu M Shin-Ya K Omura S et al. (2015) Terpene synthases are
865 widely distributed in bacteria. *Proc Natl Acad Sci USA* 112(3):857-62.
866 <https://doi.org/10.1073/pnas.1422108112>
- 867 Yin T Cook D Lawrence M (2012) Ggbio: an R package for extending the grammar of
868 graphics for genomic data. *Genome Biol* 13(8):R77. <https://doi.org/10.1186/gb-2012-13-8-r77>
- 869 Yoshida S Hiradate S Koitabashi M Kamo T Tsushima S (2017)
870 Phyllosphere *Methylobacterium* bacteria contain UVA-absorbing compounds. *J*
871 *Photochem Photobiol B* 167:168-175. <https://doi.org/10.1016/j.jphotobiol.2016.12.019>

872 Yun BR Malik A Kim SB (2020) Genome based characterization of *Kitasatospora sp.*
873 MMS16-BH015, a multiple heavy metal resistant soil actinobacterium with high
874 antimicrobial potential. Gene 5:733-144379. <https://doi.org/10.1016/j.gene.2020.144379>

875 Zhu D Xie C Huang Y Sun J Zhang W (2014) Description of *Comamonasserinivorans sp.*
876 nov., isolated from wheat straw compost. Int J Syst Evol Microbiol. 64(Pt 12):4141-4146.
877 <https://doi.org/10.1099/ijs.0.066688-0>

878

879 **Table 1** Results of the comparison of the amplicons of the strains obtained from the use of the
880 supernatant of *P. elgii* with the databases of the *NCBI* (National Center for Biotechnology
881 Information) through BLAST. Sequencing was performed by Macrogen in South Korea.

882 **Table 2** Comparative phenotypic properties between the *Methylobacterium* K003 strain,
883 obtained from the use of *P. elgii* supernatant, and the *M. radiotolerans* CBMAI 824 strain. +,
884 Positive.

885 **Table 3** Composition of fatty acids (%) in a comparative way between the *Methylobacterium*
886 K003 strain, obtained from the use of the supernatant of *P. elgii*, and its reference strain *M.*
887 *radiotolerans* CBMAI 824. The fatty acid profile was obtained through gas chromatography.
888 *Summed Feature occurs when there is imperfect peak separation. Summed feature 2
889 represents possibly 12:0 aldehyde, Summed feature 3 represents C16:1 ω 6c and/or C16:1 ω 7c
890 and summed feature 8 represents C18 : 1 ω 7c and/or C18 : 1 ω 6c. –, absent.

891 **Table 4** Results of the antibiotic susceptibility test of the strains obtained from the
892 supernatant of *P. elgii*. The method adopted was the Kirby-Bauer (disc diffusion), where filter
893 paper discs impregnated with a standard concentration of antibiotic are used and the formation
894 of an inhibition halo is observed. +, Positive; –, negative.

895 **Table 5** Data from the hybrid assembly of the genome of the *Methylobacterium* K003 strain,
896 obtained from the supernatant of *P. elgii*. The assembly was performed from the long and
897 short reading sequencing through the Unicycler program. The annotation was performed with
898 Prokka software and the quality parameters were evaluated in CheckM, Bowtie2 and Quast
899 software.

900 **Table 6** Results of the analysis of clusters of biosynthetic genes for secondary metabolites of
901 the *M. radiotolerans* K003 strain, obtained from the use of the supernatant of *P. elgii*. The
902 analysis was performed using the AntiSMASH software (v. 6.1.1) and presented in a shared
903 way with the JCM 2831 and ME94 strains of *M. radiotolerans*.

904 **Figure 1** Phylogenetic tree based on the partial sequence of the 16S ribosomal *rRNA* gene
905 obtained by cloning and Sanger sequencing of the *Methylobacterium* K003 strain, obtained

906 from the use of *P. elgii* supernatant. The evolutionary position of the strain was estimated by
 907 constructing a phylogenetic tree using the Maximum-likelihood (ML) method with 1000
 908 sampling repetitions (Bootstrap), using representative 16S ribosomal *rRNA* sequences from
 909 valid strains of the genus listed in List of Prokaryotic names with Standing in Nomenclature -
 910 LPSN.

911 **Figure 2** Circular map of the hybrid assembly of the genome of the *M. radiotolerans* K003
 912 strain, obtained using the supernatant of *P. elgii*. The illustration was generated by the
 913 DNAPlotter feature of the Artemis software demonstrating the coding regions and the GC
 914 content. Outside to center: Forward CDS (dark blue), Reverse CDS (red), GC content and GC
 915 slope (dark gray and light light gray).

916 **Figure 3** Circular visualization of the alignment of the genomes of the *M. radiotolerans* K003
 917 strain (in gray) and its closest strain according to TYGS-DSMZ, *M. radiotolerans* JCM2831
 918 (in blue). Alignment and illustration were generated using Plot Nucmer. Only genome
 919 alignment blocks with at least 1000 nucleotides and 90% identity are shown. The outer band
 920 is the size scale of the *contigs*. When the blue band is two layers deep, there are two
 921 sequences in the closest strain that overlap at the same location in the K003 strain.

922 **Figure 4** Maximum likelihood phylogenetic tree based on 233 core orthologous genes
 923 concatenated from genomic sequences of 12 *Methylobacterium* strains, including K003, the
 924 closely related strains according to the TYGS-DSMZ and the *Enterovirgarhinocerotis*
 925 outgroup. The M1CR0B1AL1Z3R web server was used for concatenating the orthologous
 926 genes and building the phylogenetic tree.

927 **Figure 5** Comparative pan-genome analysis performed on Roary. Sequences of the complete
 928 genome of *M. radiotolerans* K003 and the 10 closest sequences to it, according to TYGS-
 929 DSMZ, were used. A minimum percent identity for blastp of 80% was defined to cluster the
 930 genes encoding complete protein sequences in core genomes.

931 **Figure 6** Proteins predicted in the genome of the *M. radiotolerans* K003 strain, obtained from
 932 the use of the supernatant of *P. elgii*, classified according to the categories COGs.
 933 Classification was performed using the eggNOG mapper against the Kyoto Encyclopedia of
 934 Genes and Genomes (KEGG) database.

935 **Figure 7** Comparison between the phenotypic predictions of the hybrid assembly of the
 936 genome of the *M. radiotolerans* K003 strain, obtained from the use of the supernatant of *P.*
 937 *elgii*, and the genomes of their closest strains according to the TYGS-DSMZ obtained from
 938 the NCBI database. The zero value means that the strain is negative for the phenotype, the
 939 values 1 and 2 correspond to the positive for the phenotype in the phypat and

940 phyopat+PGLpredictors, respectively. The value 3 shows that the phenotype was positive in
941 both predictors. The analysis was performed using the Traitair software.

942 **Supplementary Table 1** Results of overlay assays of antibacterial activity of strains obtained
943 from the use of *P. elgii* supernatant. After the growth of the strains, a soft-agar inoculum with
944 *B. subtilis* was poured onto the plate and its growth was observed for up to 48 hours at room
945 temperature.

946 **Supplementary Figure 1** Efficiency test of the antibiotic capacity of *P. elgii* and its
947 supernatant, demonstrating activity of the culture indicated by the red arrow (A) and of the
948 supernatant (300µl/L) indicated by the blue arrow (B). Assays were done against *B. subtilis*
949 (indicated by yellow arrows)

950 **Supplementary Figure 2** Growth of cultures from soil samples in different concentrations of
951 *P. elgii* supernatant. The A plate (300µl/L), the B plate (600 µl/L), the C plate (1ml/L) and the
952 D plate (2ml/L). Strains of interest collected because they have zones of inhibition are
953 indicated by red arrows.

954 **Supplementary Figure 3** Overlay assays of antibacterial activity of some of the strains
955 obtained from the use of *P. elgii* supernatant. The inhibition halos are quite visible. (A) Strain
956 A3, (B) Strain A9, (C) Strain C5, and (D) Strain C12. All assays were performed against *B.*
957 *subtillis*. The backgrounds for obtaining the images were different due to the difficulty of
958 viewing some halos.

Table 1

Strain	Access number	Description	Max Score	Total Score	Query Coverage	Max Ident
A1	HQ236048.1	<i>P. elgii</i>	167	167	59%	92%
A2	JX174235.1	<i>Enterobacter sp.</i>	302	302	96%	95%
A3	AB739006.1	<i>P. ehimensis</i>	623	623	33%	92%
A4	AB739006.1	<i>P. ehimensis</i>	645	645	26%	98%
A5	HQ236085.1	<i>P. elgii</i>	1831	1831	98%	95%
A6	HM063032.1	<i>Paenibacillus sp.</i>	1943	1943	99%	97%
A7	AB739006.1	<i>P. ehimensis</i>	287	287	96%	93%
A8	AB739006.1	<i>P. ehimensis</i>	248	248	95%	90%
A9	HQ236085.1	<i>P. elgii</i>	474	474	98%	97%
A10	JQ735955.1	<i>P. elgii</i>	808	808	68%	98%
A11	HQ236048.1	<i>P. elgii</i>	1034	1034	99%	99%
A12	JX436405.1	<i>Pseudomonas sp.</i>	350	350	96%	99%
B1	JF701960.1	<i>Burkholderia sp.</i>	512	512	36%	92%
B2	GQ895737.1	<i>Sphingomonas sp.</i>	760	760	34%	95%
B3	JX090597.1	<i>P. aeruginosa</i>	22292	22292	99%	99%
B4	GU518381.1	Unculturedbacterium	115	115	83%	80%
B5	AF247494.1	<i>Burkholderia sp.</i>	642	642	97%	95%
B6	JX090597.1	<i>P. aeruginosa</i>	1000	1000	98%	98%
B7	EU139032.1	<i>Kitasatospora sp.</i>	263	263	94%	92%
B8	HQ651053.1	<i>Pseudomonas sp.</i>	1018	1018	98%	99%
B9	AF247494.1	<i>Burkholderia sp.</i>	388	388	48%	97%
B10	FR872408.1	<i>Burkholderia sp.</i>	329	329	90%	97%
B11	DQ095878.1	<i>P. aeruginosa</i>	176	176	64%	91%
B12	GU191167.1	<i>Burkholderia sp.</i>	294	294	88%	97%
C1	JX436403.1	<i>Pseudomonas sp.</i>	675	675	67%	96%
C2	JX436403.1	<i>Pseudomonas sp.</i>	497	675	63%	94%
C3	EU912440.1	<i>Methylobacterium sp.</i>	453	453	66%	88%
C4	HM771642.1	<i>P. aeruginosa</i>	990	990	99%	93%
C5	JX178938.1	<i>Xanthobacter sp.</i>	425	425	27%	95%
K003	EU912444.1	<i>Methylobacterium sp.</i>	193	193	73%	90%
C7	JX436403.1	<i>Pseudomonas sp.</i>	545	545	65%	94%
C8	JQ659918.1	<i>P. aeruginosa</i>	156	156	94%	81%
C10	JX436405.1	<i>Pseudomonas sp.</i>	279	279	98%	92%
C11	GU447238.1	<i>P. aeruginosa</i>	2017	2017	95%	96%
C12	AF247494.1	<i>Burkholderia sp.</i>	887	887	97%	95%

Table 2

Growth range (optimum):	K003	CBMAI
Temperature (°C)	15-40 (30)	15-40 (30)
PH	5-9 (8)	4-9 (8)
NaCl (%)	0-2 (0)	0-3 (0)
Catalase	+	+
Oxidase	+	+

Table 3

Fattyacid	Name	Typesoffattyacid	K003 (%)	CBMAI (%)
C12:0	Dodecanoicacid	MediumchainSaturated	-	0.67
C16:0	Hexadecanoicacid	LongchainSaturated	2.37	2.81
C18:0	Octadecanoicacid	LongchainSaturated	5.75	5.59
18:0 3OH	3-Hydroxyoctadecanoic acid	Long chain Saturated with a hydroxyl	-	0.84
SummedFeature 2	-	-	0.74	0.87
SummedFeature 3	-	-	0.76	0.90
SummedFeature 8	-	-	90.39	88.33

Table 4

Anibiótico (µg per disco)	CBMAI	K003
Amikacin (30)	+	+
Amoxicillin/Clavulanicacid (20/10)	-	-
Ampicillin (10)	+	+
Aztreonam (30)	-	-
Cephalexin (30)	+	+
Cefazolin (30)	-	-
Cefepime (30)	-	-
Ceftazidime (30)	+	+
Ceftriaxone (30)	+	+
Ciprofloxacin (5)	+	+
Clindamycin (2)	-	-
Chloramphenicol (30)	+	+
Doxycycline (30)	+	+
Erythromycin (15)	+	+
Gentamicin (10)	+	+
Imipenem (10)	+	+
Meropenem (10)	-	-
Norfloxacin (10)	+	+
Ofloxacin (5)	+	+
Oxacillin (1)	-	-
Penicillin G (10)	+	+
Piperacillin + Tazobactam (110)	+	+
Polymyxin B (300)	-	-
Sulfonamide (300)	-	-
Tetracycline (30)	+	+
Vancomycin (30)	+	-

Table 5

Assembly data	
Contigs	5
CDS	7292
<i>RRNA</i>	9
Repeatregion	0
<i>TRNA</i>	73
<i>TmRNA</i>	2
GC content (%)	70,84
Contamination (%)	2,46
Completeness (%)	100
Coverage	400x
L50	1
N50 (pb)	5241738
Total Size (bp)	7735228
Major contig	5241738

Table 6

Secondary metabolites			
Type	No. of clusters		
	K003	JCM 2831	ME94
Hserlactone	2	3	2
NAPAA	1	1	0
NRPS	1	1	0
Redox-cofator	1	1	0
<i>RIPP</i>	1	1	0
T1PKS	1	1	1
T3PKS	0	1	0
Terpeno	3	3	3
Others	0	0	2

Figure 1

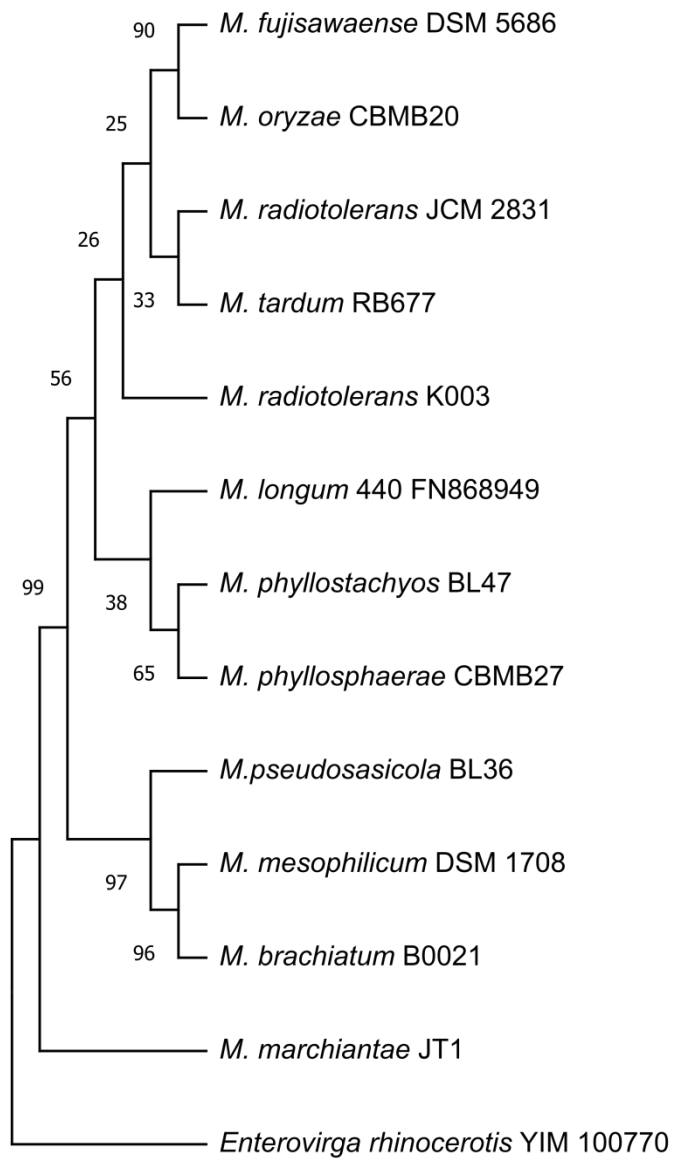


Figure 2

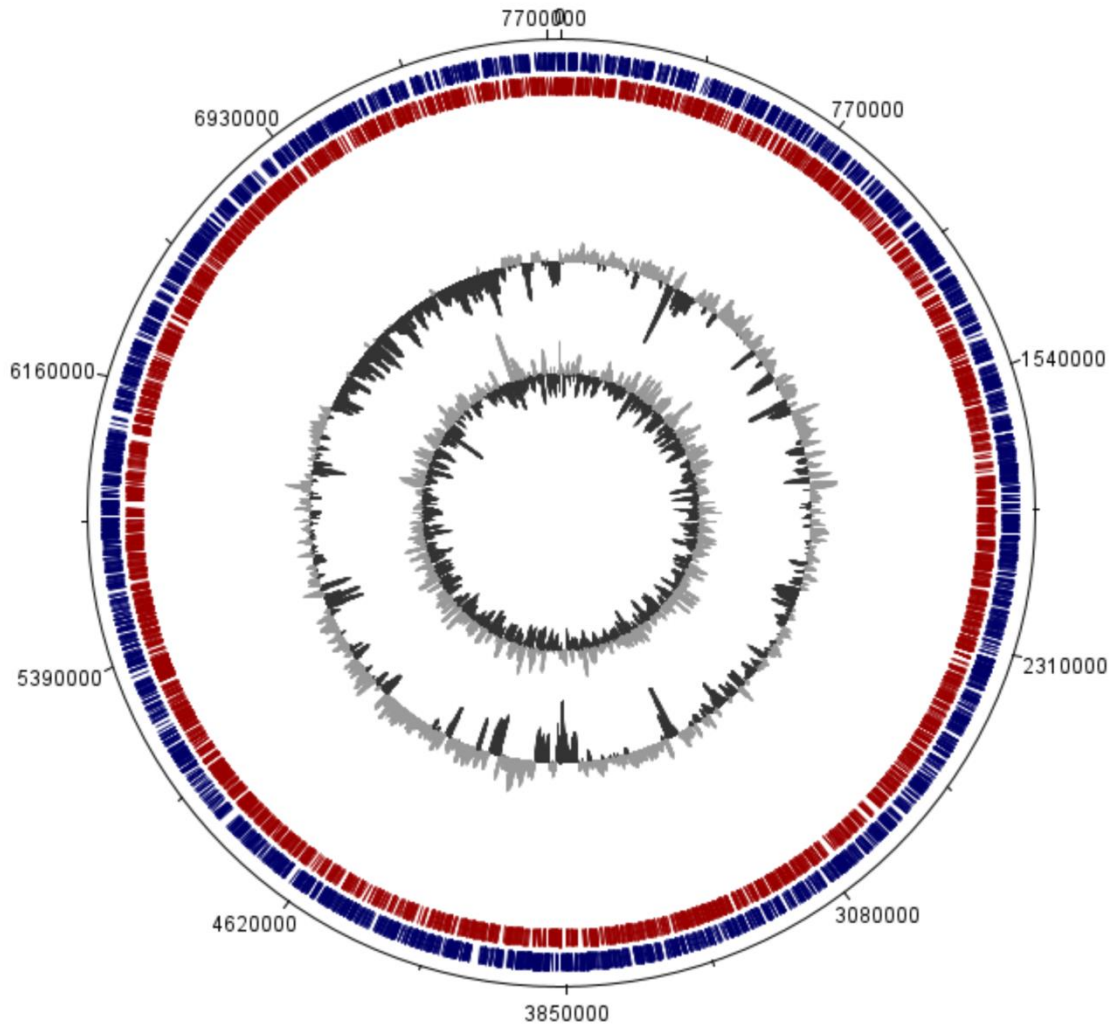


Figure 3

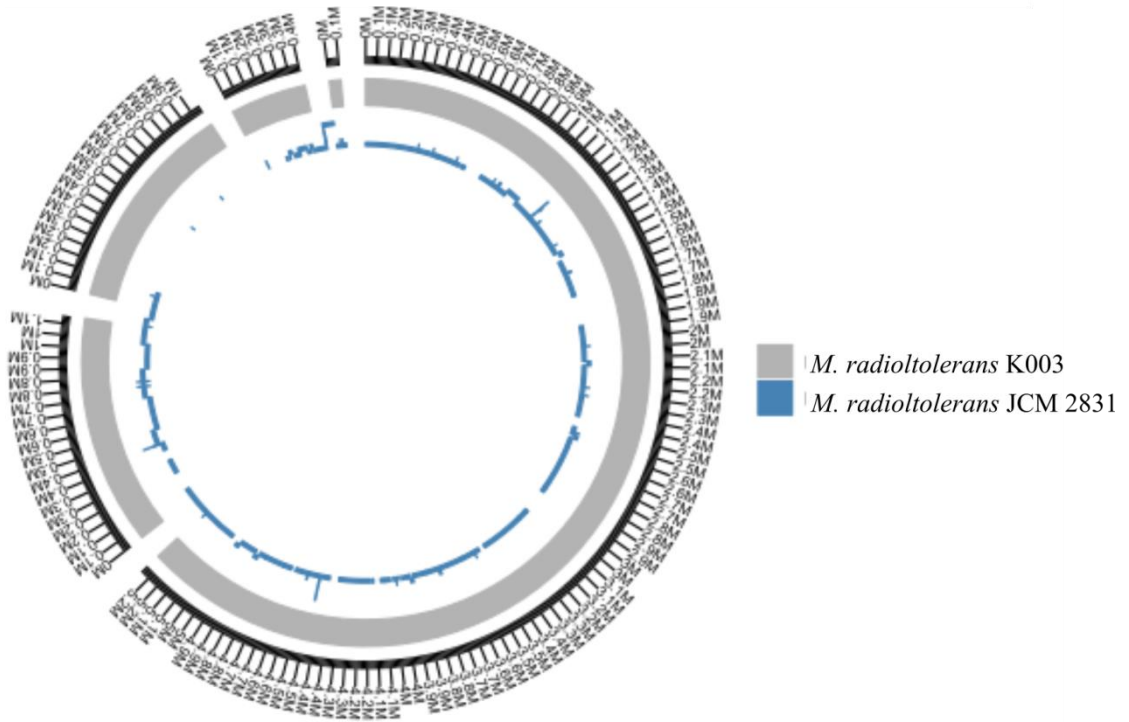


Figure 4

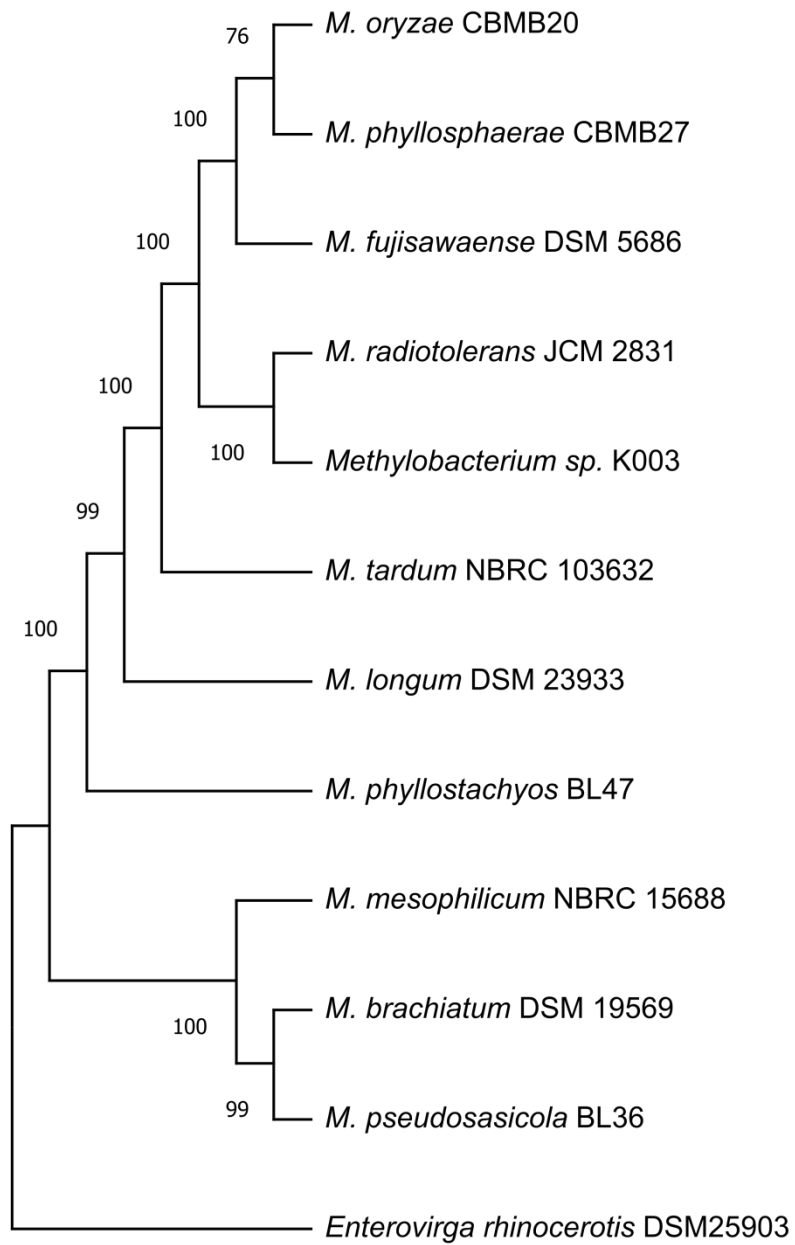


Figure 5

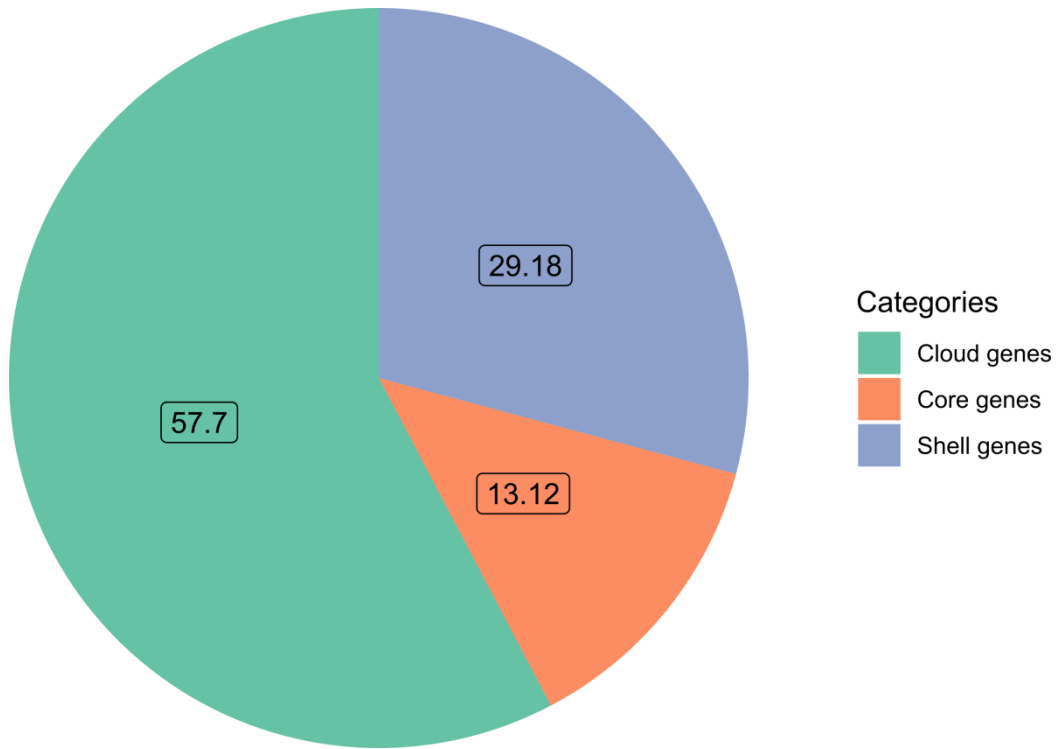


Figure 6

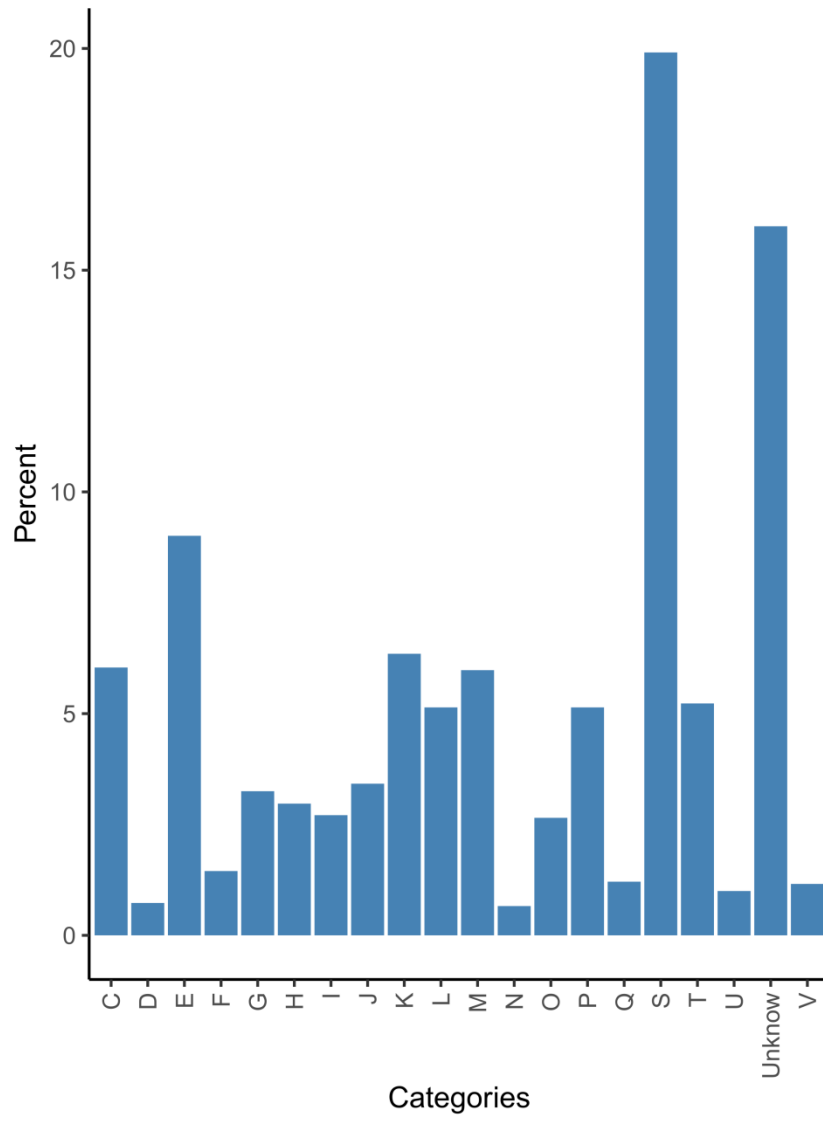
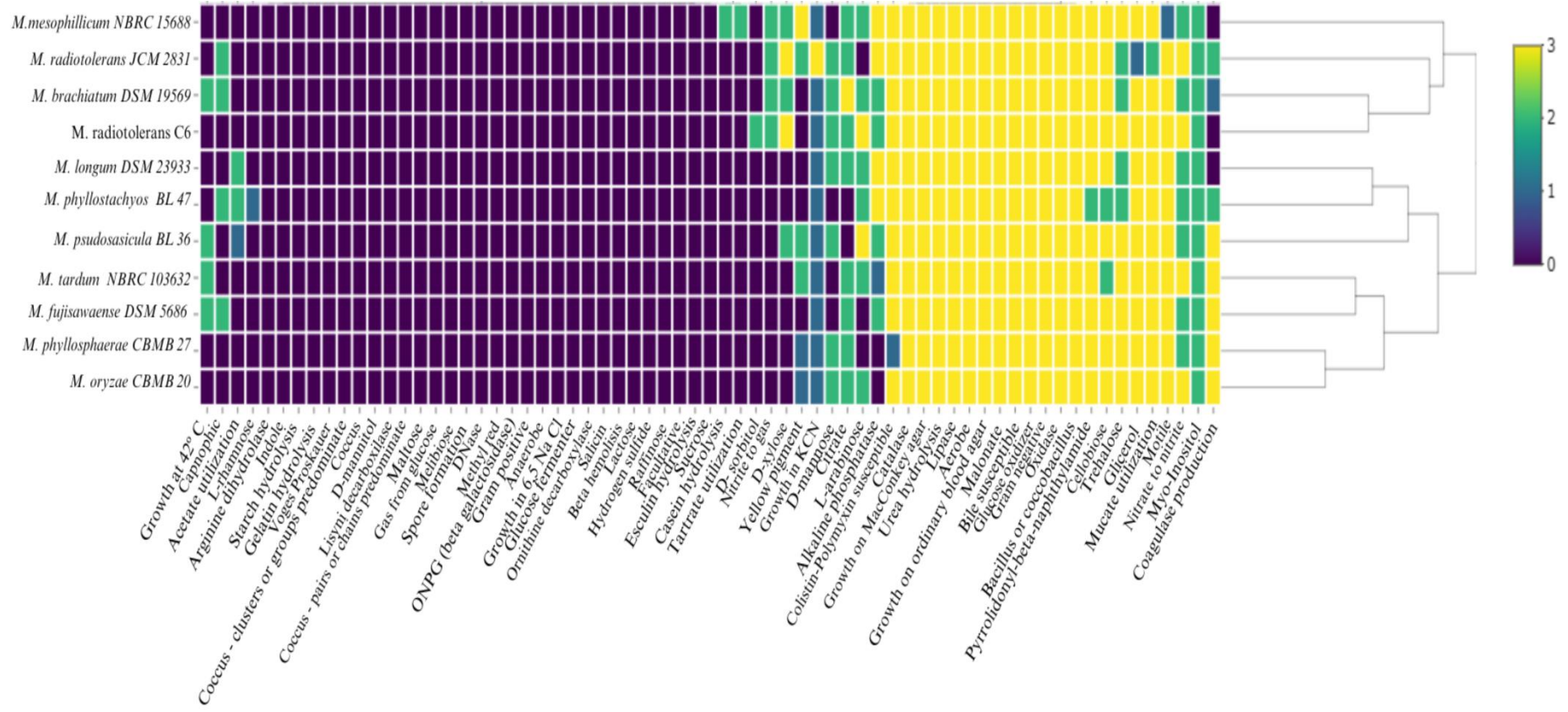


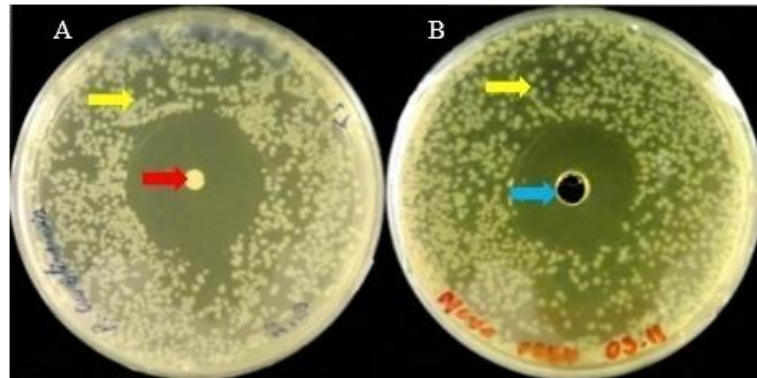
Figure 7



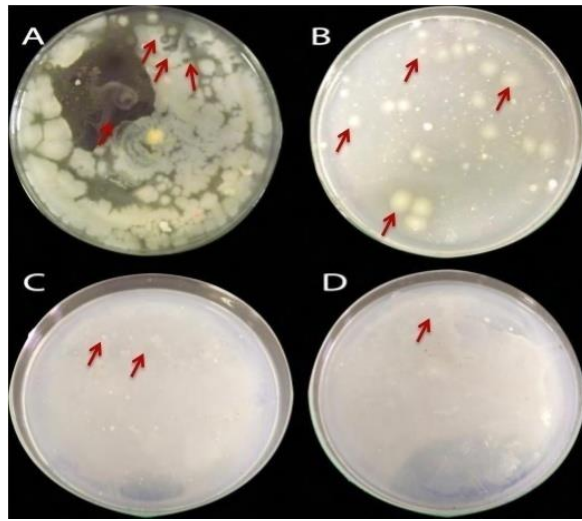
Supplementary Table 1

Strain	Inhibition halo (cm)	Strain	Inhibition halo (cm)
A1	2	B7	1
A2	3	B8	2
A3	2	B9	2
A4	3	B10	2
A5	3	B11	3
A6	2	B12	3
A7	2	C1	1
A8	3	C2	1
A9	3	C3	1
A10	3	C4	1
A11	3	C5	1
A12	3	K003	2
B1	2	C7	1
B2	2	C8	1
B3	3	C10	3
B4	3	C11	2
B5	1	C12	2
B6	1		

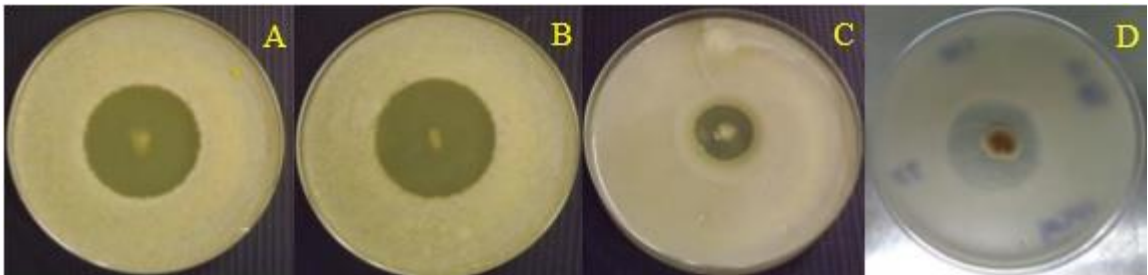
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



1 6.2 Draft do artigo do Capítulo 3

2 **Genomic and physiological characterization of *Kitasatospora* sp. nov.,**
3 **an actinobacterium with potential for biotechnological application isolated from**
4 **Cerrado soil.**

5 Cunha-Ferreira, I.C.^{1*}, Vizzotto, C.S.^{2*}, Quirino, B.F.^{3,4,5}, Freitas, M.A.M⁶, Tótola, M.R.⁷,
6 Thompson, F.L.⁶ and Krüger, R.H.¹

7
8 ¹Laboratory of Enzymology, Department of Cellular Biology, University of Brasília (UNB),
9 Brasília, Brazil.

10 ²Laboratory of Environmental Sanitation, Department of Civil and Environmental
11 Engineering, University of Brasília (UNB), Brasília, Brazil.

12 ³Genetics and Biotechnology Laboratory, Embrapa-Agroenergia, Brasília, Brazil

13 ⁴Genomic Sciences and Biotechnology Graduate Program, Catholic University of Brasília,
14 Brasília, Brazil

15 ⁵Microbial Biology Graduate Program, University of Brasília, Brasília, Brazil

16 ⁶Laboratory of Microbiology, Biology Institute, Federal University of Rio de Janeiro (UFRJ),
17 Rio de Janeiro, Brazil.

18 ⁷Laboratory of Biotechnology and Biodiversity for the Environment, Department of
19 Microbiology, Federal University of Viçosa (UFV), Viçosa, Brazil.

20
21 *Corresponding authors: kruger@unb.br

22
23 **Abstract**

24 An actinobacterial strain called *Kitasatospora* sp. K002 was isolated from Cerrado soil, a
25 well-known Brazilian biome. It was characterized and its potential for producing secondary
26 metabolites was studied. The strain is aerobic, non-motile, Gram-positive, which produces
27 grayish-white mycelium on solid culture. In liquid culture it produced submerged spores and
28 vegetative mycelia. The strain showed antibacterial potential through in vitro assays against
29 *Bacillus subtilis*, *Pseudomonas aeruginosa* and *Escherichia coli*. Based on sequence
30 similarity of the 16S *rRNA* gene, the strain showed the greatest similarities to
31 *Kitasatospora aureofaciens* (98.9%). Genome evaluation showed that the closest strain to
32 K002 is *Kitasatospora xanthocidica* JCM4862 with dDDH of 32.8-37.8% and ANI of 86.86%.
33 Pan-genome analysis identified high numbers of rare genes. A total of 60 gene clusters of 22
34 different types were detected by AntiSMASH and among them, 22 clusters showed low

35 similarity (<10%) with known compounds, which suggests the potential production of new
36 bioactive compounds. The phylogenetic analyzes and morphophysiological characteristics
37 allowed distinguishing K002 from related species, so the isolate should be recognized as a
38 new species of the genus *Kitasatospora*, for which the name *Kitasatosporabrasiliensis* sp.
39 nov. (strain type= K002) is proposed.

40

41 **Keywords:***Kitasatospora*, new specie, secondary metabolites, antimicrobial potential.

42

43 **Introduction**

44 Natural products have great potential for new drug discovery and actinobacteria are
45 among the most prolific producers of bioactive natural compounds [1]. It can be said that the
46 actinomycetes (Phylum Actinobacteria), the group to which *Kitasatospora* sp. belongs, are a
47 special group among prokaryotes for their ability to produce a broad spectrum of essential
48 chemicals and potentially useful secondary metabolites that are active even against eukaryotic
49 organisms. [2- 3].

50 The propensity to produce bioactive molecules and the richness of bacterial genomes
51 in terms of diversity of biosynthetic gene clusters were correlated with organism
52 characteristics such as complex morphology and spore formation, characteristics of
53 Actinobacteria. Furthermore, thanks to their complex metabolic regulation and ability to use
54 various nutrient sources that allow them to adapt to diverse environments and growth
55 conditions, it is not surprising that Actinobacteria are the most important producers of natural
56 compounds [1].

57 Within this phylum, Streptomycetaceae is the most diverse family in terms of potential
58 for biotechnological applications [1]. Such a family includes the genera: *Allostreptomyces*,
59 *Embleya*, *Streptacidiphilus*, *Streptomyces*, *Yinghuangia* and *Kitasatospora* [4].

60 The genus *Kitasatospora* was considered controversial for many years [5-7].
61 *Kitasatospora* was proposed by Omura et al. in 1982 for actinomycete strains that were
62 phenotypically similar to *Streptomyces* strains but contained large amounts of the mesoisomers
63 of diaminopimelic acid (DAP) and galactose in whole cell hydrolysates [5]. However,
64 Wellington et al. [6], based on phenotypic properties shared by the species of *Kitasatospora*
65 and *Streptomyces*, proposed that *Kitasatospora* be reduced to synonymy with the name
66 *Streptomyces*. However, the unification of these two genres was mistaken [8]. The
67 classification of *Kitasatospora* was resolved by applying 16S *rRNA* sequencing [9] and whole
68 genome sequencing [4,8].

69 The genus *Kitasatospora* is also distinguishable from *Streptomyces* on the basis of cell
70 wall composition [10] and by resistance to *Streptomyces* polyvalent phages [6] and
71 formation of submerged spores in liquid culture [2,5], characteristics rarely observed in
72 *Streptomyces* species.

73 So, *Kitasatospora* is a legitimate genus distinct from *Streptomyces*, based on
74 phenotypic and genetic differences. It is a genus of aerobic, Gram-positive, non-alcohol-acid
75 resistant actinobacteria that form intensely branched and non-fragmented mycelia [4].
76 Currently, the genus contains 41 validly described species
77 (<https://lpsn.dsmz.de/genus/kitasatospora> on 01/18/2023), whose type species is
78 *Kitasatospora setae*.

79 So far, at least 50 bioactive compounds have been discovered from strains of
80 *Kitasatospora*, propioxatins (enkephalinase B inhibitor), terpentecin (antitumor activity) [2],
81 setamicin (nematocidal and antifungal activity) and fosalacin (herbicide activity) [11].
82 Endophenase antibiotics were also discovered in *Kitasatospora sp.* MBT66 [12].

83 The genus *Kitasatospora* is not an easy group to isolate like the neighboring genus
84 *Streptomyces*, and therefore its biosynthetic potential has not yet been widely recognized [11].
85 Studies suggest that *Kitasatospora* is a valuable resource for natural product discovery [8]. In
86 this work, we describe the genomic and physiological characteristics of a strain of
87 *Kitasatospora* isolated from the soil of a Brazilian biome, known as Cerrado, and propose its
88 description as a new species of the genus. The present work should expand the general
89 knowledge about this taxon.

90

91 **Material and methods**

92 **Sampling and isolation**

93 Soil samples were collected using an aluminum pipe with a 15 cm depth indication in
94 a native area of the Cerrado biome, Brasília, Federal District, Brazil. The pipe was introduced
95 with the aid of a rubber mallet into the ground until reaching the 15 cm mark. Environmental
96 soil samples were previously dissolved in Middlebrook medium (Casein enzymatic
97 hydrolysate, 1(g/L); Ammonium sulfate, 0.5(g/L); Monopotassium phosphate, 1.5(g/L);
98 Disodium phosphate, 1.5(g/L); Sodium citrate, 0.4(g/L); Magnesium sulfate, 0.05(g/L); L-
99 Glutamic acid, 0.5(g/L); Ferric ammonium citrate, 0.04(g/L); Pyridoxine, 0.001(g/L); Biotin,
100 0.0005(g/L); Malachite green, 0.001(g/L); Agar, 15.0(g/L) liquid in the ratio of 1g of soil in
101 50 ml of medium and then diluted in 1:20 and inoculated in plates of culture medium. After
102 one week, the cultures were transferred to R2A medium (agar, 1.9 g/L; 15 g/L; casein acid

103 hydrolysate, 0.5 g/L; dextrose, 0.5 g/L; dipotassium phosphate, 0.3 g/L; magnesium sulfate,
104 0.024 g/L; proteose peptone, 0.5 g/L; sodium pyruvate, 0.3 g/L; starch soluble, 0.5 g/L; yeast
105 extract, 0.5 g/L), using the standard scratched plate technique to obtain pure cultures,
106 including the K002 strain studied here.

107

108 In vitro phenotypic characterization

109 Cells of the K002 strain were observed under optical microscopy to determine the
110 characteristics of the membrane, using differential Gram staining [13]. Observation was
111 performed using a Leica DM750 microscope.

112 For visualization in scanning electronic microscopy, the cells were obtained by
113 centrifugation of the culture in NB liquid medium (steak extract, yeast extract, peptone and
114 sodium chloride) and washed 3 times in phosphate saline buffer (PBS). Fixation was in
115 Karnovsky solution (2% (v/v) (paraformaldehyde, 2% (v/v) glutaraldehyde, 3% (w/v) sucrose
116 and 0.05 M cacodylate buffer). The second fixation was in solution of osmium tetroxide (1%
117 w/v) followed by adhesion on coverslips with poly-L-lysine. The cells were then dehydrated in
118 acetone solution with increasing concentration (50/70/90/100%, v/v). Critical-point drying
119 was performed with CO₂ and the cells were covered with a thin layer of gold [14]. The images
120 were generated using a JEOL JSM-7001F scanning electron microscope (JEOL Ltd., Tokyo,
121 Japan).

122 For transmission electron microscopy, the cells were washed and fixed as in scanning
123 microscopy, and then in block staining was performed in an aqueous solution of uranyl
124 acetate 0.5% for 2h, dehydrated in acetone series (30-100%) and embedded in Spurr's resin.
125 Ultrathin sections were stained with uranyl acetate and subsequently examined and
126 photographed in a Jeol 1011 Transmission Electron Microscope (JEOL Ltd., Tokyo, Japan)
127 [14].

128 The optimum growth temperature was evaluated for one week on NB agar plates at
129 temperatures of 5, 10, 15, 20, 30, 35, 40., 45, 50, 55 and 60 °C. The optimum pH range for
130 growth was examined between 4 and 10 in intervals of 1 after incubation in NB at 30°C for 3
131 days in liquid medium adjusted with HCl or NaOH containing citrate buffer/Na₂HPO₄ 50 mM
132 (pH 4–7), 100 mM NaH₂PO₄/Na₂HPO₄ buffer (pH 7–8) or 100 mM NaHCO₃/Na₂CO₃ buffer
133 (pH 9–10). Salt tolerance was tested in NB in liquid medium supplemented with NaCl (0-5%
134 in 1% intervals) and incubation at 30°C for 3 days. Both salinity and pH were examined by
135 measuring the OD600 in SpectraMax M3 with SoftMax Pro 6.0 software [15].

136

137 Physiological characterization

138 Catalase activity was observed by the formation of bubbles after application of a 3%
139 (v/v) hydrogen peroxide solution [16].

140 K002 cells grown in solid R2A medium were collected after 48h of growth,
141 transferred to glass tubes with teflon lids and then used for cellular fatty acid analysis, as
142 described by Ramaprasad et al. [17]. Fatty acids (FAME) obtained were analyzed on a gas
143 chromatograph (Agilent 7890^a) equipped with the Sherlock Microbial Identification System
144 software (Sherlock, v. 6.2) using the RTSBA6 method library [18].

145 The susceptibility to antibiotics was evaluated in solid medium by the Kirby-Bauer
146 method (disk diffusion) using filter paper discs impregnated with a standard concentration of
147 antibiotic and observation of halo formation after application of the disc on a culture spread
148 with a swab and inoculated for 48h at 30° C. The antibiotics tested were (µg per disc):
149 Amikacin (30), Amoxicillin/Clavulanic acid (20/10), Ampicillin (10), Aztreonam (30),
150 Cephalexin (30), Cefazolin (30), Cefepime (30), Ceftazidime (30), Ceftriaxone (30),
151 Ciprofloxacin (5), Clindamycin (2), Chloramphenicol (30), Doxycycline (30), Erythromycin
152 (15), Gentamicin (10), Imipenem (10), Meropenem (10), Norfloxacin (10), Ofloxacin (5),
153 Oxacillin (1), Penicillin G (10), Piperacillin + Tazobactam (110), Polymyxin B (300),
154 Sulfonamide (300), Tetracycline (30) and Vancomycin (30).

155

156 Antimicrobial activity.

157 An overlay test was performed to prove the antibacterial capacity of K002. The tests
158 were done in minimal culture media (R2A or Middlebrook). The isolate was inoculated in the
159 center of the plates and placed in growth at room temperature for a period of 6 days. On the
160 seventh day of growth, soft-agar inoculums were prepared with the Gram-positive strain
161 *Bacillus subtilis* and with the Gram-negative *Escherichia coli* and *Pseudomonas aeruginosa*
162 which were poured separately on plates containing the isolate, in order to form a thin and
163 homogeneous layer. This was the cover or overlay. After the addition of soft-agar, it was
164 expected to grow for up to 48 hours at room temperature.

165

166 DNA extraction, sequencing and genome assembly

167 DNA extraction was performed with the commercial 'GenElute Bacterial Genomic
168 DNA Kit',
169 Sigma-Aldrich, both for the analysis of the 16S *rRNA* marker and for genomic sequencing.
170 The extraction efficiency was evaluated in agarose gel (1% v/v). DNA quantification was

171 performed with a Qubit fluorometer (Invitrogen, USA) using the dsDNA BR Quant-it assay
172 kit (Invitrogen).

173 Polymerase chain reaction (PCR) (DNA) for amplification of 16S *rRNA* was
174 performed with primers 27F and 1492R. Purification of the PCR product was performed with
175 the Thermo Scientific purification kit (Gene Jet PCR Purification Kit). Bidirectional
176 sequencing was performed on a Sanger platform by Macrogen (Seoul, South Korea).
177 Comparison by similarity with sequences available in the GenBank of the NCBI (National
178 Center for Biotechnology Information) was performed using the BLAST N tool.

179 Complete sequencing of the NGS Illumina genome of K002 was performed at
180 Macrogen (Seoul, South Korea). The DNA library was prepared using the TruSeq Nano DNA
181 Kit and sequencing was performed on the NextSeq500 platform, generating paired-end reads
182 of 150 base pairs.

183 In order to improve genome assembly, NGS sequencing was performed using long
184 reads technology on the Oxford Nanopore Technologies (ONT) platform (Oxford, UK). The
185 DNA library was prepared with the Rapid Sequencing Kit (SQK-RAD004) and the
186 sequencing was performed in Flow cell FLO-MINSP6 in the MinION device (MIN-101B).

187 The de novo assembly of the genome was performed in a hybrid way with the
188 Unicycler program (V. 0.5.0) [19]. The CheckM software, through the Kbase platform, was
189 used to determine contaminating sequences and genome completeness [20]. The coverage of
190 the Illumina reads was determined using the Bowtie2 software (v. 2.4.4) [21] and the L50 and
191 N50 parameters were obtained using Quast [22]. Artemis software was used to visualize the
192 circular genome and its coding regions through the DNAPlotter resource [23]. Genome
193 annotation was performed using Prokka software (v. 1.14.6) [24].

194

195 Phylogenetic analyzes

196 The evolutionary position of the strain was estimated by constructing a phylogenetic
197 tree using the Maximum-likelihood (ML) method with 1000 sampling repetitions (Bootstrap),
198 using representative sequences for 16S *rRNA* of valid species of the genus of the strain listed
199 in the List of Prokaryotic names with Standing in Nomenclature - LPSN and deposited in the
200 GenBank/EMBL/DDBJ database (Felsenstein, 1985, Kimura, 1980). For sequence alignment
201 and construction of the phylogenetic tree based on the partial sequence of the 16S *rRNA*
202 molecular marker gene, the Mega 11 program was used (available for download at
203 <https://www.megasoftware.net/>). Clustal tool W was used to perform the alignment [25].

204 The taxonomic classification of the strain and identity confirmation were performed on
205 the TYGS – DSMZ platform, by digital DNA hybridization (DDH) on the GGDC – DSMZ
206 server [26] The ANI (Average Nucleotide Identity) was calculated on the JSpecieWS
207 platform [27].

208 The alignment of the genome of K002 and the genome of its closest strain according
209 to TYGS-DSMZ, *K. xanthocidica* JCM 4862, was performed with Plot Nucmer (available at:
210 https://github.com/fmalmeida/plot_nucmer), which is an automation of MUMmer v3 .1
211 toolkit [28] and Ggbio [29], where the former aligns filtering by length and identity, and the
212 latter generates a circular view.

213 A phylogenetic tree based on concatenated central orthologous genes from genomic
214 sequences of twelve strains, including the K002 strain, the closely related strains according to
215 the TYGS-DSMZ, and the *Streptomyces violaceorubidus* LMG 20319 outgroup available at
216 the NCBI was constructed using the M1CR0B1AL1Z3R web server (<https://microbializer.tau.ac.il/>) [30].

218 Comparative pan-genome analysis was performed using Roary version 3.13.0
219 following Prokka annotation, version 1.14.6 [24], available from Galaxy [31]. Complete
220 genome sequences of K002 and its 10 closest sequences were used, according to TYGS-
221 DSMZ. The minimum percentage identity for blastp of 80% [32] was defined to group the
222 genes that encode complete sequences of proteins in central genomes.

223

224 Functional annotation and Gene prediction

225 Functional categories based on the COG classification (Clusters of Orthologous
226 Groups) were obtained using the online version of eggNOG-mapper v.2 [33], where
227 metabolic pathways and genes of interest were compared to the database of Kyoto
228 Encyclopedia of Genes and Genomes (KEGG) data [34].

229 The Antibiotic Resistance Gene Database (ARDB) was used for annotation of
230 potential genes responsible for antibiotic resistance [35]. The CRISPR (clustered regularly
231 interspaced short palindromic repeats) loci were predicted using the CRISPRCasFinder
232 recognition tool [36].

233 Three online software were used in the in silico analysis of the biotechnological
234 potential: AntiSMASH (v. 6.1.1) [37] to identify genes and gene clusters related to secondary
235 metabolism, BAGEL4 to detect the presence of antimicrobial peptides synthesized
236 ribosomally and modified post -translationally (RiPPs) [38] and CAMPR4 for predicting
237 antimicrobial peptides (AMPs) [39], where the sequence is predicted to be antimicrobial

238 through a probability score (0 to 1). The higher the probability, the greater the possibility that
239 the peptide is antimicrobial [40].

240

241 In silico phenotypic characterization

242 The in silico phenotyping tool Traitair (Microbial Trait Analyzer) was used to predict
243 phenotypes based on K002 genome sequences in a comparative way with its closest strains
244 according to TYGS-DSMZ. This software is able to predict 67 phenotypic traits using
245 Prodigal for gene prediction and Pfam family for annotation. The software uses two
246 prediction models, the phyfat model (which predicts the presence/absence of proteins found
247 in the phenotype of 234 bacterial species) and a combination of phyfat+PGL models (which
248 uses the phyfat information combined with the acquisition or loss of protein families and
249 phenotypes through evolution), to determine phenotypic traits [41].

250

251 **Results**

252 In vitro phenotypic characterization

253 Colonies of K002 on solid medium showed grayish-white color, irregular shape and
254 leathery appearance. K002 cells were Gram positive. In electronic microscopy, vegetative
255 mycelia and submerged spores were observed from the culture in liquid medium. Submerged
256 spores sprout and elongate forming vegetative mycelia without any fragmentation (Fig. 1). No
257 aerial mycelium and spore chains were observed (Fig. 1).

258 The strain grow up aerobically at temperatures of 15–35 °C in NB medium containing
259 0-3% (w/v) NaCl at pH 7–9.0, and optimal growth occurred at 30 °C, pH 7.0 and with 0%
260 (w/v) NaCl. The differential phenotypic characteristics of the K002 strain and other species of
261 the genus [5, 42-43] are summarized in Table 1. Optimal growth at pH 8 distinguishes K002
262 from the others.

263 The fatty acid profile of strain K002 is shown in Table 2 in comparison to other
264 species of the genus [42]. Among the most abundant fatty acids (>15%) are anteiso-C15:0
265 (27.07%) and iso-C16:0 (18.44%), whereas the least abundant were C18:0 (0.52%) and
266 Summed Feature 8 (0.57%). The fatty acid profiles clearly separated the K002 strain from the
267 rest due to the iso-C18:0, C18:0 and Summed Feature 8 components.

268 Among the twenty-six antibiotics tested, K002 was resistant to seven. They are:
269 Amoxicillin/Clavulanic acid, Aztreonam, Cefepime, Ceftazidime, Meropenem, Oxacillin and
270 Sulfonamide. Table 3 shows the list of antibiotics tested, susceptibility and measurements of
271 the zone of inhibition, when it occurred.

272 The K002 strain was submitted to an overlay assay of antibacterial activity against
273 Gram-positive and Gram-negative prokaryotes, which all showed positive results. They are: *B.*
274 *subtilis*, *E. coli* and *P. aeruginosa*.

275

276 Sequencing and phylogenetic analysis

277 Sanger sequencing of the 16S *rRNA* of isolate K002 allowed the identification of the
278 sequence, where the blast indicated greater similarity (99.24%) with *Kitasatospora kinufensis*
279 JCM9081. The phylogenetic tree based on the 16S *rRNA* of K002 and the *Streptomyces*
280 *violaceorubidus* LMG 20319 outgroup is represented in Figure 2.

281 The genome sequence of the K002 strain was assembled using a combination of long
282 and short reading sequencing approaches, resulting in an assembly totaling 7,951,960 bp,
283 distributed in 14 contigs, with a GC Content of 72.8% (Tab.4 and Fig. 3). On the TYGS –
284 DSMZ platform, genome evaluation showed that the closest strain to K002 is *Kitasatospora*
285 *xanthocidica* JCM4862 with dDDH of 32.8-37.8%. Completion of the platforms automated
286 protocol indicates that K002 is potentially a new species. The ANI (Average Nucleotide
287 Identity) calculated for the K002 genome on the JspeciesWS platform was 86.86%.

288 The alignment of the genomes of K002 and its closest strain *K. xanthocidica* JCM4862
289 shows a considerable number of misaligned regions (53.45%) indicating that the genomes are
290 divergent from each other (Fig. 4). Only genome alignment blocks with at least 1,000
291 nucleotides and 90% identity were considered.

292 Expanding comparative genomic analysis to investigate the phylogenetic relationships
293 of this genus, we used the genome sequences of 12 strains (strain K002, 10 closest species
294 according to TYGS-DSMZ and the *Streptomyces violaceorubidus* LMG 20319 outgroup) to
295 extract sets of orthologous genes. Using the M1CR0B1AL1Z3R web server, a total of 520
296 gene cores (subset of genes shared among the 12 strains) were retrieved and used to construct
297 a maximum likelihood phylogenetic tree (Fig. 5).

298 It was possible to identify 36568 gene families that make up the Pan-genome of K002
299 and its closest strains and, of these, 1383 (3.7%) were present in the nucleus of the genome
300 and the rest as accessories. Rare genes were represented by 23308 families, that is, 63.73% of
301 the pan-genome (Fig. 6).

302

303 Functional features of genomes

304 The figure 7 shows the distribution of proteins in the genome of the K002 isolate in
305 their respective COG categories. Category S refers to proteins of unknown function.

306 Approximately 17% of the proteins in the K002 genome have been assigned to this group.
307 The unknow category (13%) represents proteins that were not annotated in any of the COG
308 categories. Among the well-annotated COG categories in the K002 genome, categories K
309 (11%), E (7%) and Q (5%), referring to transcription, amino acid metabolism and
310 biosynthesis/catabolism of secondary metabolites, respectively, are among the most abundant
311 in the genome.

312 With regard to antibiotic resistance, a helicase-like RNA polymerase (81.32%
313 similarity) and a rifampin glycosyltransferase (77.7% similarity) were found in K002, both
314 conferring resistance to rifamycin. Resistance mechanisms are antibiotic target protection and
315 antibiotic inactivation, respectively. An antibiotic efflux pump was also detected with 44.23%
316 similarity. Only 1 CRISPR-cas loci complete with 71 spacers and a Type IE cas protein were
317 found in K002. However, 8 CRISPR systems were found without associated cas proteins and
318 with spacers ranging in number from 1 to 22.

319 A total of sixty (60) gene clusters of twenty-two (22) different types were detected by
320 AntiSMASH for isolate K002. Among them, NRPS stands out, the most abundant (23),
321 lantipeptides of 4 different classes, PKS of the 3 types, among others. The table 5 presents
322 such clusters in comparison to those present in *K. acidiphila* [42], where 43 gene clusters of
323 11 different types were detected, and *Kitasatospora sp.* MMS16-BH015 [11], where 38 gene
324 clusters of 14 different types were detected. Altogether 141 gene clusters were identified in
325 the 3 genomes of *Kitasatospora*. While there are specific clusters, all have NRPS, terpene and
326 at least one type of lanpeptide and PKS.

327 The CAMPR4 database for predicting antimicrobial peptides (AMPs) showed fifty-
328 nine (59) predictions, in K002. In it, the sequence is predicted to be antimicrobial through a
329 probability score (0 to 1) [40]. The higher the probability, the greater the possibility that the
330 peptide is antimicrobial. K002's predictions had a probability greater than 0.68, with a higher
331 probability of 0.78. In the K002 genome, eight (8) potential areas linked to the production of
332 ribosally synthesized bacteriocins were found, they are: a gene cluster linked to the
333 production of zoocin A (bacteriocin class III), a gene cluster linked to the production of linear
334 azole-containing peptides (LAPs) (class I bacteriocins), a putative gene cluster related to the
335 biosynthesis of bacteriocins of unidentified class, three gene clusters related to the
336 biosynthesis of lantibiotics (class I bacteriocins) and two gene clusters associated with the
337 production of sactipeptides (class I bacteriocins) (Tab. 6).

338 In silico phenotypic characterization

339 The figure 8 shows the comparison of the phenotype of K002 and its closest strains
340 according to DSMZ. The K002 strain differs from the others by being negative for glucose
341 fermentation. K002 shares many characteristics with *K. xanthocidica* JCM862, its closest
342 strain according to the DSMZ, however, K002 differs in that it is negative for yellow pigment,
343 negative for growth in DNase, tartrate, glucose fermentation, cellobiose, l-arabinose, l-
344 rhamnose, melibiose, -nitrophenyl- β -D-galactopyranoside (β -galactosidase), raffinose and
345 salicin, and positive for arginine dihydrolase.

346 According to the phenotypic prediction in Traitar, the closest strains to K002 are
347 *K.aureofaciens* JCM4434 and NRRLB3447. The K002 strain differs from the others for being
348 positive for oxidase, d-xylose, d-Mannitol and indole, and for being negative for yellow
349 pigment, growth in tartrate and glucose fermentation. K002 differs only from NRRLB3447 in
350 that it does not grow on DNase and that it grows on myo-inositol. And it differs only from
351 JCM4434 by oxidizing glucose. The table 7 shows the phenotypic prediction of K002 in
352 Traitar compared to the traits of other strains of the genus [42]. The K002 strain differs from
353 the others by being positive for Arginine dihydrolase, indole production and utilization of
354 citrate and maltose.

355

356 Discussion

357 The Streptomycetaceae family is characterized by an intricate mycelial life cycle and
358 marked characteristics of multicellularity [44-45]. And unlike most Streptomycetes, the genus
359 *Kitasatospora* produces submerged spores with mycelia in liquid culture seen in micrographs
360 of K002. The varied morphology in micrographs of *Kitasatospora sp.* as observed in K002, it
361 has already been questioned, raising doubts about whether there was contamination or not, but
362 finally the variety of cell types found in liquid cultures of this genus was confirmed [2].

363 Gas chromatographic fatty acid analysis has been routinely used to identify bacteria
364 based on their unique fatty acid profiles. More than 300 fatty acids and related compounds
365 have been found in bacteria analyzed at the MIDI Research and Development Laboratory.
366 The wealth of information contained in these compounds can be estimated considering not
367 only the presence or absence of each acid, but also using the data quantitatively [18]. The
368 fatty acid profile of K002 agree with what was reported by Kämpfer [46] and verified in Kim
369 et al. [42] in *K. acidiphila* and by Klaysubun et al. [43] in *K. humi*, where iso and anteiso
370 predominate in strains of the genus *Kitasatospora*.

371 As for antibacterial activity, our results agree with the findings of Srivastava *et al.* [3],
372 where *Kitastospora* extracts showed antibacterial activity against *B. subtilis* and *E.coli*.

373 Although analysis of the K002 16S *rRNA* gene through the blast indicated great
374 similarity (99.24%) with *K. kifunensis*, genomic analysis indicated greater similarity with *K.*
375 *xanthocidica*. The values of ANI < 95% and dDDH < 70% calculated for K002, according to
376 the genomic parameters OGRI (Overall Genome Related Index) for the delineation of new
377 species, indicate that it is a different species of the genus *Kitasatospora* [47].

378 According to the study of 45 strains of *Kitasatospora* carried out by Li et al. [8], the
379 size of *Kitasatospora* genomes vary in size from 6.3 to 12.36 Mbp, with an average size of 8.7
380 ± 1.0 Mbp. This shows that K002 is not among the largest genomes of the genus (7.9 Mbp).
381 The study also shows that the GC content of the gender varies between 71.2 and 74.5% with
382 an average of 73.1% [8], so K002 corresponds to the average with 72.8%. The same work
383 shows the number of coding sequences per genome of *Kitasatospora* varies between 5,482
384 and 9,863 and K002 has 6,708.

385 In the pangenomic study also carried out by Li et al. [8], the core genome of this genus
386 was estimated at 1476 gene families and the similarity with the values of K002 (1383) suggest
387 that K002 retains the essential components of the core genome of *Kitasatospora*. Li et al. [8]
388 in their analysis found that the central genome formed 2.7% of families, while K002 has 3.7%
389 of families in its central genome and that rare genes represented 70.5% of the pangenome,
390 while in K002 it represents 63.73%. These results indicate that the gene content of the genus
391 is highly divergent.

392 As well as in *Kitasatospora sp.* MMS16-BH015, *K. griseola*, *K. mediocidica*, *K. setae*
393 and *Kitasatospora sp.* OK780, where more than 20% of unknown proteins were reported [11]
394 classified in categories S and Unknow, in K002 both fractions of categories add up to 30%
395 and point out one of the limitations of annotation of genomes with reference in databases,
396 they prospects are based only on structures already described.

397 Overall, all *Kitasatospora* genomes exhibit similar COG profiles for known proteins,
398 although some subtle differences were observed. In comparison with the genome of
399 *Kitasatospora sp.* MMS16-BH015, categories K and E are also reported to be more abundant
400 as well as in K002, however the third most abundant category in *Kitasatospora sp.* MMS16-
401 BH015 is the G for carbohydrate transport and metabolism. The abundance of category Q
402 (third most abundant in K002) agrees with what was observed in *Streptomyces* genomes [48-
403 50] genus closely related to *Kitasatospora*, where secondary metabolism represents a
404 significant portion of the genome. It is noteworthy that category T, the fifth most abundant in
405 K002 (4.76%) related to signal transduction, is known for its regulatory roles in the
406 biosynthesis of antibiotics in actinomycetes [51].

407 A CRISPR-Cas locus consists of clustered arrays of regularly interspaced short
408 palindromic repeats (CRISPR) and adjacent CRISPR associated proteins (Cas) [52]. The
409 strain K002 encoded a complete CRISPR-Cas loci. Prokaryotic cells use the CRISPR-Cas
410 system as an adaptive immune system to protect themselves from viruses and other potentially
411 harmful mobile genetic elements [53-54]. The enthusiasm for the CRISPR-Cas system also
412 lies in its DNA editing power and its discovery favors genetic engineering in plants [55] and,
413 due to its programming and simplicity, it offers a revolutionary approach to creating new
414 treatments for a wide range of genetic diseases [56].

415 Although *Streptomyces* is the model of bacterial physiology and an important resource
416 for the discovery of antibiotics, its close relatives such as *Kitasatospora* are still poorly studied
417 [8]. Several new chemicals have been discovered from strains of this genus, including
418 compounds with unique structures and interesting activities [42]. *Kitasatospora sp.* MBT66
419 has been identified as a prolific producer of antibiotics with activity against Gram-positive
420 and Gram-negative pathogens [12]. *Kitasatospora* species have also been characterized as
421 producing new secondary metabolites of varied structures, which are mainly active against
422 eukaryotic organisms [2].

423 Among the sixty gene clusters identified in K002, twenty-two clusters showed low
424 similarity (<10%) with known compounds, which suggests the potential production of new
425 bioactive compounds. They are: T1PKS (1), T3PKS (1), butyrolactone (2), NRPS (7), LAP
426 (1), Thiopeptide (1), RIPP (2), siderophore (1), NPAA (1), terpene (3), class IV lantipeptide
427 (1) and beta-lactam (1). The average for *Kitasatospora* reference sequences is 54 clusters and
428 illustrates that K002 is above average, even though it is not among the largest genomes of the
429 genus [57].

430 According to Li et al. [8], *Kitasatospora* biosynthetic clusters can be grouped into 8
431 main biosynthetic classes: RiPPs, NRPS, terpene, PKS-NRP Hybrids, “PKS other”, type I
432 PKS, saccharide and unclassified. Except for saccharides, all clusters are present in K002.
433 Compared to *Streptomyces*, *Kitasatospora* is significantly richer in RiPPs, NRPS and “other
434 PKS”, but poorer in unclassifiables” [8].

435 The presence of NRPS and PKS in an Actinobacterium is highly related to its
436 biosynthetic potential [58]. The PKS and NRPS clusters, both found in K002, represent two
437 large groups of natural products with remarkable diversity and biological activities such as the
438 antibiotics erythromycin, penicillin, daptomycin and vancomycin, the antifungals
439 amphotericin and echinocandin, the anticancer agents epothilone and bleomycin, the agent

440 cholesterol-lowering drug lovastatin, the immunosuppressants FK506 and cyclosporine, and
441 the veterinary antibiotics monensin and avermectin [59].

442 PKS are classified into Types I, II and III, all found within the K002 clusters[60].
443 Within type I PKS there is a group of products called ionophore polyethers that cause the
444 death of target cells, being of high importance for their antiparasitic activity and broad
445 spectrum antifungal, antiviral, herbicide, anti-inflammatory and immunoregulatory action
446 [59]. In the genus *Streptomyces*, belonging to the same family as *Kitasatospora* (K002), a
447 biosynthetic pathway of PKS types II and III is identified for the production of a pigment
448 characteristic of spores [61].

449 While non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS)
450 biosynthetic pathways have been well documented in actinomycetes, there are many clusters
451 of biosynthetic genes that need to be unraveled to induce the production of potentially novel
452 metabolites [62].

453 Genome sequencing efforts of the first decade of the 21st century have revealed that a
454 large class of natural products is formed by ribosomally synthesized and post-translationally
455 modified peptides (RiPP). It is a class of natural products with antimicrobial activity, whose
456 cluster is found in K002. Post-translational modifications diversify the structure of these
457 natural products, which implies better target recognition and increased chemical and
458 metabolic stability, improving their chemical functionality. They are categorized into about 20
459 different classes such as LAPs, lantipeptides and thiopeptides, whose specific clusters are also
460 present in K002 [63]. LAPs constitute a large class of RiPPs and to date, 17 LAPs are known,
461 most of which have cytolytic activity (eg streptolysin S, clostridiolysin S, listeriolysin S) or
462 antibacterial activities [64].

463 The Betalactone cluster, present in K002, refers to a protease inhibitor containing β -
464 lactone. β -lactones are structures found in many natural products. They often have anticancer
465 and/or antibiotic properties. They are a rich source for new drug research, but they are often
466 overused each by β -lactams, also present in K002 (Blactam), so their therapeutic potential
467 remains underexplored in relation to β -lactams, which comprise more than 50% of antibiotics
468 on the market today [65].

469 The Furan cluster, found in K002, refers to the furan ring, a biologically active
470 substance with high therapeutic properties, including: antibacterial, antiviral, anti-
471 inflammatory, antifungal, antitumor, antihyperglycemic, analgesic, anticonvulsant, etc. [66-
472 67].

473 Among other properties, siderophores, whose cluster is also present in K002, have
474 been related to virulence mechanisms in microorganisms. In addition, they have
475 biotechnological applications, being used as soil fertilizers and biocontrol agents, treatment of
476 anemia and malaria, in addition to being used in the selective administration of antibiotics in
477 resistant bacteria. This is a potentially powerful application that uses the iron transport
478 abilities of siderophores to transport drugs into cells through the preparation of conjugates
479 between siderophores and antimicrobial agents [68].

480 The Butyrolactone cluster related to the production of butyrolactones may be related to
481 the antibiotic activity of K002, since signaling molecules called gamma-butyrolactones are
482 produced by species of *Streptomyces*, sister genus to *Kitasatospora* (K002), in which they
483 regulate the production of antibiotics [69]. These molecules are produced by most
484 *Streptomyces* and other actinobacteria and are also involved in the activation of secondary
485 metabolism [70].

486 Fewer terpenoid metabolites have been detected in prokaryotes compared to those
487 found in fungi and plants. The first molecules found in bacteria were odoriferous terpenes,
488 mainly geosmin, a degraded sesquiterpene, which provides the characteristic odor of damp
489 earth. In K002, terpenoid clusters are present. Terpenes of microbial origin have been used as
490 potent antimicrobial agents such as pentalenene or therapeutecin [71-73].

491 Non-proteinogenic amino acids (NPAAs) are those that do not occur naturally in the
492 genetic code of an organism, but have biological importance, as many are used as metabolic
493 intermediates or are found in proteins with post-translational modifications. NPAA is a cluster
494 of polyamino acids similar to Epsilon-poly-L-lysine (ϵ -PL) present in isolate K002. ϵ -PL is a
495 natural antimicrobial peptide with a broad spectrum of activity against food spoilage and food
496 poisoning bacteria in vitro [74-75].

497 The search for valuable antibiotics begins with screening bacteria to detect the best
498 source of new bioactive metabolites. So the functional analysis of biosynthetic genes made in
499 this work confirms the genus *Kitasatospora* as an important source of compounds with
500 potential for biotechnological application.

501

502 **Conclusions**

503 The results of the phylogenetic analyzes and the morpho-physiological characteristics
504 allow us to affirm that the K002 strain belongs to the genus *Kitasatospora*. However, it is
505 clear that strain K002 represents a new species of the genus, for which the name
506 *Kitasatosporabrasiliensis* sp. nov. is proposed.

507 Our findings indicate that the genome of the K002 strain harbors several genes
508 responsible for the biosynthetic production of secondary metabolites, confirming the potential
509 of the genus to be a valuable source of bioactive compounds, which is consistent with its
510 phenotype.

511 This work broadens the general understanding and reinforces the biotechnological
512 interest in this relatively peculiar taxon.

513

514 **Description of *Kitasatospora brasiliensis* sp. nov.** (bra.si.li.en'sis N.L. masc./fem.
515 adj. *brasiliensis* refer of or pertaining to the country from where the type strain was isolated,
516 Brazil).

517 Aerobic, non-motile, Gram-stain-positive actinobacteria that produce grayish-white
518 mycelium in R2A. Producessubmergedsporesandvegetativemycelia in liquidculture. The
519 temperature range for growth is 15–35 °C (optimum, 30 °C), the pH range is pH 7–9
520 (optimum, pH 8), and the NaCl range is 0–3% (optimum, 0%). Catalase and oxidase are
521 positive. Positive for gelatin liquefaction, indole production, arginine dihydrolase, esculin
522 hydrolysis, urease, but negative for glucose fermentation, nitrate reduction and β -
523 galactosidase. Glucose, acetate, citrate, maltose, mannitol and mannose are used as the sole
524 carbon source, but not arabinose. The most abundant fatty acids (>15%) are anteiso-C15:0
525 and iso-C16:0. The type strain is K002, which was isolated from Cerrado soil.

526

527 **Acknowledgments**

528 We would like to thank the Microscopy and Microanalysis Laboratory of the
529 University of Brasília (Brazil) for their collaboration in carrying out the analyzes using
530 scanning and transmission electron microscopy. We thank the Laboratories of the Department
531 of Microbiology (Federal University of Viçosa, Brazil) for the infrastructure to carry out part
532 of the experiments in this work.

533

534 **Funding**

535 This study was funded by the National Council for Scientific and Technological
536 Development (CNPq), the Coordination for the Improvement of Higher Education Personnel
537 (CAPES), which also provided financial support to the first author, and the Federal District
538 Research Support Foundation (FAP -DF).

539

540 **Protologue**

541 The type strain is K002, isolated from Cerrado soil, Brasília, Brazil. The K002
542 GenBank accession number for the draft genome sequence is JAUOOW000000000.
543 Bioprojetct PRJNA988803. Biosample SAMN36026885.

544

545 **Declaration of interests**

546 The authors declare that they have no known competing financial interests or personal
547 relationships that could have appeared to influence the work reported in this paper.

548

549 **Author Contributions**

550 Cunha-Ferreira, Vizzotto and Krüger designed the study. The collection of sample was
551 performed by Krüger. The tests were carried out by Cunha-Ferreira, Vizzotto and Tótola.
552 Data analyzes were carried out by Cunha-Ferreira, Vizzotto and Freitas. The manuscript was
553 prepared by Cunha-Ferreira and Vizzotto and revised by all authors. Quirino and Thompson
554 critically revised and substantially edited the manuscript. All authors read and approved the
555 manuscript. Cunha-Ferreira and Vizzotto contributed equally on this paper.

556

557 **References**

- 558 1 . van Bergeijk Da Terlouw BR Medema MH van Wezel GP (2020) Ecology and genomics
559 of Actinobacteria: new concepts for natural product discovery. *Nat Rev Microbiol* 18:546–
560 558. <https://doi.org/10.1038/s41579-020-0379-y>.
- 561 2. Takahash Y (2017) Genus *Kitasatospora*, taxonomic features and diversity of secondary
562 metabolites. *J Antibiot* 70:506–513. <https://doi.org/10.1038/ja.2017.8>
- 563 3. Srivastava N Nandi I IbeyaimaA Gupta S Sarethy IP (2019) Microbial diversity of a
564 Himalayan forest and characterization of rare actinomycetes for antimicrobial compounds. *3*
565 *Biotech.* 9(1):27. <https://doi.org/10.1007/s13205-018-1556-9>
- 566 4. Nouioui I Carro L García-López M Meier- Kolthoff JP Woyke T et al. (2018) Genome-
567 based taxonomic classification of the phylum Actinobacteria. *Front Microbiol* 9:2007.
568 <https://doi.org/10.3389/fmicb.2018.02007>
- 569 5. Omura S Takahashi Y Iwai Y Tanaka H (1982) *Kitasatospora*, a new genus of the order
570 Actinomycetales. *J Antibiot* 35:1013-1019. <https://doi.org/10.7164/antibiotics.35.1013>
- 571 6. Wellington EMH Stackebrandt E Sanders D Wolstrup J Jorgensen NO G (1992)
572 Taxonomic Status of *Kitasatospora*, and Proposed Unification with *Streptomyces* on the Basis
573 of Phenotypic and 16S *rRNA* Analysis and Emendation of *Streptomyces* Waksman and

- 574 Henrici 1943, 339AL. Int J Syst Bacteriol 42:156–160. <https://doi.org/10.1099/00207713-42->
575 1-156
- 576 7. Zhang Z Wang Y Ruan J (1997) A proposal to revive the genus *Kitasatospora*(Omura,
577 Takahashi, Iwai, and Tanaka 1982). Int J Syst Bacteriol 47(4):1048-54.
578 <https://doi.org/10.1099/00207713-47-4-1048>
- 579 8. Li Y Wang M Sun ZZ Xie BB (2021) Comparative genomic insights into the taxonomic
580 classification, diversity, and secondary metabolic potentials of *Kitasatospora*, a genus closely
581 related to *Streptomyces*. Front Microbiol 12:683814. [https://doi.org/10.3389/fmicb.2021.](https://doi.org/10.3389/fmicb.2021.683814)
582 683814
- 583 9. Labeda DP Goodfellow M Brown R Ward AC Lanoot B et al. (2012) Phylogenetic Study
584 of the Species within the Family Streptomycetaceae. Antonie Van Leeuwenhoek 101:73–104.
585 <https://doi.org/10.1007/s10482-011-9656-0>
- 586 10. Ichikawa N Oguchi A Ikeda H Ishikawa J Kitani S et al.(2010) Genome Sequence of
587 *Kitasatospora setae* NBRC 14216T: an Evolutionary Snapshot of the Family
588 Streptomycetaceae. DNA Res 17:393–406. <https://doi.org/10.1093/dnares/dsq026>
- 589 11. Yun BR Malik A Kim SB (2020) Genome based characterization of *Kitasatospora sp.*
590 MMS16-BH015, a multiple heavy metal resistant soil actinobacterium with high antimicrobial
591 potential. Gene 5:733-144379. <https://doi.org/10.1016/j.gene.2020.144379>
- 592 12. Wu C van Wezel GP Hae Choi Y (2015) Identification of novel endophenaside antibiotics
593 produced by *Kitasatospora sp.* MBT66. J Antibiot (Tokyo) 68:445–452. <https://doi.org/10.1038/ja.2015.14>
- 594
- 595 13. Reddy CA Beveridge TJ Breznak JA Marzluf GA Schmidt TM Snyder LR (2007)
596 Methods for General and Molecular Microbiology. American Society of Microbiology.
597 <https://doi.org/10.1128/9781555817497>
- 598 14. Souza W (2007) Técnicas de microscopia eletrônica aplicadas às ciências biológicas.
599 Sociedade Brasileira de Microscopia. Rio de Janeiro
- 600 15. Zhu D Xie C Huang Y Sun J Zhang W (2014) Description of *Comamonasserinivorans sp.*
601 nov., isolated from wheat straw compost. Int J Syst Evol Microbiol 64(Pt 12):4141-4146.
602 <https://doi.org/10.1099/ijms.0.066688-0>
- 603 16. Raney F Oren A (2011) Taxonomy of Prokaryotes. ISBN
- 604 17. Ramaprasad EVV Mahidhara G Sasikala C Ramana CV (2018)
605 *Rhodococcus electrodiphilus sp. nov.*, a marine electro active actinobacterium isolated from
606 coral reef. Int J Syst Evol Microbiol 68(8):2644-2649. <https://doi.org/10.1099/ijsem.0.002895>

- 607 18. Sasser M (2001) Technical Note # 101 Identification of Bacteria by Gas Chromatography
608 of Cellular Fatty Acids.
- 609 19. Wick RR Judd LM Gorrie CL Holt KE (2017) Unicycler: Resolving bacterial genome
610 assemblies from short and long sequencing reads. *PLoSComput Biol* 8:13(6)-e1005595.
611 <https://doi.org/10.1371/journal.pcbi.1005595>
- 612 20. Parks DH Imelfort M Skennerton CT Hugenholtz P Tyson GW (2015) CheckM: assessing
613 the quality of microbial genomes recovered from isolates, single cells, and metagenomes.
614 *Genome Res* 25(7):1043-55. <https://doi.org/10.1101/gr.186072.114>
- 615 21. Langmead B Wilks C Antonescu V Charles R (2019) Scaling read aligners to hundreds of
616 threads on general-purpose processors. *Bioinformatics* 35(3):421-432.
617 <https://doi.org/10.1093/bioinformatics/bty648>
- 618 22. Gurevich A Saveliev V Vyahhi N Tesler G (2013) QUASt: quality assessment tool for
619 genome assemblies. *Bioinformatics* 29(8):1072-5.
620 <https://doi.org/10.1093/bioinformatics/btt086>
- 621 23. Carver T Thomson N Bleasby A Berriman M Parkhill J (2009) DNAPlotter: circular and
622 linear interactive genome visualization. *Bioinformatics* 25(1):119-20.
623 <https://doi.org/10.1093/bioinformatics/btn578>
- 624 24. Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*
625 30(14):2068-9. <https://doi.org/10.1093/bioinformatics/btu153>
- 626 25. Larkin MA Blackshields G Brown NP Chenna R McGettigan PA (2007) Clustal W and
627 Clustal X version 2.0. *Bioinformatics* 23(21):2947-8.
628 <https://doi.org/10.1093/bioinformatics/btm404>
- 629 26. Meier-Kolthoff JP Carbasse JS Peinado-Olarte RL Göker M (2022) TYGS and LPSN: a
630 database tandem for fast and reliable genome-based classification and nomenclature of
631 prokaryotes. *Nucleic Acids Res* 50(D1):D801-D807. <https://doi.org/10.1093/nar/gkab902>
- 632 27. Richter M Rosselló-Móra R Oliver Glöckner F Peplies (2016) JSpeciesWS: a web server
633 for prokaryotic species circumscription based on pairwise genome comparison.
634 *Bioinformatics* 32(6):929-31. <https://doi.org/10.1093/bioinformatics/btv681>
- 635 28. Kurtz S Phillippy A Delcher AL Smoot M Shumway M et al. (2004) Versatile and open
636 software for comparing large genomes. *Genome Biol* 5(2):R12. [https://doi.org/10.1186/gb-](https://doi.org/10.1186/gb-2004-5-2-r12)
637 [2004-5-2-r12](https://doi.org/10.1186/gb-2004-5-2-r12)
- 638 29. Yin T Cook D Lawrence M (2012) Ggbio: an R package for extending the grammar of
639 graphics for genomic data. *Genome Biol* 13(8):R77. <https://doi.org/10.1186/gb-2012-13-8-r77>

- 640 30. Avram O Rapoport D Portugez S Pupko T (2019) M1CR0B1AL1Z3R-a user-friendly web
641 server for the analysis of large-scale microbial genomics data. *Nucl Acids Res* 47:W88–W92.
642 <https://doi.org/10.1093/nar/gkz423>
- 643 31. Afgan E Baker D VAN Den Beek M Blankenberg D Bouvier D et al. (2016) The Galaxy
644 Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2016 Update.
645 *Nucleic Acids Res* 44:W3-W10. <https://doi.org/10.1093/nar/gkw343>
- 646 32. Corretto E Antonielli L Sessitsch A Compant S Höfer C et al (2017) Complete Genome
647 Sequence of the Heavy Metal Resistant Bacterium *Agromyces aureus* AR33T and
648 Comparison with Related Actinobacteria. *Stand Genomic Sci* 12:2.
649 <https://doi.org/10.1186/s40793-016-0217-z>
- 650 33. Huerta-Cepas J Forslund K Coelho LP Szklarczyk D Jensen LJ et al. (2017) Fast Genome-
651 Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol Biol*
652 *Evol* 34(8):2115-2122. <https://doi.org/10.1093/molbev/msx148>
- 653 34. Kanehisa M Araki M Goto S Hattori M Hirakawa M et al. (2008) KEGG for linking
654 genomes to life and the environment. *Nucleic Acids Res* 36(Database issue):D480-4.
655 <https://doi.org/10.1093/nar/gkm882>
- 656 35. Liu B Pop M (2009) ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res*
657 37(Database issue):D443-7. <https://doi.org/10.1093/nar/gkn656>.
- 658 36. Couvin D Bernheim A Toffano-Nioche C Touchon M Michalik J et al. (2018)
659 CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced
660 performance and integrates search for Cas proteins. *Nucleic Acids Res* 46(W1):W246-W251.
661 <https://doi.org/10.1093/nar/gky425>
- 662 37. Blin K Shaw S Kloosterman AM Charlop-Powers Z van Wezel GP et al. (2021)
663 AntiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res*
664 49(W1):W29-W35. <https://doi.org/10.1093/nar/gkab335>.
- 665 38. van Heel AJ de Jong A Song C Viel JH Kok J Kuipers OP (2018) BAGEL4: a user-
666 friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res*
667 46(W1):W278-W281. <https://doi.org/10.1093/nar/gky383>.
- 668 39. Gawde U Chakraborty S Waghu FH Barai RS Khandekar A et al. (2023) CAMPR4: a
669 database of natural and synthetic antimicrobial peptides. *Nucleic Acids Research* 51:Database
670 issue:D377–D383. <https://doi.org/10.1093/nar/gkac933>
- 671 40. Waghu FH Barai RS Gurung P Idicula-Thomas S (2016) CAMPR3: a database on
672 sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res* 44:D1094–
673 D1097. <https://doi.org/10.1093/nar/gkv1051>

- 674 41. Weimann A Mooren K Frank J Pope PB Bremges A McHardy AC (2016) From genomes
675 to phenotypes: Traitair, the microbial trait analyzer. *mSystem*. <https://doi.org/10.1101/043315>
- 676 42. Kim MJ Roh SG Kim M K Park C Kim S Kim SB (2020). *Kitasatospora acidiphila* sp.
677 nov., isolated from pine grove soil, exhibiting antimicrobial potential. *Int J Syst Evol Microbiol*
678 70:5567–5575. <https://doi.org/10.1099/ijsem.0.004450>
- 679 43. Klaysubun C Srisuk N Duangmal K (2022) *Kitasatospora humi* sp. nov., isolated from a
680 tropical peat swamp forest soil, and proposal for the reclassification of *Kitasatospora*
681 *psammotica* as a later heterotypic synonym of *Kitasatospora aureofaciens*. *Int J Syst*
682 *Evol Microbiol* 72(5). <https://doi.org/10.1099/ijsem.0.005356>.
- 683 44. Flärth K Buttner MJ (2009) *Streptomyces* morphogenetics: dissecting differentiation in a
684 filamentous bacterium. *Nat Rev Microbiol* 7(1):36-49. <https://doi.org/10.1038/nrmicro1968>.
- 685 45. Claessen D Rozen DE Kuipers OP Søgaard-Andersen L van Wezel GP (2014) Bacterial
686 solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat Rev*
687 *Microbiol* 12(2):115-24. <https://doi.org/10.1038/nrmicro3178>.
- 688 46. Kämpfer P. (2012) Genus *Incertaesedis* I. *Kitasatospora*. *Bergey's Manual of Systematic*
689 *Bacteriology* 5:1768–1777.
- 690 47. Chun J Rainey FA (2014) Integrating genomics into the taxonomy and systematics of the
691 Bacteria and Archaea. *Int J Syst Evol Microbiol* 64:316–24.
692 <https://doi.org/10.1099/ijse.0.054171-0>
- 693 48. Bentley SD Chater KF Cerdeño-Tárraga AM Challis GL Thomson NR et al. (2002).
694 Complete Genome Sequence of the Model Actinomycete *Streptomyces coelicolor* A3(2).
695 *Nature* 417: 141–147. <https://doi.org/10.1038/417141a>
- 696 49. Wang X Zhao J Liu C Wang J Shen Y et al. (2013). *Nonomura easolani* sp. nov., an
697 actinomycete isolated from eggplant root (*Solanum melongena* L.). *Int J Syst Evol Microbiol*
698 63:2418–2423. <https://doi.org/10.1099/ijse.0.045617-0>
- 699 50. Zaburanyi N Rabyk M Ostash B Fedorenko V Luzhetskyy A (2014) Insights into
700 naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics* 5:15-97.
701 <https://doi.org/10.1186/1471-2164-15-97>.
- 702 51. Rodríguez H Rico S Díaz M Santamaría RI (2013) Two-component systems in
703 *Streptomyces*: key regulators of antibiotic complex pathways. *Microb Cell Fact* 12:127.
704 <https://doi.org/10.1186/1475-2859-12-127>
- 705 52. Koonin E Makarova K (2019) Origins and evolution of CRISPR-Cas systems. *Philos*
706 *Trans R Soc Lond B Biol Sci* 374-20180087. <https://doi.org/10.1098/rstb.2018.0087>

- 707 53. Wimmer F Beisel CL (2020) CRISPR-Cas Systems and the paradox of self-targeting
708 spacers. *Front Microbiol* 10:3078. <https://doi.org/10.3389/fmicb.2019.03078>
- 709 54. Qiu Y Wang S Chen Z Guo Y Song Y (2016) An active type I-E CRISPR-Cas system
710 identified in *Streptomyces avermitilis*. *PLoS ONE* 11:e0149533. [https://doi.org/10.1371/](https://doi.org/10.1371/journal.pone.0149533)
711 [journal.pone.0149533](https://doi.org/10.1371/journal.pone.0149533)
- 712 55. Sedeek KEM Mahas A Mahfouz M (2019) Plant Genome Engineering for Targeted
713 Improvement of Crop Traits. *Front in Plant Sci* 10:114. [https://doi.](https://doi.org/10.3389/fpls.2019.00114)
714 [org/10.3389/fpls.2019.00114](https://doi.org/10.3389/fpls.2019.00114)
- 715 56. Huang J Zhou Y Li J Lu A Liang C (2022) CRISPR/Cas systems: Delivery and
716 application in gene therapy. *Front BioengBiotechnol* 22:10-942325.
717 <https://doi.org/10.3389/fbioe.2022.942325>.
- 718 57. Zin NM Ismail A Mark DR Westrop G Schniete JK Herron PR (2021) Adaptation to
719 Endophytic Lifestyle Through Genome Reduction by *Kitasatospora* sp. SUK42. *Front*
720 *BioengBiotechnol* 12:9-740722. <https://doi.org/10.3389/fbioe.2021.740722>.
- 721 58. Amin DH Abdallah NA Abolmaaty A et al. (2020) Microbiological and molecular
722 insights on rare Actinobacteria harboring bioactive prospective. *Bull Natl Res Cent* 44:5.
723 <https://doi.org/10.1186/s42269-019-0266-8>
- 724 59. Wang H Liu N Xi L Rong X Ruan J Huang Y (2011) Genetic screening strategy for rapid
725 access to polyether ionophore producers and products in actinomycetes. *Appl Environ*
726 *Microbiol* 77(10):3433-42. <https://doi.org/10.1128/AEM.02915-10>
- 727 60. Nivina A Yuet KP Hsu J Khosla C (2019) Evolution and Diversity of Assembly-Line
728 Polyketide Synthases. *Chem Rev* 119(24):12524-12547.
729 <https://doi.org/10.1021/acs.chemrev.9b00525>.
- 730 61. Funa N Funabashi M Ohnishi Y Horinouchi S (2005) Biosynthesis of
731 hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in
732 *Streptomyces griseus*. *J Bacteriol* 187(23):8149-55. [https://doi.org/10.1128/JB.187.23.8149-](https://doi.org/10.1128/JB.187.23.8149-8155.2005)
733 [8155.2005](https://doi.org/10.1128/JB.187.23.8149-8155.2005).
- 734 62. Becerril A Alvarez S Brana AF Rico S Diaz M et al.(2018) Uncovering production of
735 specialized metabolites by *Streptomyces argillaceus*: activation of cryptic biosynthesis gene
736 clusters using nutritional and genetic approaches. *PLoS One* 13(5):e0198145.
737 <https://doi.org/10.1371/journal.pone.0198145>
- 738 63. Arnison PG Bibb MJ Bierbaum G Bowers AA Bugni TS et al. (2013) Ribosomally
739 synthesized and post-translationally modified peptide natural products: overview and

- 740 recommendations for a universal nomenclature. *Nat Prod Rep* 30(1):108-60.
741 <https://doi.org/10.1039/c2np20085f>
- 742 64. Ravin DY Watson ZL Meteleev M Ward FR Osterman IA et al. (2019) Structure of
743 ribosome-bound azole-modified peptide phazolicin rationalizes its species-specific mode of
744 bacterial translation inhibition. *Nat Commun* 8:10(1)-4563. [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-019-12589-5)
745 [019-12589-5](https://doi.org/10.1038/s41467-019-12589-5)
- 746 65. Robinson SL Christenson JK Wackett LP (2019) Biosynthesis and chemical diversity of
747 β -lactone natural products. *Nat Prod Rep* 36(3):458-475. <https://doi.org/10.1039/c8np00052b>
- 748 66. Kirilmis C Ahmedzade M Servi S Koca M KizirgilA Kazaz C (2008) Synthesis and
749 antimicrobial activity of some novel derivatives of benzofuran: part 2. The synthesis and
750 antimicrobial activity of some novel 1-(1-benzofuran-2-yl)-2-mesitylethanone derivatives.
751 *Eur J Med Chem* 43(2):300-8. <https://doi.org/10.1016/j.ejmech.2007.03.023>
- 752 67. Koca M Servi S Kirilmis C Ahmedzade M Kazaz C Ozbek B Otük G (2005) Synthesis
753 and antimicrobial activity of some novel derivatives of benzofuran: part 1. Synthesis and
754 antimicrobial activity of (benzofuran-2-yl)(3-phenyl-3-methylcyclobutyl) ketoxime
755 derivatives. *Eur J Med Chem* 40(12):1351-8. <https://doi.org/10.1016/j.ejmech.2005.07.004>
- 756 68. Saha M Sarkar S Sarkar B Sharma BK Bhattacharjee S Tribedi P (2016) Microbial
757 siderophores and their potential applications: a review. *Environ Sci Pollut Res Int* 23(5):3984-
758 99. <https://doi.org/10.1007/s11356-015-4294-0>
- 759 69. Takano E (2006) Gamma-butyrolactones: *Streptomyces* signalling molecules regulating
760 antibiotic production and differentiation. *Curr Opin Microbiol* 9(3):287-94.
761 <https://doi.org/10.1016/j.mib.2006.04.003>
- 762 70. Bibb MJ (2005) Regulation of secondary metabolism in *streptomyces*.
763 *Curr Opin Microbiol* 8(2):208-15. <https://doi.org/10.1016/j.mib.2005.02.016>
- 764 71. Cane DE Ikeda H (2012) Exploration and mining of the bacterial terpenome. *Acc Chem*
765 *Res* 45(3):463-72. <https://doi.org/10.1021/ar200198d>
- 766 72. Ikeda H Ishikawa J Hanamoto AShinose M Kikuchi H Shiba et al. (2003) Complete
767 genome sequence and comparative analysis of the industrial microorganism *Streptomyces*
768 *avermitilis*. *Nat Biotechnol* 21(5):526-31. <https://doi.org/10.1038/nbt820>
- 769 73. Yamada Y Kuzuyama T Komatsu M Shin-Ya K Omura S et al. (2015) Terpene synthases
770 are widely distributed in bacteria. *Proc Natl Acad Sci U S A* 112(3):857-62.
771 <https://doi.org/10.1073/pnas.1422108112>

772 74. Chang SS Lu WY Park SH Kang DH (2010) Control of foodborne pathogens on ready-to-
773 eat roast beef slurry by epsilon-polylysine. *Int J Food Microbiol* 141(3):236-41.
774 <https://doi.org/10.1016/j.ijfoodmicro.2010.05.021>

775 75. Ongpipattanakul C Desormeaux EK DiCaprio A van der Donk WA Mitchell DA Nair SK
776 (2022) Mechanism of Action of Ribosomally Synthesized and Post-Translationally Modified
777 Peptides. *Chem Rev* 122(18):14722-14814. <https://doi.org/10.1021/acs.chemrev.2c00210>.

778

779 **Table 1** Phenotypic properties that separate the *Kitasatospora*K002 strain from related species
780 in the genus. Strains: 1, K002; 2, *K. setalba* KM-6054; 3, *K. acidiphila* MMS16-C NU292T;
781 4, *K. azatica* KCTC 9699T; 5, *K. kifunensis* IFO 15206T; 6, *K. purpeofusca* NRRL B-1817T;
782 7, *K. nipponensis* HKI 0315T and 8, *K. humi* RB6PN24.–, no data.

783 **Table 2** Fatty acid composition (%) of the *Kitasatospora* K002 strain compared to other
784 species of the genus. *Summed Feature occurs when there is imperfect peak separation.
785 Summed feature 3 represents C16:1 ω 6c and/or C16:1 ω 7c, summed feature 8 represents C18
786 : 1 ω 7c and/or C18 : 1 ω 6c, summed feature 9 represents iso-C17 : 1 w9c and/or C16 : 0 10-
787 methyl. Strains: 1, *Kitasatospora* sp. K002; 2, *K. setalba* KM-6054; 3, *K. acidiphila* MMS16-
788 C NU292T; 4, *K. azatica* KCTC 9699T; 5, *K. kifunensis* IFO 15206T; 6, *K. purpeofusca*
789 NRRL B-1817T and 7, *K. nipponensis* HKI 0315T.

790 **Table 3** Results of the antibiotic susceptibility test of the *Kitasatospora* K002 strain. The
791 method adopted was the Kirby-Bauer method (disk diffusion), where filter paper discs
792 impregnated with a standard concentration of antibiotic are used and the formation of an
793 inhibition halo is observed. +, Positive; –, negative.

794 **Table 4** Data from the hybrid assembly of the genome of *Kitasatospora* K002. The assembly
795 was performed from the long and short reading sequencing through the Unicycler program.
796 The annotation was performed using the Prokka software and the quality parameters were
797 evaluated using the CheckM, Bowtie2 and Quast software.

798 **Table 5** Results of cluster analysis of biosynthetic genes for secondary metabolites of the
799 *Kitasatospora* K002 strain. The analysis was performed using the AntiSMASH software (v.
800 6.1.1) and presented in a comparative way with other strains of the genus.

801 **Table 6** Results of the analysis of genes involved in the biosynthesis of peptides from the
802 *Kitasatospora* K002 strain. The analysis was performed using the Bage14 server.

803 **Table 7** Comparison between the phenotypic predictions of the hybrid assembly of the
804 genome of the *Kitasatospora* K002 strain and the genomes of other lineages of the genus.
805 Strains: 1, *Kitasatospora* sp. K002; 2, *K. acidiphila* MMS16-C NU292T; 3, *K. azatica* KCTC

806 9699T; 4, *K. kifunensis* IFO 15206T; 5, *K. purpeofusca* NRRL B-1817T and 6, *K.*
807 *nipponensis* HKI 0315T. +, Positive; -, negative; w, weak.

808 **Figure 1** Scanning (A and B) and transmission (C and D) electron micrographs of the
809 *Kitasatospora* K002 strain. Vegetative mycelia (elongated shape) and submerged spores
810 (rounded shape) can be observed.

811 **Figure 2** Phylogenetic tree based on the partial sequence of the 16S ribosomal *rRNA* gene
812 obtained by cloning and sequencing of Sanger strain *Kitasatospora* K002. The evolutionary
813 position of the strain was estimated by constructing a phylogenetic tree using the Maximum-
814 likelihood (ML) method with 1000 repetitions of sampling (Bootstrap), using representative
815 16S ribosomal *rRNA* sequences from valid strains of the genus listed in the List of Prokaryotic
816 names with Standing in Nomenclature - LPSN.

817 **Figure 3** Circular map of the hybrid genome assembly of the *Kitasatospora* K002 strain. The
818 illustration was generated by the DNAPlotter feature of the Artemis software demonstrating
819 the coding regions and the GC content. Outside to center: forward CDS (dark blue), reverse
820 CDS (red), GC content and GC slope (dark gray and light light gray).

821 **Figure 4** Circular visualization of the genome alignment of the *Kitasatospora* K002 strain (in
822 gray) and its closest lineage according to TYGS-DSMZ, *K. xanthocidica* JCM4862 (in blue).
823 The alignment and illustration were generated using Plot Nucmer. Only genome alignment
824 blocks with at least 1000 nucleotides and 90% identity are shown. The outer band is the size
825 scale of the contigs.

826 **Figure 5** Maximum likelihood phylogenetic tree based on 520 major orthologous genes
827 concatenated from genomic sequences of 12 *Kitasatospora* strains, including *Kitasatospora*
828 K002, the closely related lineages according to the TYGS-DSMZ, and the outgroup
829 *Streptomyces violaceorubidus*. The M1CR0B1AL1Z3R web server was used to concatenate
830 the orthologous genes and build the phylogenetic tree.

831 **Figure 6** Comparative pangenome analysis performed on Roary. The genome of
832 *Kitasatospora* K002 and the 10 genomes closest to it, according to TYGS-DSMZ, were used.
833 A minimum percentage identity for blastp of 80% was defined to cluster the genes that encode
834 complete protein sequences in core genomes.

835 **Figure 7** Predicted proteins in the genome of the *Kitasatospora* K002 strain, classified
836 according to COG categories. Classification was performed using the eggNOG mapper
837 against the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

838 **Figure 8** Comparison between the phenotypic predictions of the hybrid assembly of the
839 genome of the *Kitasatospora* K002 strain and the genomes of its closest strains according to

840 the TYGS-DSMZ obtained from the NCBI database. The zero value means that the strain is
841 negative for the phenotype, the values 1 and 2 correspond to the positive for the phenotype in
842 the phyPat and phyPat+PGL predictors, respectively. The value 3 shows that the phenotype
843 was positive in both predictors. The analysis was performed using the TraitAr software.

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

Table 1

Growth range (optimum):	1	2	3	4	5	6	7	8
Temperature (°C)	15-35 (30)	22- 37(30)	15-37 (30)	10-30 (30)	10-30 (25)	10-30 (25)	10-30 (25)	15-39 (30)
PH	7-9 (8)	-	4-9 (5)	5-8 (6)	5-9 (7)	5-9 (7) 0-2	5-9 (7) 0-2	4-8 (7)
NaCl (%)	0-3(0)	0-2(0)	0-3 (0)	0-2 (0.5)	0-1 (0)	(0.5)	(0.5)	0-2 (0)

Table 2

Fattyacid	Name	Typesoffattyacid	Percent					
			1	2	3	4	5	6
iso-C14:0	12-Methyl Tridecanoic Acid	LongchainSaturated	2.6	2.	2.	3.	1.	2.
			8	0	5	6	1	8
			1.0			2.	2.	
C14:0	Tetradecanoicacid	LongchainSaturated	9	0	0	5	0	0
iso-C15:0	13-Methyltetradecanoic acid	Longchainsaturated	13.	17	12	14	13	17
anteiso-C15:0	12-Methyltetradecanoic acid	Branched fatty acid, methyl branched, saturated fatty acid and derivatives.	11	.4	.9	.7	.8	.7
iso-C16:1			27.	29	17	25	18	21
H		longchainsaturated	07	.0	.6	.3	.8	.9
			0.7		2.			1.
	14-Methylpentadecanoic acid		2	0	5	0	0	6
iso-C16:0		LongchainSaturated	18.	14	17	6.	17	6.
			44	.8	.9	8	.0	2
			10.	6.	16	16	23	9.
C16:0	Hexadecanoicacid	LongchainSaturated	07	6	.4	.5	.8	2
anteiso-C17:1	(7Z)-13-Methyl-7-w9c Hexadecenoic acid		1.5	2.	2.	1.	1.	2.
		LongchainSaturated	4	9	7	2	0	0
iso-C17:0	Heptadecanoicacid		3.2	4.	2.	1.	3.	3.
		LongchainSaturated	6	0	5	2	8	0
anteiso-C17:0	14-Methylhexadecanoic acid	MethylBranchedSaturated	11.	11	5.	3.	7.	5.
			98	.9	6	1	0	76
cyclo-C17:0	Cis-9,10-Methylenehexadecanoic acid	Long chain saturated bearing a cyclopropane ring at C9 and C10	4.5	4.	3.	14	4.	1.
			1	6	4	.2	9	1
iso-C18:0	16-Methylheptadecanoic acid		1.1					
		Methylbranched	2	0	0	0	0	0
			0.5					
C18:0	Octadecanoicacid	LongchainSaturated	2	0	0	0	0	0
SummedFeature 3*	-	-	1.4		9.	6.	10	8.
			9	0	3	1	.8	6
SummedFeature 8*	-	-	0.5					
			7	0	0	0	0	0
SummedFeature 9*	-	-	1.8	3.	4.	2.	3.	6.
			3	1	3	2	07	7

Table 3

Antibiotic (μg per disc)	Result	Inhibition halo (cm)
Amikacin (30)	+	1,5
Amoxicillin/Clavulanicacid (20/10)	-	-
Ampicillin (10)	+	2
Aztreonam (30)	-	-
Cephalexin (30))	+	2,5
Cefazolin (30)	+	1,5
Cefepime (30)	-	-
Ceftazidime (30)	-	-
Ceftriaxone (30)	+	2
Ciprofloxacin (5)	+	3
Clindamycin (2)	+	1
Chloramphenicol (30)	+	1,5
Doxycycline (30)	+	3,5
Erythromycin (15)	+	4
Gentamicin (10)	+	1,5
Imipenem (10)	+	5
Meropenem (10)	-	-
Norfloxacin (10)	+	1,5
Ofloxacin (5)	+	2,5
Oxacillin (1)	-	-
Penicillin G (10)	+	1,3
Piperacillin + Tazobactam (110)	+	2,5
Polymyxin B (300)	+	0,8
Sulfonamide (300)	-	-
Tetracycline (30)	+	2
Vancomycin (30)	+	2

Table 4

Assembly data	
Contigs	14
CDS	6708
<i>RRNA</i>	19
Repeatregion	11
<i>TRNA</i>	90
<i>TmRNA</i>	1
GC content (%)	72,8
Contamination (%)	1,32
Completeness (%)	99,47
Coverage	300x
L50	4
N50 (pb)	761556
Total Size (bp)	7951960
Major contig	1713370

Table 5

Secondary metabolites				
Type	No. of clusters			
	<i>Kitasatospora</i> K002	<i>K. acidiphila</i>	MMS16-CNU292T	<i>Kitasatospora</i> sp. MMS16-BH015
Amglyccicl	1		0	1
Bacteriocin	0		0	1
Betalactone	1		0	0
Blactam	1		0	0
Butyrolactone	2		0	2
Furan	1		0	0
Homoserinelactone	0		0	1
Indole	0		1	0
Ladderane	0		1	0
Lantipeptideclass I	1		6	4
Lantipeptideclass II	1		0	0
Lantipeptideclass III	1		2	0
Lantipeptideclass IV	1		0	0
LAP	1		0	0
Lasso peptide	0		1	0
NPAA	1		0	0
NRPS	23		12	8
Nucleoside	0		0	1
Outro	1		0	3
PKS	1		2	0
RIPP	2		0	0
Siderophore	1		2	1
T1PKS	7		7	5
T2PKS	1		0	2
T3PKS	2		4	1
Terpene	7		5	5
Thiopeptide	1		0	3
TransAT-PKS	2		0	0

Table 6

Start (bp)	End (bp)	Class
141602	161902	<i>Zoocin A</i>
362141	382141	<i>LAPs</i>
770112	790871	<i>PutativeBacteriocinfamilyprotein</i>
62228	82228	<i>Lanthipeptideclass IV</i>
647639	667639	<i>Lanthipeptideclass II</i>
43982	63982	<i>Lanthipeptideclass I</i>
131183	151183	<i>Sactipeptides</i>
239900	259900	<i>Sactipeptides</i>

Table 7

Characteristics	1	2	3	4	5	6
Hydrolysis of:						
Casein	+	+	-	-	-	+
DNA	-	-	-	+	+	-
Starch	+	+	+	+	+	+
Enzyme:						
Acid production from glucose	-	-	-	-	+	-
Arginine dihydrolase	+	-	-	-	-	-
β -Galactosidase	-	-	+	+	+	+
Gelatin liquefaction	+	W	-	+	+	+
β -Glucosidase (aesculin hydrolysis)	+	-	+	-	+	+
Indole production	+	-	-	-	-	-
Nitrate reduction	-	-	-	-	-	-
Urease	+	-	-	-	-	+
Alkaline phosphate	+	+	+	+	+	+
Lipase	+	-	-	-	+	-
Growth on sole carbon sources:						
Arabinose	-	-	+	-	-	+
Citrate	+	-	-	-	-	-
Glucose	+	W	+	-	+	+
Maltose	+	-	-	-	-	-
Mannitol	+	-	-	-	-	+
Mannose	+	-	-	-	-	+

Figure 1

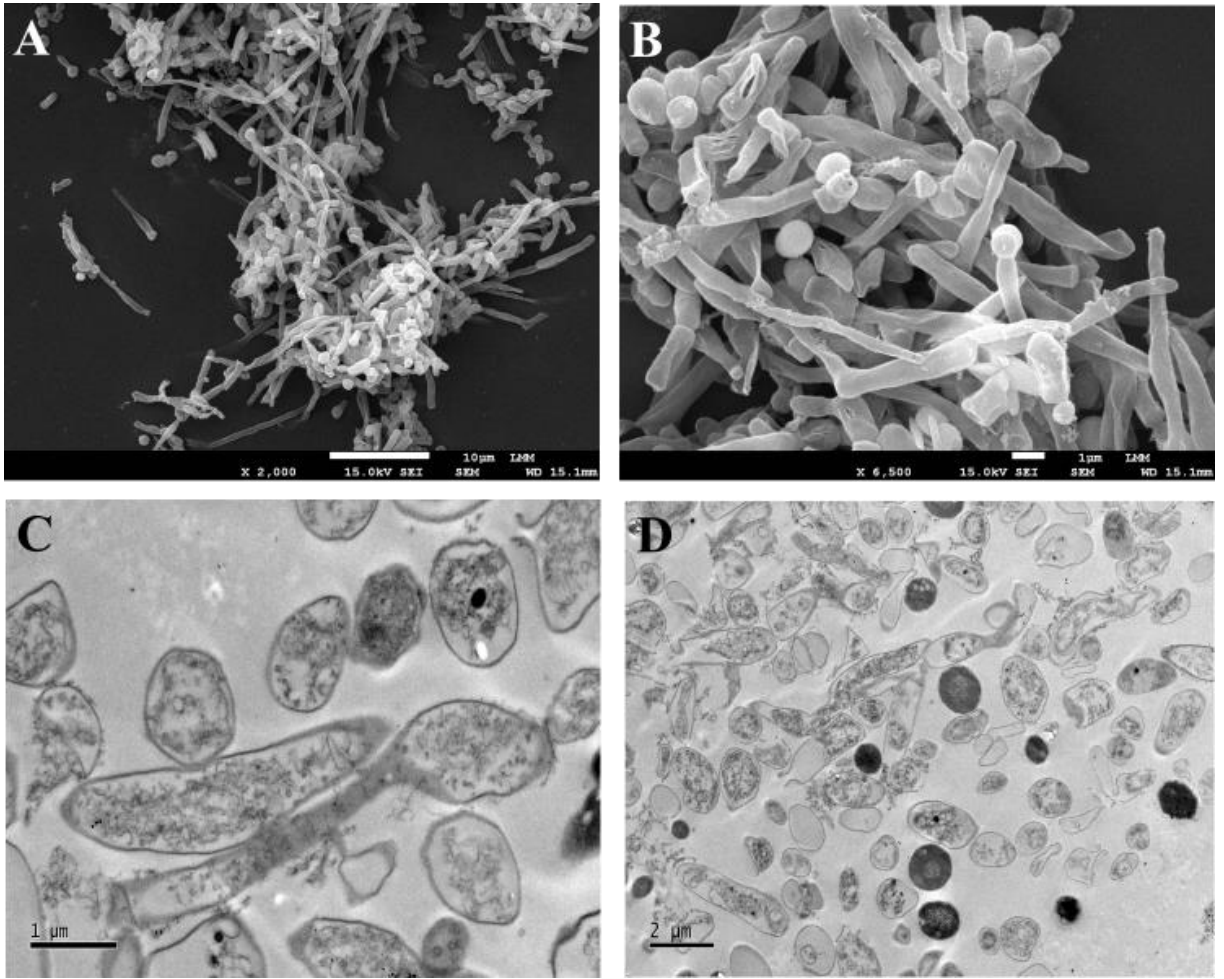


Figure 2

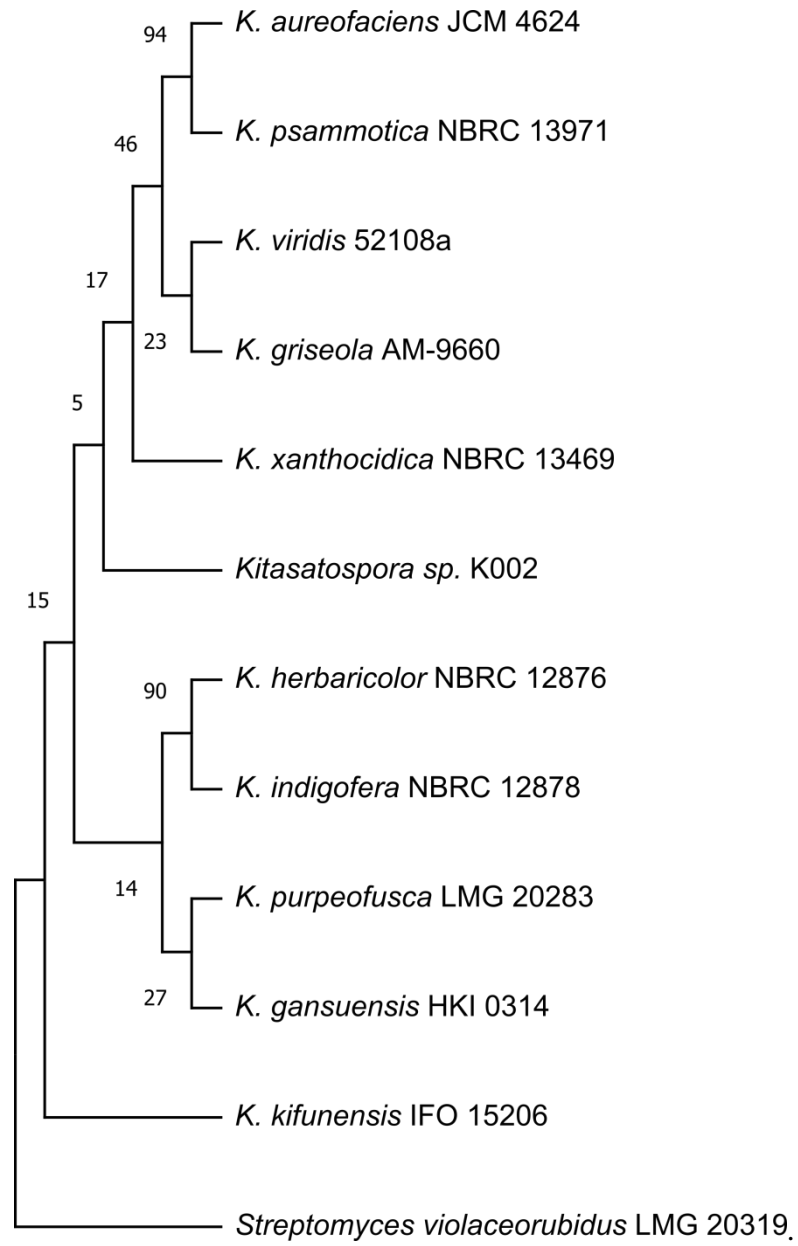


Figure 3

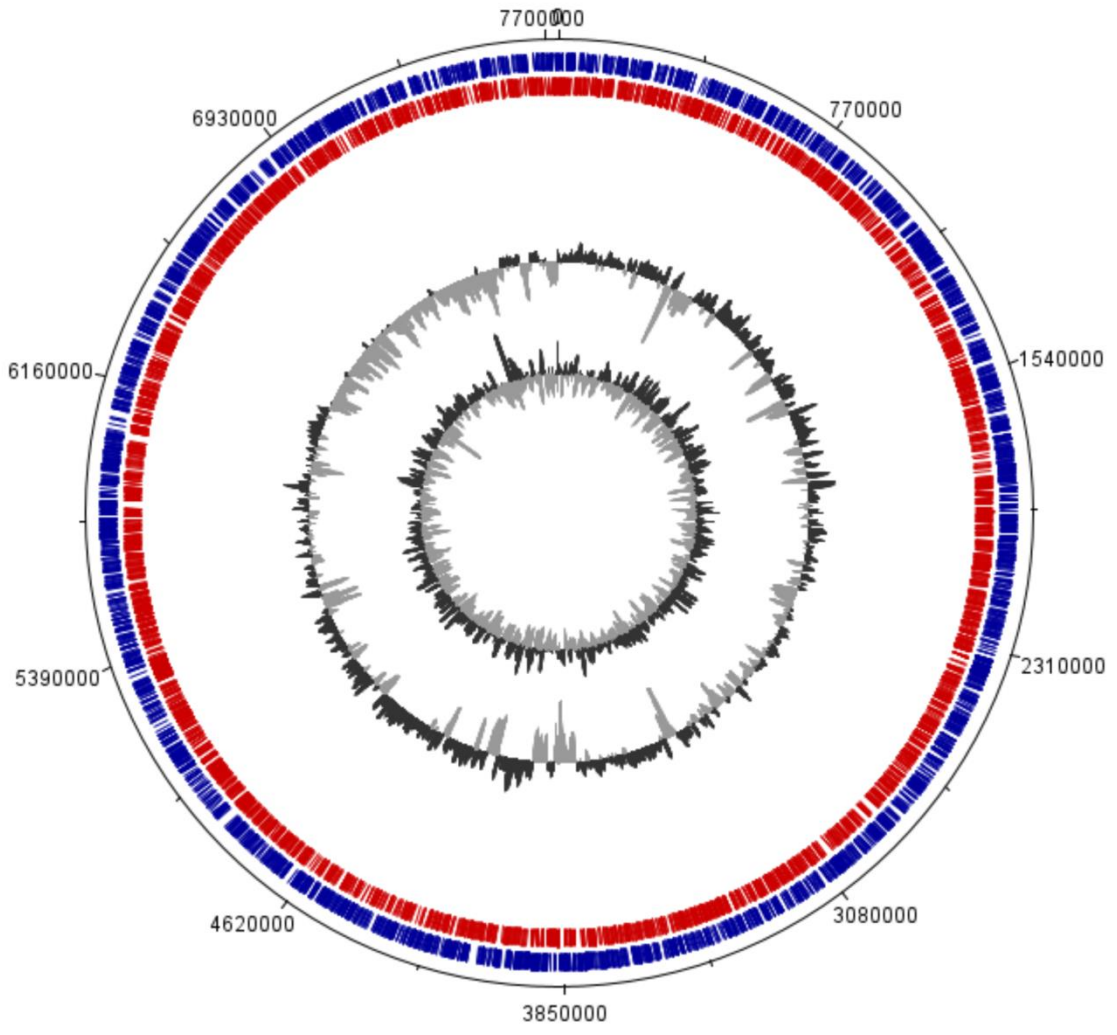


Figure 4

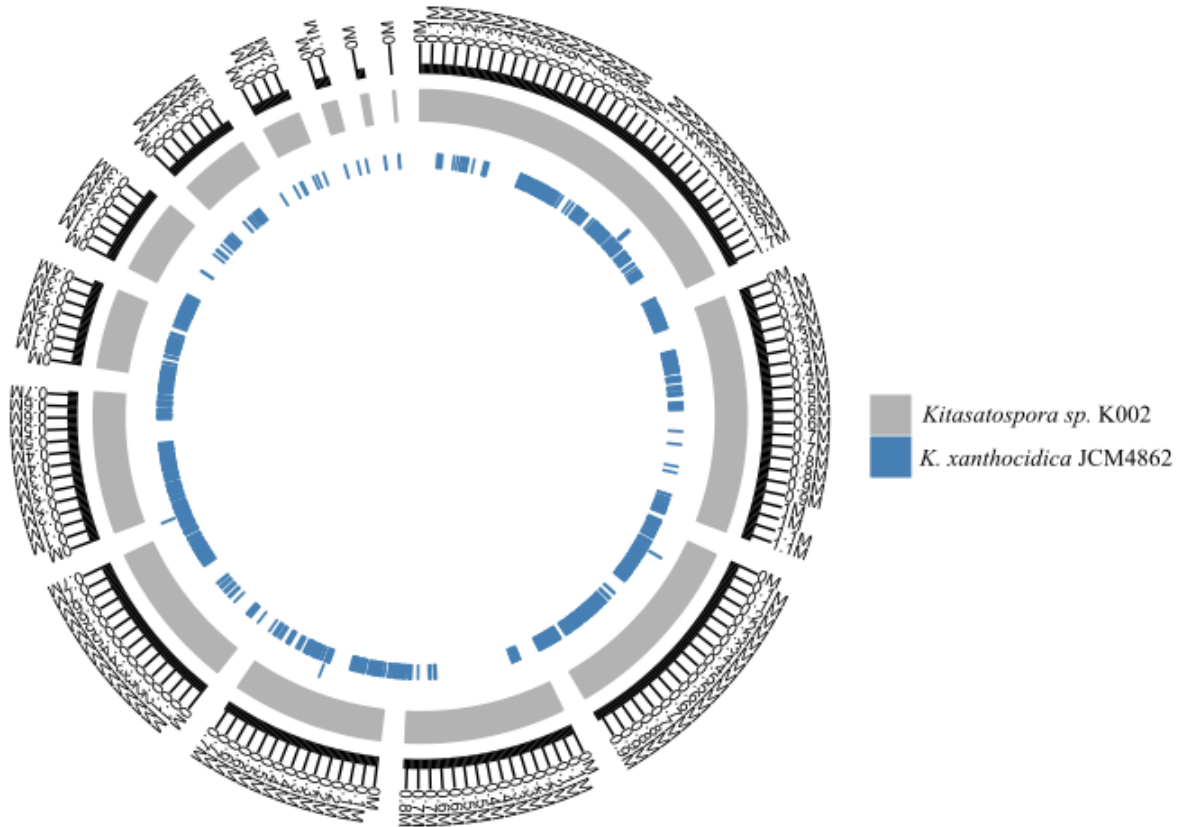


Figure 5

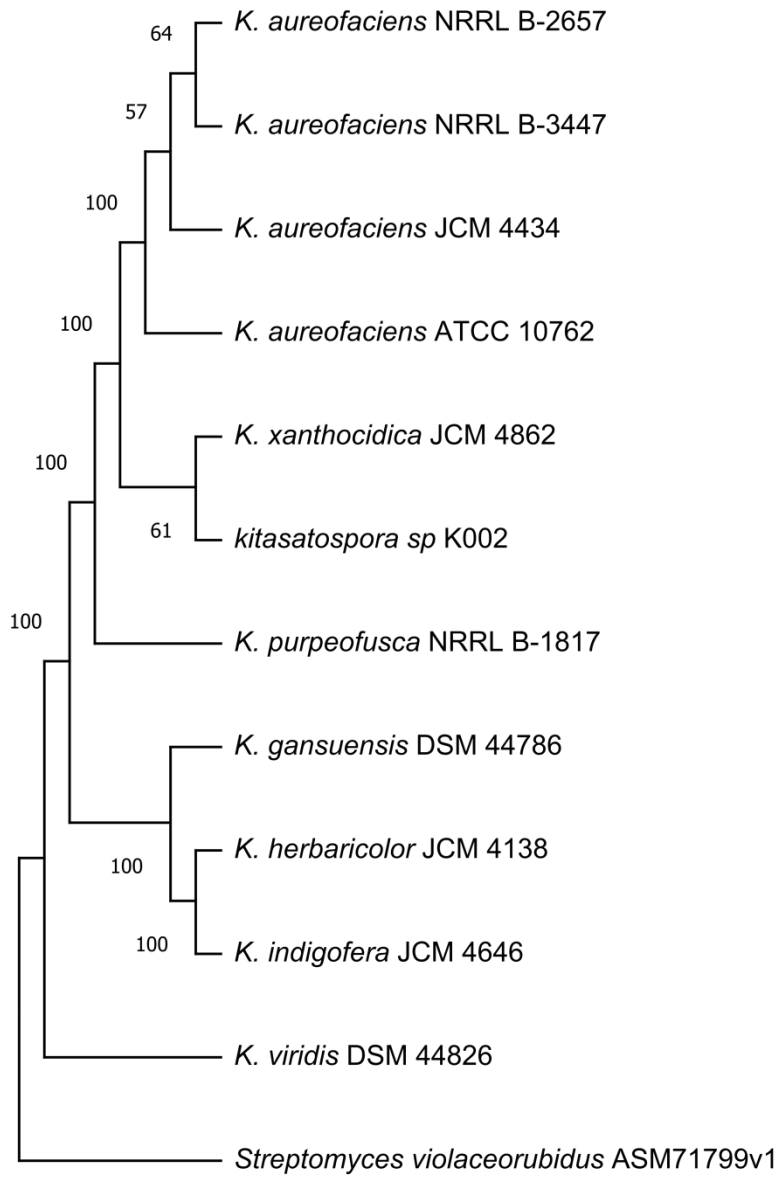


Figure 6

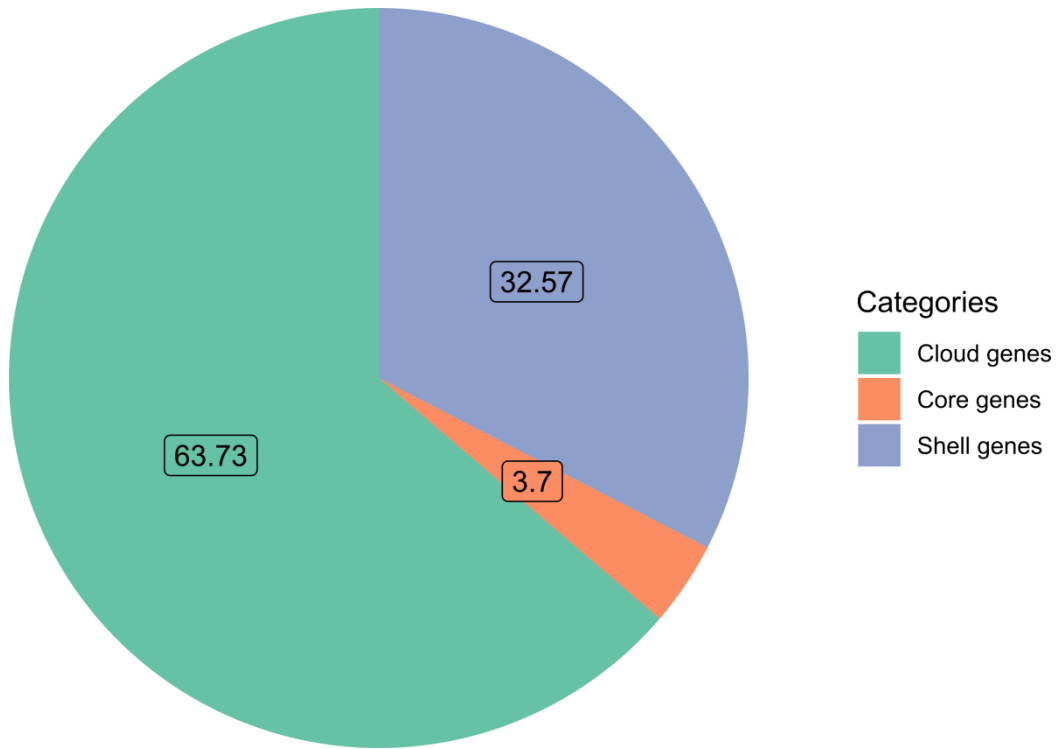
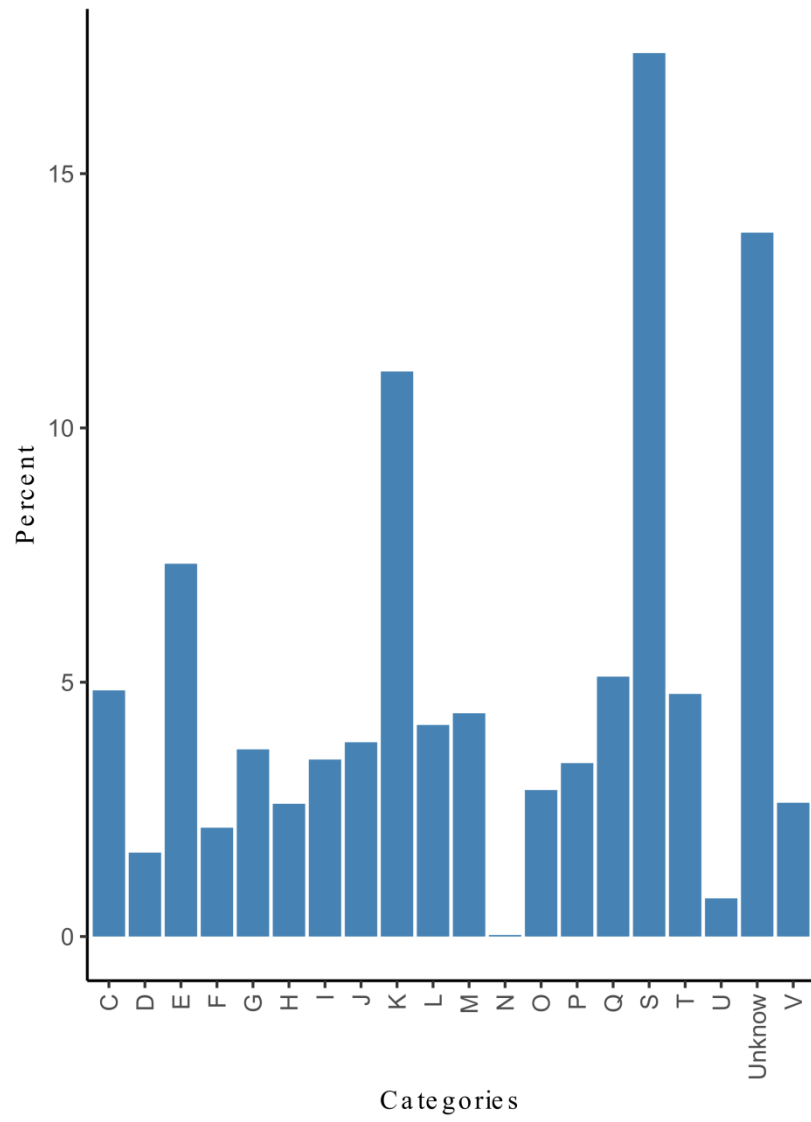


Figure 7



6.3 Artigo completo publicado em periódico



Microbiome diversity from sponges biogeographically distributed between South America and Antarctica

Mayanne A.M. de Freitas^{a,1}, Isabel C. Cunha-Ferreira^{b,1}, Camille V. Leal^a, Julio C.C. Fernandez^c, Claudia Y. Omachi^d, Lucia S. Campos^e, Bruno P. Masi^f, Ricardo H. Krüger^b, Eduardo Hajdu^g, Cristiane C. Thompson^a, Fabiano L. Thompson^{a,*}

^a Laboratory of Microbiology, Biology Institute, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil

^b Laboratory of Ecotoxicology, Biology Institute, University of Brasília (UNB), Brasília, Brazil

^c Department of Invertebrates, Natural Museum, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil

^d Laboratory of Environmental Indicators, Geospatial Institute, University of São Paulo, São Paulo, SP, Brazil

^e Department of Zoology, Biology Institute, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil

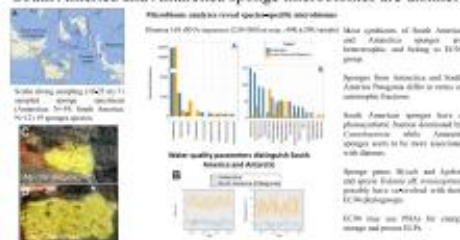
^f Laboratory of Marine Ecology and Fishery Oceanography of the Amazon (LEMOPA), Socio environmental and Water Resources Institute (ISARH), Federal Rural University of the Amazon (UFRA), Belém, PA, Brazil

HIGHLIGHTS

- Most sponge symbionts are heterotrophic (94.8%).
- EC94 is the most abundant symbiont (70–87%).
- EC94 is split into at least 10 phylogroups.
- Each EC94 group is specific to species and genus of sponge and possible coevolved with their hosts.
- Antarctic sponges had 5.5 % of chemosynthetic prokaryotes, and South America sponges had 2.3 % of photosynthetic ones.

GRAPHICAL ABSTRACT

South America and Antarctica sponge microbiomes are distinct



ARTICLE INFO

Editor: Frederic Couzin

Keywords:
Microbiome
Sponge
Antarctica
South America
EC94
Biogeography

ABSTRACT

Sponges from South America and Antarctica are evolutionarily closely related. Specific symbiont signatures that could differentiate these two geographic regions are unknown. This study aimed to investigate the microbiome diversity of sponges from South America and Antarctica. In total 71 sponge specimens were analyzed (Antarctica: N = 59, 13 different species; South America: N = 12, 6 different species). Illumina 16S rRNA sequences were generated (2.88 million sequences; 40K ± 29K/sample). The most abundant symbionts were heterotrophic (94.8 %) and belonged mainly to Proteobacteria and Bacteroidota. EC94 was the most abundant symbiont and dominated the microbiome of some species (70–87 %), comprising at least 10 phylogroups. Each of the EC94 phylogroups was specific to one genus or species of sponge. Furthermore, South America sponges had higher abundance of photosynthetic microorganisms (2.3 %) and sponges from Antarctica, the highest abundance of chemosynthetic (5.5 %). Sponge symbionts may contribute to the function of their hosts. The unique features from each of these two regions (e.g., light, temperature, and nutrients) possibly stimulate distinct microbiome diversity from sponges biogeographically distributed across continents.

* Corresponding author at: Biology Institute, CCS, UFRJ, Av. Carlos Chagas Filho, Bloco A, Anexo, Ilha do Fundão, 21941-902 Rio de Janeiro, RJ, Brazil.
E-mail address: fabianothompson1@gmail.com (F.L. Thompson).

¹ These authors contributed equally to this work.

<http://dx.doi.org/10.1016/j.scitotenv.2023.163256>

Received 7 March 2023; Received in revised form 29 March 2023; Accepted 30 March 2023

Available online 01 April 2023

0048-9697/© 2023 Published by Elsevier B.V.

6.4 Resumos publicados em anais de eventos

FERREIRA, I.C.C., VIZZOTTO, C.S., FREDERICO, T. D., KRÜGER, R.H. Genotypic characterization of bacterial isolates obtained from the use of pelgipeptins as a selection mechanism: preliminary results. In: XI Simpósio do Programa de Pós-Graduação em Biologia Molecular, 2022, Brasília.

FERREIRA, I. C. C., FREITAS, M. A. M., LEAL, C. V., THOMPSON, F., KRÜGER, R. H. Description of the nutritional profile of microbiomes associated with marine sponges from Antarctica and Chilean Patagonia. In: X Simpósio do Programa de Pós-Graduação em Biologia Molecular, 2021, Brasília.

FERREIRA, I. C. C., ARAUJO JR, S.D., KRÜGER, R. H. Prospecção de antimicrobianos em isolados bacterianos de solo do Cerrado. In: IX Simpósio do Programa de Pós-Graduação em Ciências Biológicas (Biologia Molecular), 2019, Brasília.

6.5 Participação em eventos

AANchor - Biotec Bioinformatics tools and applications workshop. 2022.

XI Simpósio do Programa de Pós-Graduação em Biologia Molecular, 2022, Brasília.

X Simpósio do Programa de Pós-Graduação em Biologia Molecular, 2021, Brasília.

IX Simpósio do Programa de Pós-Graduação em Ciências Biológicas (Biologia Molecular), 2019, Brasília.

6.6 Capítulo de livro a ser publicado

SILVEIRA, C., **FERREIRA, I.C.C.** , KRÜGER, R.H. Recifes de corais. In: MICROBIOMAS DA AMAZONIA AZUL. Org. Thompson, F. L. e Thompson, C. C. Status: aguardando publicação.

6.7 Organização de evento

KRÜGER, R. H.; GONÇALVES, P. L.; SOUSA, J. F.; FREDERICO, D. T.; **FERREIRA, I. C. C.**; PEIXOTO, J.; NEVES, C.; FIGUEIREDO, D. Comissão organizadora do *AANchor - Biotec Bioinformatics tools and applications workshop. 2022.*

FERREIRA, I. C. C. Comissão organizadora do IX Simpósio do Programa de Pós-Graduação em Ciências Biológicas (Biologia Molecular), realizado na Universidade de Brasília, de 27 a 29 de Novembro de 2019.

6.8 Atividades relevantes

Treinamento Pré-antártico (TPA). Carga horária: 60 h. Marinha do Brasil, MB, Brasil. 2019.

Participação na OPERANTAR XXXVIII a bordo do Navio Polar Maximiano, na área marítima entre Punta Arenas-Chile e Antártica, no período de 16 de janeiro a 04 de fevereiro de 2020, tendo navegado 1487 milhas náuticas e completando 18 dias no mar. Nesta expedição foram realizadas coletas em campo de amostras de água e poríferos, processamento das amostras, isolamento e cultivo de bactérias e extrações de DNA.