



Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-Graduação em Biologia Molecular

Caracterização de novos microrganismos cultivados a partir de solos do Cerrado

ALINE BELMOK DE ARAÚJO DIAS IOCCA

Tese apresentada ao Programa de Pós-Graduação em Biologia Molecular do Departamento de Biologia Celular, Instituto de Biologia, Universidade de Brasília, para a obtenção do título de Doutora em Biologia Molecular.

Orientadora: Dra. Ildinete Silva-Pereira

Co-orientadora: Dra. Cynthia Maria Kyaw

Brasília, outubro de 2020

Dedico essa tese à memória de meu querido avô Nestor Belmok, de quem herdei a paixão por livros e pelo conhecimento e que com tanto orgulho falava de sua neta que um dia seria doutora.

SUMÁRIO

Agradecimentos.....	v
Lista de Tabelas e Figuras.....	viii
Lista de Abreviações.....	xii
RESUMO	1
ABSTRACT	2
INTRODUÇÃO GERAL.....	3
CAPÍTULO 1 – Estabelecimento e caracterização de co-cultivos de archaeas e bactérias obtidos a partir de amostras de solo de Cerrado.....	4
1.1. Introdução.....	4
1.1.1. O cultivo laboratorial de microrganismos e o início da microbiologia.....	4
1.1.2. A microbiologia na era das abordagens moleculares e a proposta do domínio <i>Archaea</i>	6
1.1.3. A importância do cultivo na era das “ômicas”	10
1.1.4. A grande diversidade microbiana nos solos do bioma Cerrado.....	13
1.2. Objetivos	16
1.2.1. Objetivo geral	16
1.2.2. Objetivos específicos.....	16
1.3. Material e Métodos.....	17
1.3.1. Coleta dos solos.....	17
1.3.2. Confeção do meio de cultivo, inóculo inicial e manutenção das culturas	17
1.3.3. Extração de DNA genômico total	18
1.3.4. Ensaios de PCR.....	19
1.3.7. Transformação por choque térmico, seleção e estoque em glicerol de clones recombinantes	22
1.3.8. Extração de DNA plasmidial por lise alcalina	22
1.4. Resultados e discussão.....	29
1.4.1. Características químicas do solo e obtenção de culturas em meio sólido.....	29
1.4.2. Obtenção de culturas em meio líquido	32
1.4.3. Caracterização filogenética dos organismos cultivados a partir da análise dos produtos obtidos nos ensaios de PCR dirigidos aos genes de rRNA 16S.....	32
1.4.4. Análises microscópicas das células presentes no co-cultivo.....	41
1.4.5. Sequenciamento genômico total dos cultivos de solo de Cerrado.....	49
CAPÍTULO 2 - Caracterização morfo-fisiológica e genômica da bactéria <i>Novosphingobium terrae</i> GeG2 ^T , isolada a partir de solo de Cerrado	53
2.1. Introdução.....	53
2.1. Objetivos	55
2.1.1. Objetivo geral	55
2.1.2. Objetivos específicos.....	55
2.2. Material e Métodos.....	56
2.2.1. Isolamento e cultivo da bactéria em meios de cultura definidos	56

2.2.2. Observação das células bacterianas por microscopia óptica	57
2.2.3. Extração de DNA genômico, amplificação, clonagem, sequenciamento e análises do gene de rRNA 16S da cultura bacteriana	57
2.2.4. Comparações fisiológicas entre o isolado bacteriano GeG2 e <i>Novosphingobium rosa</i> DSM 7285	59
2.2.5. Caracterizações químicas de GeG2	61
2.2.6. Comparação do perfil proteico das bactérias GeG2 e <i>N. rosa</i> DSM 7285 por espectrometria de massa	62
2.2.7. Avaliação do perfil de susceptibilidade de GeG2 a antibióticos	64
2.2.8. Observação das células bacterianas por microscopia eletrônica de varredura (MEV) e de transmissão (MET).....	64
2.2.9. Sequenciamento, montagem e anotação do genoma completo do isolado bacteriano GeG2.....	65
2.2.10. Análises taxonômicas baseadas no genoma	67
2.3.11. Análises preliminares do crescimento de GeG2 em MM contendo o corante aromático verde malaquita	68
2.3. Resultados e Discussão	69
2.3.1. Isolamento e cultivo da bactéria presente no cultivo do solo de Cerrado em meio definido	69
2.3.2. Avaliação das células bacterianas GeG2 por microscopia óptica	70
2.3.3. Análises do gene rRNA 16S.....	72
2.3.4. Caracterizações fisiológicas de GeG2 e comparações com <i>N. rosa</i> DSM 7285.....	74
2.3.5. Caracterizações químicas de GeG2	78
2.3.6. Caracterização do perfil proteico de GeG2 por espectrometria de massa e comparação com o perfil de <i>N. rosa</i> DSM 7285	81
2.3.7. Perfil de susceptibilidade a antibióticos de GeG2.....	82
2.3.8. Observação das células de GeG2 e <i>N. rosa</i> DSM 7285 por MEV e MET.....	84
2.3.9. Dimorfismo planctônico/séssil do isolado GeG2	89
2.3.10. Sequenciamento, montagem, anotação geral e arquitetura do genoma da linhagem GeG2.....	96
2.3.11. Classificação taxonômica da linhagem GeG2 baseada no genoma completo	104
2.3.12. Análises funcionais do genoma da linhagem GeG2.....	108
2.3.13. Potencial do isolado GeG2 na degradação de compostos aromáticos	125
4.3.14. Proposta da espécie <i>Novosphingobium terrae</i> sp. nov.	129
CONCLUSÕES E PERSPECTIVAS.....	130
REFERÊNCIAS BIBLIOGRÁFICAS	132
ANEXO I – Publicações realizadas durante o período do doutorado	148

Agradecimentos

São tantas as pessoas a quem tenho que agradecer por essa jornada, que talvez poucas linhas não sejam suficientes. A trajetória percorrida nesses anos foi repleta de desafios, conquistas, alegrias e frustrações. Foram muitos colaboradores, tutores, amigos e familiares que me ensinaram, aconselharam, ajudaram, corrigiram e apoiaram. E, seguramente, essa tese não existiria se não fosse por cada um deles.

Primeiramente, agradeço à minha querida orientadora Dra. Ildinete Silva-Pereira, por me acolher como doutoranda e me auxiliar durante todo esse processo. Sem seu apoio e conselhos, nada disso teria sido possível!

Me faltam palavras para expressar a gratidão que tenho pela Profa. Dra. Cynthia Kyaw. Minha orientadora desde os tempos de iniciação científica e ao longo de toda essa pesquisa, seus ensinamentos com certeza são a razão por eu estar no caminho de me tornar pesquisadora. Tive o privilégio de receber sua orientação ativa e presente, de poder aprender com suas aulas fantásticas e de compartilhar de sua ampla experiência em microbiologia, o que certamente me tornou uma bióloga melhor. No entanto, muito além dos conceitos acadêmicos, aprendi muito com sua generosidade, paciência, empatia e amizade, o que certamente, ao longo desses quase dez anos de convivência, me tornou uma pessoa melhor. Muito obrigada por tudo, professora querida!

Aos meus pais, Elder e Luciana, serei sempre grata por essa e todas as outras conquistas em minha vida. Obrigada por serem tão amorosos e me proporcionarem todas as oportunidades e incentivos possíveis. Vocês são maravilhosos e foram fundamentais para minha formação.

Àquele que se tornou meu marido ao longo dessa trajetória, Diêgo Iocca, mas com quem já compartilho a vida desde os tempos da graduação, tenho muito o que agradecer. Sua paciência, apoio, incentivo, companheirismo e amor me motivaram e ajudaram ao longo de todos esses anos e foram essenciais para que eu chegasse até aqui.

Sou muito grata a todos os colaboradores que contribuíram com o desenvolvimento desse trabalho de alguma forma, sem os quais grande parte dos experimentos e análises não teriam sido possíveis. Agradeço aos professores e colegas da Universidade de Brasília: ao Prof. Dr. Ricardo Krüger e todo seu grupo de pesquisa, em especial a Carla Vizzotto, Fabyano Alvares, Otávio Pinto, Helena Guimarães e Julianna Peixoto, pela parceria, assistência com análises e disponibilização de recursos para

experimentos ao longo de todo projeto; à Profa. Dra. Marlene Teixeira, por seu auxílio em diversos momentos e por todo apoio prestado nos testes de antibiograma; à Profa. Dra. Eliane Noronha e à doutoranda Jéssica Pinheiro pela disponibilidade de materiais e auxílio com os experimentos de crescimento em anaerobiose; à Profa. Dra. Sônia Bão e suas colaboradoras Ingrid Martins da Silva e Ana Brígida Evangelista, pela assistência fundamental no preparo e análise das amostras por Microscopia Eletrônica de Varredura; aos Profs. Drs. Vicente Martins, Tatiana Amabile, Robert Miller e aos membros de seus grupos de pesquisa, pelo auxílio e colaboração no dia-a-dia dos trabalhos. Gostaria de agradecer especialmente ao Prof. Dr. Georgios Pappas Jr. e aos doutorandos Felipe Marques e Rodrigo Rocha, por toda a colaboração e assistência nas análises genômicas desenvolvidas nesse trabalho. Essa etapa certamente não teria sido possível sem a expertise e dedicação de vocês.

A todos os colaboradores de outras instituições que também dedicaram seu tempo e conhecimento para que esse trabalho se tornasse mais completo, meu muito obrigada: ao Prof. Dr. Marcelo Ramada, da Universidade Católica de Brasília, pela disponibilização dos equipamentos e materiais necessários e por todo o auxílio com os experimentos de MALDI-TOF; ao Prof. Dr. Marcos Tótola, da Universidade Federal de Viçosa, pela caracterização do perfil de ácidos graxos; ao Prof. Dr. Gregory Kitten e à equipe do Centro de Microscopia da Universidade Federal de Minas Gerais, pelo preparo das amostras e análises de Microscopia Eletrônica de Transmissão; à Dra. Viviane Furlan, do Laboratório Central de Saúde Pública do DF (LACEN-DF), pela disponibilidade em realizar caracterizações bioquímicas de nossa bactéria.

Também tenho muito o que agradecer ao meu companheiro de laboratório e amigo há tantos anos, Thiago Rodrigues de Oliveira. Obrigada pelas enriquecedoras conversas científicas, pela ótima parceria na escrita e publicação dos artigos e por escutar meus longos desabafos no estacionamento do IB depois de dias difíceis em que tudo parecia que ia dar errado. Obrigada por tudo e que nossa parceria acadêmica continue por muitos anos ainda, “coleguinha”! Agradeço também aos demais colegas que compartilharam comigo o dia-a-dia no laboratório ou no nosso andar, tornando-o mais agradável e alegre ao longo desses anos: Deborah, Filipe, Matheus, Lucas, MéloDi, Ikaro, Letícia Velasco, Letícia Mallmann, Gabriel Drose, Milla, Isabella, Ana Paula, Laura, Isabela, Luís, Hérick, Felipe, Paulo e Juliana, obrigada!

De maneira muito especial, gostaria de agradecer à querida Carla Vizzotto. Muito mais que uma colega de doutorado, foi companheira para todas as horas, aquela com

quem sempre pude contar ao longo de todos os altos e baixos dessa jornada, com quem compartilhei os desafios de caracterizar uma nova espécie bacteriana e as lutas contra as burocracias intermináveis, mas também aquele cappuccino de fim de dia essencial para recarregar as energias antes de seguirmos enfrentando os desafios. Nem sei o que teria sido desse doutorado sem sua eficiência, inteligência, disposição e ombro amigo. “Foco, força e fé”!

Ao Marco Oliveira, amigo querido desde os tempos de graduação e com quem pude contar do primeiro ao último minuto do desenvolvimento desse trabalho: muito obrigada! Sempre admirei a sua inteligência e, certamente, seus *insights* científicos, conselhos certos, amizade e disponibilidade em me escutar e ajudar no doutorado e em todos os outros aspectos da vida ao longo desses anos foram e são extremamente importantes.

Não poderia deixar de agradecer também aos muitos familiares e amigos que sempre me incentivaram, escutaram, ajudaram e torceram por mim ao longo da longa jornada que é o doutorado. Muito obrigada meus irmãos Paulo Vitor e Artur, meu querido padrao Bruno, meus queridos sogros Luiz e Lourdes, tias e tios, primas e primos, cunhados e cunhadas, amigos e amigas de graduação e de toda a vida: vocês são muito importantes para mim e também foram fundamentais durante toda essa trajetória.

Lista de Tabelas e Figuras

Tabela 1. Iniciadores para o gene de rRNA 16S empregados nos ensaios de PCR realizados com o DNA extraído dos cultivos de solo de Cerrado.	20
Tabela 2. Condições de ciclagem empregadas nos ensaios de PCR com os diferentes iniciadores para os genes do rRNA 16S de <i>Archaea</i> e <i>Bacteria</i>	21
Tabela 3. Valores dos parâmetros físico-químicos do solo de Cerrado utilizado para a confecção dos meios de cultura e como inóculo inicial para os cultivos.	29
Tabela 4. Classificação das sequências do gene de rRNA 16S de <i>Archaea</i> amplificadas com os três diferentes pares de iniciadores a partir dos tipos coloniais G e G2.	36
Tabela 5. Características fisiológicas das linhagens bacterianas GeG2 e <i>N. rosa</i> DSM 7285 ...	77
Tabela 6. Comparação do conteúdo de ácidos graxos (%) das linhagens GeG2 e <i>N. rosa</i> DSM 7285, ambas crescidas em MM por 72 horas.	80
Tabela 7. Diâmetros dos halos de inibição do crescimento das linhagens GeG2 e <i>N. rosa</i> DSM 7285 observados em placas contendo MM e discos impregnados com diferentes antibióticos..	83
Tabela 8. Estatísticas da montagem do genoma da linhagem bacteriana GeG2 utilizando-se o montador <i>Unicycler</i> e abordagem híbrida, incluindo dados de Pacbio (<i>long reads</i>) e Illumina (<i>short reads</i>).	96
Tabela 9. Resumo dos resultados das análises de completude da montagem dos genomas de GeG2 e de <i>N. rosa</i> NRBC15208.....	97
Tabela 10. Características dos <i>contigs</i> resultantes da montagem com o programa <i>Unicycler</i> por meio de abordagem híbrida, utilizando <i>short reads</i> e <i>long reads</i>	97
Tabela 11. Valores de dDDH (<i>digital DNA-DNA hybridization</i>) e diferença de conteúdo de GC encontrados em comparações par a par entre o genoma da linhagem GeG2 e genomas de linhagens tipo disponíveis no banco de dados identificadas como mais próximas.	108
Tabela 12. Número de proteínas preditas nos plasmídeos pGeG2a e pGeG2b da linhagem GeG2 classificadas em cada categoria funcional COG.	113
Tabela 13. Mono e dioxigenases encontradas no genoma da linhagem GeG2..	126
Figura 1. Esquema representando as regiões conservadas (em tom escuro) e as nove regiões variáveis (V1 a V9, em tom claro) do gene rRNA 16S..	20
Figura 2. Tipos coloniais G e G2, obtidos em meio de cultura inoculado com solo de Cerrado e suplementado com ampicilina, estreptomicina, cloranfenicol e itraconazol após aproximadamente um mês de crescimento.....	30
Figura 3. Colônias G e G2 observadas após vários repiques em meio contendo polimixina B.	31
Figura 4. Árvore filogenética com as sequências do gene rRNA 16S de <i>Archaea</i> obtidas com os iniciadores 21f/958r a partir do cultivo de solo de Cerrado (indicadas com	

triângulos) alinhadas com sequências de archaeas cultivadas e não cultivadas retiradas de banco de dados do NCBI.	36
Figura 5. Árvore filogenética com as sequências do gene rRNA 16S de <i>Archaea</i> obtidas com os iniciadores 109/915r a partir do cultivo de solo de Cerrado (indicadas com triângulos) alinhadas com sequências de archaeas cultivadas e não cultivadas retiradas de banco de dados do NCBI.	38
Figura 6. Árvore filogenética com as sequências do gene rRNA 16S de <i>Archaea</i> obtidas com os iniciadores Bathy442f/644r a partir do cultivo de solo de Cerrado (indicadas com triângulos) alinhadas com sequências de archaeas cultivadas e não cultivadas retiradas de banco de dados do NCBI.	40
Figura 7. Coloração de Gram dos tipos coloniais G e G2, obtidos a partir de amostras de solo de Cerrado.	42
Figura 8. Eletromicrografias de varredura de células da cultura G2 crescida em meio sólido.	43
Figura 9. Coloração de Gram dos cultivos G e G2 obtidos em meio líquido.	44
Figura 10. Análises a fresco por microscopia de contraste de fase das culturas G e G2 crescidas sobre lamínulas em meio líquido. Aumento 1000x.	45
Figura 11. Eletromicrografias de varredura da cultura G crescida em meio líquido sobre lamínulas de vidro sem tratamento e previamente tratadas com poli-L-lisina.	46
Figura 12. Eletromicrografias de varredura de células da cultura G2 crescida em meio líquido sobre lamínulas de vidro sem tratamento e previamente tratadas com poli-L-lisina.	47
Figura 13. Microscopia de fluorescência de células da cultura G2 crescidas em meio líquido contendo ampicilina, estreptomomicina, cloranfenicol e itraconazol (SACI) ou polimixina B, após incubação com o kit <i>LIVE/DEAD BacLight</i> TM	48
Figura 14. Colônias bacterianas obtidas após a primeira transferência de colônias do cultivo de solo de Cerrado em meio de extrato de solo para meio LB.	69
Figura 15. Colônias bacterianas obtidas após a transferência de colônias do cultivo de solo de Cerrado em meio de extrato de solo para meio mínimo (MM).	70
Figura 16. Análise a fresco por microscopia de contraste de fase e de coloração de Gram por microscopia de campo claro das células bacterianas GeG2, cultivadas em MM por 96 horas.	71
Figura 17. Árvore filogenética baseada em sequências do gene rRNA 16S mostrando a relação entre a linhagem GeG2 (em negrito) e outros taxa relacionados da família <i>Sphingomonadaceae</i>	73
Figura 18. Diferenças morfológicas observadas nas colônias de <i>N. rosa</i> DSM 7285 e GeG2 crescidas em MM.	75
Figura 19. Padrão de crescimento indicativo de motilidade observado para as bactérias <i>N. rosa</i> DSM 7285 e GeG2 em meios sólidos (MM) inoculados pelo método de picada e resultados obtidos nos testes de hidrólise de amido para as bactérias <i>N. rosa</i> DSM 7285 e GeG2.	76

Figura 20. Perfil de lipídeos polares totais da linhagem GeG2 obtido por cromatografia bidimensional em camada delgada e marcação com ácido fosfomolibdico.....	79
Figura 21. Dendograma construído a partir da comparação dos perfis de espectros principais (MSP) obtidos das extrações proteicas das linhagens bacterianas GeG2 e <i>N. rosa</i> DSM 7285..	82
Figura 22. Eletromicrografias de varredura das células bacterianas GeG2 e <i>N. rosa</i> DSM 7285 crescidas sobre lamínulas de vidro em MM líquido por 7 dias a 28 °C.....	85
Figura 23. Eletromicrografias de varredura das células bacterianas GeG2 e <i>N. rosa</i> DSM 7285 crescidas em MM líquido por 72 horas, sem agitação, a 28 °C.....	87
Figura 24. Eletromicrografias de transmissão das células bacterianas GeG2 crescidas por 7 dias em MM a 28 °C e posteriormente criofixadas sob alta pressão (HPC).....	88
Figura 25. Culturas da bactéria GeG2 em MM líquido incubadas por 5 dias a 28 °C sob agitação ou por 2 dias a 28 °C sem agitação	90
Figura 26. Observação a fresco dos grumos da cultura da bactéria GeG2 por microscopia de contraste de fase.....	91
Figura 27. Eletromicrografia de varredura de culturas da linhagem GeG2 apresentando grumos macroscópicos crescidas em MM líquido por 14 dias a 28 °C.....	93
Figura 28. Eletromicrografias de transmissão das células bacterianas GeG2 crescidas por MM líquido por 14 dias a 28 °C e posteriormente fixadas com aldeídos.....	95
Figura 29. Grafo da montagem do genoma do isolado bacteriano GeG2 com o programa Unicycler, por meio da abordagem híbrida, utilizando-se tanto <i>long reads</i> de PacBio quanto <i>short reads</i> de Illumina.....	98
Figura 30. Mapas circulares e aspectos genéticos do cromossomo, do cromídeo e dos dois plasmídeos (pGeG2a e pGeG2b) da linhagem bacteriana GeG2.....	100
Figura 31. Relação entre as similaridades das sequências do gene rRNA 16S e os valores de ANI ou de dDDH (d_0) entre a linhagem GeG2 e seis outras linhagens do gênero <i>Novosphingobium</i> que possuem genomas sequenciados.....	105
Figura 32. Árvore filogenética baseada em genomas completos, gerada pela ferramenta TYGS (<i>Type Strain Genome Server</i>).....	107
Figura 33. Resultado da análise de identificação taxonômica do genoma da linhagem GeG2 realizada por meio da plataforma online TrueBac ID.....	107
Figura 34. Classificação baseada em categorias funcionais COG (Cluster of Orthologous Groups) das proteínas preditas no genoma do isolado GeG2.....	110
Figura 35. Distribuição funcional das proteínas preditas no cromossomo e no cromídeo da linhagem bacteriana GeG2 baseada em classificações COG.....	112
Figura 36. Distribuição de CAZymes (enzimas ativas em carboidratos) no cromossomo, cromídeo e plasmídeo pGeG2a da linhagem GeG2..	115
Figura 37. Mapa das vias relacionadas ao metabolismo central de carbono, gerado pela ferramenta KEGG Mapper - Pathways, indicando genes de enzimas presentes no	

cromossomo, no cromídeo ou em ambos os réplicons do genoma da linhagem bacteriana GeG2.....	118
Figura 38. Representação esquemática das vias de assimilação de enxofre ambiental de bactérias evidenciando os genes identificados no cromossomo e no cromídeo da linhagem bacteriana GeG2.	120
Figura 39. Clusters gênicos de biossíntese de metabólitos secundários (BGCs) encontrados no genoma da linhagem bacteriana GeG2.....	123
Figura 40. Mapa das vias relacionadas a degradação de compostos aromáticos, gerado pela ferramenta KEGG Mapper - Pathways, indicando genes de enzimas presentes no cromossomo, no cromídeo ou em ambos os réplicons do genoma da linhagem bacteriana GeG2.....	127
Figura 41. Estrutura química do corante verde malaquita e colônias da bactéria GeG2 crescidas por sete dias em MM sólido sem glicose contendo verde malaquita.....	128

Lista de abreviações

- AA – Enzimas para atividades auxiliares
- ANI - *Average Nucleotide Identity* (Identidade Média de Nucleotídeos)
- CAZymes – *Carbohydrate Active Enzymes* (Enzimas Ativas em Carboidratos)
- CBM - *Carbohydrate Binding Module* (Módulo de Ligação a Carboidratos)
- CDS – *Coding Sequences* (Sequências Codificadoras)
- COG - *Clusters of Orthologous Groups* (Grupos de Ortólogos)
- DNA - *Deoxyribonucleic acid* (Ácido Desoxirribonucleico)
- dNTP - *Deoxyribonucleotide triphosphate* (desoxirribonucleotídeo trifosfato)
- EDTA - *Ethylenediamine tetraacetic acid* (Ácido etilenodiamino tetra-acético)
- EPS - Exopolissacarídeos
- GH - Glicosil-hidrolase
- GT - Glicosil-transferase
- HPF - *High-Pressure Cryofixation* (Criofixação sob alta pressão)
- ITS – *Intergenic spacer* (Região Espaçadora Intergênica)
- KEGG - *Kyoto Encyclopedia of Genes and Genomes*
- LB – Meio de cultura Luria-Bertani
- LPSN - *List of Prokaryotic names with Standing in Nomenclature* (Lista de Nomes Procarióticos com Base na Nomenclatura)
- MALDI-TOF - *Matrix-assisted laser desorption/ionization – Time of flight* (Ionização e dessorção a laser assistida por matriz – Tempo de voo)
- MET – Microscopia Eletrônica de Transmissão
- MEV – Microscopia Eletrônica de Varredura
- MM – Meio Mínimo
- MS – *Mass Spectrometry* (Espectrometria de massa)
- MSP - *Main spectra profiles* (Perfis de espectros principais)
- NCBI - *National Center for Biotechnology Information* (Centro Nacional de Informação Biotecnológica dos Estados Unidos)
- OD – *Optical density* (Densidade óptica)
- OGRI - *Overall Genome Relatedness Index*

OMV – *Outer Membrane Vesicles* (Vesículas de Membrana Externa)

pb – pares de bases

PBS - *Phosphate-buffered saline* (Tampão salina-fosfato)

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

PL – Polissacarídeo liase

RECOR-IBGE - Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística

RNA - *Ribonucleic Acid* (Ácido ribonucleico)

rRNA – *Ribosomal RNA* (RNA ribossomal)

tRNA – *transfer RNA* (RNA transportador)

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

TE – Tampão Tris-EDTA

RESUMO

Embora o desenvolvimento de técnicas moleculares e abordagens independentes de cultivo tenha sido essencial para o acesso à grande diversidade microbiana existente em nosso planeta, o cultivo de microrganismos ainda é imprescindível para a caracterização e entendimento de diversos aspectos de sua biologia, bem como para potenciais aplicações industriais e biotecnológicas. Nesse sentido, considerando a necessidade de ampliar a quantidade de representantes cultivados, especialmente de grupos ainda pouco explorados, o objetivo desse trabalho foi cultivar e caracterizar novos isolados microbianos a partir de amostras de solo nativo de Cerrado, importante bioma brasileiro que apresenta grande riqueza microbiana. Inicialmente, objetivou-se o enriquecimento de membros do domínio *Archaea*, um grupo ainda pouco estudado e com poucos representantes não extremófilos cultivados. A utilização de estratégias de enriquecimento para archaeas resultou na identificação de organismos pertencentes aos filos Thaumarchaeota e Bathyarchaeota em co-culturas com uma bactéria. No entanto, dificuldades relacionadas às metodologias de cultivo e detecção desses organismos impossibilitaram um expressivo enriquecimento das archaeas, culminando com crescimento mais rápido e abundante da espécie bacteriana presente nas culturas. Considerando as interessantes características da bactéria presente no cultivo, denominada GeG2 e pertencente ao gênero *Novosphingobium*, a segunda etapa do trabalho envolveu seu cultivo em meios definidos e caracterização biológica. Análises fisiológicas, moleculares e quimiotaxonômicas, bem como comparações baseadas no genoma completo, indicaram tratar-se de uma nova espécie, para a qual o nome *Novosphingobium terrae* será proposto. A formação de agregados celulares envoltos por espessa matriz extracelular e estruturas nanométricas esféricas foi observada nas culturas líquidas, indicando a produção de exopolissacarídeos (EPSs) e vesículas de membrana externa. A montagem do genoma completo revelou que GeG2 apresenta um genoma segmentado possuindo, além do cromossomo (4,1 Mb) e dois plasmídeos (212 e 68 kb), um megarréplicon (2,7 Mb) identificado como um provável cromídeo, com genes possivelmente envolvidos em funções adaptativas. Análises funcionais do genoma revelaram um amplo repertório de genes codificadores de enzimas ativas em carboidratos e envolvidos em degradação de compostos aromáticos, evidenciando o potencial biotecnológico da nova espécie, principalmente relacionado à produção de EPSs e à biodegradação e compostos recalcitrantes.

ABSTRACT

Although the development of culture-independent methods has been extremely important to access and describe our planet's great microbial diversity, laboratory cultivation of microorganisms is still essential to better describe and understand their biology and biotechnological applications. Considering the importance of increasing the number of cultured representatives, especially from poorly characterized microbial groups, the aim of this project was to cultivate and characterize new microbial isolates from native soils of Cerrado, an extremely biodiverse Brazilian biome. The initial purpose of this work was the enrichment of *Archaea*, a still poorly characterized domain of life, with few non-extremophilic cultured representatives. Specific enrichment strategies resulted in the establishment of cocultures containing archaea belonging to phyla Thaumarchaeota and Bathyarchaeota and a bacterial strain. Nevertheless, limitations regarding methodologies for cultivation and detection of the previously uncultured archaeal groups identified made it impossible to significantly enrich these microorganisms, leading to bacterial overgrowth in the cultures. Considering the interesting characteristics related to the bacteria identified in the cultures, affiliated to the *Novosphingobium* genus and denominated strain GeG2, the microorganism was further isolated in defined media and characterized. Physiological, molecular, and chemotaxonomic analyses as well as whole genome comparisons revealed that GeG2 represents a novel species of the genus *Novosphingobium*, for which the name *Novosphingobium terrae* will be proposed. Cell aggregates surrounded by extracellular matrix and nanometric spherical structures were observed in liquid cultures, suggesting the production of exopolysaccharides (EPS) and outer membrane vesicles (OMVs). The complete genome was assembled revealing a 4.1 Mb chromosome and two plasmids (212 and 68 kb), in addition to a 2.7 Mb megareplicon identified as a putative chromid that contains genes potentially involved in adaptative functions. Functional analyses showed that the genome encodes a vast repertoire of carbohydrate active enzymes and genes involved in the degradation of aromatic compounds, highlighting the biotechnological potential of the new isolate obtained from Cerrado soils, especially regarding EPS production and biodegradation of recalcitrant compounds.

INTRODUÇÃO GERAL

Apesar dos grandes avanços trazidos à microbiologia pelas técnicas moleculares independentes de cultivo desenvolvidas nas últimas décadas, a obtenção de cultivos laboratoriais ainda é imprescindível para a elucidação de diversos aspectos biológicos relacionados aos microrganismos. Um número crescente de estudos vem revelando a grande diversidade microbiana existente em nosso planeta, evidenciando que a maior parte dos microrganismos não é cultivada em laboratório, o que limita a compreensão de suas características fisiológicas, estruturais, moleculares e de seus papéis ecológicos e potenciais biotecnológicos.

No entanto, o cultivo de novos microrganismos, principalmente aqueles associados a amostras ambientais complexas, é uma tarefa laboriosa e demorada, apresentando muitos desafios. Por essa razão, juntamente com o crescente avanço de metodologias que permitem a identificação de moléculas microbianas diretamente do ambiente, as abordagens baseadas em cultivo têm sido cada vez menos empregadas pelos microbiologistas. Entretanto, o acúmulo de dados gerados pelas novas tecnologias independentes de cultivo e as inúmeras hipóteses microbiológicas advindas desse cenário também evidenciaram a importância de se estabelecer novos cultivos microbianos, o que tem sido crescentemente ressaltado na literatura.

Nesse sentido, o presente trabalho de doutorado visou o cultivo e a caracterização de novos microrganismos a partir do solo do bioma Cerrado, um hábitat que contém uma grande diversidade microbiana ainda inexplorada. Para melhor explorar os trabalhos realizados e os resultados obtidos, essa tese será dividida em dois capítulos. O primeiro capítulo abordará o estabelecimento e caracterização de co-cultivos entre organismos dos domínios *Archaea* e *Bacteria* a partir de amostras de solo de Cerrado. O segundo capítulo descreverá a caracterização genômica, morfológica e fisiológica de uma nova espécie bacteriana pertencente ao gênero *Novosphingobium*, isolada a partir dos cultivos estabelecidos.

CAPÍTULO 1 – Estabelecimento e caracterização de co-cultivos de archaeas e bactérias obtidos a partir de amostras de solo de Cerrado

1.1. Introdução

1.1.1. O cultivo laboratorial de microrganismos e o início da microbiologia

A descoberta que as bactérias podiam ser cultivadas em fluidos nutrientes, o desenvolvimento de técnicas de cultivo em meios sólidos e a obtenção das primeiras culturas microbianas puras por Robert Koch, no final do século XIX, representaram um grande avanço para o estudo dos microrganismos e contribuíram na consolidação da microbiologia (Koch, 1881; Blevins e Bronze, 2010). Nos anos que se seguiram, a adição de fontes nutricionais variadas aos meios e o desenvolvimento de novos materiais e técnicas para o cultivo permitiram o estabelecimento de culturas de diversas bactérias em laboratório (Petri, 1887; Fuller, 1895; Hitchens e Leikind, 1939; Blevins e Bronze, 2010). As culturas puras obtidas possibilitaram caracterizações morfológicas, fisiológicas e metabólicas detalhadas de microrganismos, bem como o estabelecimento dos primeiros critérios para sua classificação taxonômica.

Como o enfoque dos microbiologistas da época visava principalmente a identificação de organismos causadores de doenças, muitos agentes etiológicos foram isolados e caracterizados, contribuindo para um grande desenvolvimento da microbiologia clínica (Koch, 1876; Koch, 1882; Kitasato, 1889; MacCallum e Hastings, 1899; Brown, 1919). Gradativamente, o espectro de estudos microbiológicos foi ampliado e a diversidade microbiana em habitats naturais como solo e água, assim como as atividades metabólicas realizadas pelos microrganismos nestes ambientes, passou a ser investigada. Martinus Beijerinck e Sergei Winogradsky foram importantes pesquisadores de bactérias de ambientes naturais no final do século XIX e princípio do século XX, tendo isolado e descrito muitas espécies envolvidas nos ciclos do nitrogênio e enxofre (Dworkin, 2012). Beijerinck desenvolveu os conceitos de meios seletivos e culturas de enriquecimento, manipulando a composição nutricional dos meios de forma a favorecer o cultivo de determinados microrganismos e inibir o crescimento de outros, o que possibilitou o isolamento de bactérias fixadoras de nitrogênio a partir de nódulos

radiculares de leguminosas (Beijerinck, 1888). Winogradsky, também utilizando-se de técnicas de cultura de enriquecimento, obteve culturas puras de várias bactérias aquáticas e de solo envolvidas na nitrificação, redução de sulfato, oxidação de enxofre e ferro, evidenciando seu papel em transformações biogeoquímicas e culminando na proposta do conceito da quimiolitotrofia (Winogradsky, 1890; Dworkin, 2012).

Apesar dos rápidos avanços trazidos à microbiologia pela obtenção de cultivos laboratoriais puros, os desafios e limitações impostos pela necessidade do isolamento de microrganismos em laboratório como única forma de estudá-los são enfatizados pelos microbiologistas desde esta época (Chick, 1905; Henrici, 1933). Ao longo de seus trabalhos com solos, Winogradsky, mesmo reconhecendo que o isolamento de culturas puras era essencial para o estudo da fisiologia microbiana, passou a defender métodos de observação microscópica direta do solo como uma forma mais fiel de se investigar a sua complexidade microbiológica (Ackert Jr, 2007). A dificuldade de obtenção de culturas puras, principalmente de bactérias ambientais, e a incapacidade de se acessar toda a diversidade microbiana por esses métodos ficou cada vez mais evidente nas décadas que se seguiram.

Em 1985, Staley e Konopka cunharam um termo para descrever a diferença entre a diversidade microbiana observada nos ambientes naturais e a quantidade de microrganismos cultiváveis em meios artificiais – a “grande anomalia da contagem em placas”. Os autores relataram uma discrepância de várias ordens de grandeza entre o número de organismos de amostras ambientais quantificados por contagens microscópicas diretas e o número de colônias observadas em meios artificiais, ou seja, o número de células viáveis.

Embora tenham sido feitas estimativas de que existam cerca de $4 \text{ a } 6 \times 10^{30}$ células procarióticas em nosso planeta (Rosselló-Mora e Amman, 2001) e projeções recentes indiquem que o número de espécies microbianas pode chegar a um trilhão (10^{12}) (Locey e Lennon, 2016), sabe-se que a fração da diversidade microbiana cultivada é substancialmente menor. Dados de Amann *et al.* (1995) sugeriam que somente 0,001 a 0,1% dos microrganismos marinhos e 0,3% de microrganismos do solo quantificados por contagem direta era cultivável, enquanto estudos posteriores de Torsvik e colaboradores (2002) relataram que apenas 1% da diversidade microbiana descrita em nosso planeta tinha representantes cultivados. Assim, com a identificação e caracterização dos microrganismos ainda dependendo do seu isolamento em culturas puras, os estudos em microbiologia se encontraram limitados durante muito tempo.

No entanto, a partir da década de 70, tal cenário começou a mudar. O conceito que informações acerca da história evolutiva de um organismo poderiam ser extraídas do seu material genético, introduzido por Zuckerkandl e Pauling (1965), deu início a uma nova era na microbiologia e abordagens moleculares passaram a ser utilizadas no estudo dos procariotos (Fox *et al.*, 1977).

1.1.2. A microbiologia na era das abordagens moleculares e a proposta do domínio *Archaea*

O entendimento que o genoma carrega informações que podem ser utilizadas para o estudo de relações evolutivas entre os organismos foi um grande marco em várias áreas da biologia, mas teve importância especial na microbiologia. A taxonomia dos procariotos até então era baseada na observação de características morfológicas, bioquímicas e fisiológicas que, apesar de muito úteis, eram instáveis e não possibilitavam o estabelecimento das suas relações evolutivas (Krieg, 1988). Desta forma, a implementação de critérios genômicos e o desenvolvimento de ferramentas moleculares impulsionaram grandes mudanças e avanços na taxonomia e sistemática microbiana.

Inicialmente, a comparação entre a composição de bases dos genomas bacterianos (conteúdo de G+C) foi utilizada com propósito de classificação taxonômica, mas os valores obtidos permitiam apenas classificações superficiais (Krieg, 1988). O desenvolvimento de técnicas de hibridização DNA-DNA, no início da década de 70, trouxe novas contribuições para a sistemática bacteriana e permitiu o agrupamento dos organismos de acordo com suas semelhanças genéticas. Essa técnica baseia-se na premissa que, sob condições padronizadas, fitas de DNAs de organismos diferentes se reassociarão, dependendo do grau de similaridade de suas sequências de nucleotídeos, permitindo assim uma quantificação de similaridade entre estas (Rosselló-Mora & Amann, 2001).

No final da década de 70, um grupo de microbiologistas liderado por Carl Woese, buscando formas de inferir a relação evolutiva entre os organismos por meio de suas moléculas, desenvolveu um método inovador para agrupar e classificar microrganismos, sob uma óptica filogenética. Tal método consistia na catalogação e comparação das sequências nucleotídicas dos RNAs ribossômicos (rRNAs), moléculas consideradas excelentes cronômetros evolutivos por serem universais, ou seja, encontradas em todos os organismos, e apresentarem homologia funcional e evolutiva (Fox *et al.*, 1977; Olsen

e Woese, 1993). Além disso, Woese defendia que sequências de rRNA, por apresentarem regiões de extrema conservação e regiões hipervariáveis, possibilitavam que tanto as relações filogenéticas distantes quanto as próximas pudessem ser examinadas (Fox *et al.*, 1977). Foi constatado que o tamanho da molécula utilizada é uma característica essencial nas reconstruções filogenéticas, já que cada posição individual na sequência carrega informações de um estreito tempo evolutivo e um número maior de regiões na sequência aumenta a quantidade de informações filogenéticas que podem ser detectadas (Ludwig e Schleifer, 1994). Desta forma, dentre os diferentes RNAs ribossomais existentes, os rRNA 16S de procariotos (com aproximadamente 1500 nucleotídeos) e 18S de eucariotos (com aproximadamente 1900 nucleotídeos) foram escolhidos pelos autores por apresentarem tamanho adequado para essas análises (Olsen *et al.*, 1986).

As moléculas de rRNA 16S se mostraram muito eficientes como marcador filogenético e passaram então a ser empregadas por Woese e colaboradores nos seus estudos de sistemática procariótica. Por meio da comparação destas moléculas obtidas a partir de diversas bactérias cultivadas na época, os autores observaram que os procariotos podiam ser divididos em dois grupos filogeneticamente independentes, denominando-os então de “eubactérias” e “archaeobactérias” (Woese e Fox, 1977). Nesta divisão, o reino das eubactérias seria composto por todas aquelas bactérias até então já bem conhecidas e caracterizadas, enquanto o reino das archaeobactérias seria composto pelas bactérias produtoras de metano, genericamente chamadas de metanogênicas. A denominação “archaeobactérias” foi escolhida por acreditar-se, na época, que as bactérias metanogênicas seriam ancestrais das bactérias atuais.

Estudos posteriores levaram Woese e colaboradores a propor, em 1990, um novo sistema de classificação taxonômica para os seres vivos. Esse sistema, de natureza filogenética, propunha a criação de um táxon superior a Reino, denominado Domínio, e a divisão de todos os seres vivos nos domínios *Eukarya*, *Bacteria* e *Archaea*. Desta forma, todos os eucariotos estariam incluídos em apenas um domínio e os procariotos seriam divididos em dois grupos. Além disso, as até então denominadas archaeobactérias passariam a ser chamadas apenas de *Archaea*, evidenciando-se assim as grandes diferenças filogenéticas entre estes organismos e as bactérias. Apesar de tal proposta ter sido inicialmente alvo de muitas controvérsias, a crescente utilização de abordagens de filogenia molecular, reforçou sua solidez e coerência, sedimentando o Domínio *Archaea* como um grupo distinto de procariotos, até então pouco conhecido e estudado. Além das metanogênicas, organismos halófilos e hipertermófilos, à época considerados bactérias,

foram reconhecidos como archaeas, com características estruturais e moleculares únicas (Magrum *et al.*, 1978; Stetter *et al.*, 1990). Tais características incluíam, além dos rRNAs filogeneticamente distintos, a ausência de peptidoglicano em suas paredes celulares, maquinarias de replicação e transcrição mais semelhantes a de eucariotos e lipídeos de membrana compostos por cadeias laterais de hidrocarbonetos ligadas a moléculas de glicerol-1-fosfato por de ligações do tipo éter (Woese *et al.*, 1978; Huet *et al.*, 1983; Kates *et al.*, 1993).

A revolução trazida pela utilização de marcadores filogenéticos na identificação e classificação dos procariotos foi acompanhada pelo rápido desenvolvimento de ferramentas e técnicas moleculares, que passaram a permitir de forma relativamente simples a manipulação da ácidos nucleicos e proteínas. O desenvolvimento das técnicas de reação em cadeia da polimerase (PCR) e sequenciamento automático de DNA abriram um novo horizonte no estudo de microrganismos, possibilitando a identificação e sequenciamento de genes específicos a partir das mais variadas amostras (Weisburg *et al.*, 1991; Fuhrman *et al.*, 1993; Marchesi *et al.*, 1998). Assim, o isolamento e clonagem de genes podiam ser realizados diretamente a partir de amostras ambientais, expandindo os estudos em ecologia microbiana e permitindo que, a partir da amplificação de marcadores filogenéticos como o gene rRNA 16S de procariotos (e 18S de eucariotos), a detecção e o posicionamento filogenético de microrganismos presentes nos mais diversos ambientes naturais pudessem ser realizados sem a necessidade de cultivá-los em laboratório (Amann *et al.*, 1995).

Desde então, muitos microbiologistas passaram a utilizar essa abordagem e milhares de sequências gênicas de bactérias e archaeas foram detectadas em diversos ambientes, reforçando a noção de que as espécies não cultiváveis correspondem à principal fração da diversidade microbiana de nosso planeta (Handelsman, 2004). Por meio destes trabalhos, muitas descobertas acerca da diversidade, complexidade e distribuição das comunidades de microrganismos foram feitas, com a proposta de inúmeros novos táxons. Foi constatado, por exemplo, que as archaeas, até então consideradas restritas a ambientes extremos, são ubíquas e apresentam distribuição semelhante àquela das bactérias (DeLong, 1992; DeLong, 1998). Genes de archaeas passaram a ser detectados nos mais variados ambientes e análises do gene de rRNA 16S revelaram uma grande diversidade destes organismos em habitats mesófilos, culminando na descoberta de novos filos e expandindo significativamente o conhecimento acerca deste domínio da vida (Barns *et al.*, 1996; Brochier-Armanent *et al.*, 2008; Meng *et al.*,

2014). As abordagens moleculares também revelaram novos grupos bacterianos numericamente importantes no solo e em habitats aquáticos, mas que não haviam sido identificados anteriormente pelas metodologias tradicionais de cultivo (Forney *et al.*, 2004).

Mais recentemente, o advento das tecnologias de sequenciamento de nova geração expandiu mais ainda as fronteiras dos estudos microbiológicos independentes de cultivo, permitindo o sequenciamento do conteúdo genético de comunidades microbianas inteiras em pouco tempo e a um custo acessível, dando início a era da metagenômica (Shokralla *et al.*, 2012; Bragg e Tyson, 2014). Sequenciamentos de alto rendimento, aliados ao emprego de ferramentas computacionais, possibilitam o acesso a vários genomas microbianos diretamente do ambiente, fazendo com que um vasto repertório gênico de bactérias e archaeas de ecossistemas naturais e artificiais possa ser analisado e o seu potencial metabólico predito (Lazar *et al.*, 2017; Jungbluth *et al.*, 2017; Cai *et al.*, 2020).

Tais metodologias trouxeram muitas vantagens em relação às análises moleculares baseadas em amplificação e sequenciamento de genes individuais específicos, especialmente por serem capazes de revelar grupos microbianos completamente novos que não tiveram seus genes de rRNA 16S identificados anteriormente, por não apresentarem grau de similaridade suficiente para o pareamento efetivo com iniciadores padrão utilizados (Baker *et al.*, 2010, Brown *et al.*, 2015). Além disso, estudos metagenômicos evidenciaram que mesmo táxons relativamente bem conhecidos podem conter genes novos localizados em ilhas genômicas hipervariáveis e revelou a variabilidade genômica a nível de linhagem em populações microbianas naturais (Baker e Dick, 2013).

Análises funcionais de RNAs e proteínas também passaram a ser amplamente utilizadas nos estudos de comunidades ambientais de microrganismos. A metatranscritômica e a metaproteômica têm sido empregadas em estudos de ecologia microbiana visando responder questões relacionadas à expressão gênica, funções ecológicas e interações entre as espécies nas comunidades microbianas (Coll-Lladó *et al.*, 2011; Wang *et al.*, 2016; Cai *et al.*, 2020). Em resumo, o emprego destas novas técnicas tem auxiliado a elucidar muitos aspectos na microbiologia, trazendo informações valiosas sobre a distribuição e abundância dos microrganismos nos ambientes naturais, seus possíveis papéis nos ciclos biogeoquímicos e respostas adaptativas a mudanças ambientais, além de seu potencial biotecnológico (Mehetre *et al.*, 2016; Coughlan *et al.*, 2015; Kumar *et al.*, 2016; Green e Neufeld, 2016; Gu *et al.*, 2019). Além disso, essas

abordagens permitiram o acesso a uma grande diversidade de bactérias e archaeas previamente desconhecida, expandindo substancialmente a visão sobre a árvore da vida em nosso planeta (Hug *et al.*, 2016; Parks *et al.*, 2017; Castelle *et al.*, 2018).

No entanto, apesar das enormes contribuições trazidas à microbiologia por todas essas abordagens moleculares e de acreditar-se por algum tempo que estas superariam a necessidade do estabelecimento de cultivos laboratoriais, a importância da obtenção de culturas microbianas tem sido cada vez mais ressaltada na literatura.

1.1.3. A importância do cultivo na era das “ômicas”

Com a popularização das tecnologias de sequenciamento em larga escala, as abordagens “multi-ômicas” (metagenômica, metatranscritômica e metaproteômica) passaram a ser a principal tendência nos estudos de microbiologia ambiental e, como já mencionado, têm contribuído de forma significativa para responder muitas questões acerca da ecologia e filogenia microbiana. No entanto, ainda existem diversas dificuldades relacionadas à interpretação dos dados adquiridos por meio dessas técnicas, limitando a compreensão de muitos aspectos das comunidades microbianas. Em primeiro lugar, desafios relacionados à enorme quantidade de dados gerados, à necessidade do desenvolvimento de ferramentas computacionais avançadas para analisá-los e à falta de padronizações metodológicas e analíticas dificultam a correlação da grande diversidade detectada em muitos ambientes com os papéis metabólicos dos microrganismos nessas comunidades (Nesme *et al.*, 2016). Discrepâncias nos protocolos de extração, amplificação, sequenciamento, análises de bioinformática e bancos de dados usados para as anotações podem influenciar quantitativa e qualitativamente os resultados obtidos, levando a vieses na caracterização das populações microbianas e possíveis interpretações errôneas (Delmont *et al.*, 2013). Além disso, hipóteses ecológicas feitas a partir das abordagens de “multi-ômicas” precisam ser experimentalmente validadas, para que as capacidades metabólicas e funcionalidades microbianas preditas por esses estudos possam ser confirmadas.

Neste sentido, o cultivo dos microrganismos ainda é imprescindível. A investigação detalhada de mecanismos metabólicos e fisiológicos em nível bioquímico só pode ser realizada a partir de isolados microbianos cultivados, assim como a validação de mecanismos de interação entre espécies, princípios evolutivos e dinâmicas populacionais (Gutleben *et al.*, 2018). Outro aspecto a ser considerado é que, tendo em vista as

dificuldades e problemas ainda existentes na reconstrução de genomas a partir de amostras microbianas complexas, a disponibilidade do microrganismo em cultivo, seja em culturas puras ou de enriquecimento, possibilita o sequenciamento e reconstrução de genomas com maior qualidade e precisão (Lagier *et al.*, 2015). O cultivo também é essencial para a anotação e caracterização funcional de novos genes, já que muito do que é sequenciado atualmente não apresenta correspondência nos bancos de dados, fazendo com que sua função não possa ser predita (Muller *et al.*, 2013). De fato, a descoberta de microrganismos com capacidades metabólicas até então desconhecidas, como por exemplo archaeas metanogênicas degradadoras de compostos aromáticos metoxilados derivados de lignina, ou bactérias intestinais indutoras de diabetes, só pôde ser realizada por meio de abordagens baseadas em cultivo, já que novos fenótipos biológicos ainda não possuem seus componentes genéticos descritos, não sendo possível a sua elucidação por técnicas exclusivamente moleculares (Tamaki, 2019).

Como descrições aprofundadas de aspectos metabólicos, fisiológicos, estruturais e até mesmo genéticos não são possíveis sem a obtenção de culturas laboratoriais, este ainda é um critério fundamental para a validação de uma nova espécie microbiana (Parte, 2014). No entanto, vale ressaltar que grande parte da diversidade procariótica descrita atualmente não possui representantes cultivados, com muitos novos taxa propostos apenas a partir de sequências de DNA obtidas diretamente de amostras ambientais. Estima-se que menos de 6 % das espécies de bactérias e archaeas contidas no banco de dados SILVA tenham sido oficialmente validadas por testes fisiológicos em isolados microbianos, como previsto pela LPSN (*List of Prokaryotic Names with Standing in Nomenclature*) (Gutleben *et al.*, 2017). Estudos de Schloss *et al.* (2016) evidenciam o grande número de linhagens microbianas que não dispõem de membros cultivados, estimando que dentre 65 filos bacterianos propostos na época apenas 24 possuíam representantes cultivados, enquanto que para as archaeas esse número era ainda mais impressionante: apenas quatro dos 20 filos propostos continham representantes em cultivo.

Assim, tendo em vista o crescente número de dados trazidos pelas abordagens “multi-ômicas” e todas as hipóteses que vêm sendo levantadas a partir destas, a necessidade da obtenção de novas culturas de microrganismos, principalmente de origem ambiental, é cada vez mais evidente e urgente. Debates relacionados às dificuldades associadas ao cultivo e propostas de soluções para superá-las têm sido feitos por alguns autores (Zengler *et al.*, 2002; Alain & Querellou, 2009).

Há muito se discute que os ambientes terrestres e aquáticos provavelmente possuam muitos microrganismos oligotróficos e que sua transferência para os meios laboratoriais, ricos em nutrientes, gere um impacto metabólico que possa inibir seu crescimento (Roszak e Colwell, 1987; Watve *et al.*, 2000). Propostas da utilização de meios formulados a partir de componentes existentes nos ambientes naturais são cada vez mais frequentes e culturas de microrganismos que nunca haviam sido obtidas puderam ser estabelecidas utilizando-se métodos que mimetizavam essas condições nutricionais (Connon & Giovannoni, 2002; Rappé *et al.*, 2002; Gavrish *et al.*, 2008).

Outro aspecto importante a ser considerado é que muitos microrganismos vivem em comunidades complexas em seus ambientes naturais e parecem incapazes de crescer isoladamente, na ausência de outros microrganismos do mesmo ambiente, conhecidos como organismos auxiliares. A necessidade de co-cultivos para o crescimento microbiano foi relatada em diversos trabalhos (Suzuki *et al.*, 1988; Kaeberlein *et al.*, 2002, Tanaka *et al.*, 2004, Morris *et al.*, 2008; Park *et al.*, 2010) e alguns mecanismos que determinam a dependência de um organismo para o crescimento de outro já foram propostos. Por exemplo, a incapacidade de obtenção de colônias isoladas de bactérias do gênero *Prochlorococcus* e a percepção de que esses organismos eram capazes de crescer quando associados a outros levou à descoberta que essas bactérias são suscetíveis à ação de espécies reativas de oxigênio e que a remoção destes compostos por outros microrganismos possibilitava o seu crescimento (Morris *et al.*, 2008). Posteriormente constatou-se que o genoma de *Prochlorococcus* não possuía genes de catalase e de outras vias de proteção comuns em aeróbios e que a vasta comunidade microbiana da camada superficial oligotrófica dos oceanos exercia papel fundamental na proteção dessas bactérias contra danos oxidativos (Morris *et al.*, 2011).

Abordagens inovadoras empregando tecnologias como microchips de cultivo (Ingham *et al.*, 2007), dispositivos de microfluidos e encapsulamento de microrganismos em microgotas (Hesselman *et al.*, 2012; Ma *et al.*, 2014; Boitard *et al.*, 2015), além de manipulações de células individuais (Zengler *et al.*, 2005, Ishii *et al.*, 2010) também vêm obtendo sucesso na triagem e cultivo de novos microrganismos de comunidades ambientais complexas. No entanto, apesar da expressiva contribuição trazida por essas abordagens recentes, esforços para a obtenção de novas culturas laboratoriais ainda são fundamentais para uma melhor caracterização da grande diversidade microbiana existente em nosso planeta.

É importante ressaltar que especificamente em relação às archaeas, o fato deste domínio ter sido caracterizado principalmente por abordagens filogenéticas durante o apogeu das técnicas moleculares, a maioria das informações que possuímos atualmente sobre esses organismos é oriunda de trabalhos com esse enfoque. Desta forma, grande parte das culturas de archaeas disponíveis até hoje foram obtidas antes mesmo destes organismos serem reconhecidos como pertencentes a um domínio à parte (Pine e Barker, 1954; Edwards e McBride, 1975; Balch *et al.*, 1979; Sowers *et al.*, 1984; Mathrani *et al.*, 1988). Além disso, apesar da grande distribuição de archaeas nos mais variados ambientes naturais e artificiais, a maioria dos cultivos em meios artificiais consiste em espécies extremófilas, por serem conhecidas há mais tempo e demandarem condições de crescimento mais específicas. No entanto, excetuando-se as archaeas metanogênicas, o cultivo de archaeas de ambientes terrestres e aquáticos mesófilos continua sendo um desafio, com poucos relatos na literatura (Könneke *et al.*, 2005; Simon *et al.*, 2005; Tourna *et al.*, 2011; Kim *et al.*, 2012; Lehtorvita-Morley *et al.*, 2014) e muitas vezes consistindo de co-cultivos entre archaeas e bactérias.

Se compararmos o número de sequências de DNA de archaeas depositadas em bancos de dados com o número de organismos deste domínio cultivados, a dificuldade no estabelecimento de cultivos laboratoriais fica ainda mais evidente. Um levantamento realizado em 2019 revelou que, até aquele momento, 3634 genomas de archaeas estavam disponíveis no banco de dados do NCBI (*National Center for Biotechnology Information*), enquanto apenas 599 espécies de archaea eram reconhecidas pela LPSN (Sun *et al.*, 2019). Esses dados ressaltam o crescimento exponencial das investigações genômicas de archaeas nos últimos anos, já que apenas oito anos antes - em 2011 - a marca de 100 genomas de archaeas publicados estava sendo alcançada (Brochier-Armanet *et al.*, 2011). Com isso, a necessidade de esforços para o estabelecimento de novos cultivos de archaeas vem sendo cada vez mais frisada por pesquisadores da área, que ressaltam a importância de culturas laboratoriais para a expansão e confirmação das muitas predições metabólicas, ecológicas e evolutivas advindas das abordagens “multi-ômicas” (Adam *et al.*, 2017; Sun *et al.*, 2019).

1.1.4. A grande diversidade microbiana nos solos do bioma Cerrado

O Cerrado é um bioma extenso e diverso, com área aproximada de 2 milhões de quilômetros quadrados, o que representa cerca de 22% do território brasileiro (Oliveira-

Filho & Ratter, 2002). Este bioma é composto por um mosaico de vegetações e, devido ao grande número de espécies endêmicas e à crescente degradação de seus habitats naturais, é considerado um *hotspot* de biodiversidade e esforços para conservação (Myers et al., 2000).

A grande riqueza e diversidade microbiana encontrada em solos de Cerrado nativos e submetidos a ação antropogênica tem sido descrita (Castro *et al.*, 2008; Quirino *et al.*, 2009; Bresolin *et al.*, 2010; Araújo *et al.*, 2012; Silva, 2012; Araujo *et al.*, 2018), e a identificação de um amplo repertório gênico de bactérias e fungos nos solos desse bioma, com vias metabólicas diversas e um grande potencial biotecnológico vem sendo crescentemente reportada nos últimos anos (Alves-Prado *et al.*, 2010; Peixoto *et al.*, 2017). Já as informações acerca da diversidade e funções de organismos do domínio *Archaea* nos solos de Cerrado são escassas, com poucos relatos na literatura. Por meio da amplificação e análises filogenéticas de genes de rRNA 16S, Catão *et al.* (2013) e Araujo *et al.* (2018) descreveram e compararam a diversidade de archaeas em solos de diferentes fitofisionomias de Cerrado, revelando a presença de várias linhagens de archaeas nesses habitats. Mais recentemente, genes de grupos de *Archaea* que ainda não possuem representantes cultivados, bem como de clados associados a uma importante etapa do ciclo biogeoquímico do nitrogênio - a oxidação de amônia - foram identificados em solos de Cerrado *sensu stricto* protegidos do fogo e submetidos a queimadas bienais (Belmok *et al.*, 2019).

Neste sentido, tendo em vista o exposto acima, a obtenção de culturas de microrganismos presentes nos solos de Cerrado é fundamental para uma melhor compreensão de muitos de seus aspectos biológicos, essenciais para a elucidação de suas funções ecológicas neste importante bioma, bem como de novas vias metabólicas e potenciais biotecnológicos. Ademais, o estabelecimento de cultivos de archaeas de solo, cuja biologia permanece amplamente desconhecida por existirem poucos representantes cultivados, é essencial para uma melhor caracterização dos membros não extremófilos deste domínio da vida.

Por essas razões, parte do trabalho desenvolvido em meu mestrado visou a obtenção de culturas laboratoriais de novos isolados microbianos, principalmente de organismos pertencentes ao domínio *Archaea*, oriundos do solo do Cerrado (Dias, 2015). Para tal, meios artificiais, confeccionados a partir do próprio solo e suplementados com agentes antimicrobianos, foram inoculados e incubados, e resultados preliminares sugeriram o estabelecimento de um co-cultivo entre archaeas e bactérias (Dias, 2015).

Visando dar continuidade a essa linha de pesquisa, o trabalho descrito neste capítulo teve como principal objetivo a caracterização dos co-cultivos microbianos estabelecidos a partir do solo de Cerrado e, por meio do emprego de estratégias seletivas, o enriquecimento de membros do domínio *Archaea* nas culturas obtidas.

1.2. Objetivos

1.2.1. Objetivo geral

- Analisar e caracterizar a composição microbiana de co-cultivos obtidos a partir de solos do bioma Cerrado e empregar estratégias visando o enriquecimento nas culturas de microrganismos do domínio *Archaea* ainda não cultivados.

1.2.2. Objetivos específicos

- Enriquecer membros do domínio *Archaea* nos cultivos obtidos, por meio do emprego de condições seletivas para esses organismos;
- Caracterizar morfológica e estruturalmente os microrganismos presentes no co-cultivo;
- Identificar filogeneticamente os microrganismos cultivados, empregando como marcador molecular o gene codificador do rRNA 16S;
- Sequenciar o genoma completo dos microrganismos obtidos no cultivo;

1.3. Material e Métodos

1.3.1. Coleta dos solos

Como descrito em Dias (2015), as amostras de solo para a confecção dos primeiros meios de cultura e utilização como inóculo inicial foram coletadas no dia 29 de janeiro de 2014, na Reserva Ecológica do Instituto Brasileiro de Geografia e Estatística (RECOR-IBGE), localizada a 35 km ao sul do centro de Brasília, na BR 251 (15°55' S, 47°51' W). A RECOR pertence à Área de Proteção Ambiental (APA) Distrital Gama-Cabeça de Veado.

O solo foi coletado em uma área de Cerrado *sensu stricto* pertencente ao Projeto Fogo, coordenado pela Prof. Dr. Heloísa Sináora Miranda, do Departamento de Ecologia da Universidade de Brasília. A área de coleta é denominada parcela Controle, por ser protegida de queimadas há mais de 15 anos. As amostras foram coletadas na faixa de 0 a 10 cm de profundidade e uma parte da amostra de solo foi enviada para a avaliação de seus parâmetros químicos pela empresa SoloQuímica LTDA., enquanto outra parte foi utilizada para o preparo de meios de cultura e para a realização do inóculo inicial no mesmo dia da coleta. O restante do solo foi estocado a 4°C para o preparo de meios de cultura posteriormente. Amostras adicionais de solo foram coletadas anualmente na mesma área da primeira coleta para a confecção de novos meios de cultura e manutenção dos cultivos.

1.3.2. Confecção do meio de cultivo, inóculo inicial e manutenção das culturas

Com o intuito de mimetizar as condições nutricionais do ambiente natural, os meios de cultura foram confeccionados empregando-se o mesmo solo que foi utilizado como inóculo. Para o preparo de meio sólido, o solo foi homogeneizado em água destilada (concentração final de 5 % - p/v) e mantido em repouso por 5 a 10 minutos para permitir a deposição das partículas maiores. Em seguida, o sobrenadante foi coado em coador de papel, transferido para um novo frasco, acrescido de ágar (concentração final de 2 % - p/v) e esterilizado em autoclave. Para o preparo de meio líquido, o coado foi filtrado em membranas com poros de 0,22 µm, a fim de se retirar partículas maiores que tivessem

passado pelo coador e também para esterilizar o meio. O pH final dos meios se encontrava na faixa de 6,5 a 7. Os seguintes agentes antimicrobianos (antibióticos e antifúngicos) foram adicionados aos meios de maneira sucessiva, com a finalidade de inibir o crescimento de bactérias e fungos: ampicilina (150 µg/mL), estreptomicina (50 µg/mL), cloranfenicol (20 µg/mL), polimixina B (20 µg/mL), itraconazol (0.25 mg/mL) e cetoconazol (30 µg/mL).

O inóculo inicial foi obtido a partir da mistura do solo em água destilada estéril (concentração final de 20 % - p/v) deixada em repouso por 2 a 5 minutos para a deposição de partículas maiores. Diferentes volumes do sobrenadante (50, 75 e 100 µL) foram inoculados em placas de Petri contendo meio sólido acrescido dos agentes antimicrobianos. Diluições seriadas do inóculo inicial (1/10, 1/100, 1/1000 e 1/10000) foram preparadas em meio líquido e, após duas semanas de incubação a 28°C, alíquotas de 20, 50 e 100 µL das culturas foram semeadas no meio sólido. As placas de cultura foram mantidas em estufa a 28°C e repicadas aproximadamente uma vez por mês. Os tubos de cultura em meio líquido também foram mantidos na estufa a 28°C.

1.3.3. Extração de DNA genômico total

Para a extração de DNA das culturas crescidas em meios sólidos, as colônias foram raspadas com o auxílio de uma alça de platina e homogeneizadas em 485 µL de solução de lise, contendo tampão TE (25 mM de Tris-HCl, pH 8, e 10 mM de EDTA), proteinase K (200 µg/mL) e RNase A (100 µg/mL) em tubo tipo eppendorf. Em seguida, 15 µL de SDS 20 % foi acrescentado e os tubos incubados a 37°C por 1 h. Após este período, um volume igual de clorofane (fenol, clorofórmio e álcool iso-amílico, na proporção 25:24:1) foi adicionado às amostras, que foram vigorosamente homogeneizadas e centrifugadas a 7000 g por 3 minutos, a 4°C. A fase aquosa do sobrenadante foi então transferida para um novo tubo tipo eppendorf e um volume igual de clorofil (clorofórmio e álcool iso-amílico, na proporção 24:1) foi adicionado. Após serem mais uma vez bem misturadas, as amostras foram centrifugadas novamente a 7000 g por 3 minutos, a 4°C. Os sobrenadantes resultantes foram transferidos para novos tubos, aos quais foram adicionados NaCl (para uma concentração final de 0,3 M) e 2,5 volumes de etanol 100 % gelado. Os tubos foram incubados a -20°C e após 24 h foram centrifugados a 12000 g por 5 minutos a 4°C, o sobrenadante descartado e 500 µL de etanol 70 % foram adicionados às amostras. Os tubos foram então mais uma vez

centrifugados nas mesmas condições anteriores e então mantidos invertidos, à temperatura ambiente, para a secagem do sedimento. Após a completa secagem, estes foram ressuspensos em 50 µL de H₂O milli-Q e estocados a -20°C.

Para a extração das culturas crescidas em meio líquido, estas foram filtradas em membranas de 0,22 µm (membrana polietersulfona – PES, Kasvi), que foram então retiradas do filtro e divididas em quatro pedaços. Cada pedaço foi então inserido em um tubo tipo eppendorf contendo 500 µL da mesma solução de lise descrita acima. Todas estas etapas foram realizadas sob condições estéreis. Os tubos contendo as membranas e a solução de lise foram então agitados em vórtex por 1 minuto e 15 µL de SDS 20 % foram acrescentados. Após incubação de 1 h a 37°C, o líquido de cada tubo foi transferido para tubos novos e o restante da extração foi realizada pelo mesmo protocolo descrito acima, para as culturas em meio sólido. A qualidade e concentração dos DNAs genômicos obtidos foram analisadas por eletroforese em gel de agarose 1 %, com auxílio do marcador de massa molecular *High Mass* da Invitrogen, ou quantificada no sistema QUBIT (Thermo Fisher Scientific).

1.3.4. Ensaios de PCR

Visando a identificação filogenética dos microrganismos presentes no cultivo, foram realizados ensaios de PCR com iniciadores específicos para o gene rRNA 16S dos domínios *Archaea* e *Bacteria*. Ao todo, foram utilizados quatro pares de iniciadores: 21f/958r (DeLong, 1992) e 109f/915r (Großkopf *et al.*, 1998, Banning *et al.*, 2005), definidos como universais para o domínio *Archaea*, Bathy442f/Bathy644r, específico para o filo *Bathyarchaeota* (Yu *et al.*, 2017) e 27f/1492r (Lane, 1991) universal para o domínio *Bacteria*. Os iniciadores utilizados, suas sequências nucleotídicas e os tamanhos dos fragmentos de DNA esperados nas amplificações estão listados na Tabela 1. As regiões variáveis do gene rRNA 16S amplificadas com os diferentes iniciadores para o domínio *Archaea* empregados estão mostradas na Figura 1.

Tabela 1. Iniciadores para o gene rRNA 16S empregados nos ensaios de PCR realizados com o DNA extraído dos cultivos de solo de Cerrado.

Alvo	Iniciadores	Sequência nucleotídica (5'-3')	Fragmento esperado	Referência
Domínio <i>Archaea</i>	21f 958r	TTCCGGTTGATCCYGCCGGA YCCGGCGTTGAMTCCAATT	937 pb	DeLong, 1992
Domínio <i>Archaea</i>	109f 915r	ACKGCTCAGTAACACGT GTGCTCCCCGCCAATTCCT	806 pb	Großkopf <i>et al.</i> , 1998, Banning <i>et al.</i> , 2005
Filo <i>Bathyarchaeota</i>	Bathy442f Bathy644r	AGACTGGTGTGACCCGCC TCGCCACTGGTGGTCCTC	202 pb	Yu <i>et al.</i> , 2017
Domínio <i>Bacteria</i>	27f 1492r	AGAGTTTGATCCTGGCTCAG GGTACCTTGTACGACTT	1465 pb	Lane, 1991

Y: pirimidina, M: adenina ou citosina, K: guanina ou timina

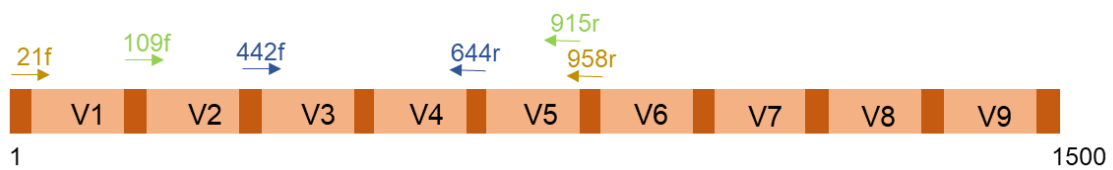


Figura 1. Esquema representando as regiões conservadas (em tom escuro) e as nove regiões variáveis (V1 a V9, em tom claro) do gene codificador do rRNA 16S. As regiões aproximadas de anelamento dos três pares de iniciadores para o domínio *Archaea* empregados neste trabalho estão indicadas, evidenciando as diferentes regiões do gene que são amplificadas por cada par de iniciadores.

As reações de PCR foram realizadas em sistemas com volume final de 30 µL, contendo 5 µL de DNA total, em concentrações que variaram de 1 a 100 ng. As concentrações finais dos demais reagentes para as reações de PCR com todos os pares de iniciadores foram: 1X de tampão de reação (Invitrogen), 1,5 µM de MgCl₂, 400 ng/µL de soroalbumina bovina (BSA), 0,5 µM de cada iniciador, 200 µM de dNTPs e 1,5 U de *Taq*DNA Polimerase (Invitrogen). Os ensaios de PCR foram realizados em um termociclador Bio Rad PTC-100® (Peltier Thermal Cycler), nas condições de ciclagem descritas na Tabela 2.

Em todos os experimentos de PCR realizados, um controle negativo foi utilizado, adicionando-se água à reação em substituição ao DNA. Nas reações com os iniciadores universais para o domínio *Archaea* (21f/958r e 109f/915r), o DNA de *Haloferax volcanii* foi usado como controle positivo. Os produtos das reações de PCR foram analisados por eletroforese em gel de agarose 1 % corado com brometo de etídio (10 mg/mL) e seu tamanho foi estimado com o auxílio do marcador de massa molecular *Ikb Plus DNA Ladder* (Invitrogen) ou *Ikb DNA Ladder* (Kasvi).

Tabela 2. Condições de ciclagem empregadas nos ensaios de PCR com os diferentes iniciadores para os genes do rRNA 16S de *Archaea* e *Bacteria*.

	21f/958r	109f/915r
1	Desnaturação a 94 °C por 1 minuto	Desnaturação a 94 °C por 1 minutos
2	Desnaturação a 94 °C por 1 minuto	Desnaturação a 94 °C por 1 minutos
3	Anelamento a 55 °C por 1 minuto	Anelamento a 57 °C por 1 minutos
4	Extensão a 72 °C por 1 minuto e 30 segundos	Extensão a 72 °C por 1 minutos e 30 segundos
5	Repetição dos passos 2 a 4 – 29 vezes	Repetição dos passos 2 a 4 – 29 vezes
6	Extensão final a 72 °C por 5 minutos	Extensão final a 72 °C por 5 minutos
7	Manutenção a 4 °C	Manutenção a 4 °C
	Bathy442f/Bathy644r	27f/1492r
1	Desnaturação a 95 °C por 3 minutos	Desnaturação a 95 °C por 5 minutos
2	Desnaturação a 95 °C por 1 minuto	Desnaturação a 95 °C por 1 minuto
3	Anelamento a 58 °C por 1 minuto	Anelamento a 55 °C por 1 minuto
4	Extensão a 72 °C por 1 minuto e 10 segundos	Extensão a 72 °C por 2 minutos
5	Repetição dos passos 2 a 4 – 29 vezes	Repetição dos passos 2 a 4 – 29 vezes
6	Extensão final a 72 °C por 5 minutos	Extensão final a 72 °C por 10 minutos
7	Manutenção a 4 °C	Manutenção a 4 °C

1.3.5. Purificação e ligação dos fragmentos de DNA amplificados no vetor *pGEM-T Easy*

Os produtos resultantes das reações de PCR com os iniciadores 21f/958r, 109f/915r e 27f/1492r foram purificados, diretamente após a amplificação, com kit *GeneJET PCR Purification* (Thermo Scientific), enquanto os fragmentos de tamanho esperado amplificados com os iniciadores Bathy442f/644r foram purificados a partir de gel de agarose pelo kit *Wizard® SV Gel and PCR Clean-Up System* (Promega). Ambos os protocolos de purificação foram realizados conforme as instruções dos fabricantes. Os fragmentos purificados foram então ligados ao vetor *pGEM-T Easy*® (Promega), na proporção 3:1 de inserto:vetor, de acordo com as instruções do fabricante. Os sistemas de ligação foram mantidos por 1 h em temperatura ambiente e, em seguida, armazenados a 4°C por uma noite.

1.3.6. Preparo e estoque de células competentes de *Escherichia coli* DH5α

Células de *E. coli* DH5α foram inoculadas em 5 mL de meio Luria Bertani (LB) e incubadas a 37°C, sem agitação, por uma noite. Após este período, 1 mL da cultura foi inoculado em 100 mL de meio LB, sob agitação de 200 rpm a 37°C, até que a cultura atingisse a densidade óptica (OD₆₀₀) entre 0,2 e 0,3. Neste momento, a cultura foi vertida em dois tubos de polipropileno e centrifugadas a 4000 g por 10 minutos a 4 °C. O

sobrenadante foi descartado e o sedimento de células ressuspensas em 10 mL de uma solução gelada de cloreto de cálcio 60 mM, HEPES 10 mM e glicerol 15 % (pH 7). As células ressuspensas nessa solução foram mantidas em gelo por 20 a 30 minutos e, em seguida, centrifugadas a 2000 g por 10 minutos, a 4°C. O sobrenadante foi mais uma vez descartado e o sedimento de células ressuspensas em 2 mL da mesma solução, aliqotado em tubos tipo eppendorf em volumes de 0,2 mL. Os tubos contendo células competentes foram então armazenados a – 80°C até o momento de sua utilização.

1.3.7. Transformação por choque térmico, seleção e estoque em glicerol de clones recombinantes

As alíquotas de células DH5 α competentes previamente preparadas e estocadas foram descongeladas em gelo e 100 μ L foram distribuídos em tubos de hemólise gelados. Em seguida, 5 μ L do sistema de ligação foram misturados às células. Os tubos contendo as células e os sistemas de ligação foram então incubados por 30 minutos em gelo e em seguida foram rapidamente transferidos para um banho a 37 °C e incubados por 5 minutos. Ao final desse tempo, 1 mL de meio LB foi adicionado a cada tubo e o sistema incubado em estufa a 37°C por 1 h.

Após esse período, diferentes quantidades do sistema de transformação (50 a 150 μ L) foram inoculadas em placas de meio LB sólido (1,5 % de ágar) suplementado com ampicilina (150 μ g/mL), Xgal (0,00625 %) e IPTG (0,5 mM). As células foram semeadas com o auxílio de pérolas de vidro estéreis e as placas foram incubadas em estufa a 37 °C por 16 a 24 h, até a observação de colônias brancas e/ou azuis.

As colônias brancas foram retiradas da placa com o auxílio de palitos de madeira estéreis e inoculadas em 5 mL de meio LB líquido suplementado com ampicilina (150 μ g/mL). Os tubos contendo as culturas recém inoculadas foram então incubados em estufa a 37°C por uma noite e alíquotas de cada cultura foram misturadas a um mesmo volume de glicerol 70 %, resultando numa concentração final de 35 % de glicerol. As culturas foram então estocadas a -20°C.

1.3.8. Extração de DNA plasmidial por lise alcalina

Os clones estocados em glicerol foram inoculados em meio LB suplementado com ampicilina (150 μ g/mL) e incubados em estufa a 37°C por uma noite. A cultura resultante

foi centrifugada a 6000 g por 5 minutos e o sobrenadante foi descartado. O sedimento celular foi então ressuspensão em 100 µL de solução I (50 mM de glicose, 25 mM de Tris-HCl, pH8, e 10 mM de EDTA) e incubado por 5 minutos à temperatura ambiente. Em seguida, foram adicionados 200 µL de solução II preparada no momento do uso (1 % SDS, 0,2 M de NaOH) e a mistura foi suavemente homogeneizada. Os tubos foram incubados no gelo por 5 minutos. Por fim, foram adicionados 150 µL de solução III (3 M de acetato de potássio, 2 M de ácido acético glacial), mais uma homogeneização foi realizada e os tubos foram novamente incubados por 5 minutos no gelo. Os tubos foram então centrifugados a 12000 g por 5 minutos, a 4°C e os sobrenadantes transferidos para novos tubos, aos quais foram adicionados 5 µL de RNase A (10 mg/mL). As amostras foram então incubadas a 37°C por 1 a 2 horas.

Após esse período, um volume igual de clorofane (fenol, clorofórmio e álcool isoamílico, na proporção 25:24:1) foi adicionado às amostras, que foram vigorosamente homogeneizadas e centrifugadas a 7000 g por 3 minutos, a 4°C. A fase aquosa do sobrenadante foi então transferida para um novo tubo tipo eppendorf e um volume igual de clorofil (clorofórmio e álcool iso-amílico, na proporção 24:1) foi adicionado. Após serem mais uma vez bem misturadas, as amostras foram centrifugadas novamente a 7000 g por 3 minutos, a 4°C. Os sobrenadantes resultantes foram transferidos para novos tubos, 30 µL de NaOH (3 M) e 3 volumes de etanol 100 % gelado foram adicionados. Os tubos foram estocados a -20°C e após 24 horas foram centrifugados a 12000 g por 5 minutos a 4°C, o sobrenadante descartado e 500 µL de etanol 70 % foram adicionados às amostras. Os tubos foram mais uma vez centrifugados nas mesmas condições anteriores e colocados invertidos à temperatura ambiente para a secagem do sedimento de DNA.

Após a completa secagem dos DNAs, estes foram ressuspensos em 50 µL de H₂O milli-Q e estocados a -20°C. A qualidade e concentração dos plasmídeos obtidos foram analisadas por eletroforese em gel de agarose 1 %, com auxílio do marcador de massa molecular *High Mass* (Invitrogen).

1.3.9. Sequenciamento de DNA e análise das sequências do gene codificador do rRNA 16S

Os DNAs plasmidiais dos clones contendo os *amplicons* obtidos a partir dos iniciadores 21f/958r, 109f/915r e 27f/1492r foram sequenciados pela empresa

Macrogen Inc., empregando-se os iniciadores 21f, 109f e 27f, respectivamente. Os plasmídeos contendo produtos amplificados pelos iniciadores Bathy442f/Bathy644f foram sequenciados utilizando-se o iniciador T7.

As sequências de DNA obtidas foram analisadas quanto à qualidade pelo algoritmo PHRED (Ewing *et al.*, 1998), através da ferramenta *Electropherogram quality analysis* disponível na página da EMBRAPA, no endereço eletrônico: <http://asparagin.cenargen.embrapa.br/phph/>. Foram selecionadas as sequências com mais de 250 nucleotídeos e que apresentaram valor de PHRED superior a 31. Sequências com valores inferiores a 31 foram descartadas.

As sequências relativas ao gene de rRNA 16S selecionadas foram então submetidas a análises comparativas no banco de dados taxonômico do *Greengenes* (DeSantis *et al.*, 2006), através do programa Mothur (Schloss *et al.*, 2009). Posteriormente, as sequências foram submetidas a alinhamentos múltiplos, utilizando-se o programa ClustalX (Larkin *et al.*, 2007) e o resultado do alinhamento foi editado manualmente, com o auxílio do programa BioEdit (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) para a obtenção de sequências de mesmo tamanho.

Diferentes árvores filogenéticas com sequências do gene de rRNA 16S foram construídas pelo programa MEGA5 (Tamura *et al.*, 2011), pelo método de *Neighbor-joining* (Saitou e Nei, 1987), com o modelo de Jukes-Cantor (Jukes e Cantor, 1969) e teste de *bootstrap* de 1000 repetições (Felsenstein, 1985). Sequências representativas em nível de espécie (97 % de identidade), obtidas por ferramentas do programa Mothur, foram comparadas a sequências disponíveis no banco de dados do NCBI (*US National Center of Biotechnology Informantion*) por meio da ferramenta BLAST (Altschul *et al.*, 1990) e aquelas que apresentaram os mais altos valores de similaridade, bem como sequências de genes de rRNA 16S de isolados, foram incluídas nas árvores.

1.3.10. Microscopia óptica das células obtidas nos diferentes cultivos

Após vários meses de manutenção das culturas, esfregaços a partir dos diferentes tipos coloniais observados nos meios de cultura foram realizados sobre lâminas histológicas de vidro, com o auxílio de alça de níquel-cromo. Os esfregaços foram então secos e fixados no fogo, submetidos à coloração diferencial de Gram, visualizados sob

aumentos de 630 ou 1000x e fotografados no microscópio de luz Axiophot (Zeiss, Alemanha), gentilmente disponibilizado pelo grupo da Prof. Dra. Sônia Nair Bão, do Laboratório de Microscopia da UnB.

Amostras provenientes do cultivo em meio líquido também foram observadas por microscopia óptica. As células foram concentradas por meio de centrifugação (6000 g, por 5 minutos) de 3 mL da cultura em tubos tipo eppendorf. Os sobrenadantes foram descartados e os sedimentos de células foram ressuspensos em 80 µL de meio líquido. Em seguida, 20 µL foram então transferidos para lâminas histológicas de vidro, que foram submetidas à secagem, fixação, coloração de Gram e observadas conforme descrito acima para os esfregaços provenientes do meio sólido.

As células provenientes das culturas em meios sólido e líquido foram também observadas a fresco, por microscopia óptica de contraste de fase. Para tal, porções de colônias foram coletadas com o auxílio da alça de níquel-cromo, homogeneizadas em 10 µL de meio líquido estéril previamente depositados em lâminas de vidro, sendo então imediatamente recobertos por lamínulas e observados sob aumento de 1000x no microscópio Axio Scope A1, gentilmente cedido pela Prof. Dra. Marlene Teixeira DeSouza, do Laboratório de Microbiologia da UnB. Para a observação das células crescidas em meios líquidos, lamínulas de vidro foram previamente introduzidas nos frascos inoculados com as diferentes culturas, os quais foram incubados por 21 dias a 28°C. Após esse período, as lamínulas foram posicionadas sobre lâminas histológicas de vidro, de forma que as células aderidas ficassem voltadas para baixo, sendo então imediatamente observadas, nas mesmas condições descritas acima.

1.3.11. Teste de viabilidade celular – kit *LIVE/DEAD*TM *BacLight*TM

Para a análise da viabilidade das células em meios líquidos, estas foram analisadas por microscopia de fluorescência, após utilização do kit *LIVE/DEAD*TM *BacLight*TM (Thermo Fisher Scientific). As células tratadas foram visualizadas sob aumento de 1000x no microscópio de fluorescência Axio Scope A1, gentilmente cedido pelo Prof. Dr. Ricardo H. Krüger, do Laboratório de Enzimologia da UnB. Para o preparo das amostras, 20 mL de culturas incubadas por aproximadamente dois meses foram centrifugadas a 8000 g por 10 minutos a 26°C, o sobrenadante foi descartado e o sedimento foi ressuspensionado em 1 mL de tampão salina-fosfato (PBS) 1X (NaCl 137 mM, KCl 2,7 mM, Na₂PO₄ 8,1 mM, KH₂PO₄ 1,5 mM) estéril. Em seguida, os reagentes do kit (Syto 9 e

iodeto de propídeo) foram adicionados à solução de células, conforme instruções do fabricante. Após a homogeneização, 15 μ L da amostra foram transferidos para uma lâmina histológica de vidro e recobertos com lamínula. Neste ensaio, as células vivas fluorescem em verde, por terem seus ácidos nucleicos marcados com o Syto 9, enquanto as células mortas fluorescem em vermelho, pois suas membranas degradadas permitem a entrada do iodeto de propídeo, que também se liga a ácidos nucleicos e altera as propriedades de emissão de fluorescência, por substituição de um fluoróforo pelo outro e mecanismos de *quenching*. Desta forma, observando-se a fluorescência em ambos espectros de emissão é possível a diferenciação de células vivas, com suas membranas plasmáticas intactas, daquelas mortas, que apresentam membranas danificadas.

1.3.12. Microscopia Eletrônica de Varredura das células obtidas nos diferentes cultivos

A preparação e análise das células por microscopia eletrônica de varredura (MEV) foram realizadas no Laboratório de Microscopia da Universidade de Brasília, coordenado pela Prof. Dra. Sônia Nair Bão. As células do cultivo em meio sólido foram retiradas das placas e lavadas em tampão salina-fosfato (PBS) 1X, fixadas em fixador Karnovsky 0,1 M (paraformaldeído 2 %, glutaraldeído 2 %, sacarose 3 %, cloreto de cálcio 0,005 M em meio tamponado por cacodilato de sódio 0,1 M) por 4 h, à temperatura ambiente. O fixador foi então retirado, o material fixado foi lavado com tampão cacodilato de sódio 0,1 M e posteriormente depositado em lamínulas de vidro recobertas com poli-L-lisina. As amostras foram então pós-fixadas com tetróxido de ósmio (1 %) durante 30 minutos, lavadas com água destilada e gradualmente desidratadas em acetona 50, 70, 90 e 100 %, por 5 minutos em cada concentração. O material foi submetido ao procedimento de secagem ao ponto crítico de CO₂ no equipamento Balzers CPD 030 (*Balzers Union AG*, Liechtenstein) e posteriormente metalizado com vapores de ouro no metalizador Balzers SCD 050 (*Balzers Union AG*, Liechtenstein).

Para o preparo e visualização dos cultivos líquidos, as células foram previamente crescidas em recipientes contendo meio líquido e lamínulas de vidro no fundo, por aproximadamente um mês. Após esse período, as lamínulas de vidro foram retiradas e colocadas em uma placa de 6 poços, onde foram fixadas, lavadas com o tampão cacodilado de sódio e pós-fixadas com tetróxido de ósmio nas mesmas condições

descritas para as células provenientes do meio sólido. As lamínulas foram então transferidas para um suporte e submetidas à desidratação gradual, secagem ao ponto crítico e metalização, também como descrito acima. Todas as amostras foram analisadas no microscópio eletrônico de varredura JEOL JSM-7001F (JEOL Ltd., Tóquio, Japão) a 15,0 kV.

1.3.13. Sequenciamento genômico total do cultivo, montagens e identificação taxonômica das sequências obtidas

Amostras de dois tipos coloniais observados no meio sólido foram selecionadas para a realização de sequenciamento genômico total. Visando uma melhor montagem dos genomas dos organismos presentes nos cultivos, duas abordagens de sequenciamento foram empregadas: 1) sequenciamento *paired-end* de biblioteca preparada com o kit *TrueSeq DNA shotgun PCR-Free* (350pb) em uma *lane* da plataforma HiSeq 2000 da Illumina e 2) sequenciamento de biblioteca Pacbio 20Kb bluepippin em duas SMRT *cells* na plataforma PacBio. Os serviços de preparo de bibliotecas e sequenciamentos foram realizados pela empresa Macrogen Inc., na Coreia do Sul. Ambas as bibliotecas submetidas aos sequenciamentos foram preparadas a partir de uma mesma amostra de DNA, gerada a partir de extrações de DNA total dos cultivos em meio sólido realizadas em diferentes tempos de crescimento: 10, 23, 48 e 70 dias. A extração de DNA de diversas placas em diferentes tempos foi realizada visando a obtenção do genoma de organismos que poderiam estar mais abundantes no cultivo em determinados momentos e menos abundantes em outros. As extrações foram feitas por fenol-clorofórmio, conforme o protocolo já descrito acima. As amostras resultantes das extrações nos diferentes dias foram misturadas e a alíquota de DNA resultante foi quantificada (conforme descrito na seção 3.3) e enviada para o preparo das duas bibliotecas e sequenciamento nas duas plataformas escolhidas.

Posteriormente, uma amostra de células proveniente do cultivo em meio líquido também foi enviada para sequenciamento genômico total. A amostra selecionada teve o DNA extraído e quantificado conforme o protocolo descrito na seção 1.3.3, após a filtração de aproximadamente 2 L de cultura. O sequenciamento *paired-end* da amostra foi realizado a partir de uma biblioteca preparada com o kit *TrueSeq Nano* na plataforma HiSeq 4000 da Illumina.

Os dados resultantes dos sequenciamentos por ambas as plataformas empregadas foram analisados com a colaboração dos Profs. Drs. Georgios Pappas Jr e Ricardo H. Krüger. Os *short-reads* resultantes do sequenciamento Illumina HiSeq, os *long-reads* oriundos do sequenciamento PacBio e os dois tipos de dados em conjunto foram inicialmente montados utilizando-se os programas *SPAdes* (Bankevich *et al.*, 2012), *Canu* (Koren *et al.*, 2017) e *hybridSPAdes* (Antipov *et al.*, 2016), respectivamente. Os *contigs* obtidos foram então comparados com as sequências genômicas disponíveis no banco de dados não redundante NCBI (*US National Center of Biotechnology Information*) por meio da ferramenta BLAST (Altschul *et al.*, 1990).

A segunda abordagem envolveu a montagem *de novo* dos *reads* resultantes do sequenciamento Illumina HiSeq, utilizando-se a ferramenta *Iterative De-Bruijn Assembler* (IDBA-UD) (Peng *et al.*, 2012), com parâmetros para a montagem de dados metagenômicos. Os *contigs* obtidos foram então submetidos à predição de genes e anotação e foram importados para o *ggKbase* (<https://ggkbase.berkeley.edu>), utilizando-se os *pipelines* e ferramentas indicados pela plataforma (<http://ggkbase-help.berkeley.edu/overview/data-preparation-metagenome/>). Visando a identificação de genomas diferentes presentes na amostra, os *contigs* foram agrupados baseando-se na sua cobertura, conteúdo GC e afiliação filogenética, levando à separação de *bins* genômicos.

1.4. Resultados e discussão

1.4.1. Características químicas do solo e obtenção de culturas em meio sólido

As características químicas do solo utilizado para a confecção dos meios de cultura e como inóculo foram determinadas e estão descritas na Tabela 3.

Tabela 3. Valores dos parâmetros físico-químicos do solo de Cerrado utilizado para a confecção dos meios de cultura e como inóculo inicial para os cultivos.

Parâmetros Analisados	Valor
pH	4,700
Fósforo (ppm)	0,800
Cálcio (cmol _c /dm ³)	0,300
Magnésio (cmol _c /dm ³)	0,100
Potássio (cmol _c /dm ³)	0,090
Sódio (cmol _c /dm ³)	0,040
Alumínio (cmol _c /dm ³)	1,000
Carbono Orgânico (g/kg)	58,40
Nitrogênio total (g/kg)	0,610
Nitrogênio nitrato	0,150
Nitrogênio nitrito	0,150
Nitrogênio amoniacal	0,050
Matéria Orgânica (g/kg)	100,4
Boro (ppm)	0,830
Cobre (ppm)	0,850
Ferro (ppm)	287,0
Manganês (ppm)	3,200
Zinco (ppm)	0,430
Enxofre	5,200

Como o objetivo inicial do trabalho era a obtenção de culturas de archaeas de solo de Cerrado, antibióticos (ampicilina, estreptomicina, cloranfenicol) e antifúngicos (itraconazol e cetoconazol) foram adicionados sucessivamente ao meio de cultura, visando a inibição do crescimento de bactérias e fungos, além do possível enriquecimento de archaeas. Por apresentarem diferenças metabólicas e estruturais em relação às bactérias, as archaeas são resistentes a vários tipos de antibióticos rotineiramente utilizados (Dridi, *et al.*, 2011; Khelaifia e Drancourt, 2012). Após alguns repiques, a contaminação por fungos foi eliminada e colônias bastante diminutas, com aspecto de crescimento procariótico, puderam ser observadas nas placas. Embora inicialmente as colônias apresentassem um crescimento lento, que possibilitava o repique apenas a cada mês, após aproximadamente 8 meses de repiques sucessivos, as culturas passaram a

apresentar um crescimento mais acelerado, sendo possível a realização de repiques em menores intervalos, geralmente a cada 10 dias.

Após vários meses de cultivo, colônias com diferentes aspectos macroscópicos foram observadas e, após a análise das sequências dos genes de rRNA 16S amplificados a partir do DNA extraído de cada uma delas (discutido à frente, na seção 1.4.3), dois tipos coloniais - denominados G e G2 - foram selecionados para as análises posteriores. O tipo colonial G (Figura 2A), com diâmetro de 1 a 2 mm após um mês de crescimento, apresentava bordas irregulares, com aspecto translúcido e rendado e uma região central protuberante, mais densa e arredondada. Já o tipo colonial denominado G2 (Figura 2B) era menor, com diâmetro aproximado de 0,5 mm, forma mais arredondada e aparência opaca. Vale ressaltar que tais colônias apresentavam morfologias variáveis em diferentes períodos do cultivo, com aspecto mais mucilaginoso em determinados momentos após os repiques (Figura 2C). Tal variação talvez possa ser decorrente de alterações fisiológicas nas células ao longo do tempo de crescimento, tais como a produção de matriz extracelular em determinados momentos por exemplo, ou ainda pela alteração na proporção de diferentes organismos presentes no cultivo.

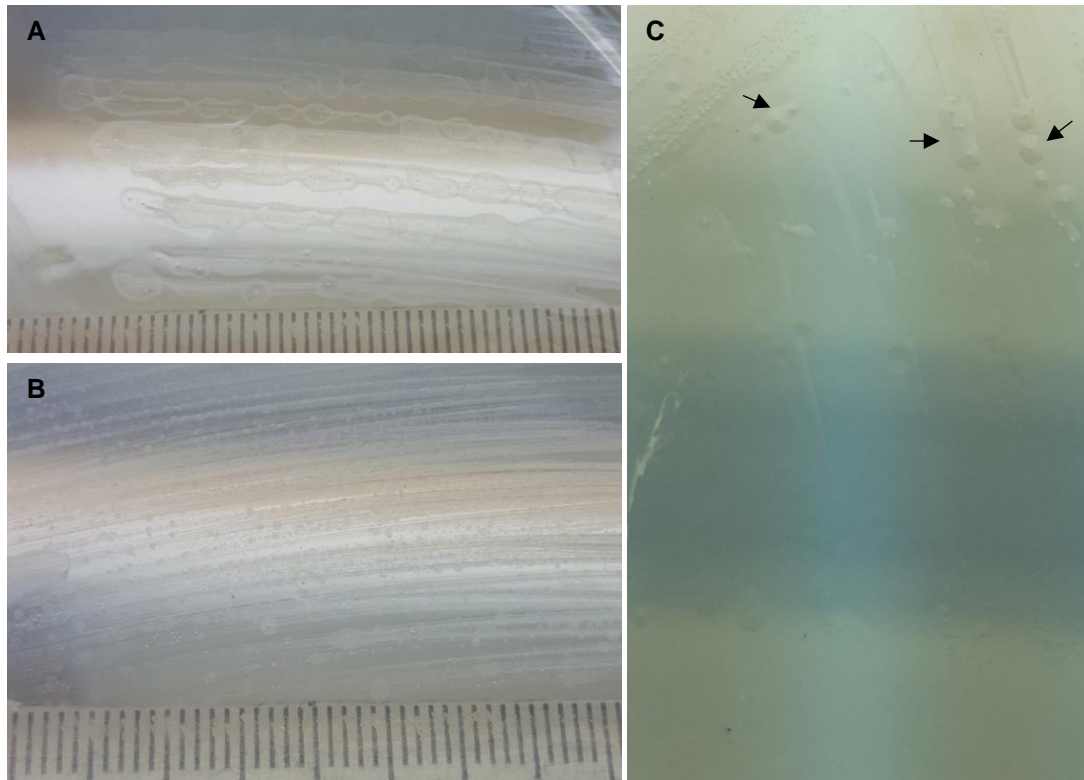


Figura 2. Tipos coloniais G (A) e G2 (B), obtidos em meio de cultura inoculado com solo de Cerrado e suplementado com ampicilina, estreptomicina, cloranfenicol e itraconazol (nas concentrações descritas na seção 3.2), após aproximadamente um mês de crescimento. Em C, aspecto mucilaginoso observado nas colônias em determinados momentos após os repiques (indicado pelas setas).

Após a constatação, por experimentos de PCR, da ocorrência de um co-cultivo entre bactérias e archaeas nas culturas crescidas em meio sólido, polimixina B foi adicionada aos meios, com o intuito de reduzir ou eliminar a população bacteriana e de enriquecer a cultura para archaeas. Após a adição deste antibiótico, inicialmente foi observado um crescimento mais lento nos meios sólidos, mas após alguns repiques nesta condição, as colônias passaram a se desenvolver no mesmo tempo observado nos meios anteriormente utilizados, havendo poucas diferenças na morfologia das colônias resultantes (Figura 3).

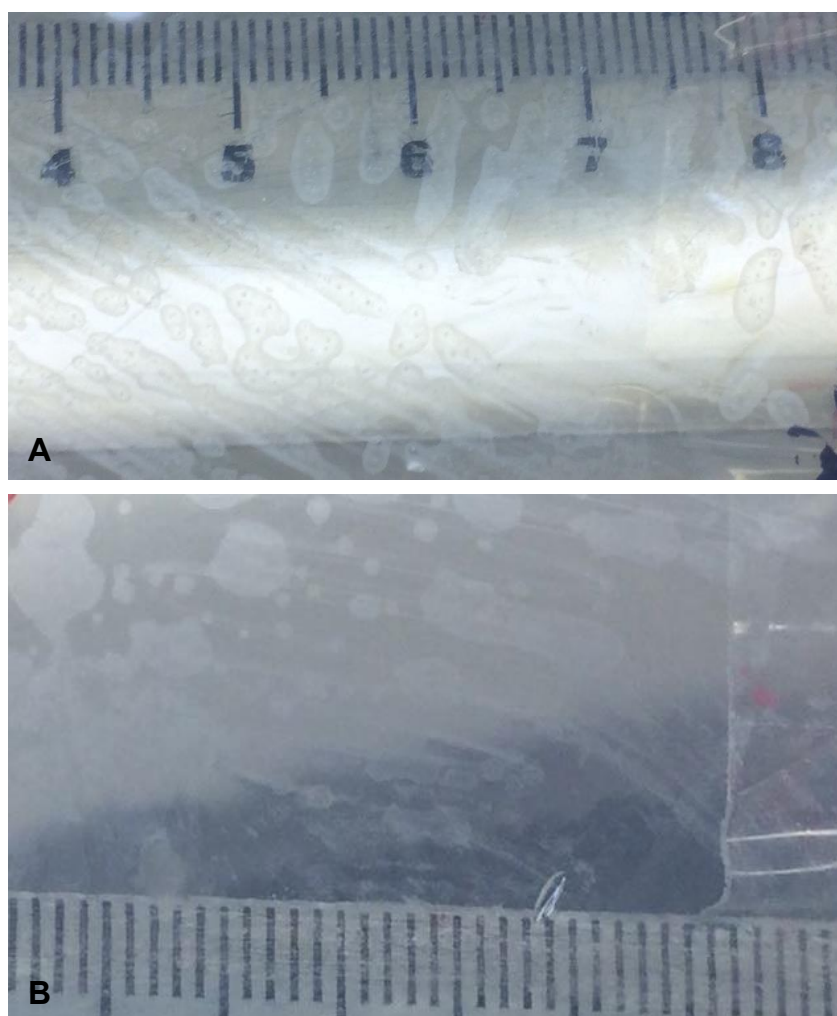


Figura 3. Colônias G (A) e G2 (B) observadas após vários repiques em meio contendo polimixina B.

1.4.2. Obtenção de culturas em meio líquido

Uma questão interessante a ser analisada refere-se ao estabelecimento das culturas em meio líquido. O inóculo inicial de solo foi semeado tanto em meio sólido quanto meio líquido, ambos confeccionados da mesma forma (a partir de um coado do solo misturado à água destilada) e adicionados dos mesmos agentes antimicrobianos. Embora o desenvolvimento de colônias tenha sido detectado após algumas semanas de incubação, não foram observados indícios de crescimento no meio líquido, levando-nos a acreditar inicialmente que os cultivos não poderiam ser estabelecidos nessa condição. Hipóteses relacionadas à incapacidade de crescimento destes organismos na ausência de superfícies para sua adesão, ou ainda a diferenças nutricionais que poderiam estar ocorrendo no meio líquido pelo fato deste ser submetido à filtração em membranas de 0,22 µm foram levantadas. Por essa razão, todos os experimentos realizados inicialmente foram conduzidos a partir dos cultivos obtidos em meio sólido.

No entanto, análises microscópicas de meios líquidos incubados por alguns meses, revelaram que, mesmo não havendo indícios macroscópicos de crescimento, diversas células podiam ser observadas nestas amostras. Colônias de G e G2 crescidas em meio sólido foram então inoculadas em tubos tipo Falcon e em frascos Erlenmeyer, incubados sem agitação e sob agitação de 150 rpm, respectivamente, a 28°C. Preparações periódicas à fresco e coradas das amostras, além de ensaios de viabilidade celular com o kit *LIVE/DEAD™ BacLight™* (todos discutidos na seção 1.4.4, adiante), revelaram a obtenção de culturas das amostras G e G2 em meio líquido. Apesar do crescimento detectado nestes meios ser inferior ao observado em meios sólidos, as culturas estabelecidas em meio líquido também puderam ser submetidas a diferentes análises a partir de então.

1.4.3. Caracterização filogenética dos organismos cultivados a partir da análise dos produtos obtidos nos ensaios de PCR dirigidos aos genes de rRNA 16S

Inicialmente, colônias de aspecto macroscópico distinto, originadas a partir do mesmo inóculo inicial, foram selecionadas para extração de DNA, PCR e análise dos genes rRNA 16S. Visando a caracterização filogenética dos microrganismos presentes nos cultivos, iniciadores para o gene do rRNA 16S dos domínios *Archaea* e *Bacteria*

foram utilizados em ensaios de PCR. Em um primeiro momento, o par de iniciadores 21f/958r, considerado universal para *Archaea* e frequentemente empregado na detecção de membros deste domínio em ambientes terrestres e aquáticos mesófilos (DeLong, 1992; Bano *et al.*, 2004; Vissers *et al.*, 2009), foi utilizado para a investigação da presença de archaeas nos cultivos de solo de Cerrado estabelecidos. Fragmentos de DNA de tamanho esperado foram obtidos nas reações de PCR, sugerindo a presença de *Archaea* nos cultivos. Os fragmentos amplificados a partir de cada tipo colonial foram clonados e submetidos ao sequenciamento de DNA. As sequências que atendiam aos critérios de qualidade definidos (descritos no item 1.3.9 da seção Materiais e métodos) foram classificadas como pertencentes predominantemente aos filos Bathyarchaeota e Thaumarchaeota.

O filo Bathyarchaeota foi proposto recentemente com base em diferenças identificadas nas sequências de genes de RNAs e proteínas ribossomais de um grupo até então denominado *Miscellaneous Crenarchaeotic Group* (Meng *et al.*, 2014). Sequências do gene de rRNA 16S afiliadas a este grupo são frequentemente detectadas em solos e sedimentos marinhos e de água doce (Etto *et al.*, 2012, Kubo *et al.*, 2012; Vaksmaa *et al.*, 2017, Winkel *et al.*, 2018). Apesar de algumas inferências acerca de seu metabolismo e papéis ecológicos nestes ambientes terem sido feitas a partir da reconstrução de genomas parciais obtidos por metagenômica (Evans *et al.*, 2015), este filo não apresenta qualquer representante cultivado descrito até o momento e, por essa razão, suas características permanecem amplamente desconhecidas. O filo Thaumarchaeota (Brochier-Armanet *et al.*, 2008) tem recebido muita atenção nos últimos anos por englobar todas as archaeas capazes de oxidar amônia descritas até o momento, consideradas essenciais no ciclo biogeoquímico do nitrogênio, e por ser o filo predominante em solos de todo o mundo e diversos ambientes aquáticos (Timonen e Bomberg, 2009, Auguet *et al.*, 2010, Pester *et al.*, 2012, Doxey *et al.*, 2015). Esse filo é atualmente dividido em três subgrupos filogenéticos principais, denominados I.1a, I.1b e I.1c, sendo que os dois primeiros já possuem alguns membros em culturas puras ou co-culturas, todos capazes de oxidar de amônia e tendo sido obtidos por meio da utilização de meios de cultura que favoreciam o enriquecimento de organismos com este metabolismo (Simon *et al.*, 2005; Hatzepichler *et al.*, 2008; Jung *et al.*, 2011, Kim *et al.*, 2012; Stieglmeier *et al.*, 2014). O subgrupo I.1c, no entanto, foi identificado apenas por meio de sequências de rDNA 16S oriundas principalmente de solos ácidos (Jurgens *et al.*, 1997; Lehtovirta *et al.*, 2009, Barton *et al.*, 2014), sem qualquer representante cultivado até o momento. Um estudo envolvendo o

estabelecimento de microcosmos *in situ* indicou que membros do grupo I.1c provavelmente não estão associados ao metabolismo de oxidação de amônia (Weber *et al.*, 2015), mas as capacidades metabólicas e demais características destes organismos seguem desconhecidas.

Por terem gerado sequências do gene de rRNA 16S afiliadas ao filo Bathyarchaeota e ao grupo I.1c de Thaumarchaeota (Figura 4 e Tabela 4) que, conforme mencionado, ainda não possuem representantes cultivados, os tipos coloniais denominados G e G2 foram selecionados para a continuação dos trabalhos. É interessante ressaltar ainda que as análises dos *amplicons* obtidos com os iniciadores 21f/958r a partir das colônias G e G2 também revelaram sequências afiliadas à classe *Methanomicrobia*, do filo Euryarchaeota. Este filo, identificado quando da proposta do domínio *Archaea* (Woese *et al.*, 1990), é constituído por organismos com um amplo espectro de fenótipos, como metanogênicos, halófilos, redutores de sulfato e hipertermófilos. Apesar das recentes sugestões resultantes de análises metagenômicas independentes de cultivo que membros dos filos Bathyarchaeota e Verstraetearchaeota também poderiam estar envolvidos no metabolismo de metanogênese (Evans *et al.*, 2015, Vanwonterghem *et al.*, 2016), todas as archaeas metanogênicas atualmente cultivadas se encontram distribuídas em seis classes de Euryarchaeota (de acordo com a classificação disponível na ferramenta *Taxonomy*, do banco de dados NCBI). As euryarchaeotas metanogênicas descritas até o momento crescem somente sob condições de anaerobiose (Edwards e McBride, 1975; Tsao *et al.*, 1994; Dridi *et al.*, 2012), fazendo com que sua presença nos cultivos obtidos a partir do solo de Cerrado - mantidos em condições aeróbicas - seja inesperada. Uma possível explicação seria que a associação dos diferentes microrganismos promoveria a formação de microambientes anaeróbicos nas placas de cultivo, possibilitando o crescimento dessas archaeas metanogênicas, ou ainda que, assim como descrito para o cultivo da bactéria *Prochlorococcus* (mencionado na Introdução deste capítulo), os demais microrganismos do co-cultivo poderiam estar atuando na proteção destas archaeas contra danos causados pelo oxigênio.

Sequências relativas ao gene de rRNA 16S de bactérias também foram amplificadas a partir do DNA extraído das amostras G e G2, indicando um co-cultivo entre archaeas e bactérias. Tais *amplicons* foram obtidos com o par de iniciadores 27f/1492r, amplamente utilizado para o domínio *Bacteria* (Frank *et al.*, 2008), e posteriormente clonados e sequenciados. Foram obtidas 58 sequências de qualidade e todas foram classificadas como pertencentes ao gênero *Novosphingobium*. Esse gênero

bacteriano ocorre em diversos ambientes naturais e artificiais, com espécies já isoladas a partir de diferentes tipos de solos (Takeuchi *et al.*, 2001; Gupta *et al.*, 2009, Kämpfer *et al.*, 2011, Niharika *et al.*, 2013; Kämpfer *et al.*, 2015). Análises realizadas com auxílio do programa Mothur revelaram que todas as sequências obtidas apresentavam mais de 99% de identidade entre si, indicando a presença de uma única espécie bacteriana nos cultivos G e G2.

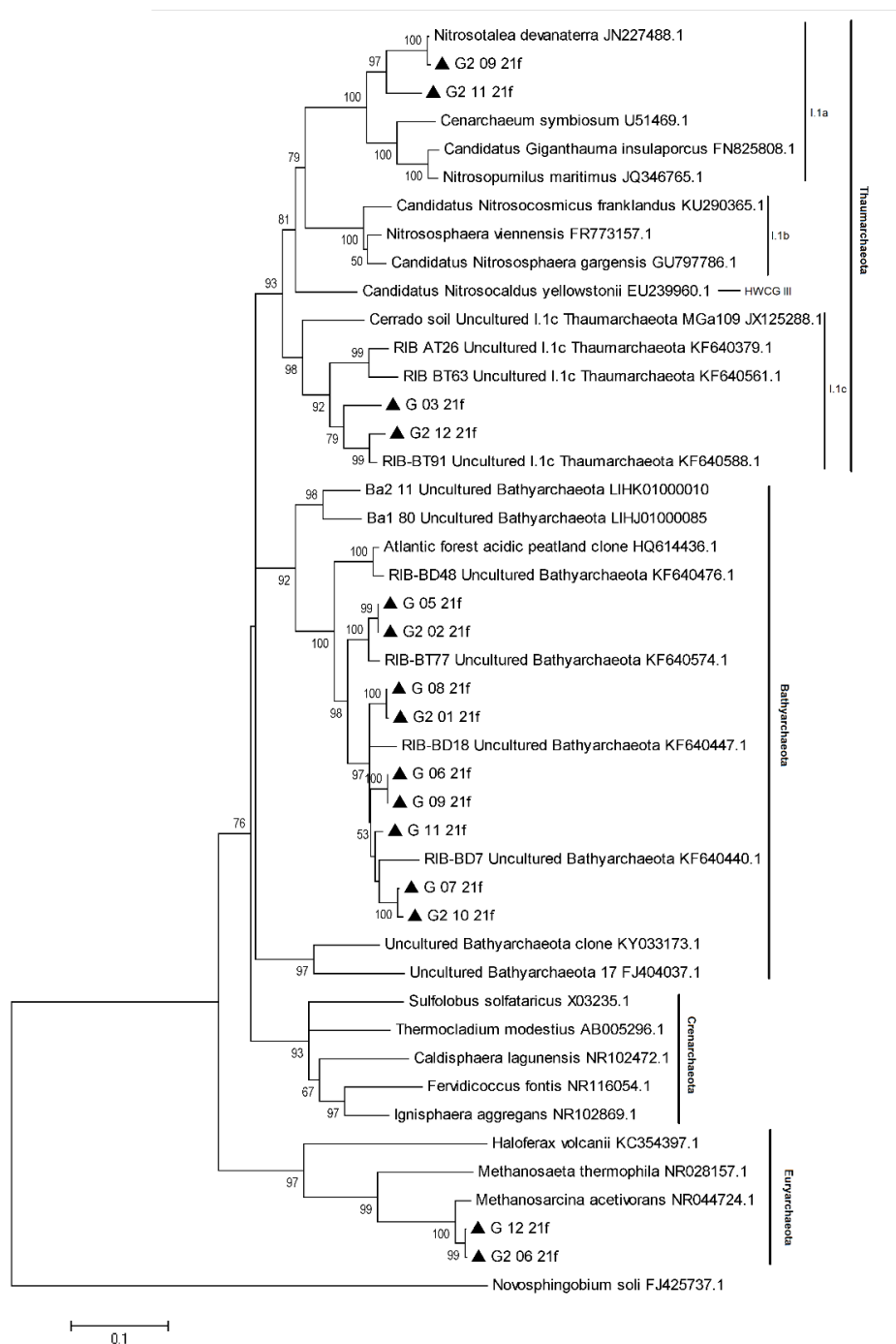


Figura 4. Árvore filogenética com as sequências do gene de rRNA 16S de *Archaea* obtidas com os iniciadores 21f/958r a partir do cultivo de solo de Cerrado (indicadas com triângulos) alinhadas com sequências de archaeas cultivadas e não cultivadas, retiradas de banco de dados do NCBI. Os números de acesso são indicados ao lado do nome de cada sequência. A árvore foi construída utilizando-se o MEGA5 (Tamura *et al.*, 2011), pelo método de *Neighbor-joining* (Saitou & Nei, 1987), com o modelo de Jukes-Cantor (Jukes e Cantor, 1969). Os valores (%) obtidos no teste de *bootstrap* (1000 réplicas) (Felsenstein, 1985) são indicados próximo aos ramos, sendo que aqueles inferiores a 50 % não são mostrados na árvore. A barra de escala indica o número de substituições por posição de nucleotídeo. Uma sequência da bactéria *Novosphingobium soli* foi utilizada como grupo externo. Os filios e grupos de *Archaea* estão evidenciados na figura. HWCG III: *Hot Water Crenarchaeotic Group III*.

Posteriormente, novas extrações de DNA foram realizadas a partir dos tipos coloniais G e G2, após cerca de um ano de repiques periódicos dessas culturas. Os DNAs extraídos foram empregados em ensaios de PCR utilizando-se os iniciadores 109f/915r, um par frequentemente utilizado na descrição da comunidades de archaeas ambientais (Galand *et al.*, 2006; Sanchez *et al.*, 2011; Staley *et al.*, 2011; Staley *et al.*, 2012). Os fragmentos de DNA amplificados foram clonados, sequenciados e a classificação das sequências em árvores filogenéticas mais uma vez revelou fragmentos do gene rRNA 16S afiliados ao filo Bathyarchaeota, ao grupo I.1c de Thaumarchaeota, assim como à classe *Methanomicrobia* (Tabela 4 e Figura 5).

Tabela 4. Classificação das sequências do gene de rRNA 16S de *Archaea* amplificadas com os três diferentes pares de iniciadores a partir dos tipos coloniais G e G2.

Iniciadores	Bathyarchaeota	Thaumarchaeota			Euryarchaeota		Parvarchaeota
		I.1a	I.1b	I.1c	<i>Methanomicrobia</i>	Outras	
21f/958r	9	2	-	2	2	-	-
109f/915r	17	43	14	13	17	8	2
Bathy442f/644r	21	-	-	14	-	1	-

No entanto, várias sequências afiliadas aos grupos I.1a e I.1b de Thaumarchaeota e outras classes de metanogênicas também foram identificadas (Figura 5). Além disso, duas sequências amplificadas por esse par de iniciadores foram classificadas como pertencentes ao filo Parvarchaeota (Tabela 4), composto por archaeas ultrapequenas ainda não cultivadas, inicialmente identificadas por reconstruções metagenômicas em drenados de minas (Baker *et al.*, 2009). Mais recentemente, foi proposto que este filo faria parte do superfilo DPANN (Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota e Nanohaloarchaeota), composto por linhagens de archaeas com tamanho e genomas diminutos, que viveriam associadas a outras archaeas e que possuem poucos representantes cultivados (Rinke *et al.*, 2013). Apesar de sequências de DNA afiliadas ao

filo Parvarchaeota serem majoritariamente encontradas em drenados ácidos de minas e fontes termais, alguns trabalhos relatam sua detecção em solos (Morais *et al.*, 2016; Essel *et al.*, 2017). Entretanto, o baixo número de sequências do gene rRNA 16S disponíveis para esse filo e o amplo desconhecimento acerca de suas características podem resultar em classificações ainda pouco acuradas nos bancos de dados.

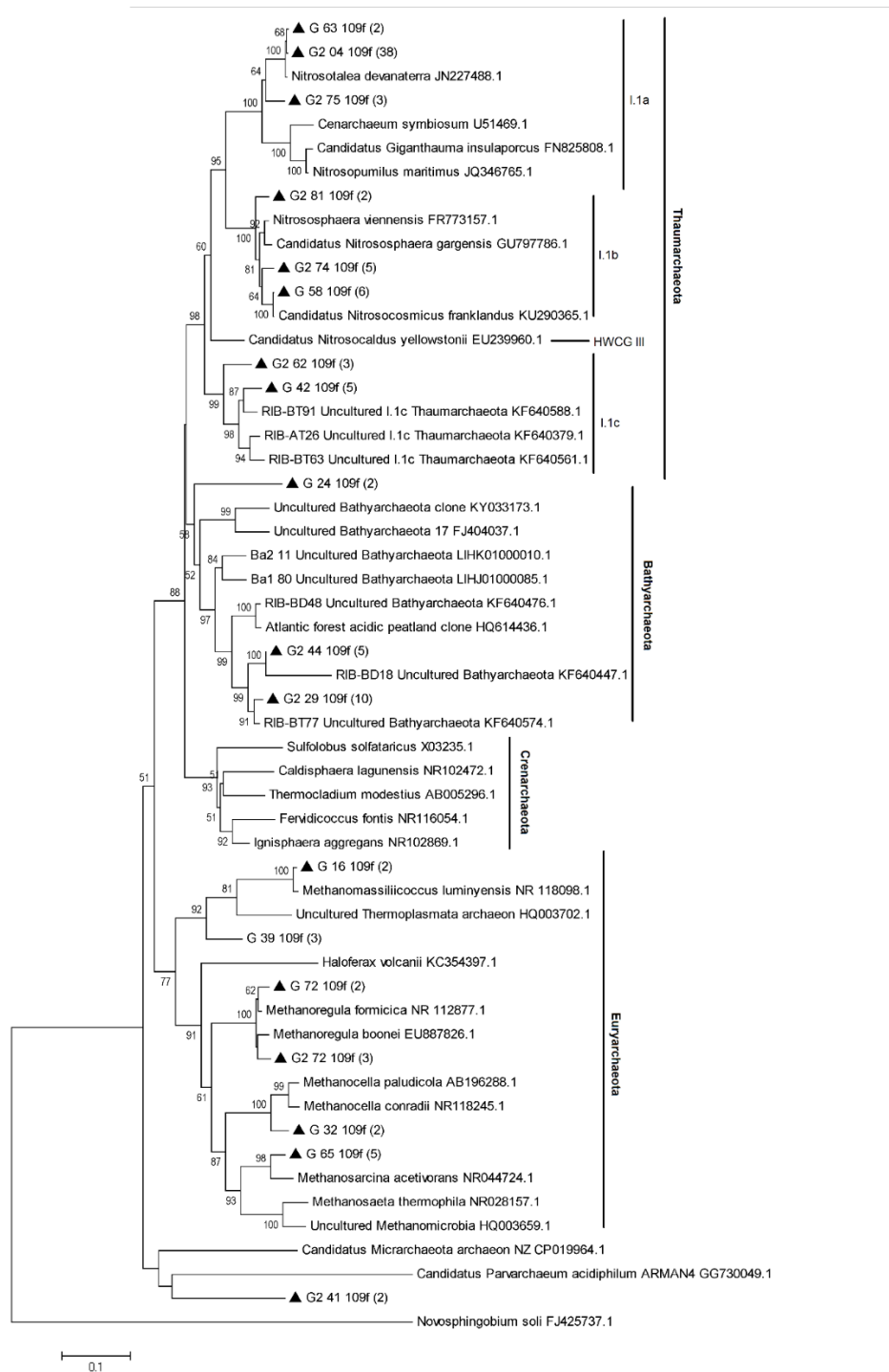


Figura 5. Árvore filogenética com as sequências do gene de rRNA 16S de *Archaea* obtidas com os iniciadores 109/915r, a partir do cultivo de solo de Cerrado (indicadas com triângulos), alinhadas com sequências de archaeas cultivadas e não cultivadas retiradas de banco de dados do NCBI. Foram incluídas na árvore apenas sequências de OTUs (Unidades Taxonômicas Operacionais) representativas a 97 % de identidade e o número de sequências representadas por cada uma está indicado entre parênteses. *Singletons* (OTUs representadas por uma única sequência) não foram incluídos. Os números de acesso das sequências retiradas do banco de dados são indicados ao lado do nome de cada sequência. A árvore foi construída utilizando-se o MEGA5 (Tamura et al., 2011), pelo método de *Neighbor-joining* (Saitou e Nei, 1987), com o modelo de Jukes-Cantor (Jukes e Cantor, 1969). Os valores (%) obtidos no teste de *bootstrap* (1000 réplicas) (Felsenstein, 1985) estão indicados próximo aos ramos, sendo que os inferiores a 50 % não são mostrados. A barra de escala indica o número de substituições por posição de nucleotídeo. Uma sequência da bactéria *Novosphingobium soli* foi utilizada como grupo externo. Os filos e grupos de *Archaea* estão evidenciados na figura. HWCG III: *Hot Water Crenarchaeotic Group III*.

A grande riqueza de archaeas identificada por meio do gene rRNA 16S nos cultivos estabelecidos foi surpreendente. Embora grande parte das archaeas mesófilas cultivadas até hoje tenham sido obtidas a partir do uso de meios enriquecidos para organismos oxidantes de amônia (Simon *et al.*, 2005; Hatzenpichler *et al.*, 2008; Jung *et al.*, 2011, Kim *et al.*, 2012; Stieglmeier *et al.*, 2014), a falta de informações acerca das necessidades metabólicas de outros grupos não extremófilos ainda representa uma grande limitação para seu cultivo. O fato de o meio de cultura utilizado neste trabalho ser confeccionado unicamente a partir do solo, a fim de mimetizar ao máximo as condições nutricionais do ambiente natural, sem qualquer tipo de enriquecimento para metabolismos específicos, poderia explicar a riqueza de archaeas observada. Assim, tal meio poderia permitir o crescimento de diferentes espécies presentes no solo, inclusive de representantes de grupos ainda não cultivados em meios artificiais.

No entanto, também foi levantada a possibilidade de estarem ocorrendo contaminações externas nos ensaios de PCR, que levariam à detecção dessa grande variedade de genes rRNA 16S de archaeas. Para avaliar essa possibilidade, iniciadores específicos para o filo Bathyarchaeota recentemente propostos (Yu *et al.*, 2017) foram sintetizados e empregados em ensaios de PCR realizados sob condições em que o risco de contaminação foi rigorosamente monitorado. Novas preparações de DNA obtidas dos tipos coloniais G e G2 crescidos em meio contendo ampicilina, estreptomina e cloranfenicol e em meio contendo polimixina B foram utilizadas e, mesmo assim foram detectados produtos de DNA de tamanho esperado em todas as amostras. Vale ressaltar que reações controle contendo, no lugar de DNA, meio de cultura líquido estéril, raspados do meio de cultura sólido estéril, ou água, foram incluídos nesses ensaios de PCR e não foram detectados produtos de amplificação em qualquer um destes.

O sequenciamento dos fragmentos amplificados a partir das amostras de DNA dos cultivos resultou em 36 sequências de qualidade. O alinhamento das sequências com o banco de dados *Greengenes* resultou na afiliação de 21 sequências com o filo Bathyarchaeota e de 14 sequências com o grupo I.1c de Thaumarchaeota, além de uma sequência classificada na classe *Methanobacteria* (Tabela 4), o que também foi observado na construção da árvore filogenética contendo essas sequências (Figura 6). Apesar desse par de iniciadores ser definido como específico para Bathyarchaeota, os autores relataram o seu pareamento com genes de rRNA 16S de outras archaeas ambientais (Yu *et al.*, 2017), o que poderia explicar a detecção destes outros grupos nas sequências amplificadas com esse par de iniciadores.

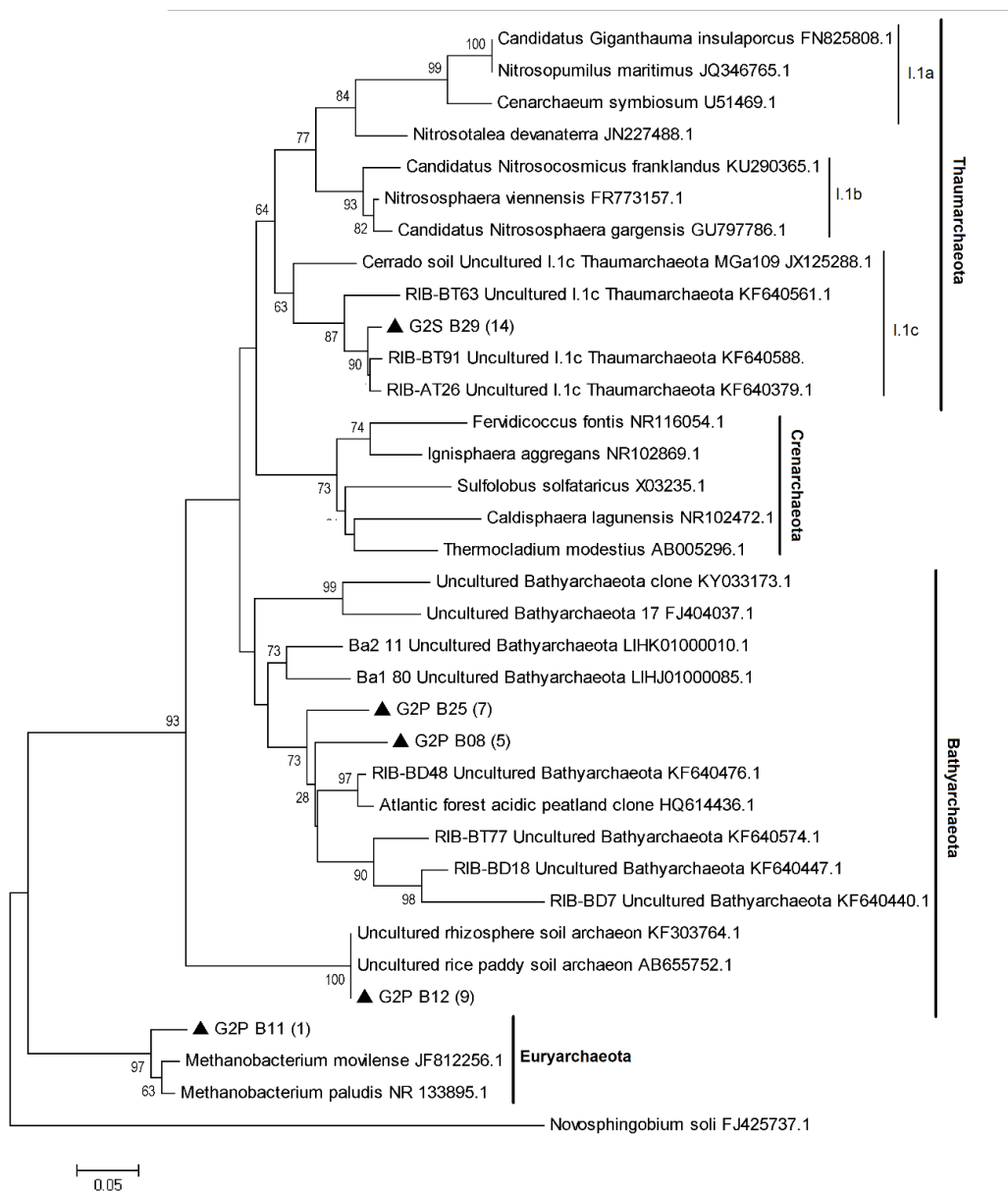


Figura 6. Árvore filogenética com as sequências do gene de rRNA 16S de *Archaea* obtidas com os iniciadores Bathy442f/644r, a partir do cultivo de solo de Cerrado (indicadas com triângulos), alinhadas com sequências de archaeas cultivadas e não cultivadas retiradas de banco de dados do NCBI. Foram incluídas na árvore as sequências de OTUs (Unidades Taxonômicas Operacionais) representativas a 97 % de identidade e o número de sequências representadas por cada uma delas está indicado entre parênteses. Os números de acesso das sequências retiradas do banco de dados são indicados ao lado do nome de cada sequência. A árvore foi construída utilizando-se o MEGA5 (Tamura et al., 2011), pelo método de *Neighbor-joining* (Saitou e Nei, 1987), com o modelo de Jukes-Cantor (Jukes e Cantor, 1969). Os valores (%) obtidos no teste de *bootstrap* (1000 réplicas) (Felsenstein, 1985) estão indicados próximo aos ramos, sendo que os inferiores a 50 % não são mostrados. A barra de escala indica o número de substituições por posição de nucleotídeo. Uma sequência da bactéria *Novosphingobium soli* foi utilizada como grupo externo. Os filós e grupos de *Archaea* estão evidenciados na figura.

Assim, os resultados obtidos a partir da análise dos fragmentos dos genes de rRNA 16S amplificados a partir de DNAs extraídos do cultivo em diferentes momentos empregando-se quatro pares de iniciadores distintos sugeriram o estabelecimento de um co-cultivo entre diferentes grupos de archaeas e bactérias pertencentes ao gênero *Novosphingobium*. Vale ressaltar que este gênero bacteriano e todos os grupos de archaeas (Bathyarchaeota, Thaumarchaeota e euryarchaeotas metanogênicas) identificados são comumente encontrados em solos (Timonen e Bomberg, 2009, Auguet *et al.*, 2010, Nguyen *et al.*, 2016) já tendo sido, inclusive, detectados em solos de Cerrado por abordagens independentes de cultivo (Catão *et al.*, 2013, Castro *et al.*, 2016; Belmok *et al.*, 2019).

É importante mencionar que, além da riqueza de filós e subgrupos de archaeas sugerida pela classificação taxonômica das sequências do gene rRNA 16S obtidas a partir do cultivo, um grande número de OTUs (Unidades Taxonômicas Operacionais) a 97 % de identidade foram identificadas a partir das análises destas sequências. Esse critério é amplamente utilizado em análises filogenéticas envolvendo a comparação de genes rRNA 16S e indicaria a presença de diversas espécies no cultivo. No entanto, heterogeneidades de genes de rRNA 16S em um mesmo genoma e análises baseadas em diferentes regiões deste gene podem levar a uma superestimação do número de OTUs (Sun *et al.*, 2013), sendo possível que as diferentes regiões do gene amplificadas por cada par de iniciadores (Figura 1) levem a tais resultados. Vale ainda ressaltar que, apesar de ser possível o alinhamento das sequências obtidas com os três pares de iniciadores escolhidos para o domínio *Archaea* por amplificarem regiões em comum do gene (Figura 1), o fato das sequências amplificadas pelos iniciadores Bathy442f/644r serem muito curtas (por volta de 200 pb) dificultou as análises de todas as sequências em conjunto, limitando a resolução filogenética obtida.

1.4.4. Análises microscópicas das células presentes no co-cultivo

Células das culturas G e G2 em meio sólido foram submetidas à coloração diferencial de Gram e observadas por microscopia de campo claro (Figura 7). Pode-se observar que em ambos os tipos coloniais houve predominância de diminutas células Gram negativas. Em ambas as amostras, algumas das células aparentam ter morfologia cocóide, enquanto outras apresentavam-se mais bacilares.

As bactérias do gênero *Novosphingobium* são bacilos Gram negativos, com 0,3-0,5 x 1-3 μm (Takeuchi *et al.*, 2001), com morfologia compatível com aquela observada nos esfregaços dos co-cultivos. Apesar da coloração de Gram ser amplamente utilizada para diferenciação de organismos do domínio *Bacteria*, as archaeas apresentam paredes celulares com estrutura e composição muito variáveis, muito diferentes daquelas encontradas em bactérias (Albers e Meyer, 2011). Desta forma, pouco se sabe sobre o padrão de coloração apresentado pelas archaeas, quando submetidas à coloração de Gram. De acordo com Beveridge e Schultze-Lam (1996), archaeas que possuem camada S como único envoltório celular e/ou paredes celulares compostas por metanocondroitina coram-se como Gram negativas e apenas a archaea *Methanobacterium formicicum*, que possui parede composta por pseudopeptideoglicano, apresenta o padrão de coloração Gram positivo. Os resultados obtidos na microscopia também poderiam sugerir a presença de archaeas, uma vez que grande parte das espécies mesófilas descritas até o momento exibem morfologia cocóide e dimensões diminutas, com diâmetros normalmente variando entre 0,5 e 1 μm (Tourna *et al.*, 2011; Lehtovirta-Morley *et al.*, 2016; Li *et al.*, 2016).

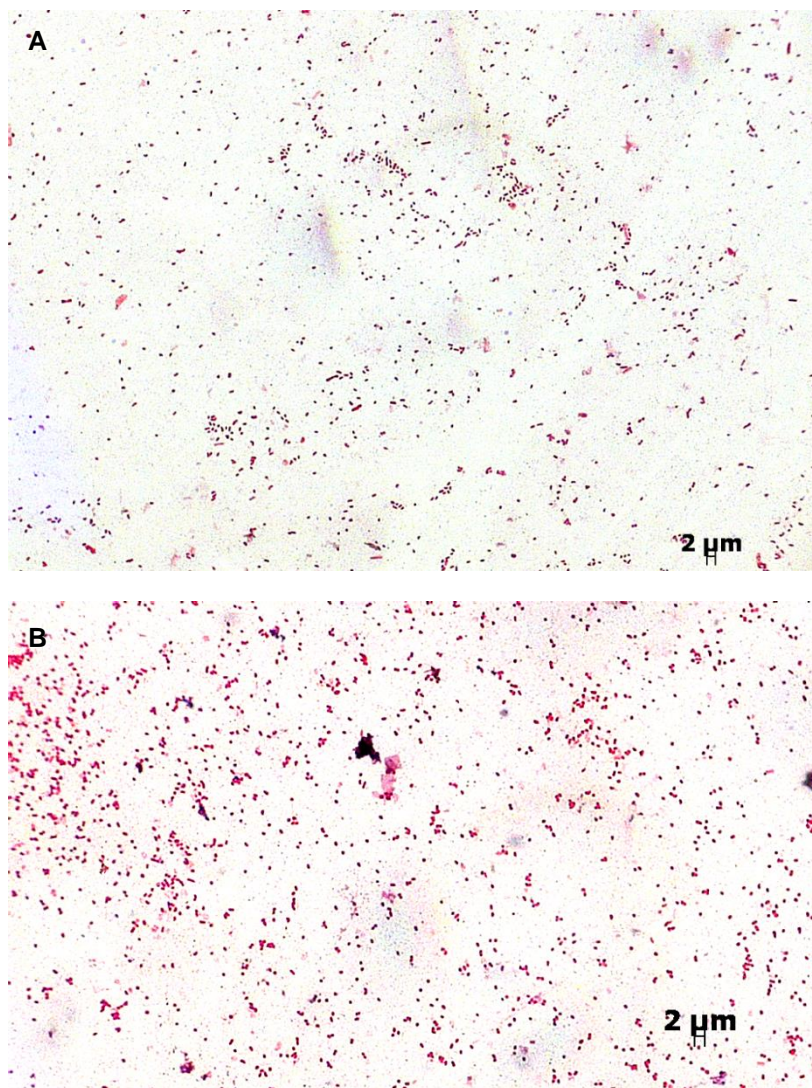


Figura 7. Coloração de Gram dos tipos coloniais G (A) e G2 (B), obtidos a partir de amostras de solo de Cerrado. Aumento 630x.

As culturas obtidas em meio sólido também foram analisadas a fresco, por microscopia de contraste de fase. Novamente, foram observadas células muito pequenas nos tipos coloniais G e G2, algumas com aspecto cocóide e outras bacilares, sendo a maioria das células observadas móveis.

Amostras do cultivo G2 em meio sólido foram ainda analisadas por MEV, onde foi possível observar dois tipos celulares distintos: bacilos maiores, de 1 a 3 μm, e células cocóides, com aspecto irregular e diâmetro inferior a 1 μm (Figura 8). Os bacilos observados apresentam morfologia e tamanho condizentes com o descrito para as bactérias do gênero *Novosphingobium* (Takeuchi *et al.*, 2001; Saxena *et al.*, 2013). Muitas das archaeas mesófilas cultivadas até o momento, quando observadas por MEV, apresentam aspecto esférico, irregular e diâmetros inferiores a 1 μm (Tournia *et al.*, 2011;

Lethorvita-Morley *et al.*, 2016), assemelhando-se aos cocos observados no co-cultivo. Além disso, experimentos de hibridização *in situ* por fluorescência (FISH) realizados com sondas específicas para o filo Bathyarchaeota a partir de amostras ambientais sugeriram que a maioria dos organismos detectados seriam cocos com tamanhos entre 0,4 e 5 μm (Kubo *et al.*, 2012).

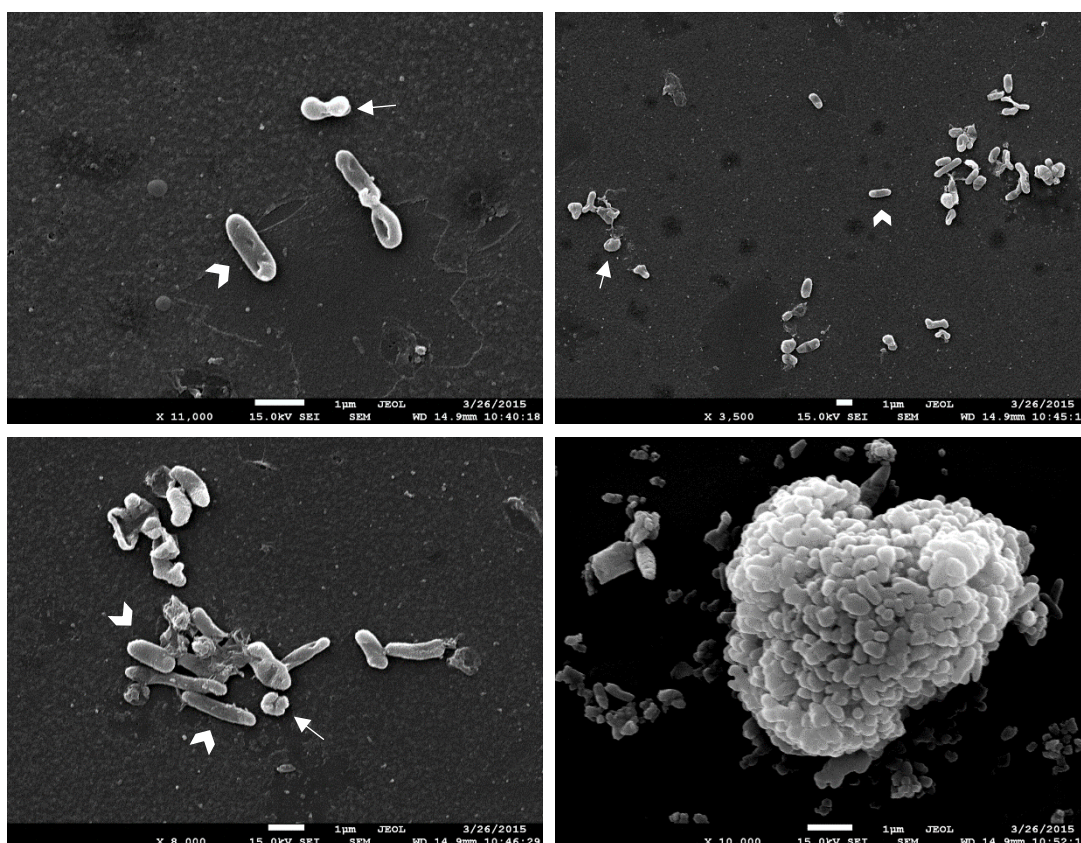


Figura 8. Eletromicrografias de varredura de células da cultura G2 crescida em meio sólido, ilustrando a ocorrência de células bacilares (cabeças de seta) e cocóides (setas). Os aumentos utilizados estão indicados em cada imagem.

Após a constatação do crescimento das culturas em meios líquidos, os tipos coloniais G e G2 foram raspados do meio sólido e inoculados em meios líquidos suplementados inicialmente com ampicilina, estreptomicina, cloranfenicol e itraconazol e, posteriormente, também com polimixina B. Amostras das culturas incubadas por aproximadamente um mês foram então concentradas, fixadas e submetidas à coloração de Gram. Assim como observado nos cultivos sólidos, diminutas células cocóides e bacilares Gram negativas foram visualizadas nas preparações. No entanto, bacilos alongados, maiores que as demais células, também estavam presentes em ambas as amostras (Figura 9). É possível que as condições nutricionais encontradas nos meios líquidos (que, diferentemente dos meios sólidos, são filtrados em membranas de 0,22 μm)

sejam mais limitadas, dificultando a divisão celular de algum dos organismos presentes no cultivo.

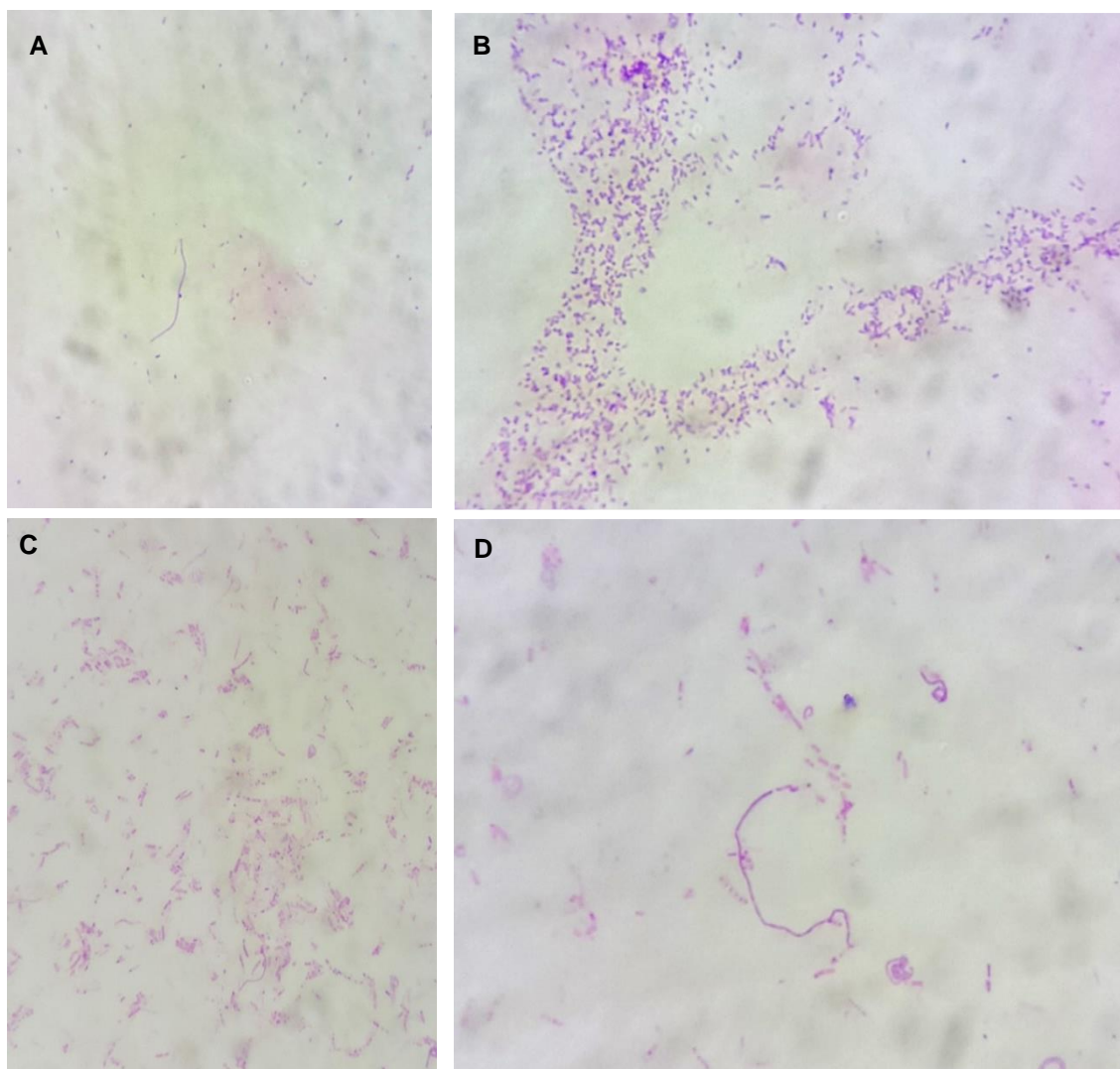


Figura 9. Coloração de Gram dos cultivos G (A e B) e G2 (C e D) obtidos em meio líquido, ilustrando a presença de células filamentosas, além das formas cocóides e bacilares visualizadas em preparações oriundas de meios sólidos. Aumento 1000x.

Visando otimizar a análise microscópica das células presentes nos meios líquidos, estas foram inoculadas sobre lamínulas de vidro, tratadas ou não com poli-L-lisina, incubadas por cerca de um mês a 28°C e observadas por microscopia de contraste de fase. Os cultivos G apresentaram uma maior proporção de pequenas células cocóides, embora longos bacilos filamentosos tenham sido detectados nesta amostra (Figura 10A e B). Algumas células encontravam-se aderidas às lamínulas, com muitas apresentando movimento ativo. As lamínulas de G2 apresentaram um maior crescimento em relação às de G, com maior proporção de bacilos. No entanto, cocos individuais muito pequenos e

arranjos semelhantes a cadeias tetracos também foram observadas nesta amostra (Figura 10C). Tanto bacilos quanto cocos móveis foram detectados.

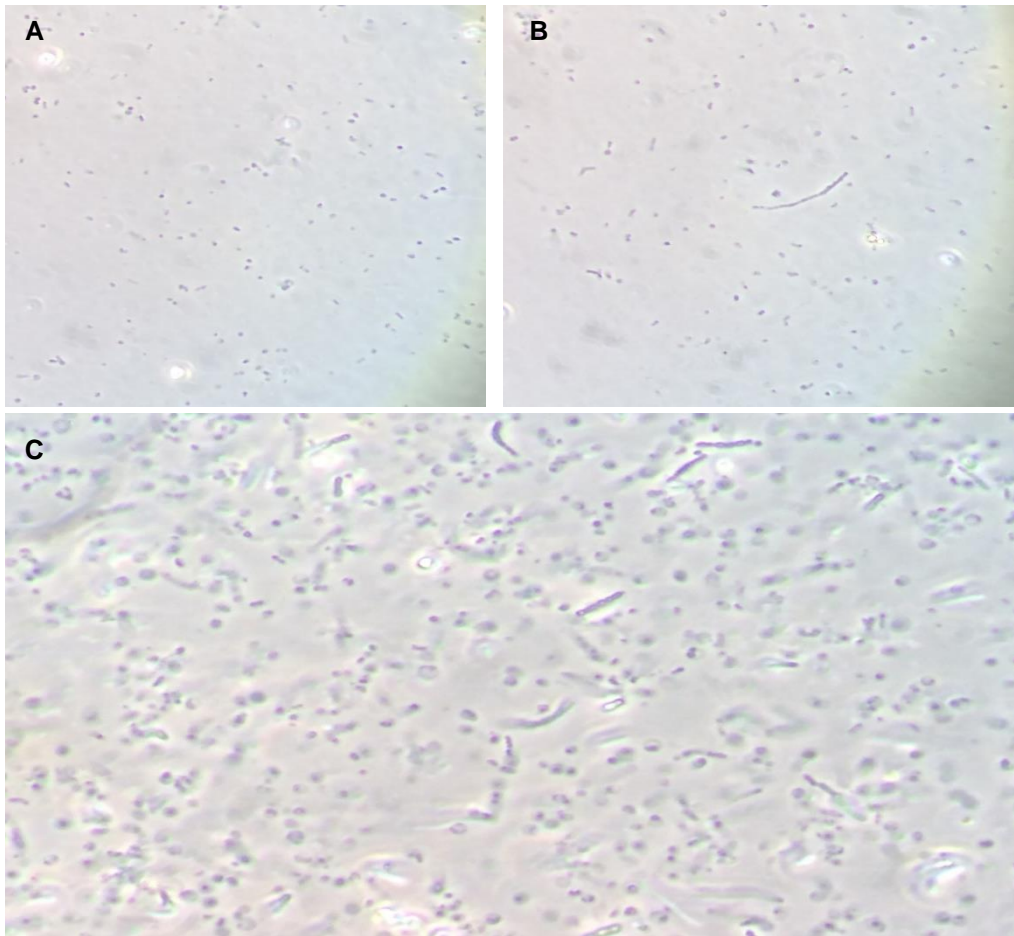


Figura 10. Análises a fresco por microscopia de contraste de fase das culturas G (A e B) e G2 (C) crescidas sobre lamínulas em meio líquido. Aumento 1000x.

Estas mesmas lamínulas foram posteriormente fixadas e preparadas para visualização por MEV. Nas amostras do cultivo G foram observados predominantemente cocos com superfícies irregulares e diâmetros em torno de $0,5 \mu\text{m}$, tanto nas lamínulas sem tratamento (Figura 11A, B e C) quanto naquelas tratadas com poli-L-lisina (Figura 11D). Nos dois tipos de lamínulas do cultivo G foram observados bacilos de aproximadamente $1 \mu\text{m}$ de comprimento, alguns com superfícies mais lisas e outros mais irregulares, evidenciando que células do cultivo G com diferentes morfologias foram capazes de crescer em meio líquido sobre ambos os tipos de lamínula. Já nas lamínulas de G2 sem tratamento prévio (Figura 12A e B) foram observados predominantemente bacilos mais longos e com superfícies mais regulares que os observados na amostra G, a maioria com cerca de $5 \mu\text{m}$ de comprimento, além de alguns com mais de $10 \mu\text{m}$ (Figura 12A). Alguns

cocos irregulares foram também visualizados nestas lamínulas. No entanto, nas lamínulas de G2 previamente tratadas com poli-L-lisina foi verificada uma quantidade maior de células, além de uma maior diversidade de tipos morfológicos. Cocos muito semelhantes aos observados nas amostras G e diferentes tipos de bacilos foram visualizados nessas lamínulas, incluindo bacilos longos, com aproximadamente 5 μm de comprimento e superfícies rugosas, além de cocobacilos, com pouco mais de 1 μm de comprimento (Figura 12C a F). Diversos bacilos com segmentações irregulares também foram observados nas amostras de G2, os quais poderiam corresponder aos tipos celulares anteriormente interpretados como cadeias de tetracocos nas análises destas lamínulas por microscopia de contraste de fase (Figura 12F).

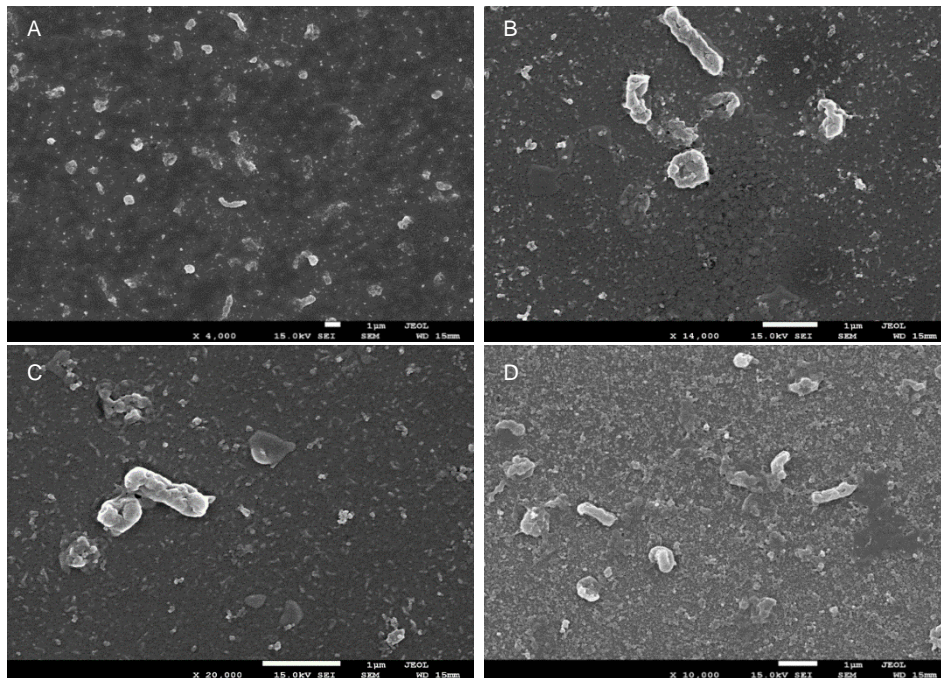


Figura 11. Eletromicrografias de varredura da cultura G crescida em meio líquido sobre lamínulas de vidro sem tratamento (A a C) e previamente tratadas com poli-L-lisina (D). Os aumentos utilizados estão indicados em cada imagem.

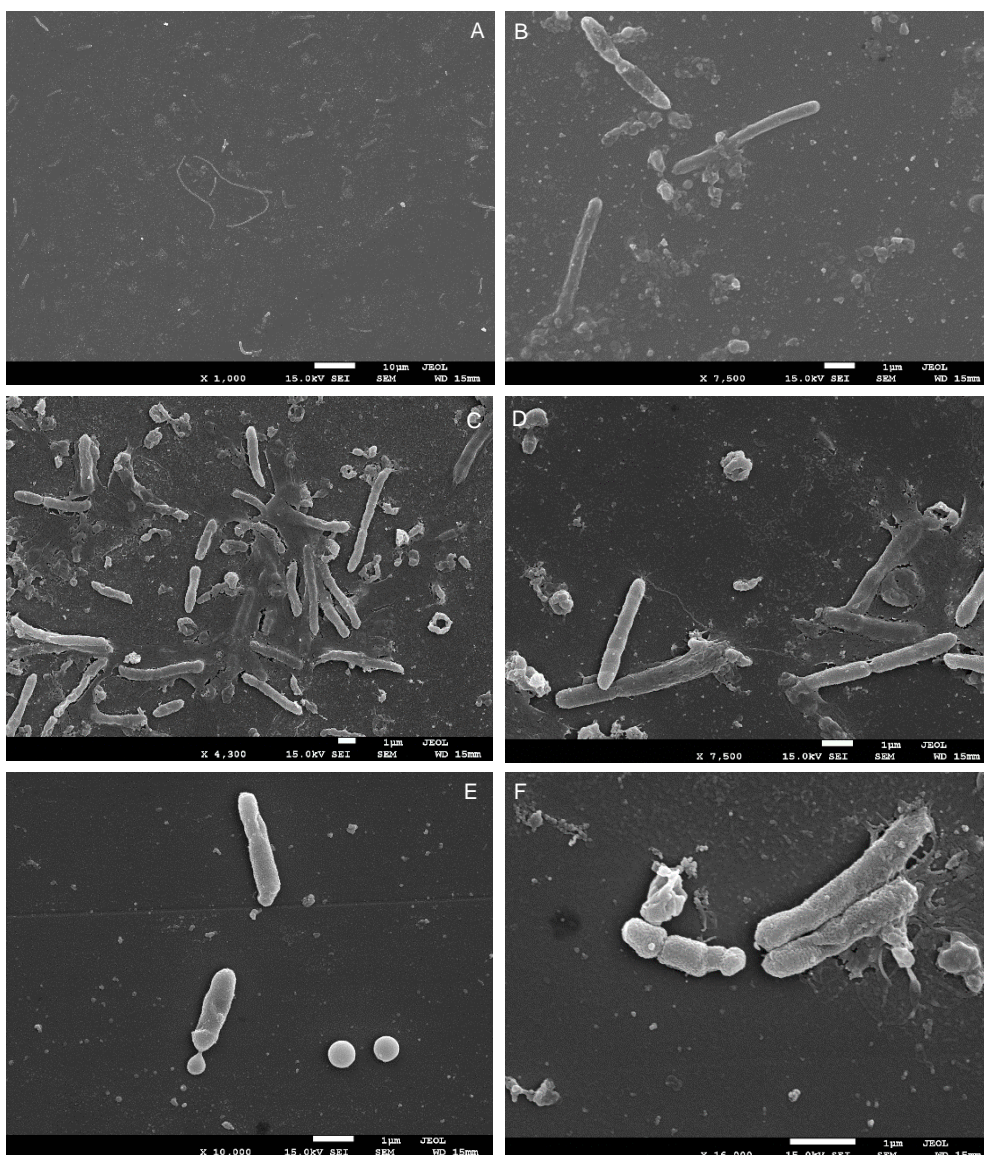


Figura 12. Eletromicrografias de varredura de células da cultura G2 crescida em meio líquido sobre lamínulas de vidro sem tratamento (A e B) e previamente tratadas com poli-L-lisina (C a F). Os aumentos utilizados estão indicados em cada imagem.

Com base nas descrições apresentadas na literatura, especulou-se que os bacilos maiores poderiam ser as bactérias do gênero *Novosphingobium*, enquanto os bacilos menores e cocos poderiam ser as archaeas, ambos domínios detectados nos cultivos pela amplificação e sequenciamento dos genes rRNA 16S. Assim, o antibiótico polimixina B, que tem como alvo a membrana celular de bactérias, frequentemente utilizado contra Gram negativas (Yuan e Tan, 2008), passou a ser adicionado em alguns repiques nos meios sólidos e líquidos. Conforme já mencionado, colônias com poucas diferenças morfológicas em relação às crescidas com os demais antibióticos foram observadas (Figura 3). Além disso, nas diferentes análises por microscopia não foram detectadas

diferenças nas células crescidas em meio sólido na presença de polimixina B em relação às células crescidas nos demais antibióticos (fotos não mostradas). No entanto, foi observado um menor crescimento nas culturas G e G2 suplementadas com polimixina B, com poucos cocos móveis observados em preparações a fresco realizadas após meses de incubação. A principal diferença foi notada nos cultivos líquidos de G2, que apresentavam grande quantidade de bacilos quando suplementados com ampicilina, estreptomicina, cloranfenicol e itraconazol (conforme evidenciado nas Figura 10 e Figura 12), mas quando adicionados de polimixina B e incubados por vários dias, passaram a exibir apenas poucas células cocóides.

Com a finalidade de se observar a viabilidade das células presentes nos cultivos líquidos de G2 suplementados com ampicilina, estreptomicina, cloranfenicol e itraconazol (que serão denominados cultivos SACI) e naqueles suplementados com polimixina B após aproximadamente 60 dias de incubação, alíquotas destes cultivos foram analisadas por microscopia de fluorescência, após incubação com os reagentes do kit *LIVE/DEAD BacLight*TM. A maioria das células observadas nos cultivos SACI, após quase dois meses de incubação estavam vivas e, como detectado anteriormente nas demais microscopias, bacilos e cocos puderam ser identificados na amostra (Figura 13A). No entanto, após a análise de vários campos e diversas alíquotas, poucas células puderam ser detectadas nos cultivos tratados com polimixina B, apenas com a observação de um pequeno número de cocos viáveis nessas amostras (Figura 13B).

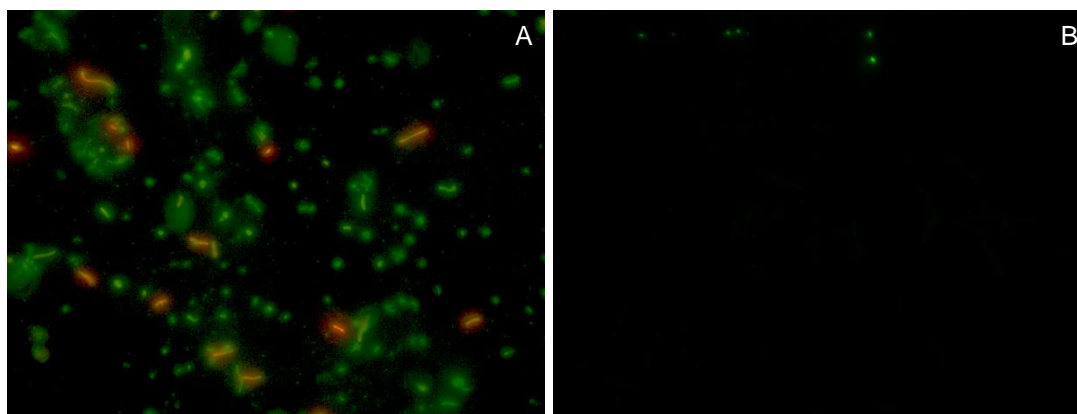


Figura 13. Microscopia de fluorescência de células da cultura G2 crescidas em meio líquido contendo ampicilina, estreptomicina, cloranfenicol e itraconazol (SACI) (A) ou polimixina B (B), após incubação com o kit *LIVE/DEAD BacLight*TM. Células vivas aparecem em verde e células mortas em vermelho.

A significativa diminuição no número de células nos meios adicionados de polimixina B poderia sugerir a eliminação das bactérias por esse antibiótico, talvez

resultando em uma limitação no crescimento das archaeas, conforme já relatado para o cultivo de archaeas de solo (Lehtovirta-Morley *et al.*, 2014). De fato, após o primeiro repique dos cultivos em meios líquidos suplementados com polimixina B, não foi mais possível a visualização de células nas análises microscópicas, não sendo detectado indícios de crescimento mesmo após alguns meses de incubação.

Já no meio sólido, apesar de inicialmente ter sido observado um crescimento microbiano mais lento nos cultivos com polimixina B, análises de esfregaços das colônias em microscópio óptico revelaram um grande número de bacilos após alguns repiques na presença do antibiótico e fragmentos do gene rRNA 16S foram amplificados com iniciadores específicos para o domínio *Bacteria* a partir de DNAs extraídos desses cultivos. O menor acesso do antibiótico às células mais internas das colônias, ou a produção de matriz extracelular pelos microrganismos podem ter limitado a ação da polimixina B nos cultivos sólidos, possibilitando o crescimento das bactérias nessa condição. Assim, apesar dos indícios iniciais da diminuição de bactérias nos cultivos após a adição de polimixina B nos meios, não foi possível o enriquecimento de archaeas nos co-cultivos de solo de Cerrado por meio dessa estratégia.

1.4.5. Sequenciamento genômico total dos cultivos de solo de Cerrado

Para o sequenciamento genômico dos organismos presentes no cultivo, diversas extrações do DNA total de placas dos cultivos G e G2 com diferentes tempos de crescimento foram realizadas, objetivando-se desta forma a obtenção de sequências genômicas de organismos que poderiam estar mais abundantes em determinados momentos e menos abundantes em outros. Após a avaliação qualitativa por eletroforese em gel de agarose e quantificação fluorimétrica, as amostras de DNA total de G e G2 foram reunidas, para o posterior preparo das bibliotecas e sequenciamento pelas plataformas Illumina HiSeq e PacBio.

Conforme mencionado na seção Material e Métodos deste capítulo (seção 1.3.13), duas abordagens distintas foram utilizadas inicialmente para a montagem e classificação taxonômicas das sequências obtidas dos cultivos em meio sólido. A primeira consistiu na utilização dos *short-reads* gerados pelo sequenciamento na plataforma Illumina e dos *long-reads* resultantes do sequenciamento PacBio, tanto separadamente como de maneira conjunta. A primeira montagem, somente com os dados de alta cobertura de Illumina HiSeq, resultou em 70 *contigs*, o maior deles de 1,3 MB, enquanto a montagem híbrida,

utilizando os *reads* gerados por ambas as plataformas, apresentou resultados mais robustos, resultando em 7 *contigs* no total. Já a montagem realizada apenas com os *long-reads* revelou-se a melhor estratégia, originando apenas 5 *contigs* e, por essa razão, tais resultados foram utilizados para as análises posteriores de classificação taxonômica das sequências. Os maiores *contigs* apresentavam aproximadamente 4,2 e 2,7 Mb e os três menores 226, 85 e 31 kb. Análises comparativas com o banco de dados não redundante do NCBI revelaram que as sequências genômicas obtidas pertenciam ao domínio *Bacteria*, com os maiores *contigs* apresentando correspondência (*hits*) com sequências de DNA de organismos da ordem *Sphingomonadales*.

A segunda abordagem realizada, que envolveu a montagem *de novo* dos *short-reads* com parâmetros metagenômicos da ferramenta IDBA-UD, resultou na montagem de 84 *contigs*. A posterior identificação de *bins* genômicos, por meio da análise de genes de cópia única e rRNAs, sugeriu que todos os *contigs* pertenciam a um único genoma de 7,11 Mb e conteúdo GC de 63,6%, também associado à ordem *Sphingomonadales*.

A análise dos dados do sequenciamento genômico das amostras de cultivo em meio líquido realizado posteriormente também revelou sequências afiliadas filogeneticamente ao domínio *Bacteria*. Todos os *reads* de qualidade obtidos nesse último sequenciamento puderam ser mapeados no genoma bacteriano montado a partir das sequências oriundas do cultivo em meio sólido, não tendo sido identificadas sequências com classificações taxonômicas diferentes. A montagem, anotação e análise do genoma bacteriano obtido serão discutidas em maior detalhe no capítulo seguinte.

Em resumo, os dados genômicos obtidos, sequenciados e analisados por diferentes estratégias, sugeriram a presença de uma única espécie bacteriana no cultivo, com todas as sequências obtidas identificadas como pertencentes a um genoma bacteriano, não sendo possível a detecção de sequências de DNA associadas a outros microrganismos. Tal resultado foi surpreendente, visto que as diferentes análises por microscopia indicavam a presença de mais de um tipo celular nas amostras. Além disso, o fato de não terem sido identificadas sequências afiliadas ao domínio *Archaea* nos sequenciamentos de alto rendimento realizados não condizem com os resultados obtidos por meio da amplificação de genes rRNA 16S discutidos anteriormente, que sugeriram a presença de archaeas e bactérias no cultivo.

Hipóteses tais como a menor eficiência dos protocolos de extração de DNA no caso de células de archaeas, ou mesmo a menor proporção destes organismos em relação às bactérias nos cultivos foram levantadas para tentar explicar os resultados obtidos no

sequenciamento genômico das células do co-cultivo. O desconhecimento acerca da estrutura e composição da parede celular da grande maioria das archaeas, principalmente de ambientes naturais não extremos, pode fazer com que os protocolos estabelecidos não sejam adequados para a extração do DNA de todos os organismos deste domínio, levando a menor eficiência na obtenção de DNA de alguns grupos (Cruaud *et al.*, 2014). Por se tratar de uma técnica altamente sensível, é possível que os ensaios de PCR tenham sido capazes de detectar os genes de rRNA 16S de archaea, mesmo que em baixa quantidade. No entanto, estas sequências de archaeas poderiam ser pouco representativas nos sequenciamentos genômicos. A hipótese de contaminação externa nos experimentos de PCR realizados também foi levantada, mas parece improvável visto que, conforme discutido na seção 1.4.3, não foram identificadas amplificações nos controles negativos realizados com água ou meios de cultura líquido ou sólido não inoculados.

Por outro lado, considerando-se os indícios da presença de mais de um microrganismo nos cultivos e visando uma boa cobertura para a montagem dos diferentes genomas esperados, as estratégias escolhidas possibilitaram uma grande profundidade de sequenciamento. Por esse motivo, a detecção de sequências associadas a outros microrganismos presentes no cultivo, por mais que estivessem em abundâncias muito pequenas no momento das amostragens de DNA total, era esperada. Desta forma, é possível que organismos do domínio *Archaea* tenham estado presentes nas culturas estabelecidas, sendo identificados no momento da realização dos ensaios de PCR, mas que o crescimento das bactérias tenha se sobressaído ao longo do tempo no cultivo, com os sucessivos repiques. Tal fenômeno é frequentemente relatado em trabalhos envolvendo o cultivo de archaeas mesófilas e dentre as possíveis causas estão a sua baixa taxa de crescimento e baixa abundância em ambientes naturais, o que resulta em um crescimento mais rápido de bactérias em culturas artificiais (Sun *et al.*, 2019). Além disso, a ausência de fatores de crescimento e moléculas sinalizadoras produzidos por outros organismos *in situ*, a competição interespecífica por nutrientes limitados e a dificuldade de se detectar o crescimento das archaeas devido à sua lenta taxa de duplicação são apontadas como grandes limitações para a obtenção de culturas de archaeas *in vitro* (Sun *et al.*, 2019).

Vale também ressaltar que as archaeas de solo cultivadas em laboratório até o momento possuem a capacidade de oxidar amônia, tendo sido empregadas estratégias específicas para a seleção deste metabolismo nos cultivos, bem como metodologias de detecção da conversão de amônia em nitrito como forma de avaliação indireta do crescimento (Lehtovirta-Morley *et al.* 2011; Jung *et al.* 2011; Tourna *et al.*, 2011;

Lehtovirta-Morley *et al.*, 2014; Jung *et al.* 2014; Lehtovirta-Morley *et al.*, 2016). No entanto, como ainda não há informações sobre as vias metabólicas existentes em linhagens sem representantes cultivados, tais como o subgrupo I.1c do filo Thaumarchaeota e o filo Bathyarchaeota, cujos genes de rRNA 16S foram identificados nos cultivos de solo do Cerrado, metodologias específicas para o enriquecimento e detecção desses grupos em cultivos ainda não estão disponíveis, dificultando a manutenção desses organismos em culturas *in vitro*.

Assim, por meio da utilização de estratégias que visavam a obtenção de culturas de archaea (meios confeccionados a partir do solo para mimetizar as condições nutricionais do ambiente natural e o emprego de agentes antibacterianos e antifúngicos), foi possível o estabelecimento de um co-cultivo de archaeas pertencentes aos filos Thaumarchaeota e Bathyarchaeota e bactérias do gênero *Novosphingobium*, a partir de amostras de solo de Cerrado. No entanto, provavelmente devido à pequena abundância das archaeas nos cultivos e considerando-se as limitações técnicas envolvidas no enriquecimento e detecção acurada desses organismos nos cultivos, é provável que o crescimento das bactérias tenha sobrepujado o das archaeas ao longo dos sucessivos repiques, levando a dificuldades na detecção das archaeas nos cultivos após alguns anos. Além disso, vale ressaltar que o isolamento de algumas espécies de *Novosphingobium* só pôde ser alcançado em meios de cultura acrescidos de antibióticos (Gupta *et al.*, 2009; Saxena *et al.*, 2013), evidenciando sua ampla resistência a antimicrobianos, o que poderia explicar a seleção e persistência dessa bactéria nos cultivos, mesmo com a constante adição destes compostos aos meios.

Por esses motivos, e considerando que as análises do genoma completo obtido da bactéria pertencente ao gênero *Novosphingobium* presente nos cultivos de solo de Cerrado sugeriram tratar-se de uma nova espécie, a próxima etapa do trabalho, apresentada no capítulo a seguir, objetivou o isolamento da cepa bacteriana em meios definidos e sua posterior caracterização morfo-fisiológica, bem como análises de aspectos funcionais de seu genoma.

CAPÍTULO 2 - Caracterização morfo-fisiológica e genômica da bactéria *Novosphingobium terrae* GeG2^T, isolada a partir de solo de Cerrado

2.1. Introdução

O gênero bacteriano *Novosphingobium*, pertencente à família *Sphingomonadaceae*, classe *Alphaproteobacteria* e filo *Proteobacteria*, foi proposto em 2001 por Takeuchi e colaboradores, a partir de uma reclassificação de membros do gênero *Sphingomonas*. Com base em filogenias do gene de rRNAs 16S e em critérios quimiotaxonômicos, que incluíam diferenças em poliaminas e perfis de ácidos graxos 2-hidroxi, foi proposto que o gênero *Sphingomonas* fosse dividido em quatro gêneros: *Sphingomonas sensu stricto* e os três novos gêneros *Novosphingobium*, *Sphingobium* e *Sphingopyxis* (Takeuchi *et al.*, 2001)

Membros de *Novosphingobium* são bacilos gram-negativos, móveis ou imóveis, não-esporulantes, com tamanho entre 0,1–1,0 x 0,6–4,0 µm e apresentando colônias esbranquiçadas, amareladas ou alaranjadas (Glaeser e Kämpfer, 2014). Bactérias pertencentes a esse gênero apresentam propriedades fisiológicas variadas, já tendo sido isoladas de diversos tipos de ambientes, tais como solos e rizosferas (Takeuchi *et al.* 1995; Kämpfer *et al.* 2011; Kämpfer *et al.* 2015), águas subterrâneas (Lee *et al.*, 2014), lagoas, lagos e fontes de água doce (Baek *et al.*, 2011; Ngo *et al.*, 2016; Sheu *et al.*, 2016), águas marinhas profundas (Yuan *et al.*, 2009; Huo *et al.*, 2015), sedimento de mangue (Lee *et al.*, 2014), estações de tratamento de águas (Tirola *et al.* 2005), emulsões lubrificantes (Kämpfer *et al.* 2018), dentre outros ambientes naturais e artificiais. Atualmente, 50 espécies de *Novosphingobium* descritas com nomenclaturas validadas constam na LPSN (<https://lpsn.dsmz.de/genus/novosphingobium>, acessado em agosto/2020).

Microrganismos desse grupo despertam interesse devido à sua habilidade de degradar xenobióticos naturais e antropogênicos recalcitrantes e compostos aromáticos diversos. Comumente encontradas em ambientes altamente contaminados, espécies de *Novosphingobium* capazes de degradar hidrocarbonetos aromáticos policíclicos (PAHs), compostos aromáticos derivados de lignina, compostos heterocíclicos, polietilenoglicóis (PEGs), disruptores endócrinos, herbicidas e pesticidas já foram identificadas, apresentando grande potencial biorremediativo (Balkwill *et al.* 1997; Fujii *et al.* 2003;

Tiirola et al., 2002; Yan *et al.*, 2007; Notomista *et al.*, 2011; Rodriguez-Conde *et al.*, 2016; Li *et al.*, 2020). Além disso, essas bactérias são produtoras de diferentes tipos de exopolissacarídeos (EPS) (Wu *et al.*, 2017), moléculas que apresentam vasta aplicabilidade industrial, podendo ser utilizadas como espessantes de alimentos e bebidas, na preservação de laticínios, na fabricação de cosméticos e fármacos, na agricultura, em processos de biorremediação, em aplicações médicas relacionadas à cura de feridas, entre outras aplicações (Sengupta *et al.*, 2018).

Muitas linhagens também são conhecidas por suas interações com espécies vegetais, podendo atuar tanto como promotores de crescimento de plantas (Zhang *et al.*, 2016; Krishnan *et al.*, 2017; Vives-Peris *et al.*, 2018,) quanto como patógenos desses organismos (Gan *et al.*, 2012). Ademais, apesar de raramente associadas a humanos, sequências de DNA associadas a *Novosphingobium* foram detectadas em amostras de pulmões saudáveis e acometidos por doenças pulmonares obstrutivas crônicas e um potencial papel de membros desse gênero como patógenos oportunistas foi sugerido com base em experimentos realizados em camundongos desafiados com *Novosphingobium panipatense* (Rutebemberwa *et al.*, 2014).

Apesar da distribuição ambiental ubíqua e das muitas funções fisiológicas e ecológicas interessantes às quais bactérias do gênero *Novosphingobium* já foram associadas, muitos aspectos relacionados a esse grupo microbiano continuam desconhecidos. Assim, trabalhos envolvendo o cultivo e caracterização de novos isolados, principalmente de ambientes ainda inexplorados como o solo de Cerrado, são fundamentais para o melhor conhecimento da biologia desse gênero, tão importante em aplicações industriais e biotecnológicas.

2.1. Objetivos

2.1.1. Objetivo geral

Isolar a linhagem bacteriana presente nos cultivos obtidos a partir do solo de Cerrado e caracterizá-la em termos morfológicos, fisiológicos, metabólicos, genômicos e taxonômicos.

2.1.2. Objetivos específicos

- Isolar e cultivar a linhagem bacteriana, pertencente ao gênero *Novosphingobium*, presente nos cultivos de solo de Cerrado;
- Avaliar o crescimento do novo isolado bacteriano em diferentes meios de cultura complexos/definidos;
- Descrever a morfologia do novo isolado bacteriano por meio de análises microscópicas;
- Identificar as espécies validadas de *Novosphingobium* mais próximas ao novo isolado bacteriano por meio da comparação de sequências do gene rRNA 16S e dos genomas completos;
- Determinar características químicas e fisiológicas da linhagem bacteriana cultivada e compará-las com as da espécie cultivada mais próxima (espécie referência);
- Identificar as faixas de temperatura, pH e salinidade de crescimento da bactéria;
- Avaliar e comparar os perfis proteicos do novo isolado bacteriano e da espécie referência por meio de espectrometria de massa;
- Descrever a estrutura genômica e analisar os potenciais metabólicos presentes no genoma do novo isolado bacteriano;
- Classificar taxonomicamente a bactéria por meio de uma abordagem polifásica e determinar se esta se trata de uma nova espécie;

2.2. Material e Métodos

2.2.1. Isolamento e cultivo da bactéria em meios de cultura definidos

Visando o isolamento da linhagem bacteriana presente nos cultivos de solo de Cerrado, colônias crescidas no meio confeccionado com solo foram inicialmente transferidas para meio LB (contendo por litro: 5 g de extrato de levedura, 10 g de NaCl, 10 g de peptona, 15 g de ágar bacteriológico). Considerando a impossibilidade da manutenção a longo prazo dos cultivos nesse meio de cultura, colônias crescidas no meio confeccionado com solo foram então transferidas para Meio Mínimo (MM, contendo por litro: 0.1 g KH_2PO_4 ; 0.2 g $(\text{NH}_4)_2\text{SO}_4$; 0.1 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$; 0.02 g $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$; 0.2 g NaCl; 0.1 g extrato de levedura; 15 g de ágar bacteriológico, pH 5,5) suplementado após esterilização em autoclave com glicose (concentração final 0,05 % - p/v), solução de elementos traço (concentração final de 0,1 % - v/v, contendo por litro: 1,5 g ácido nitriloacético; 0,556 g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$; 0,5 g MgSO_4 ; 0,5 g $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$; 0,24 g $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$; 0,1 g $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$; 0,1 g $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$; 0,1 g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$; 0,01 g $\text{CuCl}_2 \cdot 5\text{H}_2\text{O}$; 0,01 g $\text{AlK}(\text{SO}_4)_2$; 0,01 g H_3BO_3) e solução de vitaminas (concentração final 0,1 % - v/v, contendo por litro: 2 mg biotina; 2 mg ácido fólico; 10 mg cloridrato de piridoxina; 5 mg tiamina; 5 mg riboflavina; 5 mg ácido nicotínico; 5 mg ácido pantotênico; 0,1 mg vitamina B12; 5 mg ácido para-aminobenzóico; 5 mg ácido lipóico).

Por meio de sementeiras por esgotamento, colônias isoladas foram obtidas em MM sólido após 48 horas de incubação a 28°C, as quais foram então transferidas para MM líquido (com a mesma composição descrita acima, sem adição de ágar). Após a verificação de pureza por meio de observações microscópicas e análises de sequências do gene rRNA 16S amplificadas a partir da cultura (descritos nas seções 2.2.2 e 2.2.3 a seguir), alíquotas da cultura da linhagem bacteriana, denominada a partir de então de GeG2, foram estocadas em glicerol (concentração final de 20 % - v/v) a -80°C. O crescimento da bactéria em MM sem a adição das soluções de elementos traço e vitaminas também foi testado, e por não terem sido observadas diferenças no crescimento em relação ao obtido em MM com a adição destes compostos, os experimentos seguintes foram realizados em MM sem elementos traços e vitaminas.

O isolado bacteriano foi depositado em duas coleções de cultura: na CBMAI (Coleção Brasileira de Microrganismos de Ambiente e Indústria), da Universidade de

Campinas, e na CBAS (Coleção de Bactérias do Ambiente e Saúde), do Instituto Oswaldo Cruz - FIOCRUZ, sob os números de acesso CBMAI 2313 e CBAS 753, respectivamente.

2.2.2. Observação das células bacterianas por microscopia óptica

Esfregaços das colônias de GeG2 obtidas em MM sólido após 72 horas de incubação a 28°C foram preparados sobre lâminas histológicas de vidro, com o auxílio de alça de níquel-cromo e visualizados a fresco por microscopia de contraste de fase. Células bacterianas cultivadas em MM líquido por diferentes períodos de incubação (2, 3, 7 e 14 dias), sob agitação ou não, também foram observadas ao microscópio óptico. As células foram concentradas por meio de centrifugação (6000 g, por 5 minutos) de 3 mL da cultura em tubos tipo eppendorf. Os sobrenadantes foram descartados e os sedimentos de células ressuspensos em 80 µL de meio líquido. Em seguida, 20 µL foram transferidos para lâminas histológicas, recobertos com lamínula de vidro e visualizados a fresco por microscopia de contraste de fase.

Amostras bacterianas cultivadas em MM líquido por 72 horas também foram submetidas à coloração diferencial de Gram, após parte do concentrado de células, obtido da mesma forma descrita acima para as visualizações a fresco, ter sido fixado por calor em lâminas histológicas. Tanto os preparos a fresco quanto as células submetidas à coloração de Gram foram visualizadas e fotografadas sob aumentos de 1000x, no microscópio de luz Axio Scope.A1 (Zeiss, Alemanha) do Laboratório de Microbiologia da UnB.

2.2.3. Extração de DNA genômico, amplificação, clonagem, sequenciamento e análises do gene de rRNA 16S da cultura bacteriana

Para a extração de DNA total, 4,5 ml de culturas da bactéria em MM líquido incubadas por 72h a 28°C foram concentradas por centrifugação (6.000 g por 5 minutos a 4°C) e o sedimento celular resultante foi submetido ao protocolo de extração de DNA por fenol-clorofórmio, conforme previamente descrito na seção 1.3.3. Ensaios de PCR com os iniciadores 27f-1492r foram então realizados utilizando-se o DNA total obtido, nas mesmas condições descritas na seção 1.3.4 e na Tabela 2. A amplificação de fragmentos de DNA de tamanho esperado foi verificada em gel de agarose (1 %) corado

com brometo de etídio (10 mg/ml), com auxílio do marcador *Ikb Plus DNA Ladder* (Invitrogen).

Os fragmentos amplificados foram purificados com o kit *Wizard® SV Gel and PCR Clean-Up System* (Promega) e ligados ao vetor *pGEM-T easy®* (Promega), conforme as instruções dos fabricantes. Os sistemas de ligação resultantes foram utilizados para transformação por choque térmico em células de *E. coli* DH5 α e os plasmídeos dos clones recombinantes obtidos foram extraídos por lise alcalina, conforme os protocolos descritos nos itens 1.3.6 a 1.3.8 do capítulo anterior. O sequenciamento dos DNAs plasmidiais contendo os fragmentos do gene rRNA 16S foi realizado pelo método Sanger, pela empresa Macrogen Inc. (Coréia do Sul), utilizando-se os iniciadores universais T7 e SP6.

A qualidade das sequências de DNA obtidas foi avaliada de acordo com o algoritmo PHRED, por meio da ferramenta *Electropherogram quality analysis* da EMBRAPA (<http://asparagin.cenargen.embrapa.br/phph/>) e somente aquelas com valor PHRED superior a 31 foram utilizadas para as análises posteriores. As sequências completas dos *amplicons* dos genes de rRNA 16S foram obtidas por meio da montagem das sequências obtidas com os iniciadores T7 e SP6 de cada clone, utilizando-se o *software Geneious Prime* (v.2020.1.2) (<https://www.geneious.com>). A sequência consenso obtida foi depositada no *GenBank*, sob o número de acesso MT325926.1.

Análises de similaridade entre os genes de rRNA 16S de GeG2 e de outras linhagens foram realizadas na plataforma EZBioCloud (www.ezbiocloud.net) (Yoon *et al.*, 2017). Para a construção de árvores filogenéticas, sequências de genes de rRNA 16S de espécies da família *Sphingomonadaceae*, retiradas do banco de dados do NCBI, foram alinhadas com a sequência de GeG2 por meio do programa ClustalX (Larkin *et al.*, 2007) e, após edição manual com o BioEdit, as sequências foram utilizadas para a obtenção das árvores com o programa MEGA5 (Tamura *et al.*, 2011). Os métodos de *Neighbour-joining* (Saitou e Nei, 1987) e *Maximum-likelihood* (Felsenstein, 1981) foram empregados, utilizando-se os modelos de *Maximum Composite Likelihood* (Tamura *et al.*, 2004) e *Kimura 2-parameter* (Kimura, 1980), respectivamente. Em ambos os casos, valores de *bootstrap* foram calculados com base em 1000 repetições.

2.2.4. Comparações fisiológicas entre o isolado bacteriano GeG2 e *Novosphingobium rosa* DSM 7285

Como as análises do gene de rRNA 16S do isolado bacteriano GeG2 revelaram maior identidade com a espécie *Novosphingobium rosa*, uma linhagem desta bactéria – *N. rosa* DSM 7285 - foi adquirida da coleção de culturas microbianas alemã DSMZ (*Deutsche Sammlung von Mikroorganismen und Zellkulturen*) e utilizada para comparações fenotípicas e moleculares. A bactéria *N. rosa* DSM 7285 foi cultivada inicialmente em ágar/caldo nutriente (por litro: 5 g peptona; 3 g extrato de carne; 15 g ágar bacteriológico, pH 7,0), o meio de cultura indicado pela DSMZ para essa linhagem (<https://bacdiv.dsmz.de/strain/14145>).

O crescimento do isolado bacteriano GeG2, bem como de *N. rosa* DSM 7285, nos meios de cultura ágar nutriente, Mueller-Hilton (por litro: 2 g de extrato de carne; 17,5 g de casaminoácidos; 1,5 g de amido; 15 g ágar bacteriológico, pH 7,4) e TSA (*Tryptone soy agar*) por litro: 15 g tripton; 5 g digestão papaínica de Soja, 5 g NaCl; 15 g ágar bacteriológico, pH 7,3) foi avaliado. Tendo em vista a impossibilidade de manter os cultivos de GeG2 nestes meios, uma vez que as células não cresceram de forma adequada, e para que as comparações entre as linhagens bacterianas pudessem ser realizadas sob as mesmas condições de crescimento, *N. rosa* passou a ser cultivada em MM com glicose (sem a adição de elementos traços ou vitaminas), conforme descrito para o cultivo de GeG2 na seção Isolamento e cultivo da bactéria em meios de cultura definidos

Para a avaliação de motilidade, uma colônia de cada linhagem bacteriana foi semeada por picada, com auxílio de uma agulha de níquel-cromo, em tubos contendo MM sólido, que foram incubados a 28°C por 30 dias. Após esse período, o padrão de crescimento em torno da região da picada foi observado. Para a detecção da presença da enzima catalase, colônias de cada linhagem crescidas por 72 horas em MM foram raspadas com auxílio da alça de níquel-cromo e depositadas sobre lâminas histológicas de vidro. Em seguida, uma gota de H₂O₂ (3 %) foi depositada sobre a massa de células e a formação bolhas foi observada após 20 segundos.

A avaliação da hidrólise de amido foi realizada de acordo com o protocolo descrito em Tindall *et al.*, 2007. Em resumo, colônias de GeG2 e *N. rosa* foram semeadas no centro de placas de Petri contendo MM sólido com a adição de amido (0,2 %) e glicose (0,05 %) ou apenas amido. Inóculos em MM, acrescido apenas de glicose (0,05 %), ou MM sem a adição de qualquer fonte de carbono foram utilizadas como controles positivo

e negativo, respectivamente. As placas foram incubadas a 28°C por 7, 14 e 24 dias, quando foram coradas com vapor de iodo e a formação de halos ao redor dos crescimentos bacterianos, indicativos da hidrólise do amido, foi analisada. A hidrólise de gelatina também foi analisada utilizando-se o protocolo proposto em Tindall *et al.*, 2007. MM líquido acrescido de gelatina (12 %) foi distribuído em tubos com tampa de rosca. Após esterilização em autoclave, os tubos foram armazenados a 4°C e, após a solidificação do meio, colônias de GeG2 e *N. rosa* foram inoculadas, em triplicata, pelo método de picada. Os tubos foram incubados a 28°C e a ocorrência da liquefação dos meios, indicativo da hidrólise bacteriana de gelatina, foi verificada após 7, 14 e 24 dias. Tubos não inoculados, incubados a 28°C pelos mesmos períodos, foram utilizados como controle negativo.

O crescimento do isolado bacteriano GeG2, assim como o da linhagem *N. rosa* DSM 7285, também foi analisado em diferentes faixas de temperatura, pH e salinidade (% de NaCl). Para a avaliação do crescimento nas diferentes condições, alíquotas (100 µL) de diluições 1:100.000 (10^{-5}) de culturas de ambas as linhagens bacterianas crescidas em MM líquido por 48 horas (ambas com $OD_{600} \cong 0,03$) foram semeadas com o auxílio de alças de Drigalsky estéreis em placas de Petri contendo MM sólido. Para a avaliação do crescimento em diferentes temperaturas, as células bacterianas foram semeadas em MM nas condições rotineiramente usadas (pH 5,5 e 0,02 % NaCl, como descrito na seção Isolamento e cultivo da bactéria em meios de cultura definidos), sendo as placas incubadas a 4, 15, 20, 28, 33, 37 e 42°C. O crescimento das bactérias em seis diferentes faixas de pH (4,0 a 9,0, em intervalos de uma unidade) foi observado em MM tamponado com 50 mM de MES (ácido 2- (N -morfolino) etanosulfônico, para os valores de pH 4,0, 5,0 e 6,0), MOPS (Ácido 4-Morfolino Propano Sulfônico, para pH 7) ou Tris (tris(hidroximetil)aminometano, para pH 8,0 e 9,0). Já para a avaliação do crescimento em diferentes salinidades, as células foram semeadas em MM (pH 5,5) contendo diferentes concentrações de NaCl (0; 0,1; 0,3; 0,5; 0,8; 1; 1,5; 2,0 e 3,0 %). A semeadura de cada linhagem para a avaliação do crescimento sob as diferentes condições foi realizada em triplicata, e a formação de colônias observada após 48 horas e 7 dias de incubação. As placas com diferentes salinidades e faixas de pH foram incubadas a 28°C.

O crescimento em anaerobiose também foi avaliado para as duas linhagens. O experimento foi conduzido no laboratório de Enzimologia da UnB, com os equipamentos e materiais gentilmente cedidos pela Prof. Dra. Eliane Noronha. Para o preparo do meio de cultura a ser utilizado nas condições de anaerobiose, MM líquido (sem a adição de glicose) foi aquecido até a fervura e acrescido de solução redutora (para 1 L de meio:

10 mL de NaOH 0,2 M, 0,5 g de Na₂S₉H₂O e 0,125 g de cisteína) e ágar bacteriológico (concentração final 1,5 % - p/v). A glicose (concentração final 0,05 % - p/v), esterilizada por filtração, foi adicionada ao meio após a esterilização em autoclave. Os cultivos em condição de anaerobiose foram realizados em garrafas de vidro hermeticamente vedadas com rolhas de borracha e seladas com anéis de alumínio, sendo a atmosfera anaeróbica obtida pela injeção de nitrogênio gasoso (filtrado em membranas de 0,22 µm). Papéis de filtro embebidos em azul de metileno foram utilizados como indicadores de anaerobiose, já que nessa condição o reagente é reduzido e perde sua coloração azulada. Cultivos em garrafas nas quais o ar não foi substituído por gás nitrogênio foram utilizados como controle do crescimento em aerobiose. As garrafas foram incubadas por 35 dias a 28°C.

Perfis bioquímicos adicionais foram avaliados com os kits API 20NE e API ZYM (bioMérieux) pelos serviços de identificação bacteriana da DSMZ na Alemanha, conforme instruções do fabricante. Assim como para os demais testes fisiológicos, os testes bioquímicos com os kits API foram realizados com as bactérias GeG2 e *N. rosa* DSM 7285 previamente cultivadas em MM, nas mesmas condições.

2.2.5. Caracterizações químicas de GeG2

Análises da composição de lipídeos polares e de quinonas respiratórias da linhagem GeG2 foram realizadas pelos serviços de identificação bacteriana da DSMZ na Alemanha, conforme protocolos propostos em Tindall, 1990a; Tindall, 1990b; Tindall *et al.*, 2007, padronizados pela empresa (<https://www.dsmz.de/services/microorganisms/biochemical-analysis>). Em resumo, os lipídeos polares e as quinonas respiratórias foram extraídos a partir da mesma biomassa de células bacterianas liofilizadas, por meio de um protocolo de extração em duas etapas. Na primeira etapa, utilizou-se solução de hexano-metanol (1:2) e NaCl aquoso (0,3 %) para a recuperação das quinonas na fase do hexano. Em seguida, os lipídeos polares foram extraídos por meio da adição de clorofórmio à fase contendo metanol e NaCl resultante. As quinonas foram posteriormente purificadas por extrações em fase sólida à base de sílica e analisadas por cromatografia líquida de alta eficiência em fase reversa (RP-HPLC), registrando-se os espectros de absorção. Os espectros de 270 nm e 326 nm foram usados para a quantificação relativa de ubiquinonas e menaquinonas, respectivamente. Os resultados obtidos foram confirmados por espectrometria de massa, em sistema de UHPLC-ESI-qTOF (cromatografia líquida de alta eficiência acoplada a espectrometria de massa com ionização por eletrospray –

quadrupolo tempo de voo). Já os lipídeos polares obtidos foram separados por cromatografia bidimensional em camada delgada, sendo a primeira direção realizada em clorofórmio, metanol, água e a segunda em clorofórmio, metanol, ácido acético e água e posteriormente detectados utilizando-se ácido molibdatofosfórico e reagentes definidos para grupos funcionais específicos.

Os perfis de ácidos graxos das bactérias GeG2, bem como da linhagem referência de *N. rosa*, foram avaliados no Laboratório de Biotecnologia e Biodiversidade para o Meio Ambiente da Universidade Federal de Viçosa, pelo Prof. Dr. Marcos Rogério Tótola. As análises foram feitas de acordo com o protocolo padrão do sistema Sherlock® - MIDI Inc. (Sasser, 1990), como descrito a seguir. Após o cultivo de ambas as bactérias em MM sólido por 72 horas, uma massa de células de aproximadamente 40 mg foi retirada do terceiro quadrante das placas. As amostras foram transferidas para tubos de cultura de vidro com tampa de rosca e submetidas à saponificação dos ácidos graxos em 1 mL de solução metanol/NaOH (150 mL de água deionizada, 150 mL de metanol e 45 g NaOH) a 100°C por 30 minutos. Em seguida, 2 mL de solução de HCl/metanol (325 mL de HCl 6.0 mol/L e 275 mL de metanol) foram adicionados em cada tubo e a mistura foi então incubada a 80°C por 10 minutos. A separação dos metil-ésteres dos ácidos graxos na fase orgânica foi realizada por meio da adição de 1,25 mL de hexano:éter metil-terc-butílico (1:1). Após agitação por 10 minutos, a fase superior contendo os metil-ésteres dos ácidos graxos foi recuperada, lavada com NaOH (0,25 mol/L) e armazenada a 20°C até a realização das análises. A composição dos metil-ésteres dos ácidos graxos foi determinada por cromatografia gasosa utilizando-se um cromatógrafo Agilent 7890A. Os ácidos graxos foram identificados e quantificados com o Sistema de Identificação Microbiana Sherlock® (*version* 6.2, MIDI Inc.) por comparações com a biblioteca TSB Library 6.0.

2.2.6. Comparação do perfil proteico das bactérias GeG2 e *N. rosa* DSM 7285 por espectrometria de massa

Análises comparativas do perfil proteico das bactérias GeG2 e *N. rosa* DSM 7285 foram realizadas por espectrometria de massa (MS). Os experimentos foram conduzidos na Universidade Católica de Brasília (UCB), com auxílio do Prof. Dr. Marcelo Ramada. As bactérias foram repicadas e crescidas, em triplicata, em MM sólido por 48 horas a 28°C. As proteínas bacterianas foram então extraídas pelo método à base de ácido fórmico

e acetonitrila, conforme protocolo descrito em Agustini *et al.*, 2014. Em resumo: com o auxílio de ponteiras de plástico estéreis, uma porção das colônias bacterianas foi transferida para microtubos de polipropileno contendo 300 μL de água ultrapura, aos quais 900 μL de etanol absoluto foram adicionados e, após homogeneização em vórtex, os tubos foram centrifugados a 12.000 g por 2 minutos à temperatura ambiente. O sobrenadante foi descartado e, após secagem à temperatura ambiente, o sedimento foi recoberto com solução aquosa de ácido fórmico (70 %) e acetonitrila, na proporção 1:1 (v/v). A mistura foi novamente homogeneizada em vórtex e submetida à centrifugação nas mesmas condições descritas acima. Alíquotas de 1 μL dos sobrenadantes foram aplicadas, em sextuplicata para cada réplica bacteriana, nos poços da placa de aço inoxidável do MALDI (MSP 96 target). Após a secagem do material à temperatura ambiente por aproximadamente 10 minutos, 1 μL de matriz de ácido α -ciano-4-hidroxicinâmico (10 $\text{mg}\cdot\text{mL}^{-1}$, preparada em solução de 50% (v/v) de acetonitrila e 0,3% (v/v) de ácido trifluoracético) foi adicionada sobre cada poço. Porções das colônias de GeG2 e *N. rosa* também foram transferidas com o auxílio de ponteiras plásticas estéreis diretamente para poços da placa de MALDI e recobertas com 1 μL de matriz de ácido α -ciano-4-hidroxicinâmico, para a aquisição de perfis proteicos diretamente das colônias.

Os espectros foram obtidos no espectrômetro de massa do tipo MALDI-TOF/TOF (*Autoflex Speed, Bruker Daltonics, Alemanha*), no modo positivo linear, na faixa de massa/carga de 2.000-20.000 m/z , com a aquisição de 2.000 disparos por ponto de análise. Os espectros de massa obtidos para cada linhagem foram processados com o programa *MALDI Biotyper 3.0 (Bruker Daltonics)* e comparados entre si e com o banco de dados disponibilizado pela empresa. O programa realiza a correspondência de padrões dos picos obtidos e atribui pontuações (*scores*) de classificação, que variam de 0 a 3 e refletem a similaridade entre as amostras analisadas e os espectros de referência. Os critérios interpretativos propostos pelo fabricante foram empregados, sendo os *scores* maiores ou iguais a 2,0 considerados como identificações confiáveis em nível de espécie, aqueles maiores ou iguais a 1,7 e menores que 2,0 considerados como confiáveis para a identificação em nível de gênero e os menores que 1,7 não foram considerados confiáveis para a identificação (Schulthess *et al.*, 2013). Perfis de espectros principais (*Main Spectra Profiles - MSP*) foram criados, utilizando-se os perfis obtidos a partir das sextuplicatas técnicas de cada amostra bacteriana, e dendogramas foram construídos utilizando-se essas informações.

2.2.7. Avaliação do perfil de susceptibilidade de GeG2 a antibióticos

O perfil de susceptibilidade da bactéria GeG2 a diferentes antibióticos foi avaliado pelo método de disco-difusão (Bauer *et al.*, 1966, com modificações). Em síntese, suspensões bacterianas com turbidez equivalente ao padrão 1,0 da escala de McFarland foram preparadas em MM líquido, a partir de colônias crescidas por 48 horas. As amostras bacterianas foram então semeadas por espalhamento sobre a superfície de MM sólido (pH 5,5 ou 7,0) em placas de Petri de 150 mm. Discos impregnados com diferentes antibióticos (BioRad) foram depositados com o auxílio de pinças estéreis na superfície dos meios já inoculados, com uma distância de aproximadamente 2,5 cm entre um e outro e em relação à margem da placa. Os antibióticos testados foram: Ácido nalidíxico (30 µg), Amicacina (30 µg), Amoxicilina + ácido clavulânico (20 + 10 µg), Ampicilina (10 µg), Cefalotina (30 µg), Cefepima (30 µg), Cefoxitina (30 µg), Ceftazidima (30 µg), Cefuroxima (30 µg), Ciprofloxacina (5 µg), Clindamicina (2 µg), Cloranfenicol (30 µg), Eritromicina (15 µg), Gentamicina (10 µg), Levofloxacino (5 µg), Meropenem (10 µg), Nitrofurantoína (300 µg), Norfloxacino (10 µg), Oxacilina (1 µg), Penicilina (6 µg), Rifampicina (5 µg), Tetraciclina (30 µg), Trimetoprima + sulfametoxazol (1,25 + 23,75 µg), vancomicina (30 µg). Os discos foram gentilmente cedidos pela Prof. Dra. Marlene Teixeira, do Laboratório de Microbiologia da UnB. As placas foram incubadas a 28°C por 48 horas, quando a susceptibilidade/resistência aos antibióticos foi então verificada por meio da mensuração do diâmetro dos halos de inibição ao redor dos discos. Os limites de diâmetro dos halos utilizados para a classificação em susceptível ou resistente foram os propostos pelo fabricante dos discos (BioRad) e os descritos em Sha *et al.*, 2017.

2.2.8. Observação das células bacterianas por microscopia eletrônica de varredura (MEV) e de transmissão (MET)

Células das linhagens GeG2 e *N. rosa* DSM 7285 também foram analisadas por microscopia eletrônica. Para a observação por MEV, alíquotas das culturas bacterianas em MM líquido, incubadas por 72 horas ou 14 dias sem agitação, foram centrifugadas a 8.000 g por 10 minutos a 4°C. Alternativamente, células de cada uma das linhagens foram crescidas sobre lamínulas de vidro em recipientes contendo MM líquido por 7 dias. Os sedimentos celulares ou as amostras crescidas em lamínulas de vidro foram submetidas à fixação com fixador Karnovsky 0,1 M (paraformaldeído 2 %, glutaraldeído 2 %,

sacarose 3 %, cloreto de cálcio 0,005 M em meio tamponado por cacodilato de sódio 0,1 M) *overnight* a 4°C. O fixador foi então retirado e substituído por tampão cacodilato de sódio 0,1 M, no qual as amostras fixadas ficaram armazenadas, a 4°C, até a realização das próximas etapas do preparo, que foram executadas no Laboratório de Microscopia da UnB, conforme o protocolo descrito na seção 1.3.12 do capítulo anterior.

Já o preparo das amostras e as análises por MET foram realizadas no Centro de Microscopia da Universidade Federal de Minas Gerais (CM-UFGM), sob supervisão do Prof. Dr. Gregory Kitten. As imagens foram obtidas a partir de amostras da linhagem bacteriana GeG2 fixadas de duas formas: a) em fixador Karnovsky (alíquotas da amostra crescida por 14 dias também usadas para as análises por MEV descrita acima) e b) células crescidas em MM líquido ou sólido por 7 dias a 28°C, posteriormente armazenadas em geladeira por 20 dias, criofixadas sob alta pressão (*high-pressure cryofixation* – HPF) e submetidas à criosubstituição em acetona.

A partir das eletromicrografias obtidas por MEV e MET, determinações das dimensões das células e estruturas observadas foram realizadas com o auxílio da ferramenta ImageJ (Rasband, 1997).

2.2.9. Sequenciamento, montagem e anotação do genoma completo do isolado bacteriano GeG2

Conforme descrito no Capítulo 1 (itens 1.3.13 e 1.4.5), o sequenciamento genômico total do cultivo estabelecido a partir do solo de Cerrado, realizado por duas estratégias diferentes (Illumina HiSeq e PacBio), resultou em sequências identificadas como pertencentes a um único genoma bacteriano. Visando uma melhor análise do genoma desta bactéria, que posteriormente foi isolada em meio de cultura definido e denominada linhagem GeG2, os *long* e *short reads* obtidos foram submetidos a um novo processo de montagem. Todas as análises envolvendo os dados genômicos foram realizadas em colaboração com o grupo de pesquisa do Prof. Dr. Georgios J. Pappas Jr, da Universidade de Brasília.

Primeiramente, os *short reads* foram pré-processados com os programas FastQC v0.11.5 (Andrews, 2010) e TrimGalore v0.6.0 (<https://github.com/FelixKrueger/TrimGalore>) e os *reads* de PacBio foram extraídos com o pacote Pbh5Tools v0.8.0 (github.com/PacificBiosciences/pbh5tools), executados por meio do *pipeline ngs-preprocess* v1.0 (<https://zenodo.org/record/3451406#.XsLTmWhKiUk>). Após o pré-processamento, ambos

os tipos de *reads* foram utilizados para a montagem híbrida utilizando-se o programa *Unicycler* v0.4.4 (Wick *et al.*, 2017) e as estatísticas da montagem foram obtidas com o auxílio do *QUAST* v5.0.1 (Gurevich *et al.*, 2013), etapas executadas com o *pipeline* *MpGAP* v0.1 (<https://zenodo.org/record/3445486#.XsLS72hKiUk>). Mapas dos réplicons genômicos foram desenhados com a ferramenta *DNAPlotter* (Carver *et al.*, 2009). A completude do genoma montado foi então avaliada com os programas *BUSCO* v3.1.0 (Simão *et al.*, 2015) e *CheckM* v1.0.13 (Parks *et al.*, 2014), utilizando-se o conjunto de dados *sphingomonadales_odb10* e a linhagem marcadora *o_Sphingomonadales* (UID3310), respectivamente. O genoma completo da bactéria GeG2 foi depositado no *Genbank*, sob o número de acesso JABVZR000000000 (*BioProject*: PRJNA624997).

A anotação geral do genoma foi realizada com o programa *Prokka* (Seemann, 2014). As sequências de rRNA foram preditas com o *Barnap* (<https://github.com/tseemann/barnap>), enquanto a anotação de ilhas genômicas e profagos foi obtida com as ferramentas *IslandPath-DIMOB* (Bertelli *et al.*, 2018) e *Phigaro* (Starikova *et al.*, 2019), respectivamente. A avaliação de sequências plasmidiais foi realizada utilizando-se o programa *Plasflow* (Krawczyk *et al.*, 2018). Para a determinação da distância entre as abundâncias relativas de dinucleotídeos do cromossomo e do megarréplicon identificado no genoma da linhagem GeG2 (*contig 2*), inicialmente a frequência de dinucleotídeos foi calculada por meio da divisão da quantidade de vezes em que cada um dos 16 possíveis pares de dinucleotídeos ocorre em cada réplicon (calculada através da ferramenta *Genematix software suite* v3.12, disponibilizada *online* (<https://www.genomatix.de/solutions/genomatix-software-suite.html>)) pelo número total de pares de dinucleotídeos da sequência de DNA. Em seguida, a frequência de cada mononucleotídeo foi determinada, também pela ferramenta *Genematix software suite* v3.12, e frequência esperada para cada par de dinucleotídeo possível foi então calculada multiplicando-se as frequências observadas para cada combinação de nucleotídeos. A frequência obtida para cada dinucleotídeo foi dividida pela frequência esperada, resultando no perfil de abundância relativa de dinucleotídeos para cada réplicon. A diferença absoluta entre cada um dos 16 valores obtidos para o megarréplicon e para o cromossomo foi calculada e a soma dessas diferenças foi determinada, com o valor final obtido representando a distância de abundância relativa de dinucleotídeos entre o cromossomo e o megarréplicon.

As categorias funcionais COG (*Clusters of Orthologous Groups*) foram obtidas com o *eggNOG-mapper* v2 (Huerta-Cepas *et al.*, 2017). As enzimas ativas sobre carboidratos (CAZymes) foram anotadas por meio do servidor *dbCAN2* (Zhang *et al.*,

2018), que implementa três diferentes estratégias de busca em bancos de dados, tendo sido consideradas apenas as enzimas preditas por pelo menos duas ferramentas utilizadas pelo servidor. A identificação de vias metabólicas codificadas no genoma bacteriano, bem como das enzimas oxigenases, potencialmente envolvidas na degradação de compostos aromáticos, foi realizada por inspeção manual dos ortólogos funcionais identificados pelo banco de dados KEGG (*Kyoto Encyclopedia of Genes and Genomes*) (Kanehisa *et al.*, 2008). Para a detecção de agrupamentos (*clusters*) gênicos potencialmente envolvidos na síntese de metabólitos secundários utilizou-se o servidor *online* antiSMASH v5.1.2 (*antibiotics & Secondary Metabolite Analysis Shell*) (Blin *et al.*, 2019; Medema *et al.*, 2011).

2.2.10. Análises taxonômicas baseadas no genoma

Análises para a identificação e classificação taxonômica com base em critérios genômicos foram realizadas com o genoma completo da bactéria GeG2. Inicialmente, a sequência do gene de rRNA 16S predita a partir da montagem genômica foi comparada às sequências obtidas por ampliações de PCR, clonagem e sequenciamento Sanger (descrito no item Observação das células bacterianas por microscopia eletrônica de varredura (MEV) e de transmissão (MET)

. Para isso, as sequências foram alinhadas e manualmente inspecionadas quanto à correspondência de bases, com o auxílio do programa *Geneious Prime* v.2020.1.2 (<https://www.geneious.com>). Após a confirmação que se tratava da mesma sequência, a similaridade com sequências do gene de rRNA 16S depositadas nos bancos de dados do NCBI e EZBioCloud (Yoon *et al.*, 2017) foi avaliada. Análises de ANI (*Average Nucleotide Identity*) entre o genoma da bactéria GeG2 e 7.184 genomas de *Alphaproteobacteria* disponíveis no NCBI foram realizadas com o fastANI (Jain *et al.*, 2018), utilizando-se parâmetros padrão do programa. O genoma bacteriano também foi avaliado nos servidores TYGS (*Type Strain Genome Server*) (Meier-Kolthoff e Göker, 2019) e TrueBac ID (Ha *et al.*, 2019), que avaliam a classificação taxonômica com base no cálculo de índices OGRIs (*Overall Genome Relatedness Indexes*) e inferências filogenômicas realizados a partir da comparação com milhares de genomas bacterianos depositados nos bancos de dados.

2.3.11. Análises preliminares do crescimento de GeG2 em MM contendo o corante aromático verde malaquita

Considerando o repertório gênico relacionado à degradação de compostos aromáticos identificado na linhagem GeG2, análises preliminares da capacidade da bactéria crescer em um corante aromático, com estrutura semelhante a alguns fragmentos de lignina foram realizadas. Colônias crescidas em MM foram transferidas para placas contendo MM sólido sem glicose ou extrato de levedura, acrescido de 50 mg/L de verde malaquita. As placas foram incubadas em estufa a 28°C por 30 dias e o crescimento de colônias foi avaliado semanalmente.

2.3. Resultados e Discussão

2.3.1. Isolamento e cultivo da bactéria presente no cultivo do solo de Cerrado em meio definido

Visando o isolamento da bactéria identificada nos cultivos de solo de Cerrado, colônias cultivadas no meio confeccionado com extratos de solo foram transferidas inicialmente para meio LB sólido. Apesar de ter sido observado o crescimento de colônias amareladas, com aspecto bastante mucilaginoso na primeira transferência para meio LB (Figura 14), a manutenção do cultivo neste meio de cultura não foi possível, não ocorrendo crescimento em repiques posteriores, mesmo após várias tentativas.

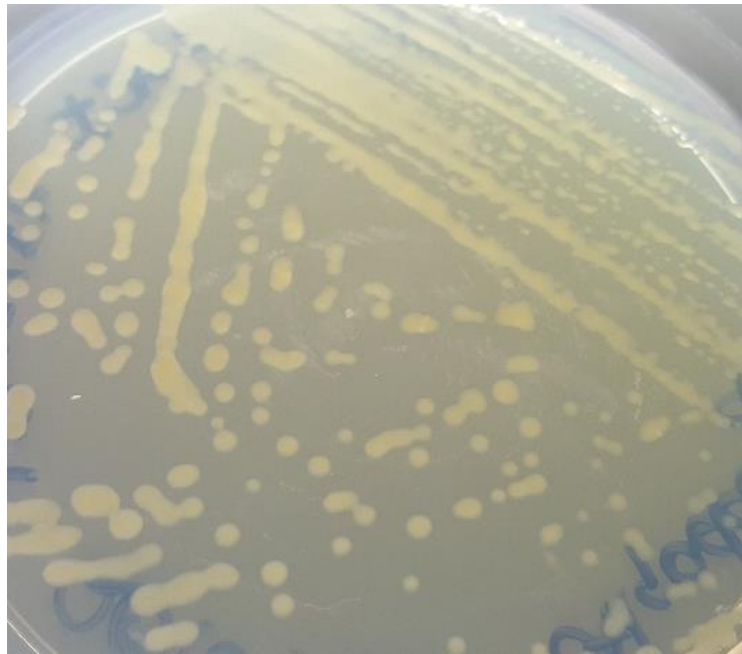


Figura 14. Colônias bacterianas obtidas após a primeira transferência de colônias do cultivo de solo de Cerrado em meio de extrato de solo para meio LB. Repiques posteriores para o mesmo meio de cultura não resultaram no desenvolvimento de colônias, não sendo possível a manutenção do cultivo nesse meio de cultura.

Conforme descrito no capítulo anterior (seção 1.4.3), a bactéria identificada nos cultivos de solo de Cerrado pertence ao gênero *Novosphingobium*, da família *Sphingomodaceae*. Apesar da capacidade da maioria dos membros dessa família crescer em diversos meios de cultura complexos, a necessidade da utilização de meios mais

oligotróficos ou versões diluídas dos meios padrão para o crescimento de algumas linhagens é usualmente relatada (Glaeser e Kämpfer, 2014). Assim, especulou-se que condições mais oligotróficas seriam necessárias para a manutenção do crescimento da bactéria. Assim, colônias do cultivo do solo de Cerrado (crescidas em meio de cultura confeccionado com extratos de solo, como descrito no item 1.3.2 do capítulo anterior) foram transferidas para meio mínimo, inicialmente suplementado com soluções de vitaminas e elementos traço, contendo glicose como fonte de carbono (item 2.2.1).

Após 48 horas de incubação, colônias de 2 a 3 mm de diâmetro, esbranquiçadas, opacas, mucilaginosas e com aspecto homogêneo foram observadas no MM (Figura 15). Colônias isoladas foram então transferidas para MM líquido e para novas placas de MM, que passou a ser utilizado como meio de cultura para a manutenção das células, uma vez que estas apresentaram crescimento exuberante neste meio. Além disso, o crescimento do isolado bacteriano, a partir de então denominado GeG2, foi avaliado em MM sem a adição de elementos traço e vitaminas e como não foi identificada diferenças no crescimento entre as duas condições, o MM sem tal suplementação passou a ser empregado para o cultivo da bactéria e para a realização dos experimentos subsequentes.

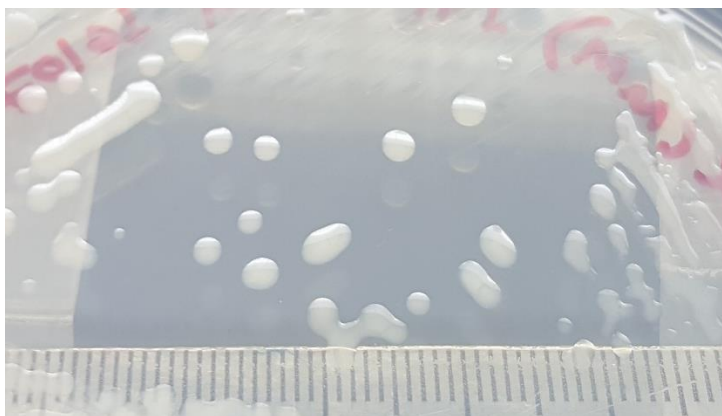


Figura 15. Colônias obtidas após a transferência da cultura em meio de solo de Cerrado para meio mínimo (MM).

2.3.2. Avaliação das células bacterianas GeG2 por microscopia óptica

Amostras de colônias crescidas em MM sólido, bem como de culturas em meio líquido, foram submetidas a análises de microscopia óptica. Observações à fresco por microscopia de contraste de fase revelaram bacilos Gram negativos móveis, com 1 a 3 μm de comprimento (Figura 16A e B). A morfologia celular detectada condiz com aquela

descrita para bactérias do gênero *Novosphingobium* (Glaeser e Kämpfer, 2014). Ademais, o aspecto homogêneo das células, observado em todas as análises de microscopia óptica, sugeriu a obtenção de culturas puras, o que foi corroborado pelas análises de identificação molecular, descritas a seguir.

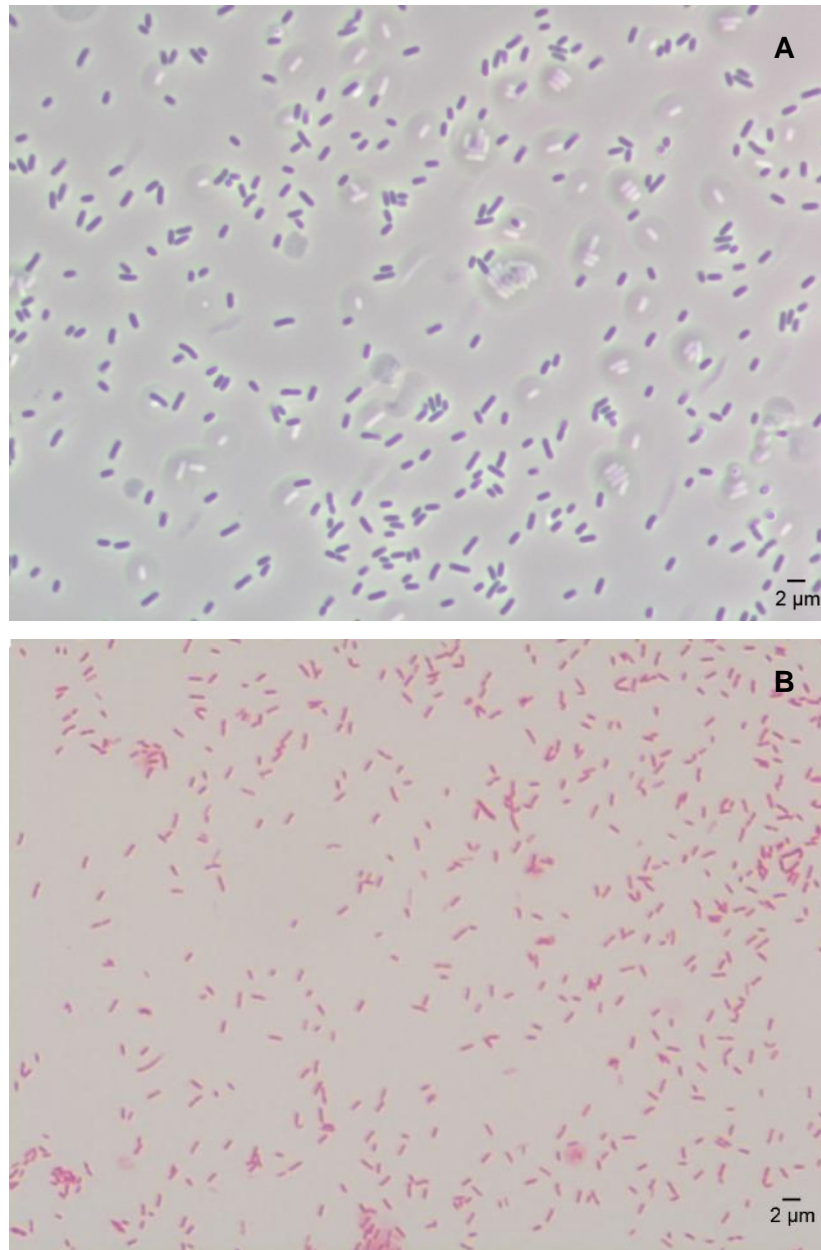


Figura 16. Análise a fresco por microscopia de contraste de fase (A) e de coloração de Gram por microscopia de campo claro (B) das células de GeG2, cultivadas em MM por 96 horas.

2.3.3. Análises do gene rRNA 16S

Um total de 22 sequências de DNA de qualidade resultaram da amplificação, clonagem e sequenciamento de fragmentos do gene de rRNA 16S, obtidos a partir do cultivo bacteriano estabelecido. A análise dessas sequências revelou que todas eram idênticas e, aliando-se ao fato de a cultura ter sido originada de uma colônia isolada, obtida a partir de sementes por esgotamento, tais resultados sugerem o estabelecimento de uma cultura bacteriana pura. A sequência completa do *amplicon* (1450 pb) do gene de rRNA 16S do isolado GeG2 foi comparada com o banco de dados EZBioCloud e, conforme previamente identificado nos cultivos de solo de Cerrado (discutido no item 1.4.3), foi classificada no gênero *Novosphingobium*, apresentando maior identidade de nucleotídeos com *Novosphingobium rosa* NRBC15208 (100%), seguido de *N. lotistagni* THG-DN6.20 (97,58 %), *N. oryzae* ZYY112 (97,02 %) e *N. barchaimii* LL02 (96,95 %).

O posicionamento do isolado bacteriano GeG2 no gênero *Novosphingobium* e o altíssimo grau de identidade de seu gene de rRNA 16S com o da espécie *N. rosa* NBRC15208 também foram evidenciados nas reconstruções filogenéticas (Figura 17). As análises realizadas apenas com este marcador filogenético sugeriam que a linhagem bacteriana isolada a partir dos cultivos de solo de Cerrado pertenceria à espécie *Novosphingobium rosa*. No entanto, ainda que comparações de sequências do gene de rRNA 16S sejam muito úteis como abordagem primária na identificação taxonômica de microrganismos, limitações do poder de resolução deste gene marcador na análise de organismos evolutivamente próximos, principalmente em nível de espécie, vêm sendo apontadas na literatura há décadas (Fox *et al.*, 1992; Stackebrandt e Goebel, 1994; Jaspers e Overmann, 2004). Apesar da observação de que espécies distintas usualmente apresentam $\geq 1,3$ % de diferenças na sequência nucleotídica de seus genes de rRNA 16S (Kim *et al.*, 2014), bactérias que possuem genes de rRNA 16S idênticos, mas que foram identificadas como pertencentes a espécies diferentes com base em hibridizações de DNA genômico e critérios fisiológicos já foram descritas (Martinez-Murcia *et al.*, 1992; Jaspers e Overmann, 2004).

Por esse motivo, abordagens polifásicas, envolvendo diferentes caracterizações moleculares e fenotípicas e, mais recentemente, critérios baseados em análises de sequências genômicas completas, são necessárias para a delimitação e descrição de novas espécies microbianas (Chun *et al.*, 2018; Raina *et al.*, 2019). Assim, conforme será discutido nas seções a seguir, tais metodologias foram empregadas visando a classificação

taxonômica mais acurada do isolado GeG2, bem como sua descrição fenotípica e genômica.

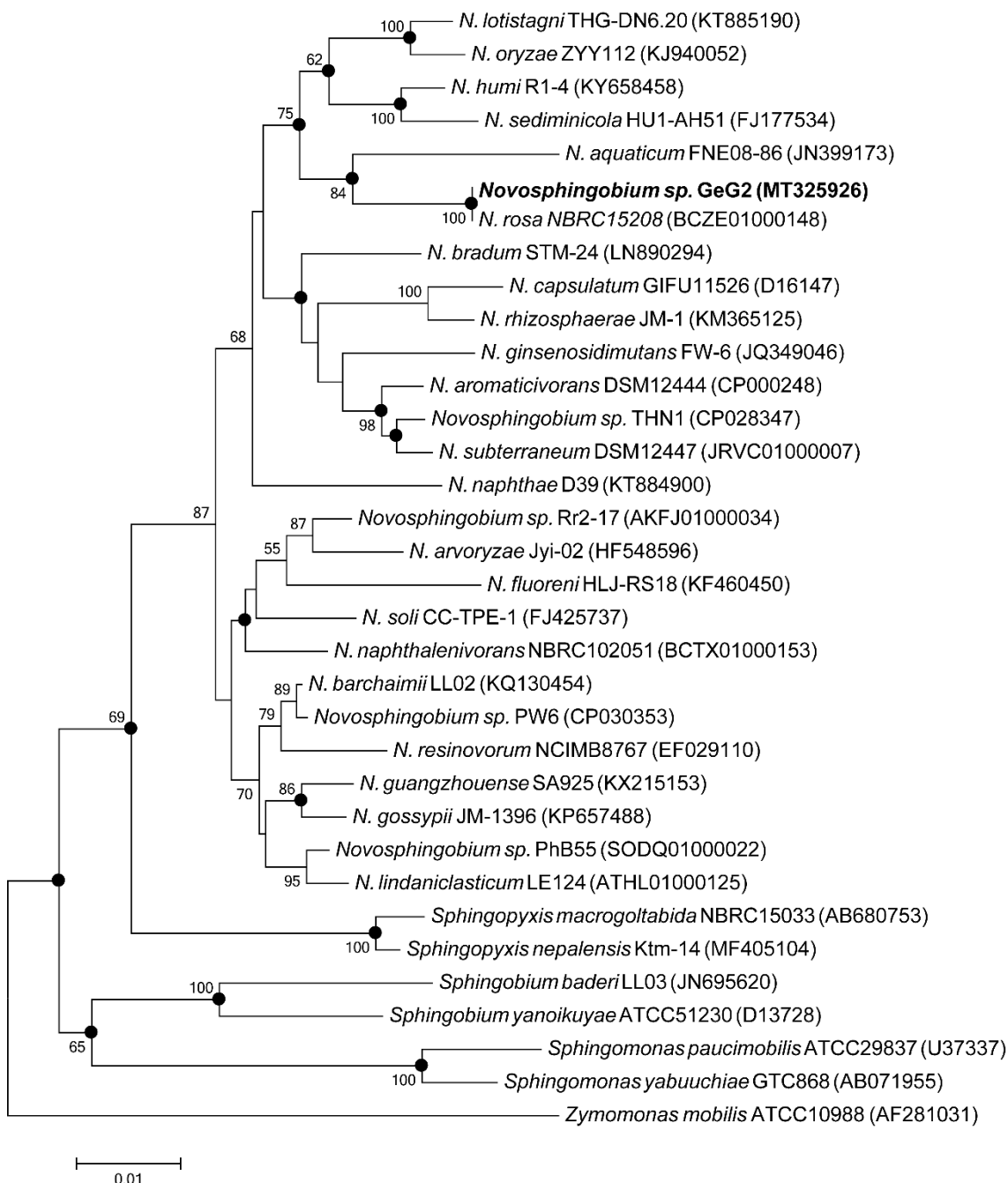


Figura 17. Árvore filogenética baseada em sequências do gene de rRNA 16S, mostrando a relação entre a linhagem GeG2 (em negrito) e outros taxa relacionados da família *Sphingomonadaceae*. A árvore foi construída utilizando-se o programa MEGA5 (Tamura *et al.*, 2011), pelo método de Neighbour-joining (Saitou e Nei, 1987), com o modelo *Maximum Composite Likelihood* (Tamura *et al.*, 2004). Os valores (%) obtidos no teste de bootstrap (1000 réplicas) (Felsenstein, 1985) são indicados próximo aos ramos, sendo que os inferiores a 50 % não são mostrados na árvore. Os círculos pretos indicam que nós correspondentes também foram obtidos na árvore construída pelo método de *Maximum-likelihood* (Felsenstein, 1981). A barra de escala indica o número de substituições por posição de nucleotídeo.

2.3.4. Caracterizações fisiológicas de GeG2 e comparações com *N. rosa* DSM 7285

Considerando a identidade de 100% entre as sequências dos genes de rRNA 16S do isolado GeG2 e de *Novosphingobium rosa*, uma linhagem tipo desta espécie (DSM 7285) foi utilizada para comparações fenotípicas. Inicialmente, o crescimento de ambas as linhagens em diferentes meios de cultura foi avaliado. Apesar de ágar/caldo nutriente ser o meio de cultura padrão indicado para o cultivo de *N. rosa*, não foi observado crescimento de GeG2 após a segunda transferência neste meio de cultura, assim como constatado no meio LB (seção 2.3.1). Além disso, o isolado GeG2 não apresentou crescimento nos meios TSA e Mueller-Hinton, nem mesmo na primeira transferência, enquanto a linhagem de *N. rosa* mostrou-se capaz de crescer em ambos os meios (Tabela 5). Tais resultados evidenciaram o caráter mais oligotrófico da linhagem GeG2, em comparação à *N. rosa*. Conforme previamente mencionado, diversos membros da família *Sphingomonadaceae*, incluindo algumas espécies de *Novosphingobium* isoladas de solo (Gupta *et al.*, 2009), exibem melhor crescimento em meios de cultura pobres em nutrientes, ou em versões diluídas dos meios de cultura complexos padrão, evidenciando adaptações a condições mais oligotróficas (Glaeser e Kämpfer, 2014). O solo de Cerrado, ambiente do qual a linhagem GeG2 foi isolada, é conhecido por sua pequena disponibilidade de nutrientes (Haridassam, 2008), o que poderia justificar as adaptações desta bactéria a condições oligotróficas.

Além de apresentar crescimento exuberante nos meios de cultura ricos, a linhagem de *N. rosa* também foi capaz de crescer no meio mínimo (MM), utilizado para o isolamento e cultivo de GeG2. Assim, este meio de cultura passou a ser utilizado para a manutenção dos cultivos de ambas as bactérias, bem como para a realização dos experimentos seguintes. Vale ressaltar que, após 48 horas de incubação a 28°C, foi possível observar diferenças morfológicas entre as colônias de GeG2 e *N. rosa*, como mostrado nas Figura 18A e B. Enquanto as colônias de GeG2 (Figura 18B) apresentavam 2 a 3 mm de diâmetro, coloração esbranquiçada e aspecto mucilaginoso, as colônias de *N. rosa* (Figura 18A) eram menores, com aproximadamente 1 mm de diâmetro, apresentando coloração amarelada e aspecto mais seco. Curiosamente, após 7 dias de incubação, apesar das colônias de *N. rosa* não apresentarem variações morfológicas significativas (Figura 18C), as colônias de GeG2 alteraram sua morfologia, passando a apresentar uma região central densa, rodeada por um anel (Figura 18D).

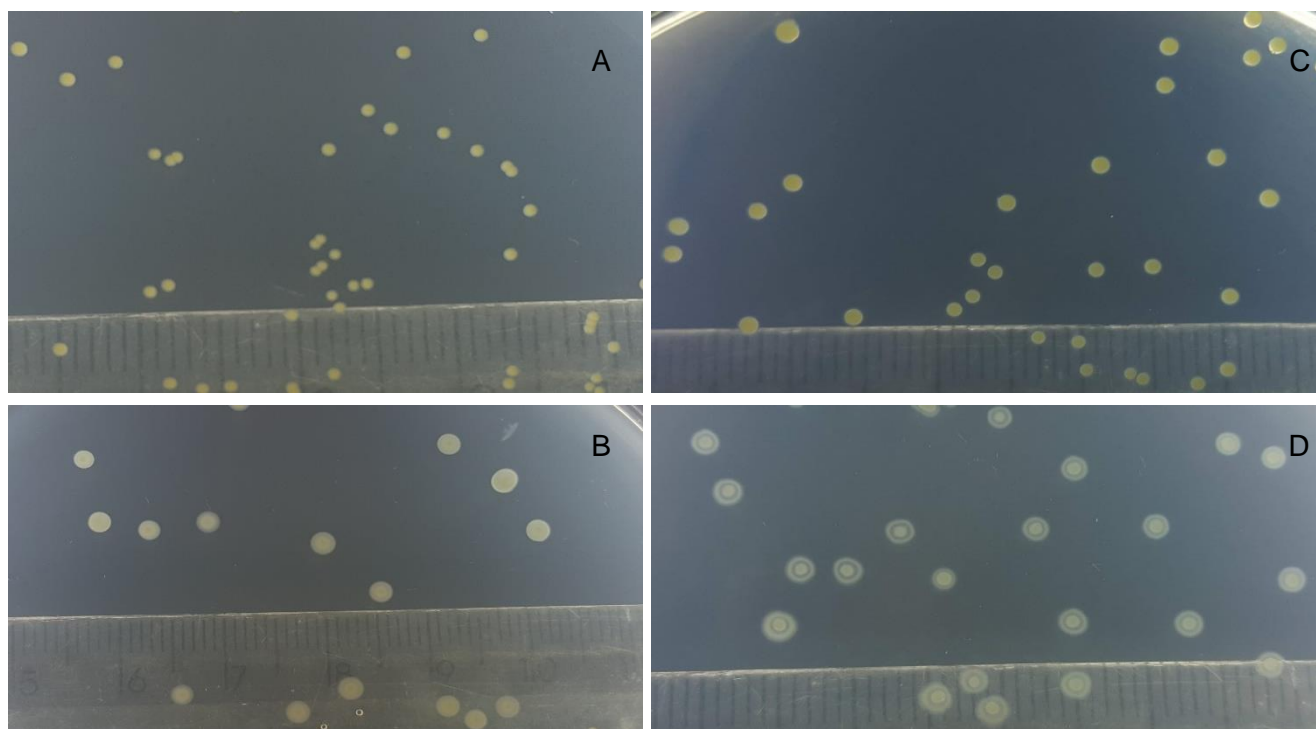


Figura 18. Diferenças morfológicas observadas nas colônias de *N. rosa* DSM 7285 (A e C) e GeG2 (B e D) crescidas em MM. Em A e B são apresentadas colônias formadas após 48 horas e em C e D após 7 dias de incubação a 28 °C.

Assim como observado para outras espécies de *Novosphingobium* (Glaeser e Kämpfer, 2014), tanto *N. rosa* DSM 7285 quanto GeG2 apresentam motilidade na fase exponencial de crescimento, evidenciada tanto nos padrões de crescimento obtidos em tubos contendo meio sólido, inoculados por picada (Figura 19A), quanto em observações a fresco por microscopia de contraste de fase. Além disso, ambas as linhagens foram positivas para a atividade de catalase e negativas para a hidrólise de gelatina (Tabela 5).

Já os resultados dos testes de hidrólise de amido revelaram que, após 14 dias de incubação, foi possível detectar um halo não corado pelo iodo ao redor do crescimento de GeG2 em placas de MM contendo amido e glicose, ou apenas amido (Figura 19D e E), sugerindo a secreção de α -amilases pela bactéria em ambas as condições. Já nas placas de *N. rosa* DSM 7285 não houve a formação de halo, mesmo após 24 dias de incubação (Figura 19B e C), indicando não ocorrer a hidrólise de amido por esta bactéria.

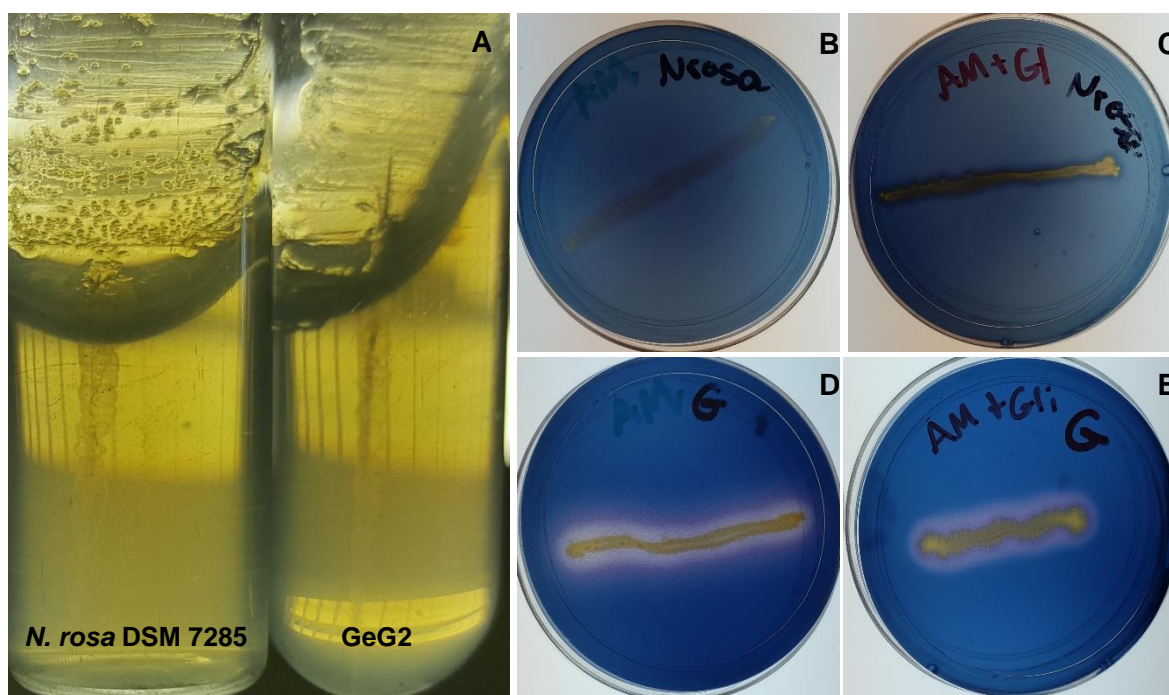


Figura 19. Padrão de crescimento indicativo de motilidade observado para as bactérias *N. rosa* DSM 7285 e GeG2 em meios sólidos (MM) inoculados pelo método de picada (A) e resultados obtidos nos testes de hidrólise de amido para as bactérias *N. rosa* DSM 7285 (B e C) e GeG2 (D e E) em placas de MM contendo apenas amido como fonte de carbono (B e D) ou amido e glicose (C e E). As placas foram submetidas à coloração com vapor de iodo após 24 dias de incubação. O halo observado ao redor dos crescimentos da bactéria GeG2 indica a hidrólise de amido em ambas as condições, o que não foi observado para *N. rosa* DSM 7285.

A avaliação do crescimento de GeG2 e *N. rosa* DSM 7285 em diferentes temperaturas revelou que as duas linhagens apresentam caráter mesófilo, crescendo entre 15 e 33°C (Tabela 5). Vale ressaltar que o crescimento de ambas as linhagens a 15°C ocorreu de maneira mais lenta, sendo possível a visualização de colônias somente após 7 dias de incubação, ao passo que nas demais temperaturas houve a formação de colônias após 48 horas. Não foi observado crescimento de qualquer uma das linhagens após sete dias de incubação a 4, 37 ou 42°C. No entanto, curiosamente, quando as placas mantidas a 4°C por sete dias foram incubadas a 28°C, houve o crescimento de colônias de *N. rosa* DSM 7285 após 72 horas, enquanto tal crescimento não foi observado para GeG2, sugerindo que a linhagem referência apresenta maior tolerância a baixas temperaturas que o novo isolado.

A mesma faixa de pH (4,0 a 7,0) também foi identificada para o crescimento de ambas as linhagens (Tabela 5), não ocorrendo crescimento em pH 8,0 ou 9,0. No entanto, foram observadas diferenças quanto à salinidade na qual ocorre o crescimento das bactérias. Enquanto *N. rosa* DSM 7285 foi capaz de crescer em meios com 0 a 1,5 % de NaCl, GeG2 cresceu em meios contendo até 1 % (Tabela 5).

Tabela 5. Características fisiológicas das linhagens bacterianas GeG2 e *N. rosa* DSM 7285

Característica	GeG2	<i>N. rosa</i> DSM 7285
Crescimento em:		
Meio Ágar Triptona de Soja (TSA)	-	+
Meio Mueller-Hinton	-	+
Meio Ágar nutriente	+/-	+
Meio Luria-Bertani (LB)	+/-	+
Meio Mínimo (MM)	+	+
Anaerobiose	Tardio (após 30 dias)	-
Coloração da colônia (em MM)	Branca	Amarela
Condições de crescimento:		
Temperatura (°C)	15 a 33	15 a 33
pH	4 a 7	4 a 7
NaCl (%)	0 a 1	0 a 1,5
Hidrólise de:		
Amido	+	-
Gelatina	-	-
Esculina	+	+
Redução de nitrato	f	-
Formação de indol	-	-
Fermentação de glicose	-	-
Assimilação de:		
Glicose	+	+
Arabinose	+	+
Manose	+	+
Manitol	-	-
N-Acetil-glicosamina	+	f
Maltose	+	+
Gluconato	+	f
Ácido caprato	-	-
Ácido adipato	-	-
Ácido malato	-	-
Citrato de trisódio	-	-
Ácido fenil-acetato	-	-
Atividade enzimática de:		
Catalase	+	+
Oxidase	+	+
Arginina dihidrolase	-	-
Urease	-	-
Fosfatase alcalina	+	+
Esterase	f	+
Esterase lipase	f	f
Lipase	-	-
Leucina-arilamidase	+	+
Valina-arilamidase	+	+
Cistina-arilamidase	f	f
Tripsina	-	-
Quimotripsina	f	-
Fosfatase ácida	+	+
Naftol-AS-BI-fosfohidrolase	+	+
α -Galactosidase	-	-
β -Galactosidase	+	+
β -Glucuronidase	f	-
α -Glicosidase	+	+
β -Glicosidase	+	+
N-Acetil- β -Glicosaminidase	-	f
α -Manosidase	-	-
α -Fucosidase	-	-

+/- indica pouco crescimento ou ausência de crescimento após o primeiro repique

f: atividade fraca (*weak positive/not clear positive*)

Apesar de membros de *Novosphingobium* terem sido descritos como aeróbios estritos quando da proposta deste gênero (Takeuchi *et al.*, 2001), posteriormente foi observado que algumas espécies são capazes de crescer facultativamente em anaerobiose (Sohn *et al.*, 2004; Sheu *et al.*, 2020). Assim como observado para *N. pentaromativorans* (Sohn *et al.*, 2004), a linhagem GeG2 apresentou crescimento em anaerobiose, ainda que bastante lento, tendo sido detectado após mais de 30 dias. Por outro lado, não foi possível detectar crescimento de *N. rosa* DSM 7285 nestas condições (Tabela 5).

Caracterizações bioquímicas adicionais, realizadas por meio do kit API 20NE, revelaram que o isolado GeG2 é positivo para a hidrólise de esculina, atividade de oxidase, assimilação de glicose, arabinose, manose, N-Acetil-glicosamina, maltose e gluconato, além de apresentar resultado positivo fraco para redução de nitrato. Quanto às atividades enzimáticas determinadas pelo painel API Zym, foram detectados resultados positivos para fosfatase alcalina, leucina-arilamidase, valina-arilamidase, fosfatase ácida, naftol-AS-BI-fosfohidrolase, β -Galactosidase, α -Glicosidase e β -Glicosidase, enquanto reações fracas de esterase, esterase lipase, cistina-arilamidase, quimotripsina e β -Glucuronidase foram observadas. De forma geral, perfis bioquímicos semelhantes foram observados para *N. rosa* DSM 7285, porém, diferentemente de GeG2, esta linhagem apresentou resultados negativos para a redução de nitrato e atividades de quimotripsina e β -Glucuronidase, bem como reação positiva para N-Acetil- β -Glicosaminidase, mesmo que de fraca intensidade (Tabela 5).

Assim, apesar da identidade nucleotídica de 100 % do gene rRNA 16S da linhagem GeG2 com o da espécie *N. rosa*, diferenças fisiológicas entre o isolado de solo de Cerrado e a linhagem tipo *N. rosa* DSM 7285 foram observadas, sugerindo que GeG2 poderia pertencer a uma nova espécie. Análises químicas e genômicas, bem como avaliações moleculares adicionais, foram então conduzidas, visando uma caracterização mais aprofundada e a determinação taxonômica do novo isolado bacteriano.

2.3.5. Caracterizações químicas de GeG2

Caracterizações quimiotaxonômicas da linhagem GeG2, incluindo a identificação de quinonas respiratórias, da composição de lipídeos polares e do perfil de ácidos graxos, foram realizadas. Assim como descrito para os demais membros da família *Sphingomonadaceae* (Takeuchi *et al.*, 2001; Glaeser e Kämpfer, 2014), a quinona respiratória predominante encontrada no isolado GeG2 foi a ubiquinona Q10 (mais de

95%), com Q9 tendo sido detectada em menores quantidades. Quanto aos lipídeos polares, foram identificados na linhagem GeG2: fosfatidilglicerol (PG), fosfatidiletanolamina (PE), esfingoglicolípido (SGL), difosfatidilglicerol (DPG), fosfatidilcolina (PC), além de um aminolípido e três lipídeos não identificados (Figura 20).

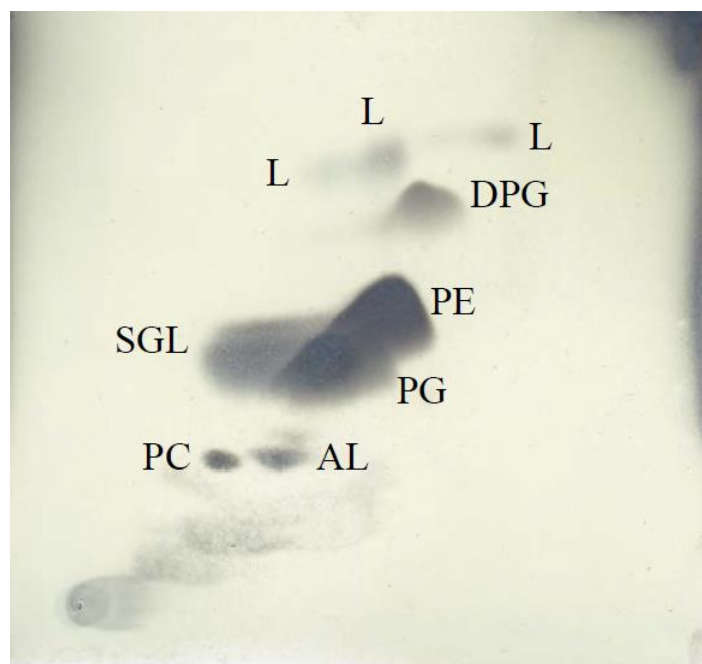


Figura 20. Perfil de lipídeos polares totais da linhagem GeG2 obtido por cromatografia bidimensional em camada delgada e marcação com ácido fosfomolibdico. PG: Fosfatidilglicerol; PE: Fosfatidiletanolamina; SGL: esfingoglicolípido; DPG: Difosfatidilglicerol; PC: Fosfatidilcolina; AL: Aminolípido não identificado; L: lipídeos não identificados (não corados com qualquer dos reagentes específicos).

De forma geral, o perfil de lipídeos polares da linhagem GeG2 está de acordo com o observado para outros membros de *Sphingomonadaceae*, em que PE, PG, DPG e SGL também são primariamente detectados (Kämpfer *et al.*, 1997 Busse *et al.*, 1999). No entanto, embora fosfatidil-monometil-etanolamina (PME) e fosfatidil-dimetil-etanolamina (PDE) sejam encontrados na maioria das espécies de *Novosphingobium* (Glaeser *et al.*, 2013), incluindo *N. rosa* (Kämpfer *et al.*, 1997), esses lipídeos polares não foram detectados na linhagem GeG2.

O perfil de ácidos graxos das linhagens GeG2 e *N. rosa* DSM 7285 foi analisado a partir de células cultivadas sob as mesmas condições (em MM sólido, por 72 horas a 28°C). Embora a composição identificada seja semelhante, a proporção dos ácidos graxos

detectada em cada linhagem variou significativamente (Tabela 6). Os principais ácidos graxos identificados em GeG2 foram C_{16:0} (24,62%), C_{18:0} (21,84 %), *summed feature* 8 (C_{18:1 ω 7c}/ C_{18:1 ω 6c}) (18,18 %), *summed feature* 3 (C_{16:1 ω 7c}/C_{16:1 ω 6c}) (11,97 %) e C_{14:0} 2-OH (7,22 %). Já em *N. rosa* DSM 7285, *summed feature* 8 foi detectado em maior quantidade (38,71 %), seguido de C_{14:0} 2-OH (19,65 %), C_{16:0} (14,84 %), *summed feature* 3 (13,88 %) e C_{18:0} (7,82 %). Além disso, enquanto o ácido graxo insaturado C_{18:1 ω 9c} foi detectado em proporções consideráveis na linhagem GeG2 (8,29 %), apenas traços desse ácido graxo foram observados em *N. rosa* (<1 %) (Tabela 6).

Tabela 6. Comparação do conteúdo de ácidos graxos (%) das linhagens GeG2 e *N. rosa* DSM 7285, ambas crescidas em MM por 72 horas. TR, traço (<1 %); -, não detectado.

Ácido graxo	GeG2	<i>N. rosa</i> DSM 7285
C _{12:0}	TR	-
C _{13:0} anteiso	TR	TR
C _{14:0}	1,85	1,30
C _{14:0} 2-OH	7,22	19,65
C _{16:0}	24,62	14,84
C _{17:0} anteiso	TR	-
C _{17:0} cyclo	TR	-
C _{17:0}	TR	-
C _{18:1ω9c}	8,29	TR
C _{18:0}	21,84	7,82
C _{19:0} iso	TR	-
C _{19:0} cyclo ω 8c	1,92	2,18
C _{20:0}	TR	-
<i>Summed feature</i> *		
3	11,97	13,88
8	18,18	38,71

**Summed feature* representa grupos de dois ou três ácidos graxos que não podem ser separados por cromatografia com o Sistema MIDI. *Summed feature* 3 inclui C_{16:1 ω 7c}/C_{16:1 ω 6c}; *summed feature* 8 inclui C_{18:1 ω 7c}/C_{18:1 ω 6c}.

Com base em uma ampla análise quimiotaxonômica, Busse e colaboradores (1999) concluíram que a presença de esfingolípídeos e de apenas ácidos graxos 2-hidroxi (ausência de ácidos graxos 3-hidroxi) é um importante indicador para a classificação na família *Sphingomonadaceae*. Além disso, os autores apontaram que a análise conjunta de lípídeos polares e ácidos graxos possibilitam identificações a nível de espécie. A maioria das espécies desse grupo bacteriano pode ser distinguida por variações qualitativas e/ou quantitativas destes compostos celulares, enquanto perfis idênticos ou muito semelhantes indicariam o posicionamento de uma nova linhagem em uma espécie já estabelecida (Busse *et al.*, 1999). Assim, as diferenças identificadas tanto na composição de lípídeos

polares quanto no perfil de ácidos graxos das linhagens GeG2 e *N. rosa* DSM 7285 sugerem que estas bactérias não pertençam à mesma espécie.

2.3.6. Caracterização do perfil proteico de GeG2 por espectrometria de massa e comparação com o perfil de *N. rosa* DSM 7285

A análise de perfis proteicos por espectrometria de massa MALDI-TOF (*Matrix-assisted laser desorption/ionization – Time of flight*) tem se mostrado uma metodologia valiosa para a identificação e classificação de novos isolados bacterianos. Tal técnica se baseia na análise de massas de biomoléculas bacterianas abundantes, especialmente proteínas ribossomais (predominantes no espectro de massa de 2 a 2000 kDa), que geram espectros característicos para cada espécie, podendo ser usados para fins de classificação taxonômica (Mellmann *et al.*, 2008). A simplicidade no preparo das amostras (perfis de espectros podem ser gerados a partir de células intactas obtidas diretamente de colônias bacterianas, ou de extrações proteicas rápidas) e a grande reprodutibilidade oferecida por essa abordagem fazem com que ela venha sendo crescentemente empregada em estudos de taxonomia microbiana comparativa (Raina *et al.*, 2019). Além disso, o desenvolvimento de programas que realizam análises e comparações dos espectros obtidos com aqueles disponíveis em bancos de dados cada vez mais robustos possibilitam resultados acurados e de grande relevância para a identificação taxonômica de novos isolados clínicos e ambientais (Hotta *et al.*, 2012; Faron *et al.*, 2015; Raina *et al.*, 2019).

Perfis de espectros de massa na faixa de 2 a 2000 kDa obtidos a partir de extrações proteicas ou diretamente de células de GeG2 e *N. rosa* DSM 7285 foram avaliados e comparados entre si e com aqueles disponíveis no banco de dados MALDI Biotyper®. Enquanto os perfis de espectros principais (MSPs) gerados a partir das extrações proteicas de *N. rosa* DSM 7285 foram identificados como espectros dessa espécie bacteriana no banco de dados, os MSPs obtidos para o isolado bacteriano GeG2 não apresentaram *scores* maiores que 2,0 com qualquer espécie, indicando tratar-se de uma espécie ainda não representada no banco de dados. Além disso, comparações dos MSPs obtidos a partir das diferentes extrações proteicas das culturas de GeG2 e de *N. rosa* nas mesmas condições também não apresentaram *scores* maiores que 2,0 entre si, sugerindo que esses microrganismos pertencem a espécies diferentes. Resultados semelhantes foram obtidos para os MSPs gerados a partir dos espectros adquiridos diretamente das colônias, com a identificação do perfil proteico de *N. rosa* nas comparações com o banco de dados,

enquanto os MSPs gerados a partir das colônias de GeG2 não apresentaram identificação (*scores* maiores que 2,0) com os espectros de qualquer espécie representada no banco de dados, sugerindo tratar-se de uma nova espécie.

Um dendrograma criado a partir da comparação dos MSPs obtidos para cada linhagem bacteriana é mostrado na Figura 21, onde é possível observar que as réplicas de GeG2 e as de *N. rosa* se agrupam em clados separados, mesmo quando grupos externos são incluídos na análise. Tais resultados corroboram os resultados que sugerem GeG2 e *N. rosa* serem espécies diferentes.

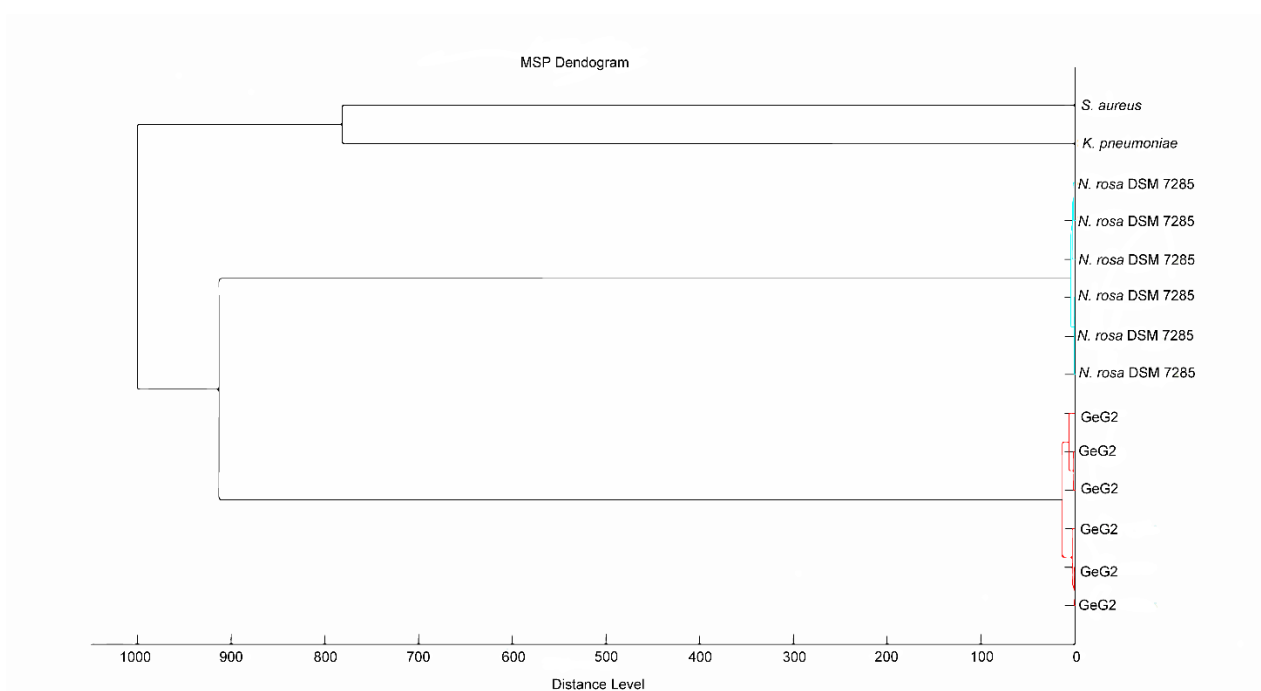


Figura 21. Dendrograma construído a partir da comparação dos perfis de espectros principais (MSP) obtidos das extrações proteicas das linhagens GeG2 e *N. rosa* DSM 7285. Cada MSP é constituído por seis espectros de massa independentes, obtidos na faixa de 2.000 a 20.000 m/z, sendo o dendrograma construído a partir da comparação dos MSPs de seis réplicas biológicas de cada bactéria. Os espectros de massa obtidos pelo método direto (diretamente das colônias bacterianas) não foram incluídos no dendrograma. Espectros das bactérias *Staphylococcus aureus* (Gram positiva) e *Klebsiella pneumoniae* (Gram negativa) retirados do banco de dados da MALDI Biotyper® (Bruker) foram utilizados como grupo externo.

2.3.7. Perfil de susceptibilidade a antibióticos de GeG2

O perfil de susceptibilidade a antibióticos do isolado GeG2 foi avaliado pela observação e mensuração de halos de inibição de crescimento ao redor de discos impregnados com diferentes antibióticos (teste de disco-difusão). Dos 23 antibióticos testados, a linhagem GeG2 apresentou susceptibilidade apenas à tetraciclina e à

rifampicina (Tabela 7). Resultado semelhante foi obtido com *N. rosa* DSM 7285, que também apresentou sensibilidade à tetraciclina e à rifampicina, porém essa bactéria também mostrou-se sensível a trimetoprima+sulfametoxazol (Tabela 7). O mesmo resultado foi obtido para a linhagem GeG2 quando o teste foi realizado em MM com pH 7,0, indicando que a ação dos antibióticos não foi influenciada pelo pH usual do meio (5,5), utilizado inicialmente nos testes.

Tabela 7. Diâmetros dos halos de inibição do crescimento das linhagens GeG2 e *N. rosa* DSM 7285 observados em placas contendo MM e discos impregnados com diferentes antibióticos. Diâmetros acima do limiar indicam que a linhagem foi sensível ao antibiótico, enquanto os valores abaixo indicam resistência (Manual dos discos para a testagem de susceptibilidade a antibióticos, BioRad; Sha *et al.*, 2017). -, ausência de halo.

Antibióticos	Diâmetros halos (mm)		
	Limiar	GeG2	<i>N.rosa</i> DSM 7285
Tetraciclina (TET)	19	30	35
Rifampicina (RIF)	18	22	25
Trimetoprima+ sulfametoxazol (SXT)	13	-	22
Vancomicina (VAN)	17	9	9

O grande número de antibióticos aos quais as linhagens GeG2 e *N. rosa* DSM 7285 apresentaram resistência foi notório. Apesar de informações relacionadas à resistência a antibióticos de *Sphingomonadaceae* ainda serem escassas, uma análise da resistência de 28 linhagens tipo dessa família a 36 diferentes antibióticos revelou que espécies oriundas de solo e rizosfera são especialmente resistentes a vários tipos de antimicrobianos, com linhagens de *N. rosa* e *N. pruni*, por exemplo, tendo apresentado resistência a mais de 50% dos antibióticos testados (Yabuuchi *et al.*, 2002). Além disso, níveis altos de resistência a beta-lactâmicos e polimixinas foram observados em membros de *Sphingomonadaceae* isolados de diferentes fontes de água potável (Vaz-Moreira *et al.* 2011). Isolados de solo classificados na ordem *Sphingomonadales* foram identificados como relevantes reservatórios naturais de resistomas ainda inexplorados (Dantas *et al.*, 2008).

Vale ressaltar ainda que, como será discutido mais à frente, espécies de *Novosphingobium* e gêneros relacionados produzem grandes quantidades de exopolissacarídeos (EPS), que formam uma espessa matriz que envolve as células e podem conferir resistência inerente a diversos agentes antimicrobianos por dificultar a penetração desses compostos (Donlan e Costerton, 2002).

2.3.8. Observação das células de GeG2 e *N. rosa* DSM 7285 por MEV e MET

Inicialmente, células do novo isolado bacteriano GeG2 e da linhagem referência *N. rosa* DSM 7285 foram crescidas sobre lamínulas de vidro, imersas em MM líquido, por 7 dias, quando foram então submetidas à fixação por aldeídos e às demais etapas de processamento para análise por microscopia eletrônica de varredura (MEV), conforme descrito na seção 2.2.8. Tanto as células de GeG2 quanto as de *N. rosa* DSM 7285 apresentaram formato bacilar, com 1,3 a 2,3 μm de comprimento e 0,3 a 0,6 μm de largura (Figura 22), assim como relatado para outros membros do gênero (Glaeser e Kämpfer, 2014). Curiosamente, enquanto as células de *N. rosa* crescidas nesta condição apresentavam-se formando pequenos agregados, com algumas células unidas por prolongamentos intercelulares, além de possuírem filamentos que parecem se aderir ao vidro (Figura 22D a F), as células de GeG2 apresentavam-se majoritariamente individualizadas e sem quaisquer projeções filamentosas (Figura 22A e B). No entanto, alguns agregados maiores, contendo muitas células envoltas por matriz extracelular e com projeções celulares foram identificados nas amostras de GeG2 crescidas sobre lamínulas de vidro (Figura 22C). Além disso, bacilos filamentosos, com mais de 8 μm de comprimento, foram observados apenas nas amostras de GeG2 (Figura 22C).

Vale ressaltar ainda que algumas células de GeG2 apresentavam superfícies rugosas, recobertas por estruturas arredondadas (indicadas por setas na Figura 22B), semelhantes a vesículas extracelulares já descritas em *Novosphingobium pentaromativorans* US6-1 (Yun *et al.*, 2016), *Novosphingobium* sp. PP1Y (De Lise *et al.*, 2019) e espécies do gênero relacionado *Sphingomonas* (Gilewicz *et al.*, 1997; Coppotelli *et al.*, 2010).

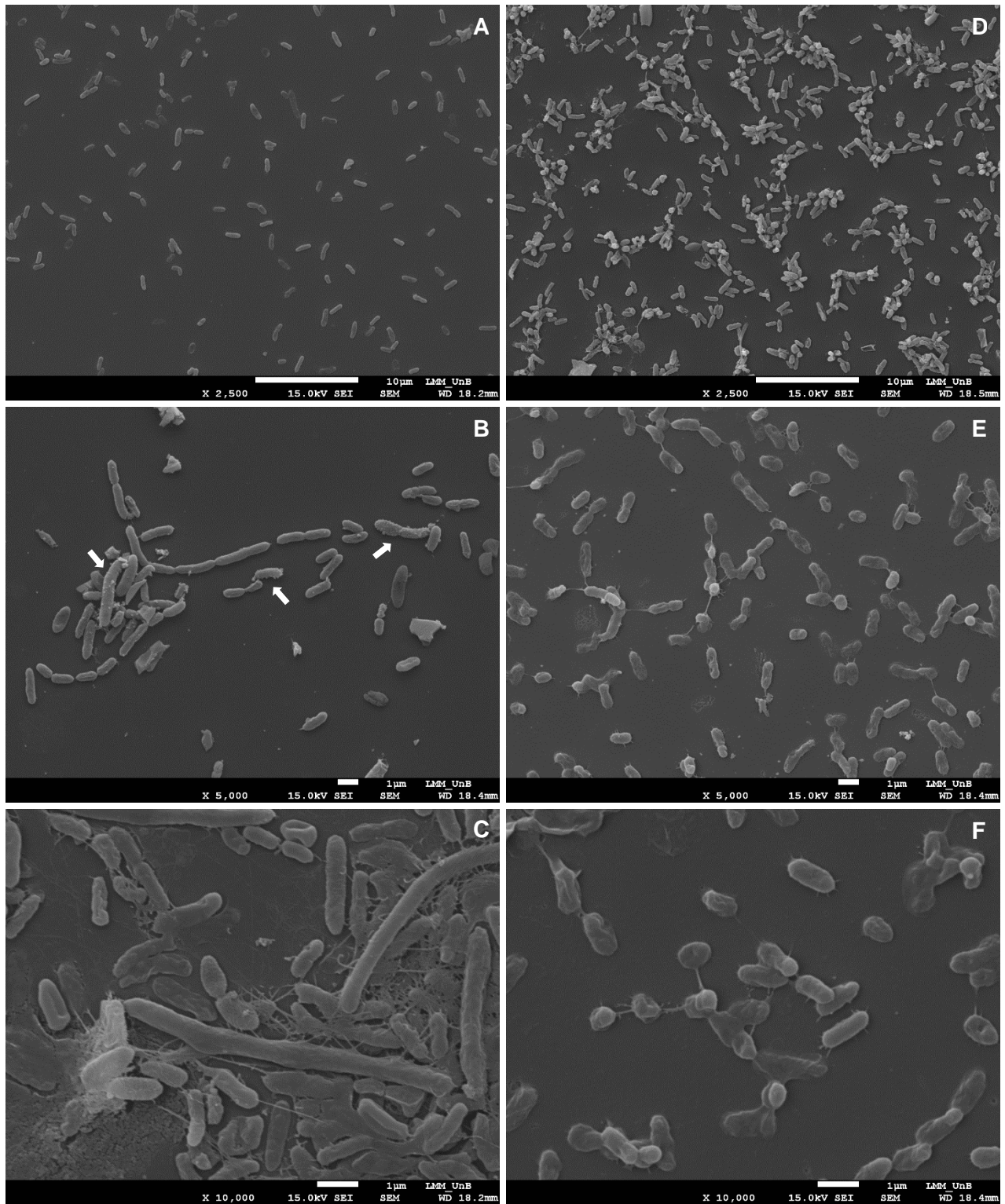


Figura 22. Eletromicrografias de varredura das células de GeG2 (A - C) e *N. rosa* DSM 7285 (D - F) crescidas sobre lamínulas de vidro em MM líquido por 7 dias a 28°C. As setas brancas em C indicam células da linhagem GeG2 que apresentavam superfície rugosa, recobertas por vesículas extracelulares. Os aumentos utilizados estão indicados na parte inferior de cada imagem.

A produção de vesículas extracelulares já foi descrita para diversas espécies de bactérias, sendo as vesículas de membrana externa (OMVs) de Gram-negativas as mais

bem caracterizadas, principalmente em bactérias patogênicas (Toyofuku *et al.*, 2019). No entanto, bactérias ambientais produtoras de OMVs vêm sendo cada vez mais relatadas e, com base no conteúdo variado dessas estruturas em diferentes condições, o seu envolvimento em diversas funções fisiológicas já foi sugerido, tais como aquisição de nutrientes, respostas a estresses, biodegradação de compostos aromáticos, ações surfactantes e formação de biofilmes (Coppotelli *et al.*, 2010; Choi *et al.*, 2014; Schwechheimer and Kuehn, 2015; Yun *et al.*, 2016; De Lise *et al.*, 2019). Conforme será mostrado na próxima seção, um grande número de células apresentando superfície rugosa, característica da produção de OMVs, foi identificado em culturas de GeG2 em meio líquido com maior tempo de crescimento (14 dias). Essas culturas apresentavam agregados macroscópicos e grandes quantidades de matriz extracelular, o que poderia indicar o envolvimento das vesículas produzidas por essa bactéria na formação de biofilmes e outras interações celulares.

Culturas de GeG2 e *N. rosa* DSM 7285 crescidas em MM líquido por 72 horas também foram analisadas por MEV. Nessas condições, a maioria das células de ambas as linhagens apresentava-se sob a forma de agregados, com poucas células individuais tendo sido detectadas (Figura 23). Além disso, uma espessa matriz polimérica pôde ser observada ao redor de muitas células de GeG2 (Figura 23B), o que não foi detectado na amostra de *N. rosa* nessas condições. Como já mencionado, membros da família *Sphingomonadaceae* são conhecidos por produzirem grandes quantidades de exopolissacarídeos (EPS), genericamente denominados de “esfinganos” (*sphingans*), biopolímeros com diversas aplicações nas indústrias farmacêutica, alimentícia e de biorremediação (Wu *et al.*, 2017). Além disso, a influência de fatores relacionados às condições de cultivo na produção e composição dos EPS por essas bactérias, tais como tipo de fontes nutricionais, níveis de oxigenação e tempo de crescimento, já foi relatada (Pollock e Armentrout, 1999; Matsuyama *et al.*, 2003). Embora as condições empregadas no cultivo a partir do qual as observações por MEV foram realizadas – meio de cultura oligotrófico com baixas quantidades de carbono (MM), incubação sem agitação, resultando em pouca oxigenação da cultura líquida, e pouco tempo de crescimento – não sejam as mais favoráveis para a produção de EPS, essas análises iniciais, aliadas ao aspecto mucilaginoso observado nas colônias, indicam que o isolado GeG2 é capaz de produzir e secretar tais polissacarídeos.

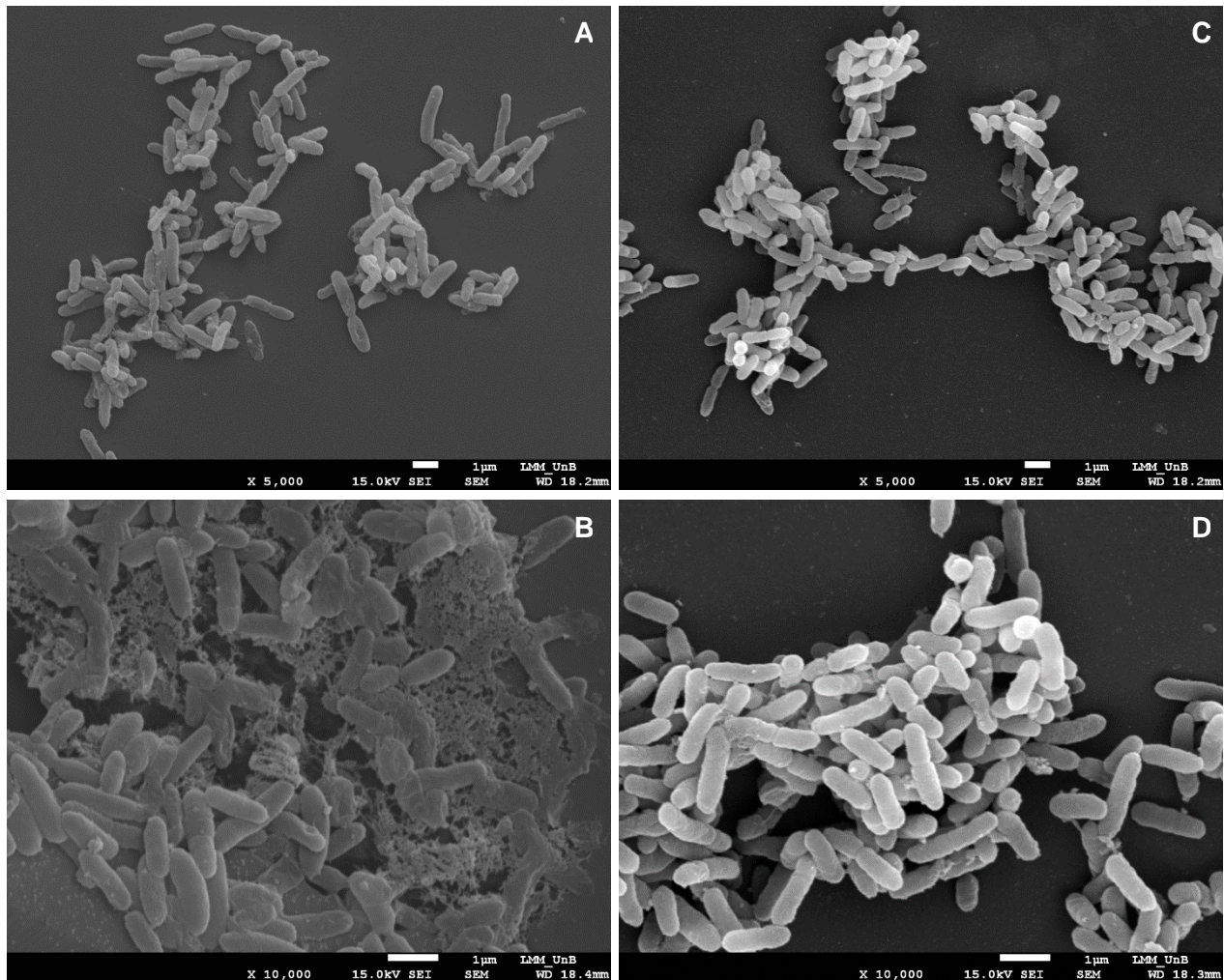


Figura 23. Eletromicrografias de varredura das células de GeG2 (A e B) e *N. rosa* DSM 7285 (C e D) crescidas em MM líquido por 72 horas, sem agitação, a 28°C. Os aumentos utilizados estão indicados na parte inferior de cada imagem.

As células de GeG2 também foram avaliadas por MET (Figura 24). Além da típica estrutura de parede celular de bactérias gram-negativas (Figura 24B), grânulos intracitoplasmáticos eletrodensos, de 70 a 150 nm de diâmetro, foram observados na maioria das células, normalmente um por célula, localizados tanto na porção central quanto na polar (Figura 24A). Em algumas células, porém, dois grânulos puderam ser observados (Figura 24C). A presença de grânulos eletrodensos também foi identificada no citoplasma de células de *Novosphingobium aquaticum* FNE08-86 e *Novosphingobium fuchskuhlense* FNE08-7, ambas isoladas de águas subsuperficiais de um lago rico em substâncias húmicas, na Alemanha (Glaeser *et al.*, 2013a; Glaeser *et al.*, 2013b). No entanto, diferentemente do observado para a linhagem GeG2, os grânulos identificados nessas espécies eram significativamente menores, não ultrapassando 60 nm, além de muito mais numerosos (Glaeser *et al.*, 2013a; Glaeser *et al.*, 2013b).

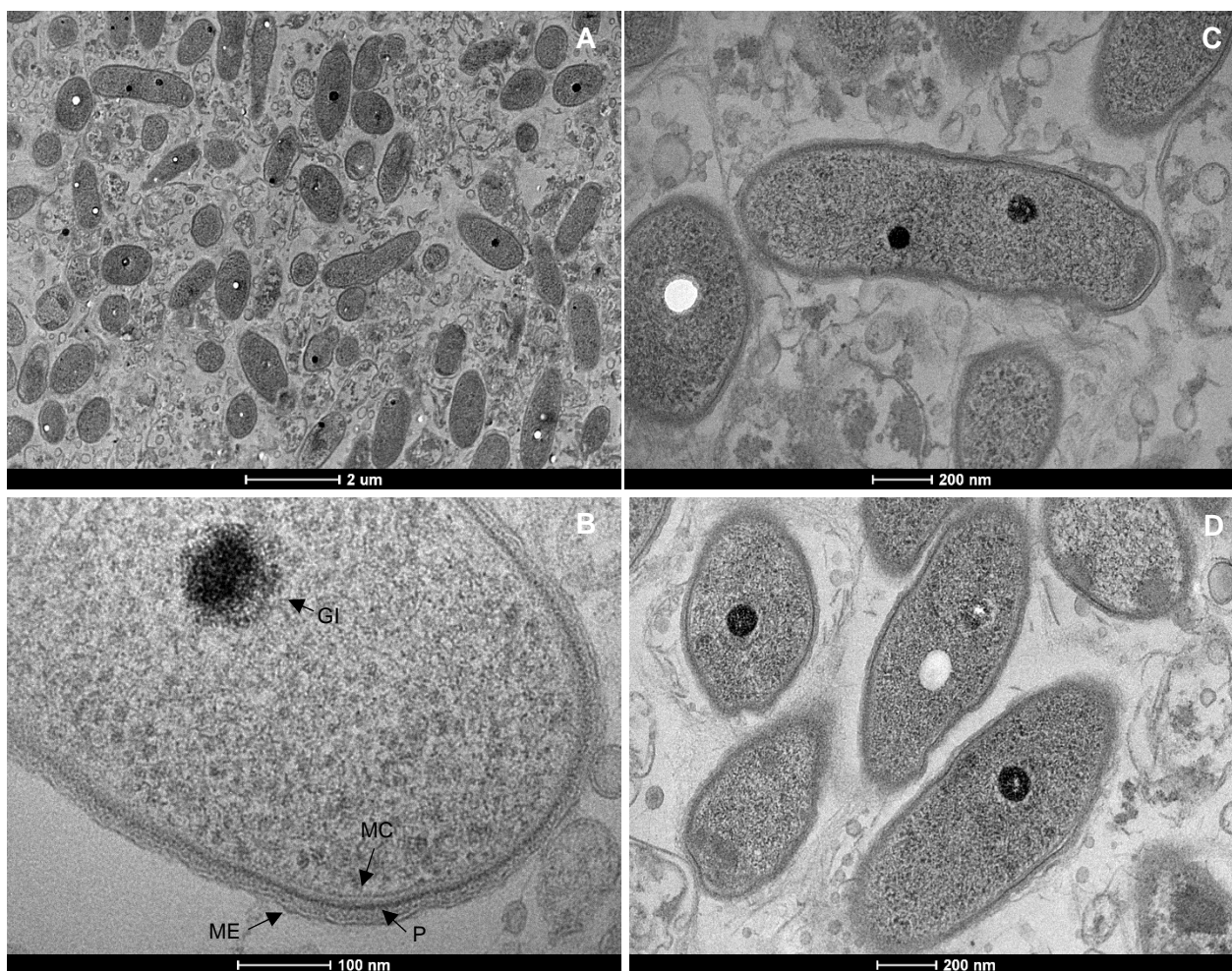


Figura 24. Eletromicrografias de transmissão das células de GeG2 crescidas por 7 dias em MM a 28°C e posteriormente criofixadas sob alta pressão (HPC). ME: membrana externa; P: peptidoglicano; MC: membrana citoplasmática, GI: Grânulo intracitoplasmático A escala está indicada embaixo de cada imagem.

As estruturas eletrodensas esféricas observadas nas células de GeG2 apresentam aspecto característico de grânulos de polifosfato (poli-P), inicialmente denominados grânulos de volutina, uma das primeiras estruturas subcelulares descritas em bactérias e já identificadas em uma grande variedade de espécies (Achbergerová e Nahálka, 2011). Tais estruturas seriam semelhantes aos acidocalcissomos, organelas eletrodensas envoltas por membranas, que apresentam grandes concentrações de polifosfato complexado com cálcio e outros elementos, inicialmente descritas em tripanossomatídeos e já identificadas em diversos eucariotos, incluindo células humanas (Docampo *et al.*, 2005). Embora tenha-se acreditado por muitos anos que, diferentemente dos acidocalcissomos eucarióticos, os grânulos de poli-P de bactérias não fossem envoltos por membranas, a presença de membranas circundantes em grânulos das alfaproteobactérias *Agrobacterium tumefaciens* e *Rhodospirillum rubrum* foi recentemente identificada, revelando que

acidocalcissomos são organelas conservadas desde bactérias até eucariotos mais complexos, o que indicaria sua origem anterior à divergência entre as linhagens procarióticas e eucarióticas (Seufferheld *et al.*, 2003; Seufferheld *et al.*, 2004). Um estudo recente identificou que, além de acidocalcissomos, grânulos de poli-P destituídos de membranas também podem ser encontrados no citoplasma da bactéria *A. tumefaciens*, evidenciando que são estruturas subcelulares diferentes, mas que podem estar presentes em uma mesma célula bacteriana (Frank e Jendrosseka, 2020).

Nas eletromicrografias de transmissão das células de GeG2 não foi possível a visualização de membranas delimitando os grânulos eletrodensos, o que poderia sugerir que as estruturas observadas no citoplasma do novo isolado bacteriano seriam grânulos de poli-P e não acidocalcissomos. No entanto, fatores relacionados ao preparo das amostras para as análises por MET podem promover alterações das estruturas biológicas, como a perda do material denso que pôde ser observado em algumas células (Figura 24A), e poderiam dificultar a visualização clara de membranas em volta dos grânulos (Docampo, 2006). Desta forma, técnicas de marcação específica para fosfolipídios ou enzimas específicas associadas às membranas de acidocalcissomos, como a H⁺-pirofosfatase vacuolar, poderiam ser empregadas para uma melhor elucidação da natureza das estruturas observadas nas células de GeG2 (Frank e Jendrosseka, 2020). Além disso, diferentes técnicas de microscopia e análises químicas poderiam auxiliar na determinação da composição específica, biogênese e potenciais funções dos grânulos citoplasmáticos do novo isolado de solo de Cerrado (Docampo, 2006; Tarayre *et al.*, 2016).

2.3.9. Dimorfismo planctônico/séssil do isolado GeG2

Após algumas transferências de GeG2 em MM líquido, a formação de grumos macroscópicos de diferentes dimensões foi observada (Figura 25A). Análises por microscopia de contraste de fase revelaram que os grumos são constituídos de agregados celulares envoltos por uma espessa matriz extracelular (Figura 26A). Ademais, células livres e móveis, com motilidade provavelmente mediada por flagelos, foram identificadas nas bordas dos grumos celulares (Figura 26B). Tal fenômeno já foi descrito para outros membros da família *Sphingomonadaceae*, tendo sido denominado dimorfismo planctônico/séssil (Pollock e Armentrout, 1999; Tirola *et al.*, 2002; Troncone, 2011).

Células de linhagens que apresentam esse fenótipo podem adotar, em meio líquido, tanto uma forma séssil, formando grandes agregados de células imóveis embebidas em matriz extracelular, quanto um estado planctônico, com células livres e móveis (Pollock e Armentrout, 1999). O dimorfismo fenotípico observado para o isolado GeG2 também já foi relatado para outras linhagens do gênero *Novosphingobium*, como *N. lentum* MT1 (Tiirola *et al.*, 2002; Tiirola *et al.*, 2005) e *Novosphingobium* sp. PP1Y (Notomista *et al.*, 2011), isoladas de um sistema de biorremediação de águas subterrâneas e de água marinha portuária, respectivamente.

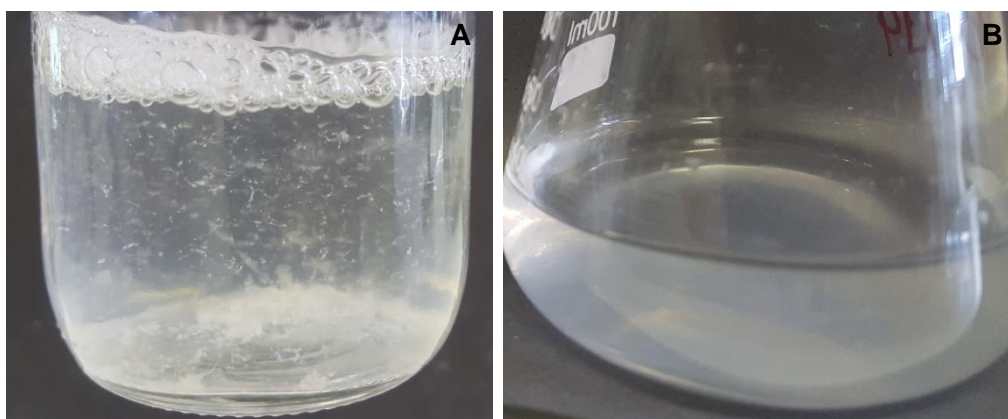


Figura 25. Culturas da bactéria GeG2 em MM líquido, incubadas por 5 dias a 28°C sob agitação (A) ou por 2 dias a 28°C, sem agitação (B).

Assim como descrito para outros membros de *Sphingomonadaceae* (Pollock e Armentrout, 1999; Troncone, 2011), o número e tamanho dos grumos nas culturas de GeG2 foram influenciados por diferentes parâmetros. A formação de grumos de GeG2 foi favorecida pela agitação da cultura, com significativo aumento após 72 horas de crescimento (Figura 25A). Por outro lado, culturas crescidas em frascos contendo volumes maiores de meio e sem agitação tenderam a apresentar poucos grumos macroscópicos e turbidez mais homogênea (Figura 25B), indicando que, assim como relatado para *Sphingomonas* sp. S88 (Pollock e Armentrout, 1999), maiores níveis de oxigênio levariam a aumentos da produção de EPS e agregação celular do isolado GeG2.

Outros fatores como temperaturas mais baixas, maiores concentrações de sal, uma alta razão carbono:nitrogênio no meio de cultura e a utilização de sacarose como substrato para o crescimento também foram correlacionados com o aumento da produção de EPS e agregação celular por bactérias da família *Sphingomonadaceae* em cultivos laboratoriais (Ashtaputre e Shah, 1995; Pollock e Armentrout, 1999; Matsuyama *et al.*, 2003).

Considerando as propriedades únicas associadas aos biopolímeros produzidos por esse grupo bacteriano, que os conferem grandes potenciais biotecnológicos e que já resultaram em dezenas de patentes (Fialho *et al.*, 2008), análises mais aprofundadas da constituição química e características reológicas dos EPS produzidos pela linhagem GeG2, bem como a determinação das condições mais favoráveis para a produção desses compostos, poderiam revelar novas características e potenciais aplicações industriais interessantes.

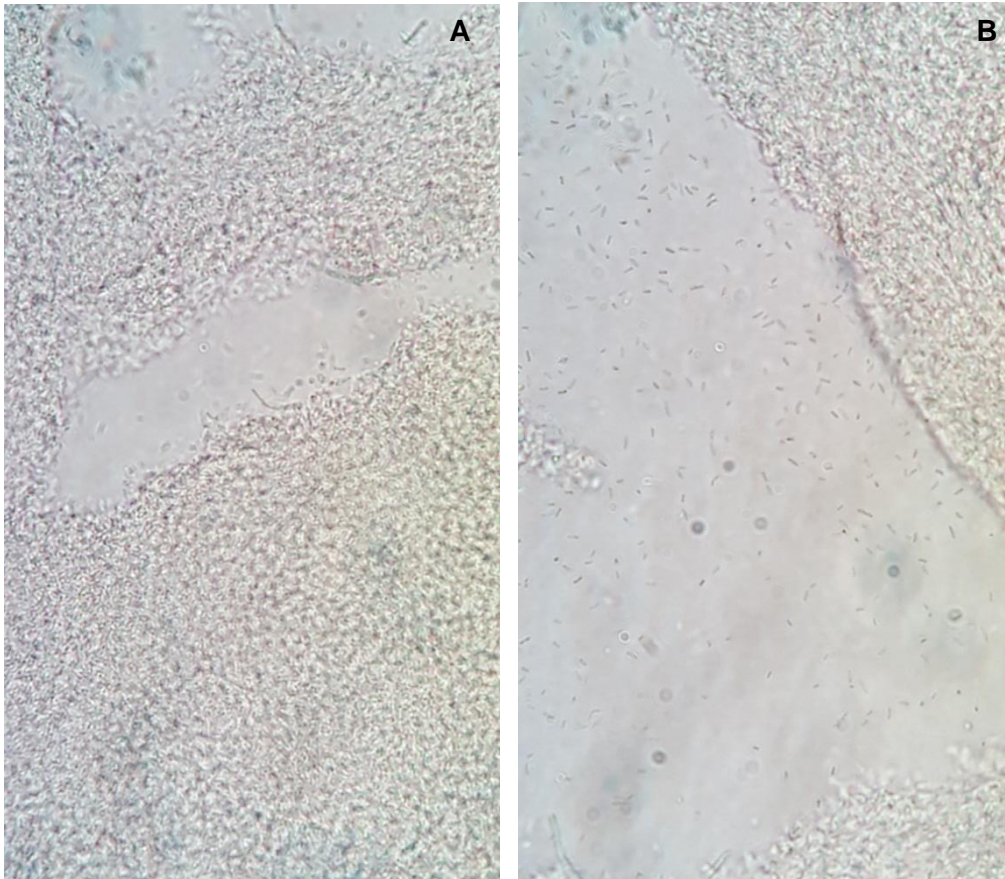


Figura 26. Observação a fresco dos grumos da cultura de GeG2 por microscopia de contraste de fase. Aumento 1000x.

Culturas de GeG2 em meio líquido, apresentando grande número de grumos macroscópicos, também foram submetidas a análises de MEV e MET. Como pode ser visto na Figura 27, além da substância amorfa característica de matriz exopolissacarídica envolvendo os agregados celulares, longos prolongamentos conectando as células em uma rede foram observados (Figura 27A e D). Além disso, estruturas globulares nanométricas, com cerca de 150 a 250 nm de diâmetro, também foram identificadas em meio aos agregados celulares (Figura 27 B e D). Curiosamente, arranjos similares foram vistos em culturas bacterianas crescidas em meios mínimos com composição semelhante

ao utilizado para o cultivo de GeG2, porém que continham hidrocarbonetos como única fonte de carbono (Goutx *et al.*, 1987; Goutx *et al.*, 1997). Análises por microscopia eletrônica utilizando marcação específica para polissacarídeos ácidos, revelaram longas fibras de exopolissacarídeos associadas a vesículas em culturas da bactéria marinha *Alcaligenes* sp. PHY 9 crescida em tetradecano (Goutx *et al.*, 1987). Redes de fibras extracelulares também foram observadas em eletromicrografias de células de *Pseudomonas nautica* 617 transferidas para meio contendo eicosano, apesar de não ocorrerem em culturas crescidas em meios contendo acetato (Goutx *et al.*, 1997). A formação dessas organizações celulares foi correlacionada ao aumento da atividade emulsificante nas culturas bacterianas, tendo sido especulado um possível papel das fibras e vesículas na otimização da degradação e assimilação dos hidrocarbonetos (Goutx *et al.*, 1987; Goutx *et al.*, 1997).

Apesar de microrganismos da família *Sphingomonadaceae* serem capazes de degradar uma grande variedade de compostos, incluindo diversos hidrocarbonetos lineares e aromáticos (Kertesz1 e Kawasaki, 2010), e genes relacionados a essas vias de degradação terem sido identificados no genoma de GeG2 (discutido na sessão 2.3.13), o meio empregado nos cultivos nos quais houve a formação de grumos continha apenas glicose e extrato de levedura como fontes de carbono (MM). O crescimento de células de *Novosphingobium* apresentando morfologia séssil, com a formação de grumos macroscópicos, foi relatada em meios mínimos contendo apenas glicose e pequenas concentrações de extrato de levedura e peptona (Tirola *et al.*, 2002), ou ácido glutâmico (Troncone, 2011), indicando que tais fenótipos não estão exclusivamente associados à degradação e assimilação de hidrocarbonetos complexos por estes organismos. No entanto, não foram realizadas análises microscópicas específicas dos agregados celulares formados por essas espécies, não sendo possível determinar se a organização e estruturas extracelulares observadas nos grumos formados pelo isolado GeG2 também ocorrem nos grumos das demais bactérias desse gênero.

Estudos morfológicos e bioquímicos detalhados dos agregados celulares formados pelas espécies de *Novosphingobium* em diferentes condições de cultivo seriam de grande relevância para a elucidação da natureza das estruturas e substâncias extracelulares envolvidas, bem como das funções associadas a esse interessante fenótipo apresentado por esse grupo bacteriano.

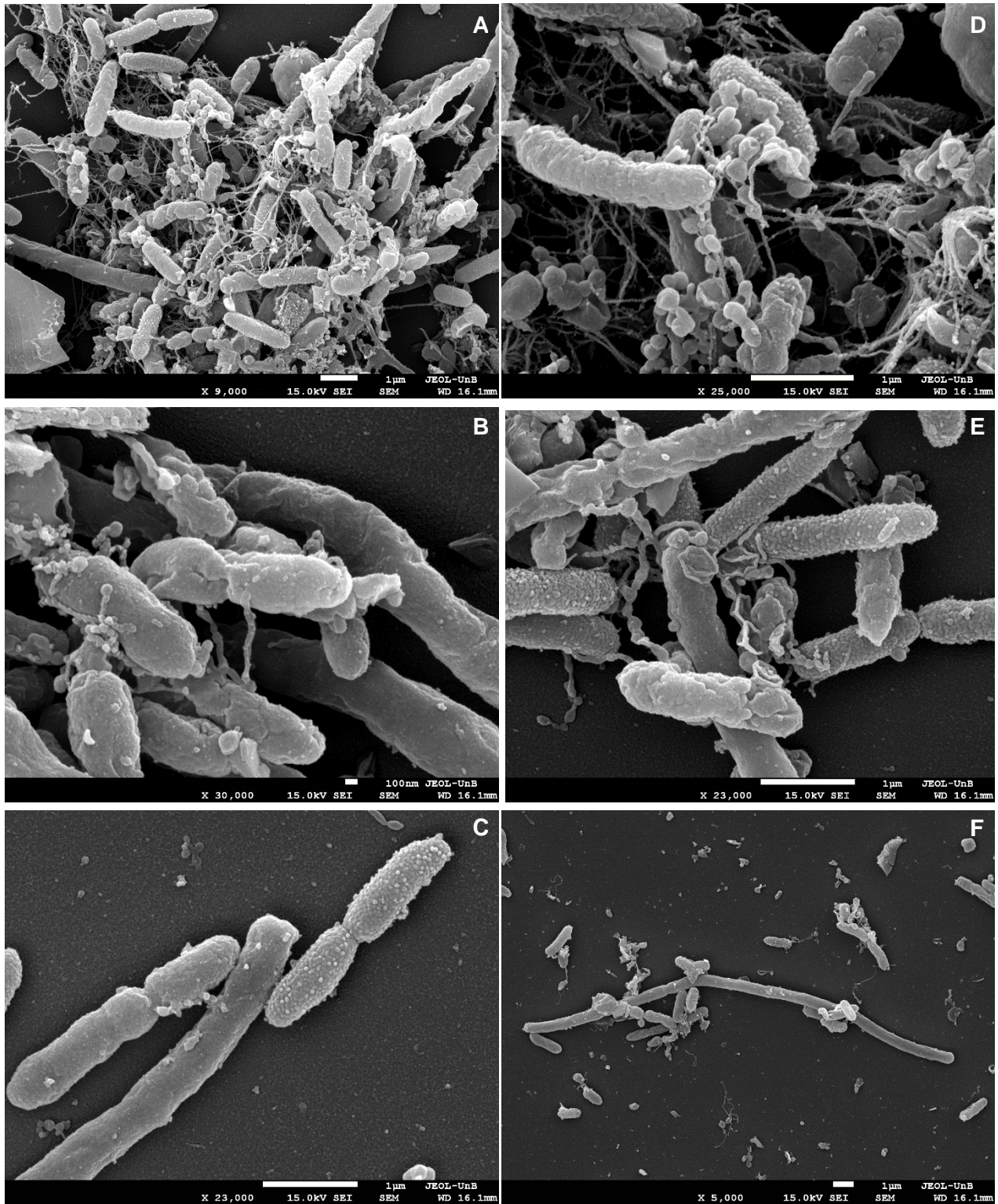


Figura 27. Eletromicrografia de varredura de culturas da linhagem GeG2 apresentando grumos macroscópicos, após crescimento em MM líquido por 14 dias a 28°C. Os aumentos utilizados estão indicados na parte inferior de cada imagem.

Vale ressaltar ainda que, assim como identificado nas eletromicrografias de varredura obtidas a partir de culturas crescidas sobre lamínulas de vidro, células de GeG2 com superfícies de aspectos diferentes também foram observadas nos cultivos líquidos

nos quais houve a formação dos grumos, com algumas células apresentando superfícies lisas e uniformes, enquanto outras exibiam aspecto rugoso, recobertas com diminutas estruturas esféricas, característica de células produtoras de vesículas de membrana externa (OMVs) (Figura 27C e E) (Gilewicz *et al.*, 1997; Coppotelli *et al.*, 2010). Como já mencionado, OMVs foram previamente identificadas e caracterizadas em culturas de bactérias pertencentes ao gênero *Novosphingobium* (Yun *et al.*, 2016; De Lise *et al.*, 2019). Estruturas globulares nanométricas de diferentes tamanhos, as maiores possuindo entre 100 e 200 nm de diâmetro e outras bem menores, foram detectadas por meio de análises de microscopia de força atômica de sobrenadantes livres de células, recolhidos de culturas em meio mínimo de *Novosphingobium* sp. PP1Y no meio da fase exponencial (De Lise *et al.*, 2019). Curiosamente, não foram identificadas vesículas em sobrenadantes de culturas desta bactéria no final da fase exponencial de crescimento em meio mínimo e em nenhum momento do crescimento em meio LB, tendo sido sugerido que as OMVs se agregariam nos grumos macroscópicos amorfos formados nessas condições e por isso não seriam detectadas livres no sobrenadante. As imagens de MEV obtidas a partir dos grumos formados por GeG2 após 14 dias de incubação (Figura 27A e D) parecem corroborar tal hipótese.

Estudos proteômicos de OMVs purificadas a partir de culturas de *Novosphingobium* sp. PP1Y crescidas em meio mínimo contendo glutamato como única fonte de carbono revelaram uma maior porcentagem de proteínas hidrolíticas em comparação à distribuição proteômica de células inteiras ou frações da membrana externa, além da presença de proteases do tipo IV ativas, sugerindo um possível papel dessas vesículas na otimização da aquisição de substratos em condições nutricionais limitadas (De Lise *et al.*, 2019). Considerando que, assim como *Novosphingobium* sp. PP1Y, a linhagem GeG2 foi mantida em meio mínimo, com baixa disponibilidade de nutrientes, as vesículas identificadas associadas aos grumos celulares também poderiam estar relacionadas à melhor aquisição de nutrientes em condições de escassez. Além disso, a presença de células bastante alongadas nas culturas (Figura 27F) também poderia indicar uma resposta a estresses nutricionais associados ao período de incubação estendido em meios com baixa disponibilidade de nutrientes, que podem levar a filamentação celular causada por interrupções nos ciclos de divisão celular, enquanto o metabolismo e crescimento celulares continuam ativos (Heinrich *et al.*, 2019).

Eletromicrografias de transmissão também foram obtidas a partir das culturas do isolado GeG2 contendo grumos macroscópicos (Figura 28). Estruturas arredondadas com

tamanho compatível ao observado por MEV foram identificadas, aparentando serem envoltas por membranas duplas e contendo materiais eledrondensos em seu interior (Figura 28F). Além disso, um aspecto ondulado da membrana externa foi observado em algumas células, que poderia estar relacionado ao processo de evaginação e formação de OMVs (Figura 28B e C). A formação de uma estrutura no citoplasma que aparenta ser secretada pela célula foi observada (Figura 28E). No entanto, após intensa busca na literatura, não foi possível associar tal formação a qualquer estrutura ou processo caracterizado em bactérias ou archaeas.

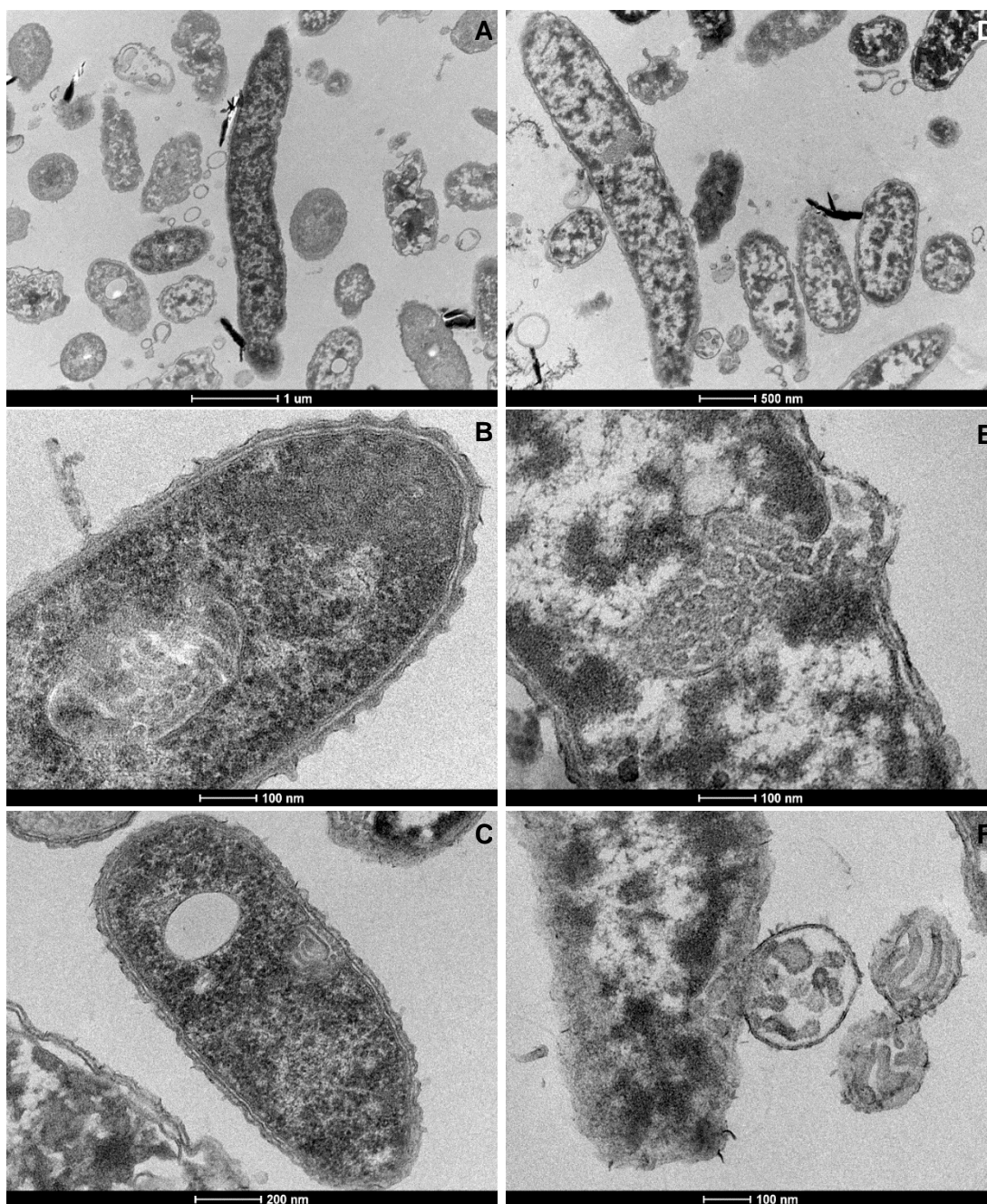


Figura 28. Eletromicrografias de transmissão das células bacterianas GeG2 crescidas por MM líquido por 14 dias a 28°C e posteriormente fixadas com aldeídos. A escala está indicada embaixo de cada imagem.

2.3.10. Sequenciamento, montagem, anotação geral e arquitetura do genoma da linhagem GeG2

Como descrito nas seções 1.3.13 e 1.4.5 do Capítulo 1, o sequenciamento genômico dos cultivos de solo de Cerrado realizado por meio de duas plataformas (Illumina HiSeq e PacBio) resultou em *reads* pertencentes a um único genoma bacteriano, pertencente ao gênero *Novosphingobium*. Após o isolamento da bactéria (GeG2) em meios definidos e das caracterizações morfológicas, fisiológicas, quimiotaxonômicas e moleculares, o genoma bacteriano foi novamente montado a partir dos dados de sequenciamento previamente obtidos, utilizando-se um *pipeline* diferente do descrito no Capítulo 1 (item 1.3.13). Dentre diferentes abordagens de montagem avaliadas, a montagem genômica híbrida, utilizando-se os dados de *long reads* de PacBio e de *short reads* de Illumina, realizada com o programa *Unicycler* gerou os melhores resultados, que se encontram resumidos na Tabela 8. Pelo fato de inicialmente ter sido assumido que se tratava de uma amostra metagenômica, as estratégias escolhidas para o sequenciamento dos cultivos de solo de Cerrado resultaram em uma cobertura bastante alta do único genoma bacteriano identificado: 336x para as sequências de PacBio e 4235x para as de Illumina.

Tabela 8. Estatísticas da montagem do genoma da linhagem bacteriana GeG2 utilizando-se o montador *Unicycler* e abordagem híbrida, incluindo dados de Pacbio (*long reads*) e Illumina (*short reads*).

Característica	Valor
Nº de <i>contigs</i>	6
Maior <i>contig</i> (pb)	4.164.843
Tamanho total (pb)	7.162.928
N50	4.164.843
L50	1
GC (%)	63,57

A completude da montagem do genoma de GeG2 foi avaliada por meio de ferramentas que analisam a presença de conjuntos de genes esperados, com base em genomas de linhagens relacionadas. O programa CheckM utiliza um amplo conjunto de genes específicos à posição de um genoma dentro de uma árvore genômica de referência, para estimar a porcentagem de completude em montagens genômicas (Park *et al.*, 2014). Já o programa BUSCO realiza uma avaliação quantitativa da completude de montagens genômicas, com base em expectativas evolutivas de conjuntos gênicos de ortólogos de

cópia única praticamente universais, denominados de categorias BUSCO (*Benchmarking Universal Single-Copy Orthologs*) (Simão *et al.*, 2015). Os resultados obtidos com ambas as ferramentas sugerem que a montagem híbrida do genoma do isolado GeG2 foi de alta qualidade e completude, como pode ser visto na Tabela 9. O genoma de *N. rosa* NRBC15208, retirado do banco de dados *GenBank*, foi utilizado como referência nas análises.

Tabela 9. Resumo dos resultados das análises de completude da montagem dos genomas de GeG2 e de *N. rosa* NRBC15208 (retirado do banco de dados *GenBank*).

Ferramenta	Parâmetro	Linhagem	
		GeG2	<i>N. rosa</i> NBRC15208
CheckM ^a	Completude (%)	99,42	99,59
	Total	1018	1018
	Completos	1003	996
BUSCO ^b (categorias)	Completos cópia única	987	986
	Completos duplicados	16	10
	Fragmentados	0	3
	Faltando	15	19

^a Linhagem marcadora: o__Sphingomonadales (UID3310)

^b Banco de dados de referência: sphingomonadales_odb10

Ao todo, o genoma completo de GeG2 foi montado em seis *contigs*, sendo os quatro maiores circulares (Tabela 10). Como mostrado na Figura 29, apesar de parecerem relacionados ao maior *contig* (*contig* 1) na visualização do grafo da montagem, os dois *contigs* menores (*contigs* 5 e 6, com 5.558 e 359 pb, respectivamente) foram montados separadamente. Ao analisar-se os genes presentes nesses *contigs*, identificou-se que o *contig* 5 contém o operon rRNA completo, com os genes codificadores dos rRNAs 16S, 23S e 5S, além de duas sequências de RNAs transportadores - tRNA^{ILE(GAT)} e tRNA^{ALA(TGC)} – contidas na região espaçadora intergênica (ITS) entre os genes rRNA 16S e rRNA 23S. Já o *contig* 6, com apenas 359 pb, contém um único gene - tRNA^{MET(CAT)}.

Tabela 10. Características dos *contigs* resultantes da montagem com o programa *Unicycler* por meio de abordagem híbrida, utilizando *short reads* e *long reads*.

<i>Contig</i>	Tamanho (pb)	Conteúdo GC (%)	Cobertura*
1	4.164.906	64,23	1x
2	2.710.928	62,93	1,08x
3	212.687	61,41	0,906x
4	68.553	55,61	2,03x
5	5.558	53,54	8,44x
6	359	59,89	7,03x

* Normalizada pelo montador *Unicycler*, em relação ao maior *contig* (*contig* 1).

Um clado bacteriano pertencente à classe *Alphaproteobacteria*, com o operon rRNA (genes 16S, tRNA^{ILE}, tRNA^{ALA}, 23S e 5S) situado em um pequeno plasmídeo ao invés do cromossomo foi reportado recentemente, revelando que tais sequências gênicas não estão exclusivamente localizadas em cromossomos principais e secundários em todas as bactérias, como se acreditava até então (Anda *et al.*, 2015). O operon rRNA desse clado, pertencente ao gênero *Aureimonas*, foi identificado exclusivamente no menor réplicon circular (9,4 kb) presente no genoma, denominado pAU20rrn, o que foi confirmado por análises de *Southern blot*. Além disso, genes relacionados a sistemas de replicação plasmidial (p. ex. *repA*) foram identificados no réplicon pAU20rrn, o que, juntamente com o alto número de cópias estimado nas diferentes fases do crescimento celular (18 a 34) e seu menor conteúdo GC em relação ao cromossomo, indicam sua natureza plasmidial (Anda *et al.*, 2015). Apesar de não terem sido feitas análises experimentais que poderiam confirmar sua localização no genoma, o fato do *contig* 5 conter somente os genes associados ao operon rRNA, não apresentando genes plasmidiais característicos, aliado aos resultados do grafo da montagem do genoma que sugerem sua associação com o *contig* 1 (Figura 29), indicam que é provável que o *contig* 5, assim como o *contig* 6, façam parte do *contig* 1 e que sua separação seja decorrente de dificuldades relacionadas à montagem dessa região genômica repetitiva.

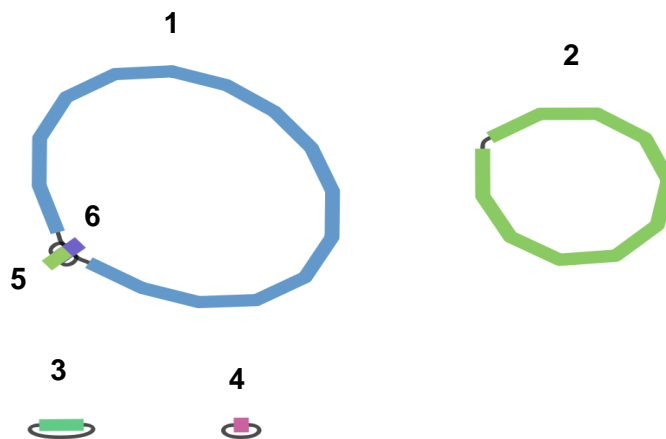


Figura 29. Grafo da montagem do genoma do isolado GeG2 com o programa *Unicycler*, por meio da abordagem híbrida, utilizando tanto *long reads* de PacBio quanto *short reads* de Illumina. A numeração atribuída aos *contigs* está indicada. O grafo foi desenhado com a ferramenta Bandage (Wick *et al.*, 2015).

Regiões genômicas repetitivas podem gerar problemas na montagem, muitas vezes levando à não identificação de sequências que ocorrem em múltiplas cópias no genoma, fazendo com que colapsem em um único *contig* e dificultando o processo de *scaffolding* (Tørresen *et al.*, 2019). Por apresentarem muitas vezes tamanhos ainda maiores que as sequências geradas a partir de tecnologias de sequenciamento de terceira geração (p. ex. PacBio), regiões repetitivas muito longas usualmente levam a erros de multiplicidade, mesmo em montagens realizadas a partir de dados de *long reads* (Wick *et al.*, 2017). A cobertura normalizada (considerando a cobertura do cromossomo como 1x) estimada pelo *Unicycler* para o *contig* 5 foi de 8,44x enquanto para o *contig* 6 foi de 7,03x. Tal resultado poderia indicar a ocorrência de 8 cópias seguidas da região genômica contendo o operon rRNA e 7 cópias da região contendo o gene codificador do tRNA^{MET(CAT)}, o que justificaria a impossibilidade da montagem dessas sequências no *contig* 1. Remontagens futuras dos dados genômicos, incluindo alterações de parâmetros e acabamentos manuais, poderiam ajudar a resolver melhor a montagem dessa porção do genoma.

O genoma completo da linhagem GeG2 apresenta 7.162.928 pb, conteúdo de GC geral de 63,57 %, com 6063 sequências codificadoras (CDSs), 3 rRNAs, 57 tRNAs e 1 tmRNA (RNA transferência-mensageiro). A Figura 30 apresenta o mapa dos quatro réplicons circulares que constituem o genoma (*contigs* 1 a 4), bem como alguns de seus aspectos genéticos. O maior réplicon (*contig* 1) possui 4.164.843 pb e conteúdo de GC de 64,23 %, contendo a maior parte das CDSs do genoma. Os *contigs* 3 e 4 são os menores réplicons, apresentando tamanhos de 212.687 e 68.553 pb e conteúdo de GC de 61,41 e 55,61 %, respectivamente. Curiosamente, um megaréplicon de 2.710.928 pb, com conteúdo GC de 62,93 % (*contig* 2), também foi identificado no genoma de GeG2.

Com o grande aumento no número de sequenciamentos e montagens de genomas completos, observou-se que vários genomas procarióticos apresentam, além do cromossomo e pequenos plasmídeos, réplicons circulares ou lineares de grande tamanho, potencialmente essenciais para a células. Genomas contendo um ou mais réplicons grandes, adicionais ao cromossomo, são chamados de genomas multipartidos ou segmentados (diCanzo e Finan, 2017), uma característica comumente encontrada em membros de *Sphingomonadaceae* (Aylward *et al.*, 2013; Wang *et al.*, 2018). Com base em aspectos funcionais e evolutivos, além de determinadas assinaturas gênicas, as moléculas de DNA presentes em uma célula bacteriana podem ser classificadas como cromossomos, cromossomos secundários, cromídeos, megaplasmídeos e plasmídeos.

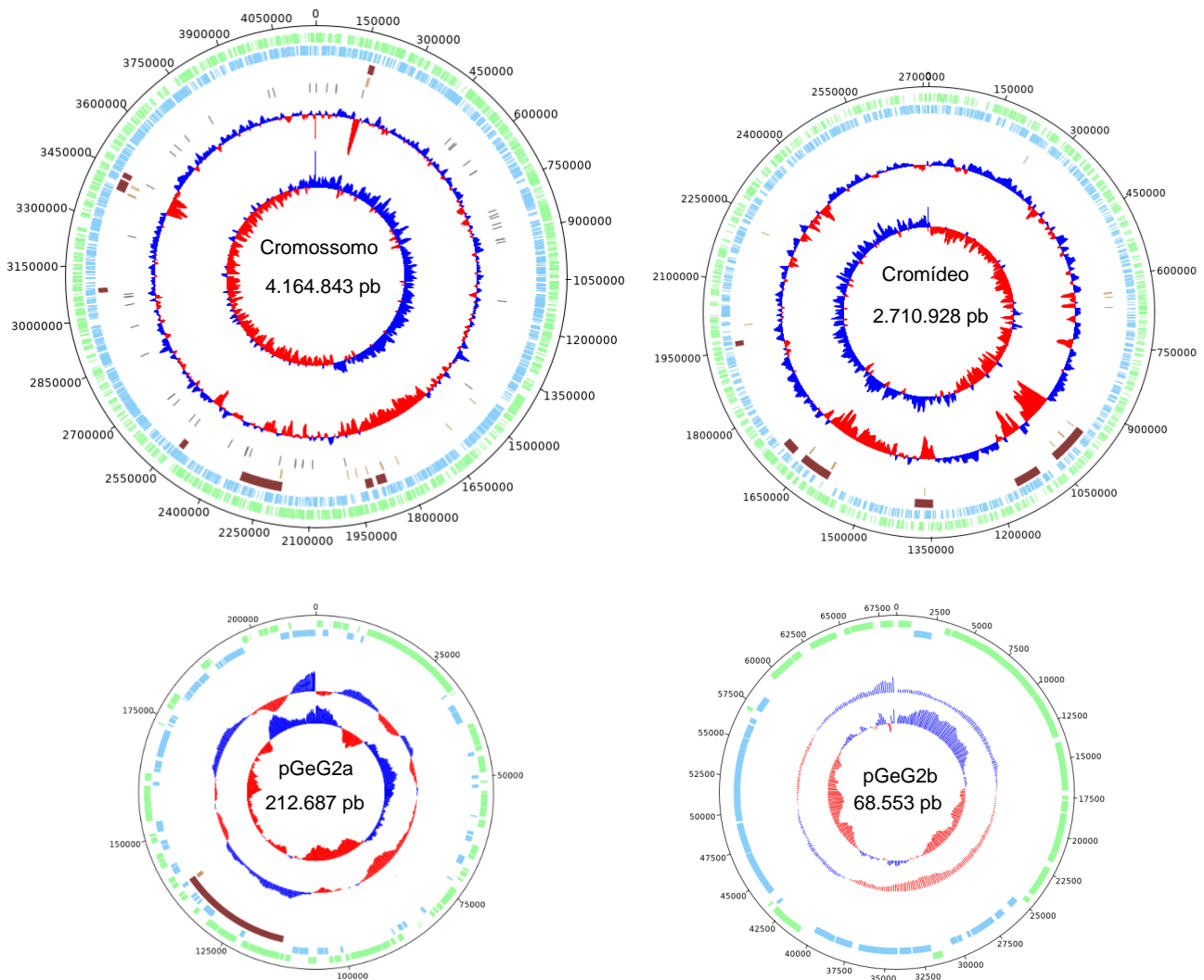


Figura 30. Mapas circulares e aspectos genéticos do cromossomo, do megaréplicon (provável cromídeo) e dos dois plasmídeos (pGeG2a e pGeG2b) da linhagem GeG2. Estão indicados, de fora para dentro: numeração das bases (preto) seqüências codificadoras (CDSs) da fita senso (verde), CDSs da fita anti-senso (azul claro), ilhas genômicas (marrom), transposases (rosa), tRNAs (roxo), conteúdo de GC e GC *skew* (azul escuro e vermelho). Os mapas foram gerados com a ferramenta *DNAPlotter* (Carver *et al.*, 2009).

Os cromossomos são sempre o maior réplicon, contendo a grande maioria dos genes essenciais, sendo a única molécula de DNA de muitos genomas bacterianos (Harrison *et al.*, 2010). Já os cromossomos secundários ocorrem raramente e seriam oriundos da quebra do cromossomo ancestral em dois diferentes réplicons, distribuindo os genes essenciais de forma relativamente aleatória entre eles (diCenzo e Finan, 2017). Os cromídeos são estruturas intermediárias entre cromossomos e plasmídeos, contendo sistemas de manutenção e replicação típicos de plasmídeos, mas uma composição

nucleotídica similar à do cromossomo, além de possuírem pelo menos um gene essencial, normalmente encontrado no cromossomo de outras espécies (Harrison *et al.*, 2010). Plasmídeos são definidos pela ausência de genes essenciais e, portanto, seriam dispensáveis para a viabilidade celular, apesar de conferirem vantagens adaptativas importantes diante das diversas variáveis ambientais (Smalla *et al.*, 2015). Além disso, apresentam sistemas de replicação, partição e manutenção específicos, característicos desses réplicons extracromossomais (del Solar *et al.*, 1998; Cevallos *et al.*, 2008). A montagem dos primeiros genomas completos revelou a existência de plasmídeos com uma grande variedade de tamanhos, sendo aqueles com tamanhos maiores que o habitualmente observado denominados megaplasmídeos (Vedler *et al.*, 2009). No entanto, tal classificação é arbitrária, com diferentes limiares de tamanho já tendo sido usados (Vedler *et al.*, 2009; Zheng *et al.*, 2013). Recentemente, diCenzo e Finan (2017) propuseram um limiar de 350 kb para a distinção de plasmídeos e megaplasmídeos, considerando que esse valor equivale a aproximadamente 10 % do tamanho do genoma bacteriano médio. No entanto, os autores reconhecem a fragilidade de se utilizar o tamanho do réplicon como único parâmetro e ressaltam que o aumento na disponibilidade de genomas completos poderá elucidar critérios mais específicos para a distinção entre plasmídeos e megaplasmídeos. É importante ressaltar ainda que muitas moléculas de DNA bacteriano apresentam características que ultrapassam tais definições, dificultando sua identificação e evidenciando que as classificações propostas não apresentam delimitações rígidas.

O *contig* 1 é o maior réplicon identificado no genoma da linhagem GeG2 (4,1 Mb), no qual foram detectados grande parte dos genes *housekeeping*, estruturais e funcionais, essenciais à viabilidade celular. Por essa razão, tal *contig* foi definido como o cromossomo bacteriano (Figura 30). Análises dos réplicons menores (212 e 68 kb) realizadas com o *Plasflow* (Krawczyk *et al.*, 2018), uma ferramenta que se baseia em assinaturas genômicas e emprega modelos de redes neurais para a identificação de plasmídeos bacterianos, identificaram assinaturas plasmídias no *contig* 3, mas não foram capazes de classificar o *contig* 4. Buscas manuais posteriores revelaram genes relacionados ao sistema de replicação e partição plasmidial da família ABC (p. ex: *repA*), comumente encontrados em plasmídeos de bactérias da classe *Alphaproteobacteria* (Pinto *et al.*, 2012), tanto no *contig* 3 quanto no *contig* 4. Por esse motivo, e considerando que possuíam menos de 350 kb, ambos os réplicons foram classificados como plasmídeos, de acordo com o critério proposto por diCenzo e Finan (2017), e denominados pGeG2a e

pGeG2b (Figura 30). Vale ressaltar que genes do sistema conjugativo *tra* foram identificados em ambos os plasmídeos, indicando sua natureza conjugativa.

Por sua vez, o megarréplicon circular (2,7 Mb) representado pelo *contig 2* apresenta características interessantes. Apesar de assinaturas cromossômicas terem sido identificadas neste réplicon por meio das análises com o *Plasflow*, a presença de genes relacionados à replicação e partição de plasmídeos da família ABC de alfa-proteobactérias foi observada, sugerindo assim que o megarréplicon extracromossomal presente no genoma de GeG2 apresentaria tanto características plasmidiais quanto cromossomais, podendo tratar-se de um cromídeo. A determinação da presença de um ou mais genes essenciais seria necessária para a confirmação da classificação do megarréplicon como um cromídeo e não um megaplasmídeo, já que essa é a principal característica discriminatória entre esses réplicons bacterianos. No entanto, considerando-se o amplo repertório genético que pode ser encontrado entre os diferentes táxons bacterianos e a variabilidade de condições ambientais a que podem estar submetidos, a determinação de um gene como essencial não é trivial, tornando-se tarefa difícil de ser realizada por anotações genômicas. Assim, evidências experimentais seriam necessárias para determinar se o réplicon de fato carrega genes indispensáveis para a viabilidade celular (diCenzo e Finan, 2017). Contudo, como tais procedimentos laboratoriais são extremamente laboriosos e dispendiosos, alguns critérios genômicos que poderiam auxiliar na distinção entre cromídeos putativos e megaplasmídeos vêm sendo sugeridos, com base em amplas análises comparativas de genomas completos disponíveis em bancos de dados.

Cromídeos apresentam composição nucleotídica mais semelhante ao cromossomo do que os plasmídeos, com o conteúdo de GC usualmente diferindo em torno de 1 % em relação ao cromossomo, enquanto plasmídeos costumam apresentar conteúdos de GC bem menores que os cromossomos (Harrison *et al.*, 2010, diCenzo e Finan, 2017). Além disso, diCenzo e Finan (2017) propuseram que a distância da abundância relativa de dinucleotídeos observada entre o cromossomo e o megarréplicon de um determinado genoma poderia ser utilizada como um indicativo para a classificação de réplicons maiores que 350 kb em cromídeos putativos ou megaplasmídeos. A abundância relativa de dinucleotídeos refere-se à soma das diferenças nas frequências de cada par de dinucleotídeo entre duas fontes de DNA, sendo um limiar de 0,4 proposto, com valores iguais ou menores a esse usualmente detectados em cromídeos, enquanto valores maiores são observados para megaplasmídeos. Com base nesses critérios, os valores obtidos tanto para a variação de conteúdo de GC entre o cromossomo e o megarréplicon secundário

(1,3 %) quanto para a distância da abundância relativa de dinucleotídeos entre esses réplicons (0,257) sugerem que o megaréplicon extracromossomal presente no genoma de GeG2 é possivelmente um cromídeo. Vale ressaltar que análises incluindo as sequências nucleotídicas dos *contigs* 5 e 6 no cromossomo e considerando o número de cópias esperadas com base na cobertura observada para esses *contigs* (8 vezes para o *contig* 5 e 7 vezes para o *contig* 6, conforme mencionado acima) também resultaram em valores de variação de conteúdo de GC e distância de abundância relativa de dinucleotídeos dentro dos limiares esperados para cromídeos - 1,18 % e 0,308, respectivamente.

É interessante notar que megaréplicons secundários, com mais de 350 kb, foram identificados em todos os seis genomas completos de membros do gênero *Novosphingobium* disponíveis em bancos de dados até o momento. Apesar de inicialmente identificado como um plasmídeo (Bell *et al.*, 2010), o megaréplicon pNL2 de *N. aromaticivorans* DSM 12444, de 487,3 kb, foi classificado como um cromídeo putativo em análises realizadas por diCenzo e Finan (2017), que consideraram os critérios de variação de conteúdo de GC e distância de abundância relativa de dinucleotídeos entre o cromossomo e o réplicon secundário para a classificação. Já os megaréplicons pLA3 de *N. pentaromativorans* US6-1 (756,8 kb) e Mpl de *Novosphingobium* sp. PP1Y (1,16 Mb) foram identificados como megaplasmídeos por meio desses parâmetros genômicos (diCenzo e Finan, 2017). Análises mais aprofundadas visando a determinação da natureza dos grandes réplicons secundários identificados nos genomas de *N. resinivorum* SA1 (pSA1 com 1,756 Mb e pSA2 com 960,8 kb) e *Novosphingobium* sp. THN1 (1,161 Mb) não foram realizadas, com esses réplicons extracromossomais, que foram denominados genericamente de plasmídeos (Wang *et al.*, 2018). O genoma de *Novosphingobium* sp. P6W, além de dois réplicons de 720,5 e 188,6 kb identificados como plasmídeos (pP6W1 e pP6W2), apresenta um megaréplicon de 2,246 Mb que foi proposto como um cromossomo secundário devido a seu grande tamanho (Gogoleva *et al.*, 2019). O cromídeo putativo identificado no genoma da linhagem GeG2 é ainda maior (2,710 Mb), sendo o maior réplicon secundário identificado até então em genomas de *Novosphingobium*. É possível que o aumento da disponibilidade de genomas com a arquitetura completamente resolvida resultante da crescente utilização de sequenciamentos de *long reads* leve à identificação mais genomas bacterianos multipartidos e auxilie na elucidação da natureza, da classificação e das características específicas dos grandes réplicons secundários presentes nos genomas de diferentes grupos bacterianos.

2.3.11. Classificação taxonômica da linhagem GeG2 baseada no genoma completo

Inicialmente, a sequência do gene rRNA 16S completo predita a partir do genoma do isolado GeG2 (1486 pb) foi comparada com as sequências dos fragmentos do gene amplificadas por PCR diretamente da cultura (discutido na seção 2.3.3), revelando-se idênticas. Conforme previamente descrito para as sequências de *amplicons*, a sequência do gene rRNA 16S completo extraída do genoma apresenta 100 % de identidade com a sequência da linhagem *N. rosa* NBRC15208 disponível nos bancos de dados (95 % de *query coverage*).

Conforme recentemente proposto por Chun *et al.*, 2018, análises baseadas em dados de genomas completos podem ser utilizadas para a classificação taxonômica de procaríotos e o cálculo de índices que avaliam o grau de semelhança geral entre genomas, denominados OGRI (*Overall Genome Relatedness Index*), é necessário para a determinação de novas espécies, especialmente para isolados que apresentam identidades do gene rRNA 16S $\geq 98,7$ % com espécies já validadas. Análises de ANI (*Average Nucleotide Identity*) e dDDH (*digital DNA-DNA hybridization*) são os OGRI mais utilizados, com valores de ANI de 95~96 % e de dDDH de 70 % propostos e amplamente aceitos para a delimitação de espécies diferentes (Chun *et al.*, 2018). Como pode ser visto na Figura 31, apesar da identidade de 100 % observada entre os genes rRNA 16S de GeG2 e *N. rosa* NBRC15208, tanto o valor de ANI quanto o de dDDH encontrados para essas linhagens estão muito abaixo do limiar preconizado para espécie, indicando que pertencem a espécies diferentes.

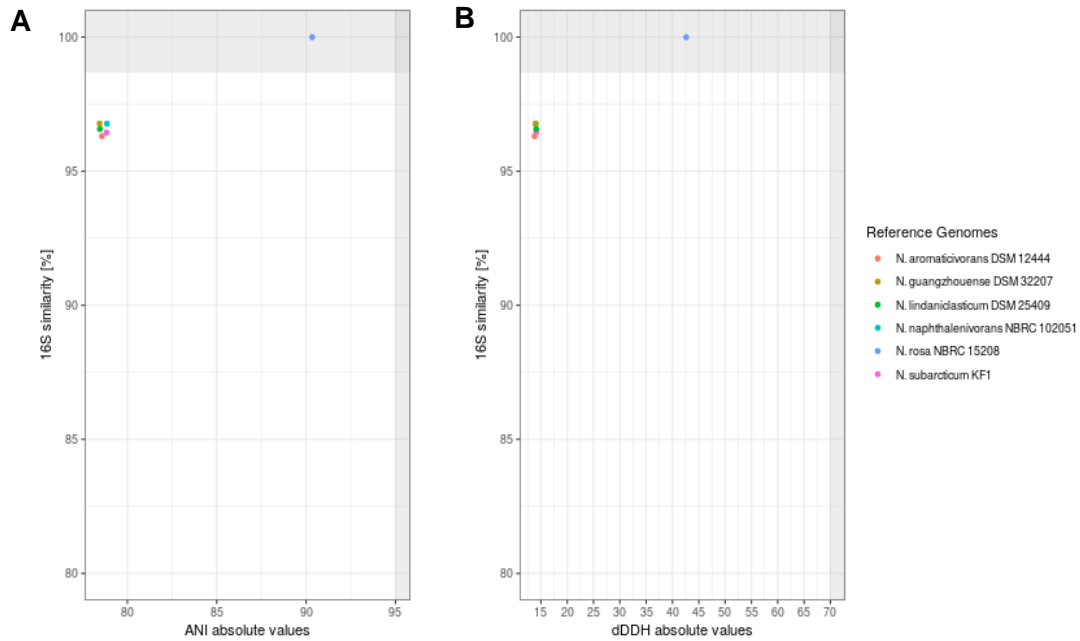


Figura 31. Relação entre as similaridades das sequências do gene rRNA 16S e os valores de ANI (A) ou de dDDH (d_0) (B) entre a linhagem GeG2 e seis outras linhagens do gênero *Novosphingobium* que possuem genomas sequenciados. Os limiares dos valores de ANI e dDDH propostos para a delimitação de espécies estão indicados a 95 e 70 %, respectivamente (Chun *et al.*, 2018).

Recentemente foram disponibilizados dois servidores automatizados que realizam análises de taxonomia de procariotos baseadas em genomas completos e utilizam bancos de dados abrangentes, constantemente atualizados e taxonomicamente validados (Ha *et al.*, 2019; Meier-Kolthoff e Göker, 2019). O TYGS (*Type Strain Genome Server*) utiliza uma abordagem integrada, inicialmente identificando os genomas mais próximos do genoma a ser classificado por meio da extração do gene rRNA 16S e comparações com um amplo banco de dados curado de genomas de qualidade de linhagens tipos. Em seguida, a distância entre o genoma avaliado e os genomas identificados como mais próximos é calculada por meio do método GBDP (*Genome BLAST Distance Phylogeny*) e empregada em inferências de árvores filogenômicas. Além disso, valores de dDDH e diferenças no conteúdo de GC genômico são calculados e utilizados, juntamente com os resultados filogenômicos, para a classificação taxonômica do genoma, por vezes até o nível de subespécie (Meier-Kolthoff e Göker, 2019). Já o *TrueBac ID* realiza a identificação taxonômica de espécies procarióticas com base em estimativas de valores e coberturas de ANI calculadas entre o genoma a ser classificado e genomas disponíveis no banco de dados *EZBioCloud* (Yoon *et al.*, 2017), bem como em identidades entre sequências dos genes rRNA 16S, *repA* e *rplC* (Ha *et al.*, 2019).

Análises genômicas realizadas tanto na plataforma TYGS quanto na *TrueBac ID* identificaram o isolado GeG2 como uma nova espécie bacteriana, não pertencendo a qualquer espécie cujo genoma esteja atualmente presente nos bancos de dados (Figura 32 e Figura 33). Vale ressaltar que o genoma da linhagem *N. rosa* NBRC15208 está presente nos bancos de dados de ambas as plataformas e, apesar de ter sido identificado como o genoma mais próximo de GeG2 (Tabela 11, Figura 32 e Figura 33) as linhagens foram classificadas como espécies distintas. A Tabela 11 apresenta os valores de dDDH, estimados de três diferentes maneiras, entre os genomas de GeG2 e das seis linhagens identificadas como sendo as mais próximas pelo TYGS, enquanto a Figura 32 mostra a árvore filogenômica inferida pela mesma plataforma, incluindo algumas características dos genomas relacionados e evidenciando o posicionamento de GeG2 como uma espécie distinta. O resultado obtido na análise realizada pelo sistema *TrueBac ID*, baseado em estimativas de valores de ANI e identidades de genes específicos, está apresentado na Figura 33.

Por conta do amplo acesso a tecnologias de sequenciamento genômico cada vez mais aprimoradas e eficientes, abordagens baseadas em comparações de sequências de genomas completos vêm sendo crescentemente empregadas para o posicionamento taxonômico de novos isolados procarióticos, principalmente por possibilitarem um novo nível de padronização resultante da sua rápida reprodutibilidade e validação, que levam a um grande ganho de tempo e precisão em relação às técnicas de hibridização DNA-DNA tradicionalmente consideradas o padrão ouro para a delimitação de novas espécies (Chun *et al.*, 2018). No entanto, é importante ressaltar que a ampla utilização da taxonomia de procariotos baseada em genomas depende do aumento da disponibilidade de genomas completos de espécies validadas e linhagens tipos. Nesse sentido, iniciativas de sequenciamentos genômicos em larga escala, como o GEBA (*Genomic Encyclopedia of Bacteria and Archaea*) (Whitman *et al.*, 2015), têm contribuído significativamente para a expansão da cobertura genômica de microrganismos, aumentando a relevância prática da utilização de critérios genômicos para a taxonomia e sistemática microbiana. Além disso, a inclusão e disponibilização de dados e análises de genomas completos recentemente passaram a ser obrigatórias para a descrição e validação de novas taxa procarióticos (*International Journal of Systematic and Evolutionary Microbiology* - <https://www.microbiologyresearch.org/journal/ijsem/scope>; *Antonie van Leeuwenhoek* - <https://www.springer.com/journal/10482/updates/17239594>, acessados em julho 2020).

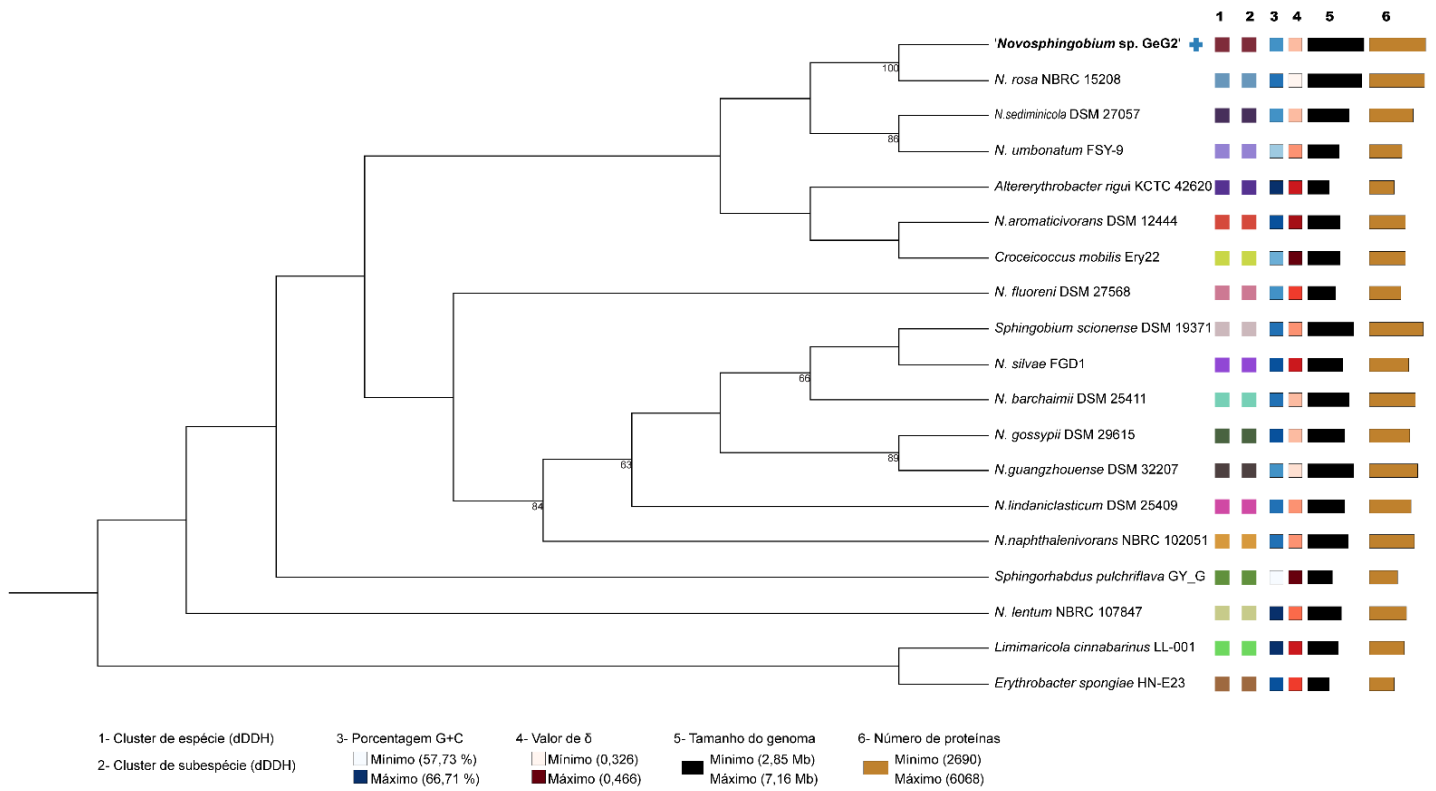


Figura 32. Árvore filogenética baseada em genomas completos, gerada pela ferramenta TYGS (*Type Strain Genome Server*) (Meier-Kolthoff e Göker, 2019). A árvore foi inferida por meio do programa FastME 2.1.6.1 (Lefort *et al.*, 2015) utilizando distâncias GBDP (*Genome BLAST Distance Phylogeny*) calculadas a partir das seqüências genômicas de GeG2 (em negrito, indicada pela cruz azul) e das linhagens mais próximas identificadas pelo servidor TYGS. A escala dos ramos está de acordo com a fórmula de distância GBDP d_5 . Os números acima dos ramos representam valores de pseudo-bootstrap GBDP, sendo mostrados apenas os > 60 % (100 réplicas), com um suporte médio dos ramos de 49,6 %. A árvore foi enraizada no ponto médio (Faris, 1972). As marcações em cores ao lado dos nomes das espécies indicam a afiliação a *clusters* de espécies (1) e subespécies (2), conteúdo de GC genômico (3), valores de δ (4), tamanho dos genomas (5) e número de proteínas identificado (6).

TRUEBAC

alinebelmok@aluno.unb.br

Sample name: **Novosphingobium_sp_G...**

Decision: **DEFINITIVE**

Genome size: 7,162,928 bp No. of contigs: 6

GC content: 63.57% NSI: 4,164,843 bp

DB ver: 20200206 Pipeline ver: 1.92

Sequencing depth: N/A No. of LBGG (paralogs): 92/92 (0)

Download

TrueBac™ ID - Genome Result

Novosphingobium_sp_GeG2 is identified as: **Novosphingobium sp. nov.** genomic evidence

Candidate species

No.	Hit taxon	ANI (%)	ANI coverage (%)	16S (%)	rncA (%)	rplC (%)	Taxonomy
1	Novosphingobium rosa	90.10	61.0	100	94.77	97.38	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
2	Novosphingobium lotistagni	N/A	N/A	97.59	N/A	N/A	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
3	Novosphingobium oryzae	N/A	N/A	97.02	N/A	N/A	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
4	Novosphingobium barchaimii	83.80	4.9	96.95	N/A	N/A	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
5	CP030353_s	83.54	4.8	96.81	N/A	N/A	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
6	Novosphingobium capsulatum	83.57	6.7	96.44	87.52	84.89	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
7	FWXL_s	83.92	4.8	96.37	N/A	85.71	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
8	Novosphingobium umbonatum	83.84	10.4	96.36	N/A	85.46	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
9	CP028347_s	83.53	5.3	96.30	N/A	83.33	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
10	Novosphingobium subterraneum	83.89	4.7	96.16	N/A	85.71	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
11	Novosphingobium aromaticivorans	83.79	5.0	96.09	85.93	85.71	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium
12	FPEH_s	83.68	5.0	95.93	86.10	85.95	Bacteria;Proteobacteria;Alphaproteobacteria;Sphingomonadales;Sphingomonadaceae;Novosphingobium

Figura 33. Resultado da análise de identificação taxonômica do genoma da linhagem GeG2 realizada por meio da plataforma online *TrueBac ID* (<https://www.truebacid.com/>) em abril de 2020.

Tabela 11. Valores de dDDH (*digital DNA-DNA hybridization*) e diferença de conteúdo de GC encontrados em comparações par a par entre o genoma da linhagem GeG2 e genomas de linhagens tipo disponíveis no banco de dados identificadas como mais próximas. Os valores de dDDH estimados com três diferentes fórmulas (d_0 , d_4 e d_6), bem como os intervalos de confiança (C.I.) para cada uma delas, estão apresentados. d_0 = extensão de todos pares de segmentos com maiores pontuações (HSPs) divididos pela extensão total do genoma. d_4 = soma de todas as identidades achadas nos HSPs dividido pela extensão geral de HSP. d_6 = soma de todas as identidades encontradas nos HSPs dividido pela extensão total do genoma. Assim, enquanto d_0 e d_6 medem a similaridade em termos de conteúdo gênico, d_4 reporta a similaridade baseada em identidades de sequências. As análises foram realizadas por meio da ferramenta TYGS (*Type Strain Genome Server*), disponível *online* (<https://tygs.dsmz.de/>), em julho de 2020.

Linhagem	dDDH (d_0, %)	C.I. (d_0, %)	dDDH (d_4, %)	C.I. (d_4, %)	dDDH (d_6, %)	C.I. (d_6, %)	Diferença G+C (%)
<i>N. rosa</i> NBRC15208	42,6	39,2-46	37,3	34,9 - 39,8	40,8	37,9 - 43,9	0,94
<i>N. silvae</i> FGD1	13,6	10,8-16,9	21,0	18,8 - 23,5	13,9	11,5 - 16,7	1,52
<i>N. naphthalenivorans</i> NBRC102051	14,0	11,2 – 17,3	20,2	18,0 – 22,6	14,2	11,8 - 17,0	0,26
<i>N. aromaticivorans</i> DSM 12444	13,8	11,1-17,2	20,1	17,8 - 22,5	14,1	11,7 - 16,9	1,55
<i>N. lindaniclasticum</i> DSM 25409	14,1	11,3 – 17,5	20,1	17,9 - 22,5	14,4	11,9 - 17,2	0,99
<i>N. barchaimii</i> DSM 25411	14	11,2-17,3	20,0	17,8 - 22,4	14,2	11,8-17,0	0,45

2.3.12. Análises funcionais do genoma da linhagem GeG2

O grande número de sequenciamentos genômicos realizado nos últimos anos resultou no aumento exponencial no depósito de sequências de proteínas preditas em bancos de dados públicos, com apenas uma pequena fração dessas avaliadas experimentalmente ou estudadas mais detalhadamente. Dessa forma, abordagens computacionais tornam-se essenciais para a atribuição das prováveis funções do conjunto de proteínas codificado em genomas microbianos (Galperin *et al.*, 2015). Dentre as diferentes abordagens existentes para a classificação de proteínas, o banco de dados COG (*Clusters of Orthologous Groups*) tem sido bastante utilizado para a análise funcional de genomas procarióticos. Criada em 1997, essa abordagem é baseada na comparação de proteínas conhecidas e preditas em genomas microbianos completos visando inferir conjuntos de ortólogos, ou seja, homólogos funcionais em diferentes genomas que descendem de uma linhagem vertical ininterrupta e tipicamente retêm suas funções fisiológicas (Tatusov *et al.*, 1997). Desde sua proposta, diversas atualizações foram feitas no banco de dados COG (Tatusov *et al.*, 2003; Galperin *et al.*, 2015), que passou a ser implementado em várias ferramentas de anotação funcional de genomas microbianos e

amplamente utilizado em estudos de genômica comparativa, sendo adotado como uma das métricas necessárias para a descrição de novos genomas (Galperin e Kolker, 2006; Galperin *et al.*, 2019).

Uma interessante característica do sistema COG é a divisão dos *clusters* de grupos ortólogos de proteínas em 26 categorias funcionais, denominadas por letras, que permitem a rápida visualização e comparação da distribuição funcional das proteínas preditas em genomas procarióticos e eucarióticos (Galperin *et al.*, 2019). Análises funcionais do genoma da linhagem GeG2 realizadas pela ferramenta *eggNOG-mapper* (Huerta-Cepas *et al.*, 2017) atribuíram 5.255 CDSs a uma ou mais categorias COG, distribuídas em 22 categorias ao todo (Figura 34). Apesar de anotados pela ferramenta, outros 332 genes não puderam ser atribuídos a COGs. Considerando que a classificação das proteínas preditas em uma categoria funcional COG é realizada caso apresente pelo menos três melhores correspondências (*best hits*) com proteínas classificadas na mesma categoria COG, mas originando de três clados taxonômicos distintos, uma mesma proteína pode ser atribuída a mais de uma categoria funcional. Do total de CDS do genoma de GeG2 atribuídos a COGs, um pouco mais de 90 % (4807) foram classificadas em uma única categoria funcional, enquanto 351 (6,67 %) foram classificadas em duas categorias e 96 (1,83 %) em três categorias diferentes.

Ao todo, proteínas preditas no genoma de GeG2 atribuídas a categorias relacionadas com metabolismo (C, E, F, G, H, I, P, Q) representaram a maior parte (47,2 %), enquanto aquelas classificadas em categorias associadas a processos celulares (D, M, N, O, T, U, V, Z) e a armazenamento e processamento de informações (A, B, J, K, L) somaram por volta de 20 %, cada (Figura 34A). Além disso, vale ressaltar que 1094 proteínas foram classificadas na categoria de função desconhecida (S) (Figura 34A). A distribuição das categorias funcionais identificadas em cada um dos quatro réplicons do genoma de GeG2 está apresentada na Figura 34B.

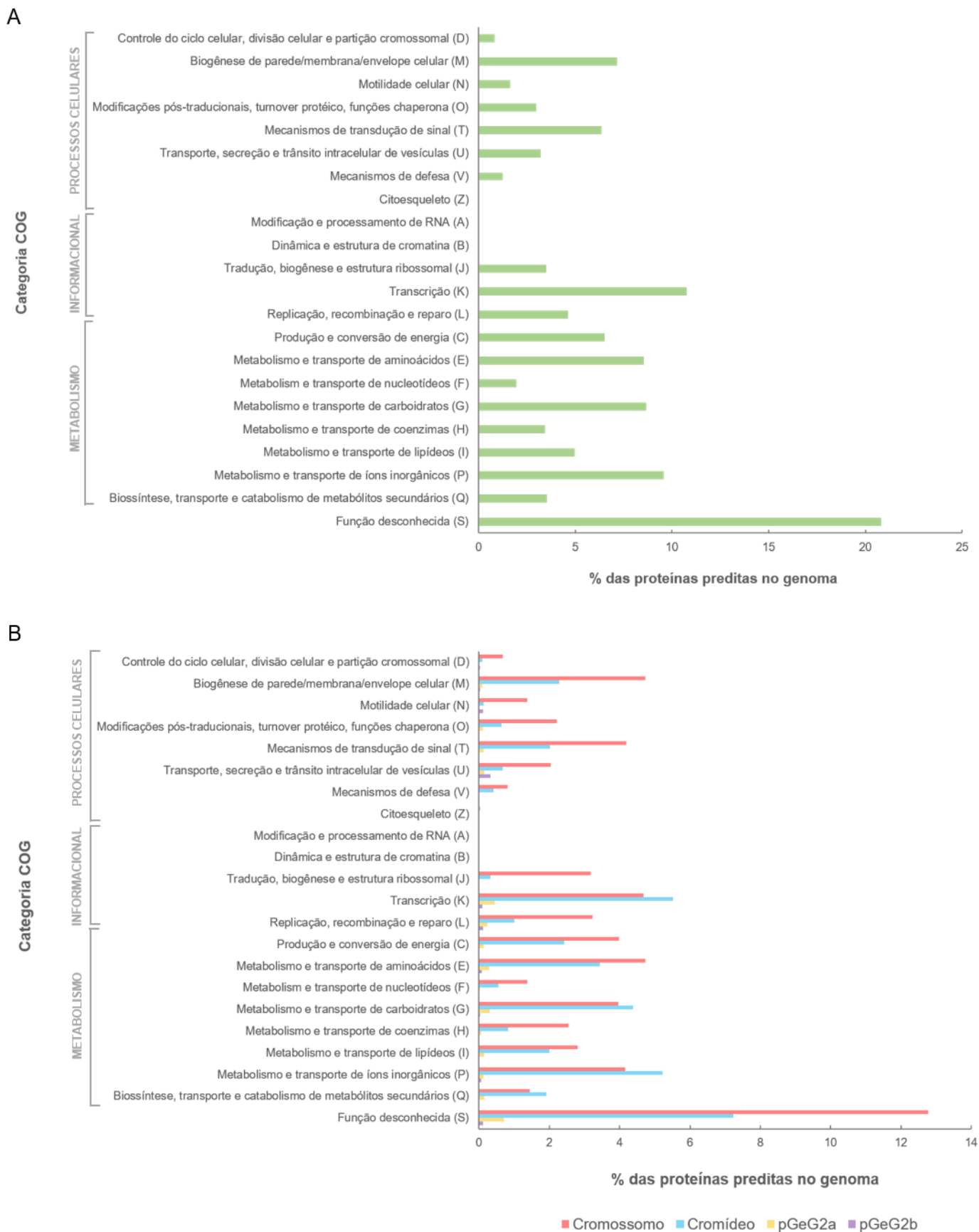


Figura 34. Classificação baseada em categorias funcionais COG (*Cluster of Orthologous Groups*) das proteínas preditas no genoma do isolado GeG2. Em **A** está apresentada a distribuição das categorias COG no genoma como um todo e, em **B**, a distribuição em cada um dos 4 réplicons do genoma (cromossomo, cromídeo e plasmídeos pGeG2a e pGeG2b).

Vieses funcionais entre os réplicons de genomas segmentados vêm sendo recorrentemente identificados em diversas espécies bacterianas por meio de abordagens baseadas em COG. Conforme esperado, categorias COG associadas a funções centrais (*core*) como replicação e reparo de DNA (L), tradução (J), modificações pós-traducionais (O), divisão celular (D) e biogênese de membrana e parede celulares (M) são consistentemente identificadas como enriquecidas nos cromossomos (Mackenzie *et al.*, 2001; Chain *et al.*, 2006; Jansen *et al.*, 2010; diCenzo e Finan, 2017). Por outro lado, genes de proteínas associadas a transporte e metabolismo de aminoácidos (E), de lipídeos (I), de íons inorgânicos (P), biossíntese e catabolismo de metabólitos secundários (Q), bem como transcrição (K) e transdução de sinal (T), costumam ser enriquecidos em megaréplicons secundários (Chain *et al.*, 2006; Jansen *et al.*, 2010; diCenzo *et al.*, 2019).

Análises da distribuição do conteúdo funcional nos diferentes réplicons do genoma de GeG2 revelaram que, de acordo com o observado em outros genomas bacterianos segmentados, proteínas pertencentes às categorias L, J, D, M e O, envolvidas em processamento de informações e processos celulares essenciais, estão majoritariamente presentes no cromossomo, enquanto que as categorias funcionais Q, P, G, E, I e K, relacionadas ao metabolismo e transporte de compostos secundários, íons inorgânicos, carboidratos, aminoácidos, lipídeos, bem como transcrição, são super-representadas no cromídeo (Figura 35). O enriquecimento dessas classes funcionais no cromídeo pode sugerir seu papel adaptativo frente a variabilidades ambientais, provavelmente apontando um maior número de fatores de transcrição neste réplicon, que poderiam permitir regulações gênicas em resposta a estímulos ambientais, assim como de genes envolvidos em armazenamento energético e comportamento celular responsivo (Jansen *et al.*, 2010; diCenzo e Finan, 2017). No entanto, curiosamente, embora genes associados à motilidade celular (N) e a mecanismos de transdução de sinal (T) encontrem-se super-representados em cromídeos de algumas espécies bacterianas ambientais, como *Cupriavidus metallidurans* CH34 (Jansen *et al.*, 2010) e *Marinovum algicola* DG898 (Frank *et al.*, 2015), o mesmo não foi observado no genoma do isolado GeG2, que contém a maior parte dos genes relacionados a essas funções no cromossomo (Figura 35). Além disso, proteínas de função desconhecida foram encontradas em maiores quantidades no cromossomo de GeG2, ao contrário do observado em genomas de outras bactérias da família *Sphingomonadaceae*, nos quais proteínas hipotéticas são majoritariamente codificadas nos réplicons secundários (Aylward *et al.*, 2013).

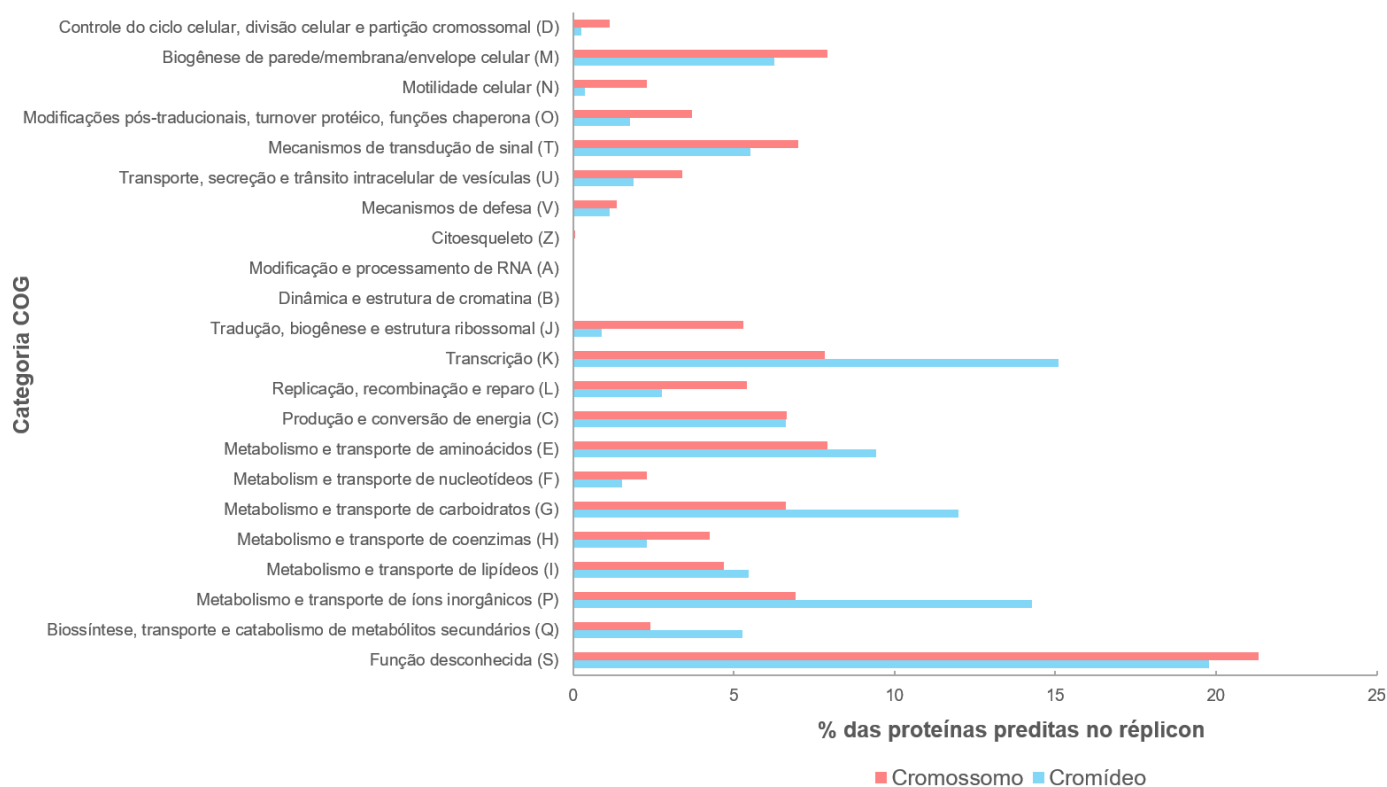


Figura 35. Distribuição funcional das proteínas preditas no cromossomo e no cromídeo da linhagem bacteriana GeG2 baseada em classificações COG. A porcentagem de proteínas preditas em cada réplicon classificada em cada categoria COG é apresentada.

Uma abrangente análise funcional de genomas segmentados, incluindo genomas representativos de mais de 1.700 espécies bacterianas, indicou que vieses funcionais entre diferentes réplicons genômicos não estão restritos a taxa bacterianos específicos, com cada classe de réplicon apresentando enriquecimentos funcionais específicos independentemente da filogenia (diCenzo e Finan, 2017). Além disso, os autores identificaram que categorias funcionais associadas a transporte de aminoácidos (E), carboidratos (G) e íons inorgânicos (P) são primariamente enriquecidas em cromídeos e não em megaplasmídeos (diCenzo e Finan, 2017). Assim, o fato desses grupos funcionais terem sido identificados em maior abundância no megaréplicon secundário de GeG2 também poderia corroborar sua classificação como um cromídeo.

Quanto aos plasmídeos identificados no genoma de GeG2, é interessante notar que a maior parte dos genes de pGeG2b atribuídos a COGs foram classificados na categoria funcional de transporte, secreção e trânsito intracelular de vesículas (U), enquanto que em pGeG2a, além de muitas proteínas preditas de função desconhecida (S), as classes funcionais predominantes foram aquelas associadas a transcrição (K), metabolismo e transporte de carboidratos (G) e de aminoácidos (E) (Tabela 12). Além

disso, uma porção considerável das proteínas de ambos os plasmídeos foram atribuídas a funções de replicação e reparo de DNA (L). O enriquecimento das categorias L e K já foi identificado em plasmídeos de diferentes espécies bacterianas, estando provavelmente relacionado a replicação e conjugação desses réplicons (Zheng *et al.*, 2015; diCenzo e Finan, 2017). Proteínas pertencentes a categoria U também são frequentemente identificadas como abundantes em plasmídeos bacterianos, tendo sido sugerido seu envolvimento em resistência a compostos tóxicos (Zheng *et al.*, 2015; diCenzo e Finan, 2017; Suzuki *et al.*, 2019). Considerando os indícios de produção de vesículas identificados por MEV em culturas de GeG2 discutidos acima (Figura 27), é tentador especular que proteínas dessa categoria funcional presentes especialmente no plasmídeo pGeG2b, bem como no cromossomo e cromídeo (Figura 35), poderiam estar relacionadas a esse fenótipo.

Tabela 12. Número de proteínas preditas nos plasmídeos pGeG2a e pGeG2b da linhagem GeG2 classificadas em cada categoria funcional COG.

Categoria COG	Descrição	pGeG2a		pGeG2b	
		Total ^a	% ^b	Total ^a	% ^b
(D)	Controle do ciclo celular, divisão celular	1	0,70	2	4,55
(M)	Biogênese de envelope celular	5	3,50	2	4,55
(N)	Motilidade celular	1	0,70	6	13,64
(O)	Modificações pós-traducionais, <i>turnover</i> protéico	6	4,20	1	2,27
(T)	Mecanismos de transdução de sinal	7	4,90	1	2,27
(U)	Transporte e secreção de vesículas	8	5,59	17	38,64
(V)	Mecanismos de defesa	1	0,70	0	0,00
(K)	Transcrição	24	16,78	5	11,36
(L)	Replicação, recombinação e reparo	13	9,09	6	13,64
(C)	Produção e conversão de energia	7	4,90	0	0,00
(E)	Metabolismo e transporte de aminoácidos	15	10,49	4	9,09
(F)	Metabolismo e transporte de nucleotídeos	2	1,40	0	0,00
(G)	Metabolismo e transporte de carboidratos	16	11,19	2	4,55
(H)	Metabolismo e transporte de coenzimas	3	2,10	0	0,00
(I)	Metabolismo e transporte de lipídeos	8	5,59	1	2,27
(P)	Metabolismo e transporte de íons inorgânicos	7	4,90	3	6,82
(Q)	Biossíntese e catabolismo de metabólitos secundários	8	5,59	0	0,00
(S)	Função desconhecida	37	25,87	6	13,64

^a Proteínas classificadas em mais de uma categoria foram contabilizadas em cada categoria

^b Em relação ao número total das proteínas do plasmídeo atribuídas a COG

O genoma da linhagem bacteriana GeG2 codifica um extenso repertório de proteínas relacionadas ao metabolismo de carboidratos. Anotações específicas de enzimas envolvidas na síntese, degradação e modificação de carboidratos complexos e glicoconjugados, coletivamente denominadas de CAZymes (*Carbohydrate-Active enZymes*) (Cantarel *et al.*, 2009), revelaram 41 glicosil transferases (GTs), 131 glicosil hidrolases (GHs), 11 carboidratos esterases (CEs), duas polissacarídeo liases (PLs) e sete enzimas para atividades auxiliares (AAs), totalizando 192 CAZymes identificadas no genoma de GeG2 (Figura 36). Além disso, 10 módulos de ligação a carboidratos (CBMs – *carbohydrate-binding modules*), proteínas não-catalíticas que se ligam aos carboidratos e os direcionam para a maquinaria catalítica de CAZymes às quais estão associadas, aumentando sua eficiência (Ficko-Blean e Boraston, 2012), foram identificados associados a GHs.

Apesar de serem enzimas comumente encontradas nesse grupo bacteriano, o número de CAZymes detectados no genoma do isolado GeG2 é bem maior do que o previamente observado em outros membros do gênero *Novosphingobium* e gêneros de *Sphingomonadaceae* relacionados, nos quais foram identificadas entre 24 e 118 enzimas ativas em carboidratos (Aylward *et al.*, 2013; D'Argenio *et al.*, 2014). Já análises do genoma da espécie mais próxima *N. rosa* NRBC15208 revelaram uma quantidade de CAZymes mais semelhante ao observado para GeG2, apesar de um número menor dessas enzimas também ter sido detectado nessa linhagem (116 GHs, 38 GTs, 6 CEs, 4 PLs e 3 AAs, totalizando 173 CAZymes). Considerando a grande produção de exopolissacarídeos observada em culturas destas bactérias em diferentes condições, bem como sua capacidade de formar biofilmes, o vasto repertório de CAZymes detectado nos genomas não é surpreendente, com muitas dessas enzimas provavelmente estando relacionadas a esse fenótipo. Além disso, genomas de muitos membros de *Sphingomonadaceae* possuem vias de degradação de biopolímeros variados, incluindo polissacarídeos complexos, importantes para o aproveitamento de fontes de carbono e energia nos ambientes oligotróficos a que muitas espécies estão adaptadas (Glaeser e Kämpfer, 2014), característica observada para o isolado GeG2.

Assim como previamente identificado pelas análises da distribuição de categorias funcionais COG entre os réplicons que compõe o genoma de GeG2 (Figura 35), grande parte das CAZymes estão codificadas no cromídeo (47 %), especialmente GHs e AAs (Figura 36), sugerindo um papel importante deste réplicon no metabolismo de diferentes carboidratos. Diferentemente do observado em outros genomas bacterianos, nos quais

PLs foram identificadas exclusivamente no cromídeo (Hernández *et al.*, 2018), ambas as PLs encontradas no genoma de GeG2 são codificadas no cromossomo. Vale ressaltar ainda que peptídeos sinais foram identificados em 75 % das sequências de GHs presentes no cromídeo (53 de 70), indicando que a maioria das enzimas relacionadas a carboidratos deste réplicon são potencialmente secretadas. Já dentre as 101 CAZymes anotadas no cromossomo, sequências de peptídeos sinais foram preditas em apenas 37, sendo em 30 GHs, 3 CEs, uma AA, além de em ambas as PLs identificadas no réplicon.

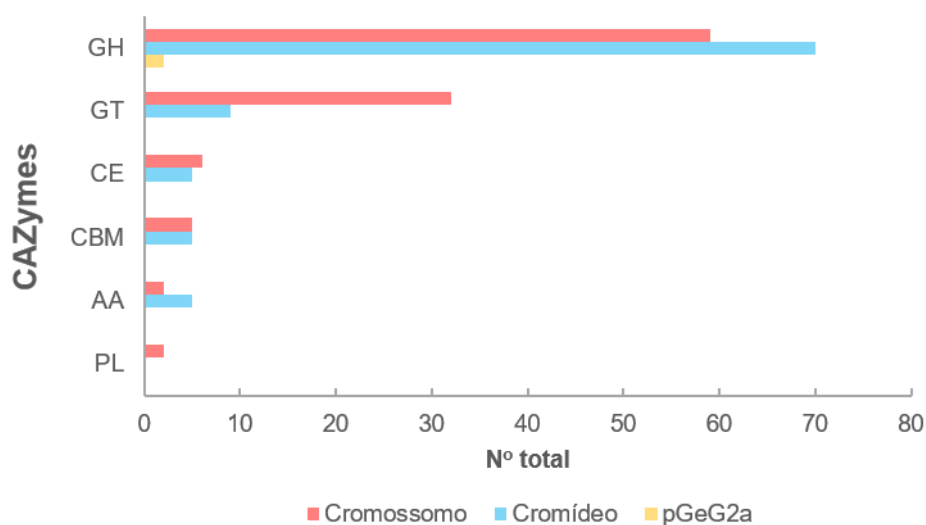


Figura 36. Distribuição de CAZymes (enzimas ativas em carboidratos) no cromossomo, cromídeo e plasmídeo pGeG2a da linhagem GeG2. Nenhuma CAZyme foi identificada no plasmídeo pGeG2b. GH: glicosil hidrolases, GT: glicosil transferases, CE: carboidrato esterases, CBM: módulo de ligação a carboidrato, AA: enzimas para as atividades auxiliares PL: polissacarídeo liases.

As diferentes CAZymes são classificadas em várias famílias funcionais, inferidas com base na comparação das sequências de aminoácido e similaridades estruturais com enzimas bioquimicamente caracterizadas, disponíveis no banco de dados CAZY (<http://www.cazy.org/>) (Cantarel *et al.*, 2009). Ao todo, 47 famílias de GHs, 10 de GTs, cinco de AAs, quatro de CEs e duas de PLs foram identificadas no genoma de GeG2.

GHs são enzimas responsáveis pela quebra de ligações glicosídicas e, juntamente com CEs, PLs e AAs, atuam na degradação de polissacarídeos, incluindo importantes componentes da biomassa vegetal, como celulose, hemicelulose e amido, desempenhando assim papéis cruciais na ciclagem do carbono nos ecossistemas terrestres (Nguyen *et al.*, 2018). Além disso, GHs possuem muitas aplicações biotecnológicas e industriais, sendo amplamente utilizadas, por exemplo, na produção de biocombustíveis, papel, tecidos,

insumos farmacêuticos, dentre muitos outros produtos (Singh *et al.*, 2016). Dentre as GHs presentes no genoma do isolado GeG2, 40 (23 codificadas no cromídeo e 17 no cromossomo) foram classificadas em famílias que contém enzimas relacionadas a degradação de xilanas e outros componentes de hemicelulose (GH3, GH5, GH8, GH10, GH16, GH30, GH43, GH51, GH67, GH115) (Nguyen *et al.*, 2018). Além disso, CEs envolvidas na remoção das cadeias laterais das xilanas e que possibilitam o acesso das xilanases à cadeia principal (CE1 e CE4) também foram identificadas tanto no cromídeo quanto no cromossomo.

Hidrolases atribuídas às famílias GH13 (12) e GH15 (2), que incluem endo e exoenzimas, respectivamente, que atuam na quebra de ligações α -glicosídicas, como α -amilases, pululanases, amiloglicosidases e glicoamilases (Møller e Svensson, 2016), foram detectadas exclusivamente no cromossomo. Seis enzimas de famílias relacionadas a quebra de componentes de peptideoglicano (GH23, GH24 e GH25) foram detectadas no cromossomo, enquanto apenas uma está presente no cromídeo. Por outro lado, GHs que tem oligossacarídeos como substrato (famílias GH1, GH28, GH29, GH31, GH35, GH36, GH39, GH42, GH92) encontram-se codificadas majoritariamente no cromídeo.

Cinco GHs das famílias GH106 e GH78, que agrupam α -L-ramnosidases, enzimas que catalisam resíduos terminais de α -L-ramnose em diversos compostos naturais (Manzanares *et al.*, 2007), também foram identificadas no genoma da bactéria GeG2, sendo três GH106 no cromídeo e uma GH78 (associada a CBM67, módulo ligante de ramnose) e uma GH106 no cromossomo. L-ramnose é um açúcar amplamente distribuído em plantas e microrganismos, constituindo polissacarídeos pécticos, flavonóides, antocianinas, triterpenóides, além de heteropolissacarídeos e ramnolipídeos bacterianos (Manzanares *et al.*, 2007). Por esse motivo, L-ramnosidases são enzimas que desempenham importantes funções biológicas e despertam grande interesse como biocatalisadores de processos industriais, principalmente na fabricação de bebidas e fármacos (Yadav *et al.*, 2010). Recentemente, uma α -L-ramnosidase identificada em extratos celulares de *Novosphingobium* sp. PP1Y, pertencente à família GH106 e codificada no megaplasmídeo dessa linhagem, foi isolada e caracterizada, sendo capaz de hidrolisar flavonoides glicosilados variados e apresentando aspectos bioquímicos interessantes para potenciais aplicações biotecnológicas (DeLise *et al.*, 2016).

Já as GTs, enzimas que catalisam a formação de ligações glicosídicas e estão relacionadas à biossíntese de oligossacarídeos, polissacarídeos e glicoconjugados, estão majoritariamente codificadas no cromossomo de GeG2 (Figura 36). Assim como

observado para outras bactérias produtoras de EPS (Deo *et al.*, 2019; Wang *et al.*, 2019), as maiores famílias de GTs identificadas no genoma de GeG2 foram GT2 e GT4, contendo 13 e 18 sequências, respectivamente. Essas famílias de GT são abrangentes, incluindo enzimas de diversas origens e funções, como celulose sintases, quitina sintases, sacarose sintases, manosiltransferases, glicosiltransferases, ramnosiltransferases, dentre outras (Breton *et al.*, 2006). Uma enzima de cada uma das famílias GT5, GT20, GT28, GT32, GT35 e GT39, que incluem majoritariamente amido fosforilases e glicosiltransferases, α -manosiltransferases, trealose-6-fosfato fosfatases e glicano sintases, também foram detectadas no cromossomo do isolado GeG2, além de três mureínas polimerases (GT51).

Predições de perfis metabólicos realizadas com as ferramentas relacionadas ao banco de dados KEGG revelaram que, conforme anteriormente verificado nas análises de COG (Figura 35) e observado em outros genomas bacterianos segmentados (Dziewit *et al.*, 2014; diCenzo e Finan, 2017), todos os genes associados ao processamento de informação genética, incluindo as vias de biossíntese de aminoacil-tRNAs, exportação de proteínas, degradação de RNA, replicação de DNA, reparo por excisão de bases, reparo de pareamentos errados (*mismatch repair*) e recombinação homóloga, foram identificados no cromossomo. As únicas exceções foram os genes *dnaQ* e *dnaE*, codificadores das subunidades ϵ e α da DNA polimerase III, respectivamente, que estão presentes tanto no cromossomo quanto no cromídeo e o gene *mug* (uracila DNA glicosilase), detectado exclusivamente no cromídeo. Além disso, todos os genes necessários para a biossíntese de peptideoglicano, montagem do flagelo e quimiotaxia também se encontram somente no cromossomo.

Assim como identificado em outros 22 genomas de *Novosphingobium* (Wang *et al.*, 2018), o genoma da linhagem GeG2 contém as vias completas de glicólise, pentose fosfato e ciclo do ácido tricarboxílico. Curiosamente, apesar da maioria dos genes envolvidos no metabolismo central de carbono estarem localizados no cromossomo de GeG2, como observado para outras *Novosphingobium* (Wang *et al.*, 2018), muitos são encontrados tanto no cromossomo quanto no cromídeo da bactéria, principalmente os envolvidos na via da pentose fosfato e na oxidação de piruvato (Figura 37). Genes relacionados ao metabolismo do nitrogênio são encontrados exclusivamente no cromossomo, incluindo transportadores de nitrito e nitrato extracelular, nitrito redutases (*nirB/nirD*), nitronato monooxigenase (*ncd2*), assim como transportadores de amônia da família Amt.

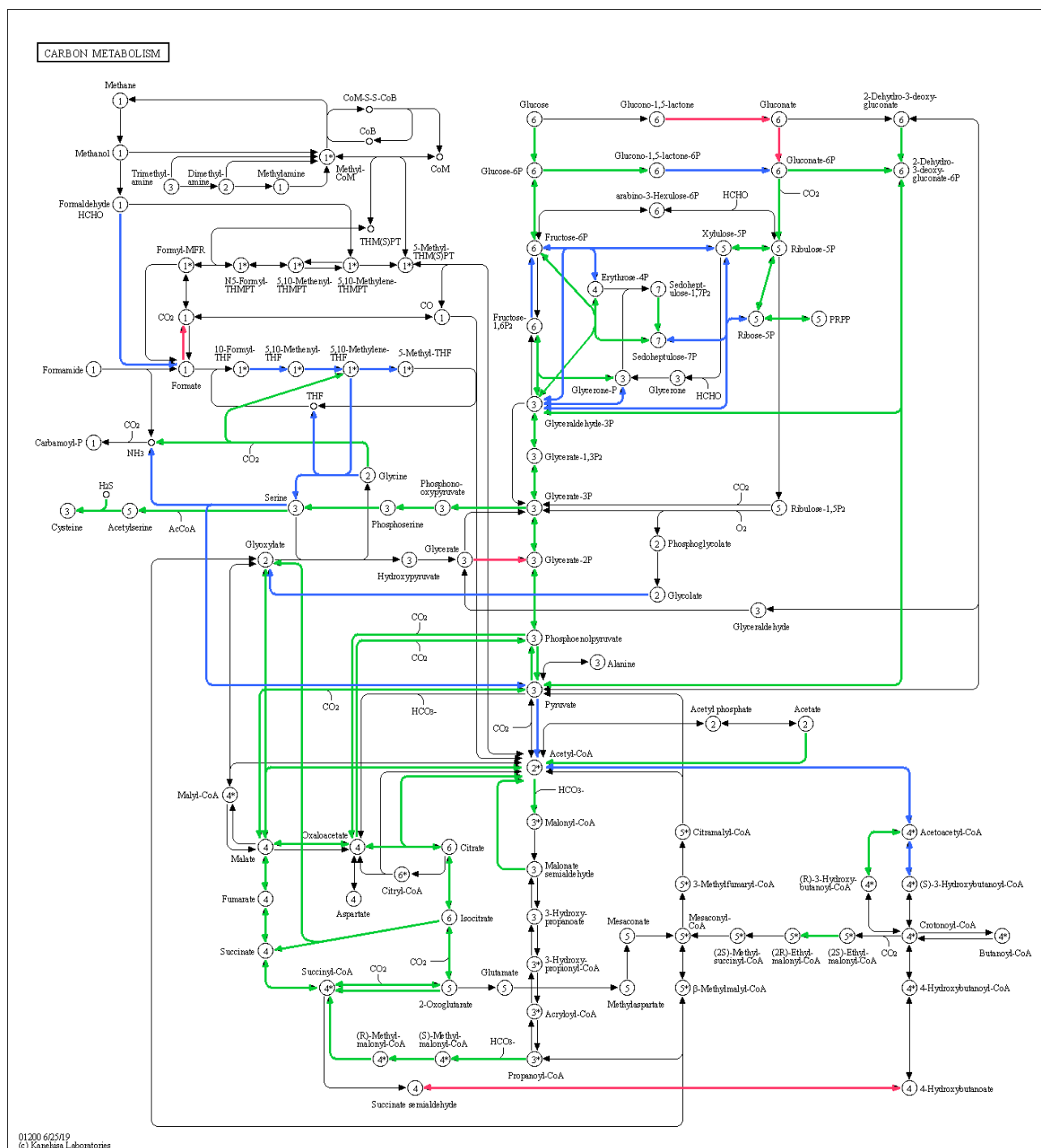


Figura 37. Mapa das vias relacionadas ao metabolismo central de carbono, gerado pela ferramenta *KEGG Mapper - Pathways*, indicando genes de enzimas presentes no cromossomo (em verde), no cromídeo (em rosa) ou em ambos os replicons (em azul) do genoma da linhagem bacteriana GeG2.

Existem três vias de assimilação de enxofre descritas para bactérias: 1) assimilação de enxofre inorgânico (sulfatos e tiosulfatos) por meio do transportador codificado pelo sistema *cysPAUW*, seguido pela ativação e utilização do sulfato pelas enzimas *CysD/CysNC* e pela biossíntese da cisteína pelas enzimas *CysK, CysE* e *CysQ*; 2) assimilação de sulfonatos ambientais por meio do transportador codificado pelo

sistema *ssuABC* e posterior conversão a sulfito pela monooxigenase *SsuD*; 3) transporte de taurina pelo transportador *tauABC* e metabolização pela taurina dioxigenase (*TauD*) (Kumar *et al.*, 2017) (Figura 38). No genoma do isolado GeG2, os únicos genes de assimilação de enxofre extracelular identificados são o do transportador *ssuABC*, juntamente com os genes *ssuD* (alcanosulfonato monooxigenase) e *ssuE/sfnG* (dimetilsulfona monooxigenase), todos presentes exclusivamente no cromídeo (Figura 38). É interessante notar que, em um abrangente trabalho de genômica comparativa entre 27 genomas de *Novosphingobium* oriundos de diferentes habitats, a via completa de assimilação de alcanosulfonatos (genes *ssu*) foi identificada exclusivamente em isolados de solo (Kumar *et al.*, 2017), mesmo hábitat do qual a linhagem GeG2 foi isolada, sugerindo uma especialização na assimilação desses compostos em espécies associadas a ambientes terrestres. Curiosamente, apesar dos genes para o sistema transportador *cysPAUW* não terem sido identificados no genoma de GeG2, os genes *cysNC/cysD*, responsáveis pela conversão de sulfato a fosfoadenosina fosfosulfato (PAPS) e *cysJ/cysI* e *cysK/cysO*, que realizam a conversão a sulfito a sulfeto e de sulfeto a L-cisteína, respectivamente, foram identificados exclusivamente no cromossomo, indicando um papel complementar entre os réplicons da linhagem GeG2 na assimilação e metabolismo de enxofre (Figura 38). Além disso, embora genes *tauD* tenham sido detectados em todos os 27 genomas de *Novosphingobium* analisados por Kumar e colaboradores (2017), nenhum gene associado à via de metabolismo da taurina pôde ser identificado no genoma de GeG2.

Complementariedades entre o cromossomo e o cromídeo da linhagem GeG2 também foram identificadas em relação a outras vias metabólicas. Com exceção do gene *hisC*, que codifica a histidinol-P aminotransferase e foi detectado em ambos os réplicons, os genes relacionados à biossíntese de histidina (*hisG*, *hisE*, *hisJ*, *hisA*, *hisF*, *hisB*, *hisN*, *hisD*) encontram-se exclusivamente no cromossomo, enquanto os genes relacionados à degradação da histidina em glutamato (*hutH*, *hutU*, *hutI*, *hutF* e *hutG*) são codificados apenas no cromídeo. Além disso, enquanto todos os genes das vias de biossíntese *de novo* de purinas estão presentes no cromossomo, incluindo os envolvidos na síntese de inosina monofosfato (IMP) e ribonucleotídeos adenina e guanina (ADP, ATP, GDP, GTP), os genes adenina deaminase (*ade*), adenosina deaminase (*add*), guanina deaminase (*guaD*), purina nucleosidase (*iunH*), urato oxidase (*hpxO*), 5-hidroxi-isourato hidrolase (*uraH*), entre outros envolvidos na degradação e reaproveitamento das bases púricas são codificados exclusivamente no cromídeo.

O isolado GeG2 apresentou resistência a um grande número de antibióticos em análises de disco-difusão (seção 2.3.7). Além de dois genes de β -lactamases, preditos como sendo das classes A e B3, associadas à hidrólise de penicilinas/cefalosporinas e carbapenêmicos respectivamente (Bush e Jacoby, 2010), diversos genes codificadores de proteínas que poderiam estar associadas à resistência a antimicrobianos foram detectados no genoma de GeG2. Ao todo, 63 genes relacionados a mecanismos de resistência a múltiplas drogas foram anotados no cromossomo, no cromídeo e em pGeG2a, incluindo genes de bombas de efluxo das famílias RND (*Resistance, Nodulation, Division*) – *mdtABC*, *mexAB*, *acrB*, MATE (*Multidrug and toxic compounds extrusion*) – *norM* e MFS (*Major Facilitator Super-family*) – *emrB*, *emrKY* (Moreira *et al.*, 2004). Esses transportadores podem estar relacionados ao efluxo de uma ampla variedade de antibióticos, além de metais pesados, detergentes, corantes e solventes simples, participando de respostas a estresses e conferindo resistência a diversos agentes antimicrobianos (Li *et al.*, 2015). Embora análises mais aprofundadas sejam necessárias para a elucidação da natureza das bombas de efluxo encontradas, bem como o tipo de moléculas às quais conferem resistência, os genes detectados podem estar relacionados ao fenótipo observado no isolado GeG2 de resistência aos vários antibióticos testados.

Outro aspecto interessante do genoma da linhagem bacteriana GeG2 está relacionado aos sistemas de secreção identificados. O transporte de moléculas do citoplasma para o ambiente ou para o interior de outras células é uma função essencial de células procarióticas, mediada por estruturas proteicas denominadas sistemas de secreção (Green e Mecsas, 2016). Nove tipos de sistemas de secreção já foram identificados até o momento, denominados tipos I a IX (T1SS a T9SS), sendo classificados de acordo com seus componentes, o número de membranas que atravessam (bactérias Gram-positivas ou negativas e se penetram ou não membranas de outras células) e o tipo de moléculas que transportam (Green e Mecsas, 2016, Lasica *et al.*, 2017). Além disso, vias de secreção amplamente conservadas em todos os domínios da vida, conhecidas como Sec e Tat (translocação de arginina dupla), realizam o transporte de proteínas do citoplasma para outros ambientes celulares, incluindo o periplasma e membrana interna, no caso de bactérias Gram-negativas (Green e Mecsas, 2016). Dentre os sistemas de secreção que podem ser encontrados em Gram-negativas (tipos I a VI, VIII e IX), já foram detectados em membros da família *Sphingomonadaceae* os sistemas T1SS, T2SS, T3SS, T4SS e T6SS, além de vias de secreção Sec e Tat (Wu *et al.*, 2017, Sheu *et al.*, 2020). No genoma do isolado GeG2, genes associados ao sistema de secreção tipo IV (T4SS) e com as vias

de Tat e Sec foram detectados no cromossomo, enquanto os sistemas de secreção T1SS e T6SS estão codificados apenas no cromídeo bacteriano.

O T4SS consiste em um mecanismo especializado de secreção de macromoléculas relacionado ancestralmente aos sistemas de conjugação bacterianos, apresentando grande versatilidade, sendo capaz de mediar o transporte tanto de proteínas monoméricas quanto de complexos de proteínas e de nucleoproteínas (proteínas associadas a DNA) (Walden *et al.*, 2010). Apesar de usualmente associados a plasmídeos, genes de T4SS também foram identificados em cromossomos de diversas espécies bacterianas (Alegria *et al.*, 2005; Hamilton *et al.*, 2005; Walden *et al.*, 2010; Fischer *et al.*, 2020), inclusive em membros do gênero relacionado *Sphingomonas* (Wu *et al.*, 2017), tendo sido associado principalmente a transferências horizontais de DNA e secreção de mediadores de interações bacterianas cooperativas e antagonistas (Alegria *et al.*, 2005; Hamilton *et al.*, 2005).

Já o T1SS é um sistema de secreção amplamente distribuído entre bactérias Gram-negativas, com estrutura relativamente simples, composta de três subunidades proteicas: um transportador ABC (HlyB) localizado na membrana interna, uma proteína periplasmática de fusão de membrana (MPF, denominada HlyD) e uma porina na membrana externa (TolC) (Morgan *et al.*, 2017). Responsável pela secreção de diversas hidrolases e toxinas protéicas, esse sistema é comumente identificado em bactérias da família *Sphingomonadaceae*, muitas vezes relacionado a mecanismos de efluxo, conferindo resistência a diversos tipos de compostos (Wu *et al.*, 2017). Vale ressaltar que, além de compor sistemas T1SS, a porina TolC pode estar envolvida em diversos outros processos celulares, especialmente no efluxo de antibióticos (Iyer *et al.*, 2018) e, além de estar presente no cromídeo de GeG2 juntamente com os demais componentes do T1SS, um gene desta porina foi também detectado no cromossomo.

O T6SS, inicialmente identificado como um clássico fator de virulência responsável pela injeção de proteínas efetoras em células eucarióticas (Pukatzki *et al.*, 2006), é atualmente reconhecido como um sistema de secreção bastante versátil, com funções relacionadas à absorção de íons escassos, mediação de comportamentos “sociais”, aquisição de material genético por transferência horizontal, bem como a competição entre bactérias por meio da injeção de toxinas antibacterianas no periplasma de células rivais (Coulthurst, 2019). É um complexo multiprotéico, análogo à cauda contrátil de fagos, ubíquo em proteobacteria, presente no genoma de muitas espécies provenientes de solos (Bingle *et al.*, 2008), sugerindo um importante papel fisiológico e

ecológico nesses ambientes. O fato desse sistema de secreção, bem como do sistema do tipo I, ter sido identificado exclusivamente no cromídeo pode corroborar seu papel adaptativo, podendo estar associado, assim como outros genes detectados nesse réplicon, a respostas a variações ambientais e interações com outros organismos no solo.

Análises de agrupamentos gênicos envolvidos na biossíntese de metabólitos secundários (*Biosynthetic Gene Clusters* - BGCs) realizadas com a ferramenta antiSMASH (Blin *et al.*, 2019) revelaram oito possíveis BGCs no genoma da linhagem GeG2, sendo quatro localizados no cromossomo e quatro no cromídeo (Figura 39A). Classificações baseadas na homologia de genes centrais da biossíntese de metabólitos secundários indicaram que o cromossomo possui dois BGCs, com genes relacionados à síntese de terpenos, um envolvido na produção de lassopeptídeos/homoserina-lactona e um semelhante ao que contém sintetases de peptídeos não ribossomais (NRPS-*like* - *Non-ribosomal peptide synthetase like*) (Figura 39A). Já o cromídeo apresenta um BGC contendo policetídeo-sintases do tipo 1 (T1PKS), um com genes de síntese de bacteriocinas, um de lasso peptídeos/homoserina-lactona, além de um BCG híbrido contendo sintetases de peptídeos não ribossomais e policetídeo-sintases do tipo 1 (NRPS/T1PKS) (Figura 39A).

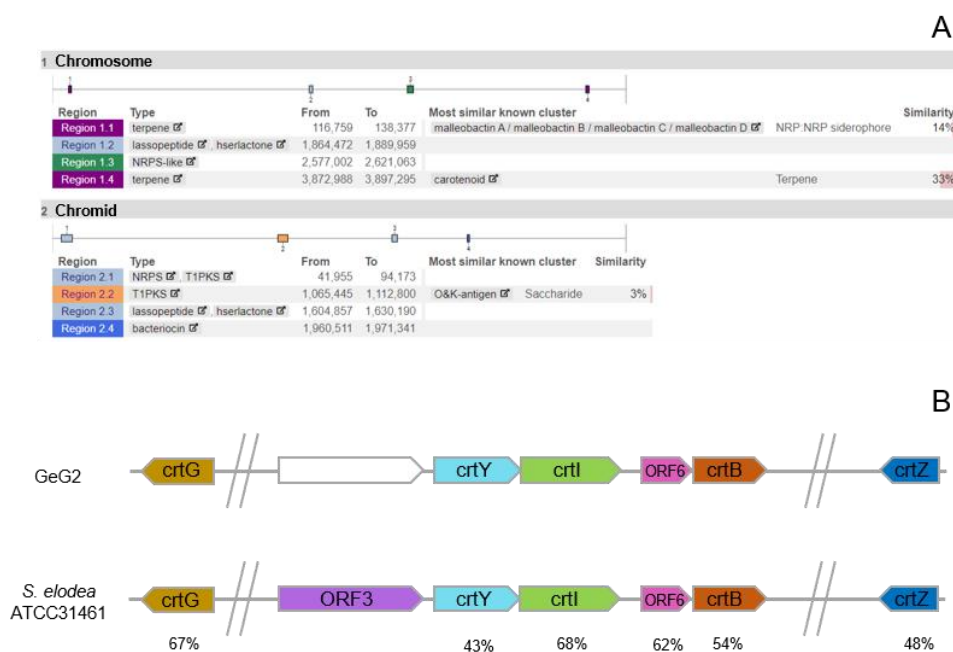


Figura 39. Clusters gênicos de biossíntese de metabólitos secundários (BGCs) encontrados no genoma da linhagem bacteriana GeG2. Em A, estão apresentados os resultados obtidos com a ferramenta antiSMASH (Blin *et al.*, 2019), indicando a identificação de quatro BGCs no cromossomo e quatro no cromídeo. Em B, comparação do *cluster* gênico de biossíntese de carotenóide (1.4) identificado no genoma de GeG2 com o *cluster* gênico de biossíntese de nostoxantina de *Sphingomonas elodea* ATCC31461 (Zhu *et al.*, 2012). As identidades (%) observadas entre cada um dos genes que compõe os *clusters* de GeG2 e *S. elodea* ATCC31461 estão indicadas.

Bactérias de solo usualmente produzem metabólitos secundários que desempenham diversas funções ecológicas importantes, incluindo a aquisição de nutrientes e a mediação de comunicação, competição e outras interações com diferentes microrganismos e com o ambiente (Sharrar *et al.*, 2020). Além disso, algumas dessas moléculas representam fontes importantes de produtos naturais de interesse médico e biotecnológico, principalmente antibióticos (Cragg e Newman, 2013).

Apesar da maioria dos BCGs identificados no genoma da linhagem GeG2 não terem apresentado similaridade com qualquer *cluster* conhecido (Figura 39A), impossibilitando a pedição de seus produtos específicos, a identificação de BCGs contendo NRPSs e PKSs no cromídeo de GeG2 é interessante, já que produtos dessas enzimas possuem uma ampla gama de atividades biológicas e propriedades farmacológicas, incluindo atividades antibacterianas, antifúngicas, de sideróforos e imunossupressoras (Wang *et al.*, 2014). Já as bacteriocinas, antimicrobianos peptídicos produzidos por bactérias, são geralmente ativas contra espécies próximas e provavelmente atuam na redução da competição em um mesmo nicho (Tyc *et al.*, 2017). Essas três classes de BCGs, juntamente com os relacionados à síntese de terpenos, foram as mais comumente encontradas em uma abrangente análise de metagenomas de solo, sugerindo o importante papel ecológico desses compostos nesses habitats terrestres e evidenciando o potencial ainda inexplorado de produção de valiosas moléculas de interesse biotecnológico por microrganismos de solo (Sharrar *et al.*, 2020).

Análises mais específicas do BCG 1.4 encontrado no cromossomo da linhagem GeG2, que apresentou maior similaridade com BCGs de síntese de carotenoides (Figura 39A), revelaram homologia com os genes de síntese de nostoxantina caracterizados na bactéria *Sphingomonas elodea* ATCC31461 (Zhu *et al.*, 2012) (Figura 39B). Nostoxantina é um derivado poli-hidroxi de β -carotenos, sendo o principal carotenoide encontrado em membros de *Sphingomonadacea* (Glaeser e Kämpfer, 2014). A estrutura do BCG identificado linhagem GeG2 é bastante semelhante ao BCG identificado como produtor de nostoxantina em *S. elodea* (Zhu *et al.*, 2012), contendo os quatro genes carotenogênicos *crtG*, *crtY*, *crtI* e *crtB*, além de um gene de β -caroteno hidroxilase (*crtZ*) localizado fora do *cluster* (Figura 39B). Esses resultados sugerem que, apesar de identificações químicas dos carotenoides presentes nas células da linhagem GeG2 ainda não terem sido realizadas, esta bactéria provavelmente produz nostoxantina, assim como outras *Novosphingobium* (Glaeser e Kämpfer, 2014).

2.3.13. Potencial do isolado GeG2 na degradação de compostos aromáticos

Bactérias do gênero *Novosphingobium* são conhecidas por seu potencial de degradação de diversos tipos de compostos aromáticos recalcitrantes e, como observado nos genomas de espécies deste e de outros gêneros relacionados (Aylward *et al.*, 2013; Wang *et al.*, 2018), um grande número de monooxigenases e dioxigenases foram detectadas no genoma da linhagem GeG2 (Tabela 13). Essas enzimas catalisam o passo crítico de clivagem do anel na degradação de compostos aromáticos em condições de aerobiose, estando envolvidas na metabolização de substâncias como benzeno, tolueno, naftaleno, fluoreno, fenantreno, dentre muitas outras (Ladino-Orjuela *et al.*, 2016).

Um total de 31 genes de monooxigenases e 30 de dioxigenases no genoma de GeG2 (Tabela 13) foram encontrados, um número semelhante ao observado no genoma de espécies de *Novosphingobium* capazes de degradar compostos aromáticos como hexaclorocicloexano, pentaclorofenol, bifenilo, fenantreno, pireno, benzo(a)pireno, naftaleno, fluoreno, antraceno, entre outros (Sohn *et al.*, 2004; Tirola *et al.*, 2005; Notomista *et al.*, 2011; Aylward *et al.*, 2013; Saxena *et al.*, 2013; Wang *et al.*, 2018). Ao contrário do relatado para *N. aromaticivorans* DSM12444, *Novosphingobium* sp. P6W, *Novosphingobium* sp. PP1Y, *N. pentaromativorans* US6-1 e *Novosphingobium* sp. THN1, em que a maioria das mono e dioxigenases encontram-se codificadas nos cromossomos (Wang *et al.*, 2018), a maior parte das monooxigenases da linhagem GeG2 foi detectada no cromídeo, enquanto um número semelhante de dioxigenases foi identificado no cromossomo e no cromídeo (Tabela 13).

Análises de vias de degradação de compostos aromáticos realizadas pelo KEGG revelaram que, apesar de não terem sido identificadas as enzimas tolueno monooxigenase (*tmoA*) e tolueno metil-monooxigenase (*xylM*), usualmente envolvidas nas primeiras etapas da degradação bacteriana de tolueno e isômeros de xileno, outras enzimas relacionadas à biodegradação desses compostos foram identificadas no genoma da linhagem GeG2 (Figura 40). Vale ressaltar que, enquanto enzimas da via superior de degradação de compostos aromáticos a intermediários de catecol foram identificadas tanto no cromossomo quanto no cromídeo, todos os genes relacionados a orto-clivagem de catecóis a ácidos tricarbóxicos foram identificadas no cromossomo (Figura 40, Tabela 13). Além disso, apesar da via completa não ter sido detectada pela ferramenta, algumas enzimas responsáveis pela meta-clivagem de catecóis foram identificadas no cromídeo (Figura 40).

Tabela 13. Mono e dioxigenases encontradas no genoma da linhagem GeG2. Enzimas também encontradas no genoma de espécies de *Novosphingobium* capazes de degradar diferentes compostos aromáticos (Aylward *et al.*, 2013; Wang *et al.*, 2018) estão indicados com *.

<u>Monooxigenases</u>			
KO	Nome	Cromossomo	Cromídeo
K00459*	NMO; nitronate monooxygenase [EC 1.13.12.16]	2	0
K00481*	pobA; 4-hydroxybenzoate 3-monooxygenase [EC 1.14.13.2]	1	0
K03380	phenol 2-monooxygenase (NADPH) [EC 1.14.13.7]	0	1
K03862*	vanA; vanillate monooxygenase [EC:1.14.13.82]	0	8
K14974	nicC; 6-hydroxynicotinate 3-monooxygenase [EC:1.14.13.114]	0	1
K22027	iacA; indole-3-acetate monooxygenase [EC:1.14.13.235]	0	1
K04091*	ssuD; alkanesulfonate monooxygenase [EC:1.14.14.5]	0	2
K16047	hsaA; 3-hydroxy-9,10-secoandrosta-1,3,5(10)-triene-9,17-dione monooxygenase [EC:1.14.14.12]	0	4
K00500	phhA; phenylalanine 4-monooxygenase [EC:1.14.16.1]	0	1
K22553	4-methoxybenzoate monooxygenase (O-demethylating) [EC:1.14.99.15]	0	2
K20942	graA; resorcinol 4-hydroxylase (FADH2) [EC:1.14.14.27]	0	1
K05712	mhpA; 3-(3-hydroxy-phenyl)propionate hydroxylase [EC:1.14.13.127]	1	1
K18242	nagG; salicylate 5-hydroxylase large subunit [EC:1.14.13.172]	0	1
K18243	nagH; salicylate 5-hydroxylase small subunit [EC:1.14.13.172]	0	1
K03185	ubiH; 2-octaprenyl-6-methoxyphenol hydroxylase [EC:1.14.13.-]	1	0
K23464	tgnB; flavin-dependent trigonelline monooxygenase [EC:1.14.14.-]	0	1
K06134*	COQ7; 3-demethoxyubiquinol 3-hydroxylase [EC:1.14.99.60]	1	0
Total		6	25
<u>Dioxigenases</u>			
KO	Nome	Cromossomo	Cromídeo
K05549*	benA-xylX; benzoate/toluato 1,2-dioxygenase subunit alpha [EC:1.14.12.10 1.14.12.-]	1	1
K05550*	benB-xilY; benzoate/toluato 1,2-dioxygenase subunit beta [EC:1.14.12.10 1.14.12.-]	1	0
K03381*	catA; catechol 1,2-dioxygenase [EC:1.13.11.1]	1	0
K07104	catE; catechol 2,3-dioxygenase [EC:1.13.11.2]	0	1
K00450*	gtDA; gentisate 1,2-dioxygenase [EC:1.13.11.4]	0	2
K00451*	hmgA; homogentisate 1,2-dioxygenase [EC:1.13.11.5]	0	2
K04100	ligA; protocatechuate 4,5-dioxygenase, alpha chain [EC:1.13.11.8]	2	1
K04101	ligB; protocatechuate 4,5-dioxygenase, beta chain [EC:1.13.11.8]	2	1
K06911	PIR; quercetin 2,3-dioxygenase [EC:1.13.11.24]	3	2
K00457*	hpdD; 4-hydroxyphenylpyruvate dioxygenase [EC:1.13.11.27]	1	2
K08967*	mtnD/Z; acireductone dioxygenase (Ni ²⁺ -requiring) [EC:1.13.11.53]	1	0
K08967*	mtnD/Z ;acireductone dioxygenase [iron(II)-requiring][EC:1.13.11.54]	1	0
K21822	8'-apo-carotenoid 13,14-cleaving dioxygenase [EC:1.13.11.82]	1	0
K00472	P4HA; procollagen-proline 4-dioxygenase	1	0
K18068	pht3; phthalate 4,5-dioxygenase [EC:1.14.12.7]	0	1
K05916*	hmp; nitric oxide dioxygenase [EC:1.14.12.17]	1	0
K15777*	DOPA; 4,5-DOPA dioxygenase extradiol [EC:1.13.11.-]	1	0
Total		17	13

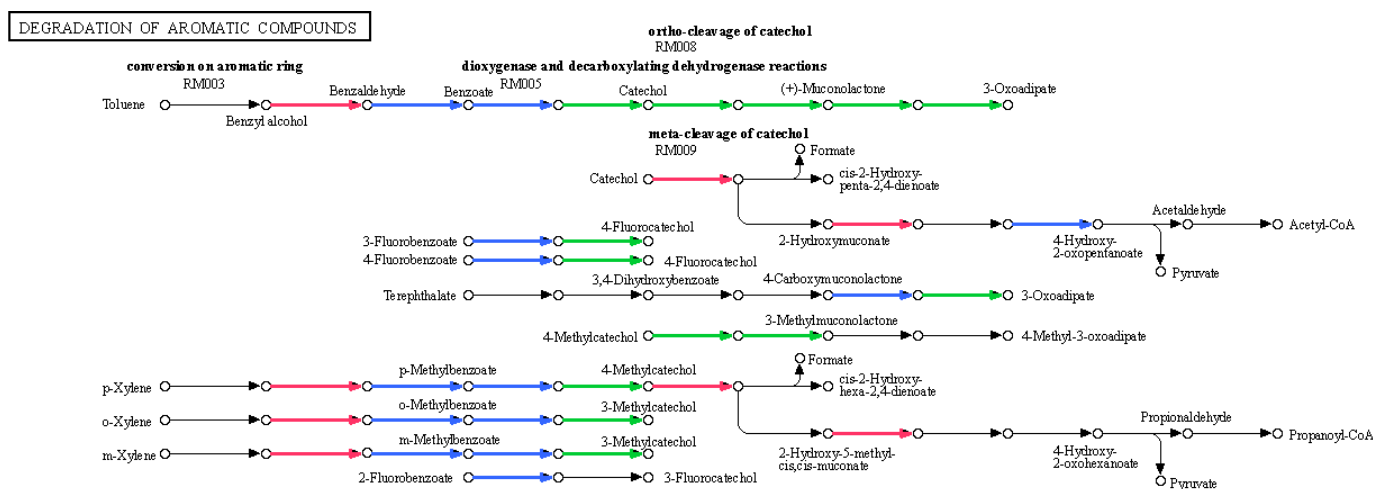


Figura 40. Mapa das vias relacionadas a degradação de compostos aromáticos, gerado pela ferramenta *KEGG Mapper - Pathways*, indicando genes de enzimas presentes no cromossomo (em verde), no cromídeo (em rosa) ou em ambos os réplicons (em azul) do genoma da linhagem bacteriana GeG2.

A degradação aeróbica de compostos aromáticos normalmente envolve a sua ativação e oxidação a alguns poucos intermediários centrais, que podem ser catecóis (dihidrodióis) ou ácidos carboxílicos aromáticos hidróxi-substituídos (Ladino-Orjuela *et al.*, 2016). Catecol é o intermediário usual na degradação de um grande número de compostos contaminantes, incluindo por exemplo benzoato, benzeno, fenol, naftaleno e salicilato, que podem então ser convertidos a ácidos tricarboxílicos por clivagem intradiol (orto) pela catecol 1,2-dioxygenase ou por clivagem extradiol (meta) pela catecol 3,4-dioxygenase (Cámara *et al.*, 2007), ambas encontradas no genoma de GeG2 (Tabela 13). Os mecanismos genéticos envolvidos na biodegradação de muitos compostos aromáticos recalcitrantes ainda são pouco elucidados, fazendo com que a identificação por meio de dados genômicos de substratos potencialmente degradados e das vias completas relacionadas seja tarefa difícil. No entanto, a detecção de enzimas chaves usualmente relacionadas a vias clássicas de degradação aeróbica de inúmeros compostos aromáticos no genoma da linhagem GeG2 certamente sugere o potencial dessa bactéria para esses processos. Análises experimentais de cultivos em meios contendo diferentes compostos aromáticos como fonte de carbono poderiam auxiliar na elucidação do potencial degradativo do isolado GeG2.

Análises preliminares sugeriram que a bactéria GeG2 é capaz de crescer em meio mínimo contendo como única fonte de carbono o composto verde malaquita, um corante aromático do tipo trifenilmetano (Figura 41). Diferentes bactérias contendo um amplo repertório de enzimas relacionadas a degradação de compostos aromáticos e capazes de

degradar esse importante poluente ambiental já foram caracterizadas (Wang *et al.*, 2012; Wanyonyi *et al.*, 2017), incluindo espécies do gênero *Sphingomonas* (Ayed *et al.*, 2009). Genes usualmente encontrados em genomas de bactérias capazes de degradar esse e outros corantes que mimetizam fragmentos de lignina, incluindo glutathione S-transferases (EC 2.5.1.18), glutathione peroxidases (EC 1.11.1.9), maleilacetoacetato isomerases (EC 5.2.1.2), além de uma DyP (*dye-decolorizing peroxidase*) (Melo-Nascimento *et al.*, 2018), foram detectados na anotação do genoma de GeG2.

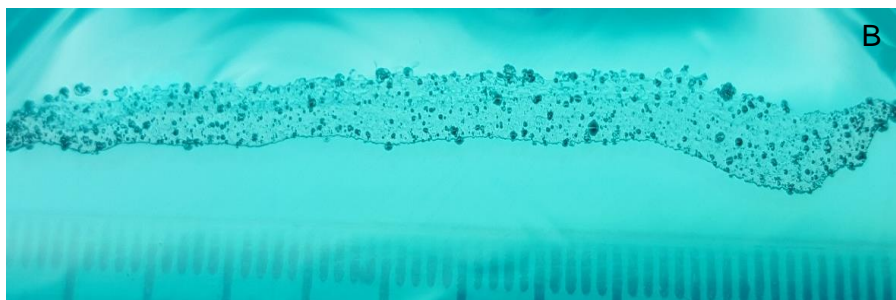
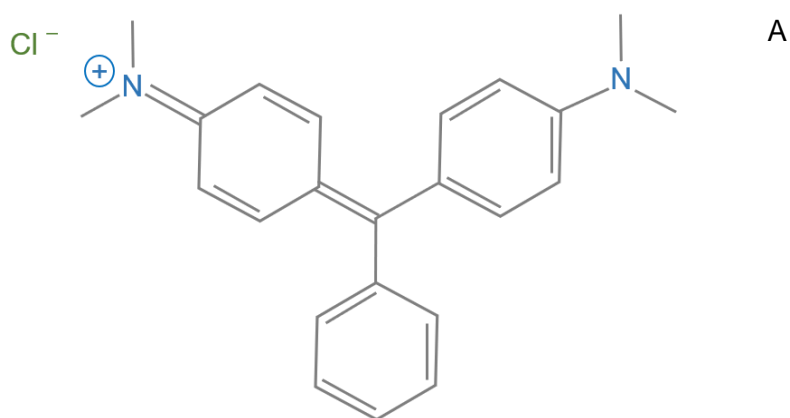


Figura 41. Estrutura química do corante verde malaquita (A) e colônias da bactéria GeG2 crescidas por sete dias em MM sólido sem glicose contendo verde malaquita (B).

No entanto, é importante ressaltar que estudos mais aprofundados, contendo controles e visando avaliar a descoloração do corante associado ao crescimento bacteriano em meios líquidos e sólidos, são necessários para determinar se a bactéria GeG2 é realmente capaz de degradar verde malaquita e outras moléculas relacionadas. Considerando o caráter bastante oligotrófico da bactéria GeG2, o crescimento observado em MM sólido acrescido de verde malaquita poderia estar sendo sustentado por impurezas advindas do ágar utilizado no meio, que poderiam ser empregadas como fonte de carbono pela bactéria, e a coloração observada nas colônias poderia ser resultante da

adsorção do corante pelas células e não necessariamente da sua degradação (Bandounas *et al.*, 2011).

4.3.14. Proposta da espécie *Novosphingobium terrae* sp. nov.

Com base nas análises fisiológicas, moleculares e quimiotaxonômicas, além de comparações de genomas completos, a linhagem GeG2, isolada do solo nativo de Cerrado, pode ser considerada uma nova espécie do gênero *Novosphingobium*, para a qual o nome *Novosphingobium terrae* sp. nov. será proposto.

Novosphingobium terrae (ter'rae. L. gen. n. *terrae* do solo, da terra, substrato do qual a linhagem tipo foi isolada).

CONCLUSÕES E PERSPECTIVAS

O cultivo de microrganismos ainda é imprescindível para a caracterização de seus diversos aspectos morfológicos, fisiológicos, bioquímicos e até genômicos. Neste trabalho, co-cultivos entre archaeas afiliadas aos filos Thaumarchaeota e Bathyarchaeota e uma bactéria pertencente ao gênero *Novosphingobium* foram estabelecidos a partir de solos nativos do bioma Cerrado. As características associadas a microrganismos do domínio *Archaea* presentes em solos e outros ambientes mesófilos continuam amplamente desconhecidas e, apesar do emprego de estratégias seletivas que resultaram na obtenção e detecção desses organismos em co-culturas com bactérias ao longo de alguns anos, o seu enriquecimento e isolamento não foi possível, com o crescimento bacteriano se sobressaindo.

A bactéria presente nos cultivos oriundos de solo de Cerrado foi então transferida e isolada em meios artificiais definidos. Análises filogenéticas baseadas no gene de rRNA 16S revelaram a afiliação da linhagem, denominada GeG2, ao gênero *Novosphingobium*, tendo como espécie descrita mais próxima *Novosphingobium rosa*, que foi então utilizada como referência na caracterização polifásica realizada. Comparações bioquímicas, moleculares e quimiotaxonômicas entre o novo isolado e a espécie referência indicaram diferenças fisiológicas e nos perfis de lipídeos ácidos graxos, além de variações no perfil proteico avaliado por MALDI-TOF. Ademais, análises taxonômicas comparativas baseadas em genomas completos revelaram que a linhagem GeG2 não pertence a qualquer espécie atualmente disponível em bancos de dados, incluindo *N. rosa*. Desta forma, os resultados indicam que o isolado de solo de Cerrado obtido neste trabalho pertence a uma nova espécie do gênero *Novosphingobium*, para qual o nome *Novosphingobium terrae* será proposto.

Cultivos prolongados em meio mínimo revelaram que a nova espécie bacteriana apresenta dimorfismo planctônico/séssil, com células imóveis formando agregados celulares macroscópicos envoltos por espessa matriz celular, além de um menor número de células livres móveis. Além disso, indícios da produção de vesículas extracelulares, projeções celulares com morfologia variada e grânulos de polifosfato pela bactéria foram também observados.

O sequenciamento genômico total por meio de duas abordagens possibilitou a montagem do genoma completo da linhagem GeG2, revelando uma estrutura segmentada,

contendo um cromossomo, dois plasmídeos e um megaréplicon identificado como um provável cromídeo. Vieses funcionais foram observados entre o cromossomo e o cromídeo e avaliações do conteúdo gênico presente em cada um dos réplicons sugerem um papel adaptativo do cromídeo, com genes que poderiam estar relacionados a respostas celulares a variações nutricionais e outros estresses ambientais tendo sido encontrados.

Apesar de avaliações experimentais futuras serem necessárias para a confirmação dos substratos e características bioquímicas específicas das enzimas relacionadas ao metabolismo de carboidratos codificadas no genoma de GeG2, indícios genômicos sugerem um amplo potencial do novo isolado para a degradação e síntese de diversos polissacarídeos, com um amplo repertório de CAZymes tendo sido detectado. Além disso, assim como comumente observado para membros do gênero *Novosphingobium*, genes codificadores de um grande número de monooxigenases, dioxigenases e outras enzimas usualmente envolvidas na degradação de diversos compostos aromáticos foram identificados no genoma da nova espécie. Em conjunto, os dados genômicos sugerem um amplo potencial de *N. terrae* quanto à sua aplicabilidade industrial e biotecnológica, principalmente na produção de EPSs e a degradação de compostos aromáticos recalcitrantes.

Futuramente, análises comparativas com genomas de outras espécies de *Novosphingobium* e gêneros relacionados poderão auxiliar no melhor entendimento tanto das propriedades do cromídeo quanto no potencial funcional codificado no genoma de GeG2 como um todo. Além disso, avaliações experimentais das condições de produção e composição dos EPSs, do conteúdo das vesículas produzidas e da capacidade de degradação de diferentes compostos aromáticos poderão ser realizadas e, além de melhor elucidarem a biologia da bactéria, poderão revelar a aplicabilidade da nova espécie isolada do solo de Cerrado em vários processos biotecnológicos importantes.

REFERÊNCIAS BIBLIOGRÁFICAS

- Achbergerová L; Nahálka J. (2011). Polyphosphate--an ancient energy source and active metabolic regulator. **Microb Cell Fact.**, 10:63. <https://doi.org/10.1186/1475-2859-10-63>
- Ackert Jr LT (2007). The "cycle of life" in ecology: Sergei Vinogradskii's Soil Microbiology, 1885-1940. **J. Hist. Biol.**, 40:109–145.
- Adam PS; Borrel G; Brochier-Armanet C; Gribaldo S. (2017). The growing tree of *Archaea*: new perspectives on their diversity, evolution and ecology. **ISME J.**, 11(11):2407-2425. <https://doi.org/10.1038/ismej.2017.122>
- Agustini BC; Silva LP; Bloch Jr C. *et al.* (2014). Evaluation of MALDI-TOF mass spectrometry for identification of environmental yeasts and development of supplementary Database. **Appl. Microbiol. Biotechnol.**, 98:5645–5654. <https://doi.org/10.1007/s00253-014-5686-7>
- Alain, K. e Querellou, J. (2009). Cultivating the uncultured: limits, advances and future challenges. **Extremophiles**, 13:583-594. <https://doi.org/10.1007/s00792-009-0261-3>
- Albers, S. V. e Meyer, B. H. 2011. The archaeal cell envelope. **Nature**, 9(6):414-427. <https://doi.org/10.1038/nrmicro2576>
- Alegria, M. C., Souza, DP; Andrade MO. *et al.* (2005). Identification of new protein-protein interactions involving the products of the chromosome- and plasmid-encoded type IV secretion loci of the phytopathogen *Xanthomonas axonopodis* pv. citri. **J. Bacteriol.** 187, 2315–2325. <https://doi.org/10.1128/JB.187.7.2315-2325.2005>
- Altschul SF, Gish W; Miller W *et al.* (1990). Basic local alignment search tool. **J. Mol. Biol.**, 215:403-410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Alves-Prado HF; Pavezzi FC; Leite RS *et al.* (2010). Screening and production study of microbial xylanase producers from Brazilian Cerrado. **Appl. Biochem. Biotechnol.**, 161(1-8):333-46. <https://doi.org/10.1007/s12010-009-8823-5>
- Amann RI; Ludwig W; Schleifer KH. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. **Microbiologic. Rev.**, 59(1):143-169.
- Anda M; Ohtsubo Y; Okubo T *et al.* (2015). Bacterial clade with the ribosomal RNA operon on a small plasmid rather than the chromosome. **Proc. Natl. Acad. Sci. USA**, 112(46):14343-14347. <https://doi.org/10.1073/pnas.1514326112>
- Antipov D; Korobeynikov A; McLean JS; Pevzner PA. (2016). hybridSPAdes: an algorithm for hybrid assembly of short and long reads. **Bioinformatics**, 32(7): 1009–1015. <https://doi.org/10.1093/bioinformatics/btv688>
- Araujo ASF; Mendes L; Bezerra WM *et al.* (2018). Archaea diversity in vegetation gradients from the Brazilian Cerrado. **Braz. J. Microbiol.**, pii: S1517-8382(17)30177-6. <https://doi.org/10.1016/j.bjm.2017.08.010>
- Araújo JF; Castro AP; Costa MMC *et al.* (2012). Characterization of soil bacterial assemblies in Brazilian savanna-like vegetation reveals acidobacteria dominance. **Microb. Ecol.**, 64(3):760-770. <https://doi.org/10.1007/s00248-012-0057-3>
- Ashtaputre AA; Shah AK. (1995). Studies on a Viscous, Gel-Forming Exopolysaccharide from *Sphingomonas paucimobilis* GS1. **Appl. Environ. Microbiol.**, 61(3):1159-1162. <https://doi.org/10.1128/AEM.61.3.1159-1162.1995>
- Auguet JC; Barberan A; Casamayor EO. (2010). Global ecological patterns in uncultured *Archaea*. **ISME J.**, 4(2):182-190. <https://doi.org/10.1038/ismej.2009.109>
- Ayed L; Chaieb K; Cheref A; Bakhrouf A. (2009). Biodegradation of triphenylmethane dye Malachite Green by *Sphingomonas paucimobilis*. **World J. Microbiol. Biotechnol.**, 25(4):705-711. <https://doi.org/10.1007/s11274-008-9941-x>
- Aylward FO; McDonald BR; Adams SM *et al.* (2013). Comparison of 26 sphingomonad genomes reveals diverse environmental adaptations and biodegradative capabilities. **Appl. Environ. Microbiol.**, 79(12):3724-3733. <https://doi.org/10.1128/AEM.00518-13>
- Baek SH; Lim JH; Jin L; Lee HG; Lee ST. (2011). *Novosphingobium sediminicola* sp. nov. isolated from freshwater sediment. **Int. J. Syst. Evol. Microbiol.**, 61(Pt 10):2464-2468. <https://doi.org/10.1099/ijs.0.024307-0>
- Baker B.J; Dick GJ. (2013). Omic approaches in microbial ecology: charting the unknown. **Microbe**, 8: 353–359. <https://doi.org/10.1128/microbe.8.353.1>

- Baker BJ; Comolli LR; Dick GJ *et al.* (2010). Enigmatic, ultrasmall, uncultivated *Archaea*. **Proc. Natl. Acad. Sci. USA**, 107(19): 8806–8811. <https://doi.org/10.1073/pnas.0914470107>
- Balch WE; Fox E; Magrum LJ *et al.* (1979). Methanogens: reevaluation of a unique biological group. **Microbiol. Rev.**, 43(2):260-296.
- Balkwill DL; Drake GR; Reeves RH *et al.* (1997). Taxonomic study of aromatic-degrading bacteria from deep-terrestrial-subsurface sediments and description of *Sphingomonas aromaticivorans* sp. nov., *Sphingomonas subterranea* sp. nov., and *Sphingomonas stygia* sp. nov. **Int. J. Syst. Bacteriol.**, 47(1):191-201. <http://doi.org/10.1099/00207713-47-1-191>
- Bandounas L; Wierckx NJ; de Winde JH; Ruijssenaars HJ. (2001). Isolation and characterization of novel bacterial strains exhibiting ligninolytic potential. **BMC Biotechnol.**, 11:94. <https://doi.org/10.1186/1472-6750-11-94>
- Bankevich A; Nurk S; Antipov D. *et al.* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **J. Comput. Biol.**, 19(5), 455-477. <https://doi.org/10.1089/cmb.2012.0021>
- Banning N; Brock F; Fry JC *et al.* (2005). Investigation of the methanogen population structure and activity in a brackish lake sediment. **Environ. Microbiol.**, 7(7):947-960. <https://doi.org/10.1111/j.1462-2920.2004.00766.x>
- Bano N; Ruffin S; Ransom B; Hollibaugh JT (2004). Phylogenetic composition of Arctic Ocean archaeal assemblages and comparison with Antarctic assemblages. **Appl. Environ. Microbiol.**, 70(2):781-789. <https://doi.org/10.1128/AEM.70.2.781-789.2004>
- Barns SM; Delwiche CF; Palmer JD; Pace NR. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. **Proc. Natl. Acad. Sci. USA**, 93(17):9188-9193.
- Barton HA; Giarrizzo JG; Suarez P *et al.* (2014). Microbial diversity in a Venezuelan orthoquartzite cave is dominated by the Chloroflexi (Class Ktedonobacterales) and Thaumarchaeota Group I.1c. **Front. Microbiol.**, 5(615):1-14. <https://doi.org/10.3389/fmicb.2014.00615>
- Bauer AW; Kirby WMM; Sherris JC Turck, M. (1966). Antibiotic Susceptibility Testing by a Standardized Single Disk Method. **Am. J. Clin. Pathol.**, 45(4):493–496. https://doi.org/10.1093/ajcp/45.4_ts.493
- Beijerinck M. 1888. The root-nodule bacteria *in Milestones in Microbiology: 1556 to 1940*, Washington, USA: ASM Press. 1998, p. 220-224.
- Bell SG; Dale A; Rees NH; Wong LL. (2010). A cytochrome P450 class I electron transfer system from *Novosphingobium aromaticivorans*. **Appl. Microbiol. Biotechnol.**, 86(1):163-175. <https://doi.org/10.1007/s00253-009-2234-y>
- Belmok A; Rodrigues-Oliveira T; Lopes FAC *et al.* (2019). Long-term effects of periodical fires on archaeal communities from Brazilian Cerrado soils. **Archaea**, 2019(11):1-11. <https://doi.org/10.1155/2019/6957210>
- Bertelli C; Brinkman FSL. (2018). Improved genomic island predictions with IslandPath-DIMOB. **Bioinformatics**, 34(13):2161–2167. <https://doi.org/10.1093/bioinformatics/bty095>
- Beveridge TJ; Schultze-Lam S. (1996). The response of selected members of the *Archaea* to the Gram stain. **Microbiol.**, 142:2887-2892. <https://doi.org/10.1099/13500872-142-10-2887>
- Bingle LE; Bailey CM; Pallen MJ. (2008). Type VI secretion: a beginner's guide. **Curr. Opin. Microbiol.**, 11(1):3-8. <https://doi.org/10.1016/j.mib.2008.01.006>
- Blevins SM; Bronze MS. (2010). Robert Koch and the ‘golden age’ of bacteriology. **Int. J. Infect. Dis.**, 14(9):e744-51
- Blin K; Shaw S; Steinke K; Villebro R *et al.* (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. **Nucleic Acids Res.**, 47(W1):W81-W87. <https://doi.org/10.1093/nar/gkz310>
- Boitard L; Cottinet D; Bremond N *et al.* (2015). Growing microbes in millifluidic droplets. **Engineer. Life Sci.**, 15(3):318-326. <https://doi.org/10.1002/elsc.201400089>
- Bragg L; Tyson GW (2014). Metagenomics Using Next-Generation Sequencing *in Environmental Microbiology: Meth. Molec. Biol.*, 1096: 183-201.
- Bresolin JD; Bustamante MMC; Krüger RH *et al.* (2010). Structure and composition of bacterial and fungal community in soil under soybean monoculture in the Brazilian Cerrado. **Braz. J. Microbiol.**, 41:391-403. <https://doi.org/10.1590/S1517-83822010000200021>
- Breton C; Snajdrová L; Jeanneau C; Koca J; Imberty A. (2006). Structures and mechanisms of glycosyltransferases. **Glycobiology**, 16(2):29R-37R. <https://doi.org/10.1093/glycob/cwj016>
- Brochier-Armanet C; Boussau B; Gribaldo S; Forterre P. (2008). Mesophilic Crenarchaeota: proposal for a) third archaeal phylum, the Thaumarchaeota. **Nature Rev. Microbiol.**, 6(3):245-252. <https://doi.org/10.1038/nrmicro1852>
- Brochier-Armanet, C; Forterre P; Gribaldo S. (2011). Phylogeny and evolution of the *Archaea*: one hundred genomes later. **Curr. Opin. Microbiol.**, 14(3):274-81. <https://doi.org/10.1016/j.mib.2011.04.015>

- Brown CT; Hug LA; Thomas BC *et al.* (2015). Unusual biology across a group comprising more than 15% of domain *Bacteria*. **Nature**, 523:208–211. <https://doi.org/10.1038/nature14486>
- Brown JH. (1919) The use of blood agar for the study of streptococci. **Rockefeller Institute for Medical Research Monograph No.9**. The Rockefeller Institute for Medical Research, New York.
- Bush K; Jacoby GA. (2010). Updated functional classification of beta-lactamases. **Antimicrob. Agents Chemother.**, 54(3), 969–976. <https://doi.org/10.1128/AAC.01009-09>
- Busse HJ; Kämpfer P; Denner EB. (1999). Chemotaxonomic characterisation of *Sphingomonas*. **J. Ind. Microbiol. Biotechnol.**, 23(4-5):242-251. <https://doi.org/10.1038/sj.jim.2900745>
- Cai M; Liu Y; Yin X; Zhou Z. *et al.* (2020) Diverse Asgard archaea including the novel phylum Gerdararchaeota participate in organic matter degradation. **Sci. China Life Sci.**, 63. <https://doi.org/10.1007/s11427-020-1679-1>
- Cámara B; Bielecki P; Kaminski F *et al.* (2007). A gene cluster involved in degradation of substituted salicylates via ortho cleavage in *Pseudomonas* sp. strain MT1 encodes enzymes specifically adapted for transformation of 4-methylcatechol and 3-methylmuconate. **J. Bacteriol.**, 189(5):1664-1674. <https://doi.org/10.1128/JB.01192-06>
- Cantarel BL; Coutinho PM; Rancurel C; Bernard T; Lombard V; Henrissat B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. **Nucleic Acids Res.**, 37(Database issue):D233-D238. <https://doi.org/10.1093/nar/gkn663>
- Carver T; Thomson N; Bleasby A *et al.* (2009). DNAPlotter: circular and linear interactive genome visualization. **Bioinformatics**, 25;1;119-20 <https://doi.org/10.1093/bioinformatics/btn578>
- Castelle CJ; Banfield JF. (2018). Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. **Cell**, 172(6):1181-1197. <https://doi.org/10.1016/j.cell.2018.02.016>
- Castro AP; Da Silva MRSS; Quirino BF *et al.* (2016). Microbial diversity in cerrado biome (neotropical savanna) soils. **PLoS ONE**, 11(2):1-16. <https://doi.org/10.1371/journal.pone.0148785>
- Castro AP; Quirino BF; Pappas Jr GJ *et al.* (2008). Diversity of soil fungal communities of Cerrado and its closely surrounding agriculture fields. **Arch. Microbiol.**, 190:129-139. <https://doi.org/10.1007/s00203-008-0374-6>
- Catão E; Castro AP; Barreto CC *et al.* (2013). Diversity of *Archaea* in Brazilian savanna soils. **Arch. Microbiol.**, 195(7):507-512. <https://doi.org/10.1007/s00203-013-0882-x>
- Cevallos MA; Cervantes-Rivera R; Gutiérrez-Ríos RM. (2008). The repABC plasmid family. **Plasmid**. 60(1):19-37. <https://doi.org/10.1016/j.plasmid.2008.03.001>
- Chain PS; Denev VJ; Konstantinidis KT *et al.* (2006). *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. **Proc. Natl. Acad. Sci. USA.**, 103(42):15280-15287. <https://doi.org/10.1073/pnas.0606924103>
- Chaudhary DK; Kim, J. (2016). *Novosphingobium naphthae* sp. nov., from oil-contaminated soil. **Int. J. Syst. Evol. Microbiol.**, 66(8):3170-3176.
- Chick H. (1905). The biological limitations of the method of pure culture. **New Phytol. J.**, 4:120–124.
- Choi CW; Park EC; Yun SH *et al.* (2014). Proteomic characterization of the outer membrane vesicle of *Pseudomonas putida* KT2440. **J. Proteome Res.**, 13(10):4298-4309. <https://doi.org/10.1021/pr500411d>
- Choi DH; Kwon YM; Kwon KK; Kim SJ. (2015). Complete genome sequence of *Novosphingobium pentaromativorans* US6-1. **Stand. Genomic Sci.**, 10: 107.
- Chun J; Oren A; Ventosa A *et al.* (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. **Int. J. Syst. Evol. Microbiol.**, 68(1):461-466. <https://doi.org/10.1099/ijsem.0.002516>
- Coll-Lladó M; Acinas SG; Pujades C; Pedrós-Alió C. (2011). Transcriptome fingerprinting analysis: An approach to explore gene expression patterns in marine microbial communities. **PLoS ONE**, 6(8): e22950. <https://doi.org/10.1371/journal.pone.0022950>
- Connon SA; Giovannoni SJ. (2002). High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. **Appl. Environ. Microbiol.**, 68:3878-3885. <https://doi.org/10.1128/aem.68.8.3878-3885.2002>
- Coppotelli BM; Ibarrolaza A; Dias RL *et al.* (2010). Study of the degradation activity and the strategies to promote the bioavailability of phenanthrene by *Sphingomonas paucimobilis* strain 20006FA. **Microb. Ecol.**, 59(2):266-276. <https://doi.org/10.1007/s00248-009-9563-3>
- Coughlan LM; Cotter PD; Hill C; Alvarez-Ordóñez A. (2015). Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. **Front. Microbiol.**, 6:1-22. <https://doi.org/10.3389/fmicb.2015.00672>
- Coulthurst S. (2019). The Type VI secretion system: a versatile bacterial weapon. **Microbiology**, 165(5):503-515. <https://doi.org/10.1099/mic.0.000789>

- Cragg GM; Newman DJ. (2013). Natural products: a continuing source of novel drug leads. **Biochim. Biophys. Acta.** 1830(6):3670-3695. <https://doi.org/10.1016/j.bbagen.2013.02.008>
- Cruaud P; Vigneron A; Lucchetti-Miganeh C *et al.* (2014). Influence of DNA extraction method, 16S rRNA targeted hypervariable regions, and sample origin on microbial diversity detected by 454 pyrosequencing in marine chemosynthetic ecosystems. **Appl. Environ. Microbiol.**, 80(15):4626-4639. <https://doi.org/10.1128/AEM.00592-14>
- Dantas G; Sommer MO; Oluwasegun RD; Church GM. (2008). Bacteria subsisting on antibiotics. **Science**, 320(5872):100-103. <https://doi.org/10.1126/science.1155157>
- D'Argenio V; Notomista E; Petrillo M *et al.* (2014). Complete sequencing of *Novosphingobium* sp. PP1Y reveals a biotechnologically meaningful metabolic pattern. **BMC Genomics**, 15(1):384. <https://doi.org/10.1186/1471-2164-15-384>
- De Lise F; Mensitieri F; Rusciano G *et al.* (2019). *Novosphingobium* sp. PP1Y as a novel source of outer membrane vesicles. **J Microbiol.**, 57(6):498-508. <https://doi.org/10.1007/s12275-019-8483-2>
- De Lise F; Mensitieri F; Tarallo V *et al.* (2016). RHA-P. Isolation, expression and characterization of a bacterial α -1-rhamnosidase from *Novosphingobium* sp. PP1Y. **J. Mol. Catalysis B: Enzymatic**, ;134:136–147. <https://doi.org/10.1016/j.molcatb.2016.10.002>
- del Solar, G; Giraldo R; Ruiz-Echevarría MJ *et al.* (1998). Replication and control of circular bacterial plasmids. **Microbiol. Mol. Biol. Rev.**, 62(2), 434–464.
- Delmont TO; Simonet P; Vogel TM. (2013). Mastering methodological pitfalls for surviving the metagenomic jungle. **Bioessays**, 35:744–754. <https://doi.org/10.1002/bies.201200155>
- DeLong EF. (1992). *Archaea* in costal marine environments. **Proc. Natl. Acad. Sci. USA**, 89:5685-5689. <https://doi.org/10.1073/pnas.89.12.5685>
- DeLong EF. (1998). Everything in moderation: *Archaea* as ‘non-extremophiles’. **Curr. Opin. Genet. Dev.**, 8(6):649-654.
- Deo D; Davray D; Kulkarni R. (2019). A Diverse Repertoire of Exopolysaccharide Biosynthesis Gene Clusters in *Lactobacillus* Revealed by Comparative Analysis in 106 Sequenced Genomes. **Microorganisms**, 7(10), 444. <https://doi.org/10.3390/microorganisms7100444>
- DeSantis TZ; Hugenholtz P; Larsen N *et al.* (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. **Appl. Environ. Microbiol.**, 72(7):5069-5072. <https://doi.org/10.1128/AEM.03006-05>
- Dias, A. B. A. (2015). Filogenia molecular e cultivo de *Archaea* de solos de Cerrado sensu stricto. **Dissertação de Mestrado**, Departamento de Biologia Celular, Instituto de Ciências Biológicas, Universidade de Brasília.
- diCenzo GC; Finan TM. (2017). The Divided Bacterial Genome: Structure, Function, and Evolution. **Microbiol. Mol. Biol. Rev.**, 81(3):e00019-17.. <https://doi.org/10.1128/MMBR.00019-17>
- diCenzo GC; Mengoni A; Perrin E. (2019). Chromids Aid Genome Expansion and Functional Diversification in the Family *Burkholderiaceae*. **Mol. Biol. Evol.**, 36(3):562–574. <https://doi.org/10.1093/molbev/msy248>
- Docampo R. (2006). Acidocalcisomes and Polyphosphate Granules. *In: Inclusions in Prokaryotes - Microbiology Monographs*, vol 1. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-33774-1_3
- Docampo R; de Souza W; Miranda K *et al.* (2005) Acidocalcisomes - conserved from bacteria to man. **Nat. Rev. Microbiol.**, 3(3):251-261. <https://doi.org/10.1038/nrmicro1097>
- Donlan RM; Costerton JW. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. **Clin. Microbiol. Rev.**, 15(2), 167–193. <https://doi.org/10.1128/cmr.15.2.167-193.2002>
- Doxey AC; Kurtz DA; Lynch MD *et al.* (2015) Aquatic metagenomes implicate Thaumarchaeota in global cobalamin production. **ISME J.**, 9(2):461-71. <https://doi.org/10.1038/ismej.2014.142>
- Dridi B; Fardeau M; Ollivier B; Raoult D; Drancourt M. (2011). The antimicrobial resistance pattern of cultured human methanogens reflects the unique phylogenetic position of archaea. **J. Antimicrob. Chemother.**, 66:2038-2044. <https://doi.org/10.1093/jac/dkr251>
- Dridi B; Fardeau; ML; Ollivier B *et al.* (2012). *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. **Intern. J. System. Bacteriol.** 62(8):1902-1907. <https://doi.org/10.1099/ijs.0.033712-0>
- Dworkin, M. (2012). Sergei Winogradsky: A founder of modern microbiology and the first microbial ecologist. **FEMS Microbiol. Rev.**, 36(2): 364-379.
- Dziewit L; Czarnecki J; Wibberg D *et al.* (2014). Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. **BMC Genomics**, 15:124. <https://doi.org/10.1186/1471-2164-15-124>
- Edwards T McBride BC (1975). New Method for the Isolation and Identification of Methanogenic Bacteria. **Appl. Microbiol.**, 29(4): 540-545.

- Epstein W. (2003). The roles and regulation of potassium in bacteria. **Prog. Nucleic Acid Res. Mol. Biol.**, 75:293-320. [https://doi.org/10.1016/s0079-6603\(03\)75008-9](https://doi.org/10.1016/s0079-6603(03)75008-9)
- Essel E; Li L; Deng C *et al.* (2017). Effect of Conservational and Conventional Tillage Systems on Functional Soil Archaea Diversity in Wheat-Pea Rotation Field. **J. Agric. Sci.**, 9(11): 34-44. <https://doi.org/10.5539/jas.v9n11p34>
- Etto RM; Cruz LM; Jesus EC *et al.* (2012). Prokariotic communities of acidic peatlands from the Southern Brazilian Atlantic Forest. **Braz. J. Microbiol.**, 661-674. <https://doi.org/10.1590/S1517-83822012000200031>
- Evans PN; Parks DH; Chadwick G.L *et al.* (2015). Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. **Science**, 350(6259):434-438. <https://doi.org/10.1126/science.aac7745>
- Ewing B; Hillier L; Wendl MC; Green P (1998). Base-calling of automated sequencer traces using Phred. I. accuracy assessment. **Genome Res.**, 8:175-185. <https://doi.org/10.1101/gr.8.3.175>
- Faron ML; Buchan BW; Hyke J *et al.* (2015) Multicenter Evaluation of the Bruker MALDI Biotyper CA System for the Identification of Clinical Aerobic Gram-Negative Bacterial Isolates. **Plos One**, 10(11): e0141350. <https://doi.org/10.1371/journal.pone.0141350>
- Farris JS. (1972). Estimating phylogenetic trees from distance matrices. **Am. Nat.**, 106: 645–667.
- Felsenstein J (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. **J. Mol. Evol.**, 17, 368–376. <https://doi.org/10.1007/BF01734359>
- Felsenstein J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. **Evolution**, 39:783-791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>
- Fialho AM; Moreira LM; Granja AT; Popescu AO; Hoffmann K; Sá-Correia I. (2008). Occurrence, production, and applications of gellan: current state and perspectives. **Appl. Microbiol. Biotechnol.**, 79(6):889-900. <https://doi.org/10.1007/s00253-008-1496-0>
- Ficko-Blean E; Boraston AB. (2012). Insights into the recognition of the human glycome by microbial carbohydrate-binding modules. **Curr. Opin. Struct. Biol.**, 22(5):570-577. <https://doi.org/10.1016/j.sbi.2012.07.009>
- Fischer W; Tegtmeyer N; Stingl K; Backert S. (2020). Four Chromosomal Type IV Secretion Systems in *Helicobacter pylori*: Composition, Structure and Function. **Front. Microbiol.**, 11, 1592. <https://doi.org/10.3389/fmicb.2020.01592>
- Forney LJ; Zhou X; Brown CJ. (2004). Molecular microbial ecology: land of the one-eyed king. **Curr. Opin. Microbiol.**, 7(3):210-220. <https://doi.org/10.1016/j.mib.2004.04.015>
- Fox GE; Pechman KR; Woese CR. (1977). Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to procaryotic systematics. **Intern. J. Syst. Bacteriol.**, 27(1):44-57.
- Fox GE; Wisotzkey JD; Jurtshuk Jr P. (1992) How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. **Int. J. Syst. Bacteriol.**, 42(1):166-170. <https://doi.org/10.1099/00207713-42-1-166>
- Frank C; Jendrossek D. (2020). Acidocalcisomes and Polyphosphate Granules Are Different Subcellular Structures in *Agrobacterium tumefaciens*. **Appl. Environ. Microbiol.**, 86(8):e02759-19. <https://doi.org/10.1128/AEM.02759-19>
- Frank JA; Reich CI; Sharma S *et al.* (2008). Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. **Appl. Environ. Microbiol.**, 74(8):2461-2470. <https://doi.org/10.1128/AEM.02272-07>
- Frank O; Göker M; Pradella S; Petersen J. (2015). Ocean's twelve: flagellar and biofilm chromids in the multipartite genome of *Marinovum algicola* DG898 exemplify functional compartmentalization. **Environ. Microbiol.**, 17:4019–4034. <https://doi.org/10.1111/1462-2920.12947>
- Fuhrman JA; McCallum, K; Davis AA. (1993). Phylogenetic Diversity of Subsurface Marine Microbial Communities from the Atlantic and Pacific Oceans. **Appl. Environ. Microbiol.**, 59, (5):1294-1302.
- Fujii K; Satomi M. Morita N; *et al.* (2003). *Novosphingobium tardaugens* sp. nov., an oestradiol-degrading bacterium isolated from activated sludge of a sewage treatment plant in Tokyo. **Int. J. Syst. Evol. Microbiol.**, 53(Pt 1):47-52. <https://doi.org/10.1099/ijs.0.02301-0>
- Fuller GW. (1895). On the proper reaction of nutrient media for bacterial cultivation. **Papers and proceedings of the convention of bacteriologists.** 381-399.
- Galand PE; Lovejoy C; Vincent WF. (2006). Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. **Aquat. Microb. Ecol.** 44(2):115-126. <https://doi.org/10.3354/ame044115>
- Galperin MY; Kristensen DM; Makarova KS *et al.* (2019). Microbial genome analysis: the COG approach. **Brief Bioinform.** 20(4), 1063–1070. <https://doi.org/10.1093/bib/bbx117>

- Galperin MY; Makarova KS; Wolf YI; Koonin EV. (2015). Expanded microbial genome coverage and improved protein family annotation in the COG database. **Nucleic Acids Res.**, 43(Database issue):D261-D269. <https://doi.org/10.1093/nar/gku1223>
- Galperin, MY; Kolker, E. (2006). New metrics for comparative genomics. **Curr. Opin. Biotechnol.**, 17(5), 440–447. <https://doi.org/10.1016/j.copbio.2006.08.007>
- Gan HM; Chew TH; Hudson AO; Savka MA. (2012). Genome sequence of *Novosphingobium* sp. strain Rr 2-17, a nopaline crown gall-associated bacterium isolated from *Vitis vinifera* L. grapevine. **J. Bacteriol.**, 194(18):5137-5138. <https://doi.org/10.1128/JB.01159-12>
- Gavriš E; Bollmann A; Epstein S; Lewis K. (2008). A trap for *in situ* cultivation of filamentous actinobacteria. **J. Microbiol. Methods.**, 72:257–262. <https://doi.org/10.1016/j.mimet.2007.12.009>
- Gilewicz M; Ni'matuzahroh; Nadalig T *et al.* (1997). Isolation and characterization of a marine bacterium capable of utilizing 2-methylphenanthrene. **Appl. Microbiol. Biotechnol.**, 48, 528–533. <https://doi.org/10.1007/s002530051091>
- Glaeser SP, Bolte K, Busse HJ, Kämpfer P, Grossart HP, Glaeser J. (2013b). *Novosphingobium aquaticum* sp. nov., isolated from the humic-matter-rich bog lake Grosse Fuchskuhle. **Int. J. Syst. Evol. Microbiol.**, 63(Pt 7):2630-2636. <https://doi.org/10.1099/ijs.0.048280-0>
- Glaeser SP; Bolte K; Martin K *et al.* (2013a). *Novosphingobium fuchskuhlense* sp. nov., isolated from the north-east basin of Lake Grosse Fuchskuhle. **Int. J. Syst. Evol. Microbiol.**, ;63(Pt 2):586-592. <https://doi.org/10.1099/ijs.0.043083-0>
- Glaeser SP; Kämpfer P (2014) The Family *Sphingomonadaceae*. In: **The Prokaryotes**. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-30197-1_302
- Gogoleva NE; Nikolaichik YA; Ismailov TT *et al.* (2019). Complete genome sequence of the abscisic acid-utilizing strain *Novosphingobium* sp. P6W. **3 Biotech.**, 2019;9(3):94. <https://doi.org/10.1007/s13205-019-1625-8>
- Goutx M; Mutaftshiev S; Bertrand J (1987). Lipid and exopolysaccharide production during hydrocarbon growth of a marine bacterium from the sea surface. **Mar. Ecol. Prog. Ser.**, 40(3), 259-265. <https://doi.org/10.3354/meps040259>
- Green ER; Mecsas J. (2016). Bacterial Secretion Systems: An Overview. **Microbiol. Spectr.**, 4(1):10.1128/microbiolspec.VMBF-0012-2015. <https://doi.org/10.1128/microbiolspec.VMBF-0012-2015>
- Green SJ; Neufeld JD. (2016). Introduction to “Microbial Community Analysis of Environmental Samples with Next-Generation Sequencing” in **Manual of Environmental Microbiology**, 4^a ed. Washington, ASM Press, 2.4.1-1-2.4.1-6.
- Großkopf R; Stubner S; Liesack W. (1998). Novel euryarchaeotal lineages detected on rice roots and in the anoxic bulk soil of flooded rice microcosms. **Appl. Environ. Microbiol.**, 64(12):4983-4989.
- Gu C; Kim GB; Kim WJ *et al.* (2019). Current status and applications of genome-scale metabolic models. **Genome Biol.**, 20(1):121. <https://doi.org/10.1186/s13059-019-1730-3>
- Gupta SK; Lal D; Lal, R. (2009). *Novosphingobium panipatense* sp. nov. and *Novosphingobium mathurensense* sp. nov., from oil-contaminated soil. **Intern. J. Syst. Bacteriol.**, 59, 156–161. <https://doi.org/10.1099/ijs.0.65743-0>
- Gurevich A; Saveliev V; Vyahhi N; Tesler G. (2013). QUASt: quality assessment tool for genome assemblies. **Bioinformatics**, 29(8):1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Gutleben J; Chaib De Mares M; van Elsas JD *et al.* (2018). The multi-omics promise in context: from sequence to microbial isolate. **Crit. Rev. Microbiol.**, 44(2):212-229. <https://doi.org/10.1080/1040841X.2017.1332003>
- Ha SM; Kim CK; Roh J *et al.* (2019). Application of the Whole Genome-Based Bacterial Identification System, TrueBac ID, Using Clinical Isolates That Were Not Identified With Three Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) Systems. **Ann. Lab. Med.**, 39(6), 530–536. <https://doi.org/10.3343/alm.2019.39.6.530>
- Hamilton HL; Domínguez NM; Schwartz KJ *et al.* (2005). *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. **Mol. Microbiol.**, 55(6):1704-1721. <https://doi.org/10.1111/j.1365-2958.2005.04521.x>
- Handelsman, J. (2004). Metagenomics: application of genomics to uncultured microorganisms. **Microbiol. Mol. Biol. Rev.**, 68:669-685.
- Haridasan M. (2008). Nutritional adaptations of native plants of the cerrado biome in acid soils. **Braz. J. Plant Physiol.**, 20(3), 183-195. <https://doi.org/10.1590/S1677-04202008000300003>
- Harrison PW; Lower RP; Kim NK; Young JP (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. **Trends Microbiol.**, 18(4):141-148. <https://doi.org/10.1016/j.tim.2009.12.010>

- Hatzenpichler R; Lebedeva EV; Spieck E *et al.* (2008). A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. **Proc. Nat. Acad. Sci. USA**, 105(6):2134-2139. <https://doi.org/10.1073/pnas.0708857105>
- Heinrich K; Leslie DJ; Morlock M; Bertilsson S; Jonas K. (2019). Molecular Basis and Ecological Relevance of Caulobacter Cell Filamentation in Freshwater Habitats. **mBio**, 10(4):e01557-19. <https://doi.org/10.1128/mBio.01557-19>
- Henrici AT. (1933). Studies of freshwater bacteria. **J. Bacteriol.**, 25(3): 277–287.
- Hesselman MC; Odoni DI; Ryback BM *et al.* (2012). A multi-platform flow device for microbial (co-) cultivation and microscopic analysis. **PLoS One**, 7(5): e36982. <https://doi.org/10.1371/journal.pone.0036982>
- Hitchens AP; Leikind MC (1939). The introduction of agar-agar into Bacteriology. **J. Bacteriol.**, 37(5): 485–493.
- Hotta Y; Sato H; Hosoda A; Tamura H. (2012). MALDI-TOF MS analysis of ribosomal proteins coded in S10 and spc operons rapidly classified the Sphingomonadaceae as alkylphenol polyethoxylate-degrading bacteria from the environment. **FEMS Microbiol. Lett.**, 330(1):23-29. <https://doi.org/10.1111/j.1574-6968.2012.02525.x>
- Huerta-Cepas J; Forslund K; Coelho LP *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>
- Huet J; Schnabel R; Sentenac A; Zillig, W. (1983). Archaeobacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. **EMBO J.**, 2(8):1291-1294.
- Hug LA; Baker B; Anantharaman K *et al.* (2016). A new view of the tree of life. **Nature Microbiol.**, 1(5):1-6. <https://doi.org/10.1038/nmicrobiol.2016.48>
- Huo YY; You H; Li ZY *et al.* (2015). *Novosphingobium marinum* sp. nov., isolated from seawater. **Int. J. Syst. Evol. Microbiol.**, 65(Pt 2):676-80. <https://doi.org/10.1099/ijs.0.070433-0>
- Husain DR; Goutx M; Bezac C; Gilewicz M; Bertrand JC. (1997). Morphological adaptation of *Pseudomonas nautica* strain 617 to growth on eicosane and modes of eicosane uptake. **Let. Appl. Microbiol.**, 24: 55-58. <https://doi.org/10.1046/j.1472-765X.1997.00345.x>
- Ingham CJ; Sprenkels A; Bomer J; Molenaar D. *et al.* (2007). The micro-Petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. **Proc. Natl. Acad. Sci. USA**, 104(46): 18217-18222. <https://doi.org/10.1073/pnas.0701693104>
- Ishii S; Tago K; Senoo K. (2010). Single-cell analysis and isolation for microbiology and biotechnology: methods and applications. **Appl. Microbiol. Biotechnol.**, 86(5):1281-92. <https://doi.org/10.1007/s00253-010-2524-4>
- Iyer R; Moussa SH; Tommasi R; Miller AA. (2019). Role of the *Klebsiella pneumoniae* TolC porin in antibiotic efflux. **Res. Microbiol.**, 170(2):112-116. <https://doi.org/10.1016/j.resmic.2018.11.003>
- Janssen PJ; Van Houdt R; Moors H *et al.* (2010). The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. **PLoS One**, 5(5):e10433. <https://doi.org/10.1371/journal.pone.0010433>
- Jaspers E; Overmann J. (2004). Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysologies. **Appl. Environ. Microbiol.**, 70(8):4831-4839. <https://doi.org/10.1128/AEM.70.8.4831-4839.2004>
- Jukes TH; Cantor, CR. (1969). Evolution of proteins molecules. In: **Mammalian protein metabolism**. Nova York: Academic Press. p. 21-132.
- Jung MY; Park SJ; Min D *et al.* (2011). Enrichment and characterization of an autotrophic ammonia-oxidizing archaeon of mesophilic crenarchaeal group I.1a from an agricultural soil. **App. Environ. Microbiol.**, 77(24):8635-8647. <https://doi.org/10.1128/AEM.05787-11>
- Jungbluth SP; Amend JP; Rappé MS. (2017). Metagenome sequencing and 98 microbial genomes from Juan de Fuca Ridge flank subsurface fluids. **Scientific Data**, 4:170037.
- Jurgens G; Lindstrom K; Saano, A. (1997). Novel group within the kingdom crenarchaeota from boreal forest soil. **App. Environ. Microbiol.**, 63(2):803-805. <https://doi.org/10.1128/AEM.63.2.803-805.1997>
- Kaeberlein T; Lewis K; Espstein SS. (2002). Isolating “Uncultivable” microorganisms in pure culture in a simulated natural environment. **Science**, 296:1127-1129. <https://doi.org/10.1126/science.1070633>
- Kämpfer P; Busse HJ; Glaeser SP. (2018). *Novosphingobium lubricantis* sp. nov., isolated from a coolant lubricant emulsion. **Int. J. Syst. Evol. Microbiol.**, 68(5):1560-1564. <https://doi.org/10.1099/ijsem.0.002702>
- Kämpfer P; Denner EB; Meyer S *et al.* (1997) Classification of “*Pseudomonas azotocolligans*” in the genus *Sphingomonas* as *Sphingomonas trueperi* sp. nov. **Int. J. Syst. Bacteriol.** 47:577–583. <https://doi.org/10.1099/00207713-47-2-577>

- Kämpfer P; Young CC; Busse H. *et al.* (2011). *Novosphingobium soli* sp. nov., isolated from soil. **Int J Syst Evol Microbiol.**, 61(Pt 2):259-263. <https://doi.org/10.1099/ijs.0.022178-0>
- Kämpfer; P., Martin K; McInroy JA; Glaeser SP. (2015). Proposal of *Novosphingobium rhizosphaerae* sp. nov., isolated from the rhizosphere. **Int. J. Syst. Evol. Microbiol.**, 65(Pt 1):195-200. <https://doi.org/10.1099/ijs.0.070375-0>
- Kanehisa M; Araki M; Goto S *et al.* (2008). KEGG for linking genomes to life and the environment. **Nucleic Acids Res.**, 36:D480-D484. <https://doi.org/10.1093/nar/gkm882>
- Kates, M. 1993. Membrane lipids of archaea. In: *The Biochemistry of Archaea (Archaeobacteria)* (Kates, M; Kushner, D. J.; Matheson, A. T., eds). Amsterdam: Elsevier. Capítulo 9, p. 261-295.
- Kertesz MA; Kawasaki A. (2010) Hydrocarbon-Degrading Sphingomonads: *Sphingomonas*, *Sphingobium*, *Novosphingobium*, and *Sphingopyxis*. In: **Handbook of Hydrocarbon and Lipid Microbiology**. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-77587-4_119
- Khelaifia; Drancourt M. (2012). Susceptibility of archaea to antimicrobial agents: applications to clinical microbiology. **Clin. Microbiol. Infect.**, 18(9): 841-848. <https://doi.org/10.1111/j.1469-0691.2012.03913.x>
- Kim J; Jung M; Park S *et al.* (2012). Cultivation of a highly enriched ammonia-oxidizing archaeon of thaumarchaeotal group I.1b from an agricultural soil. **Environ. Microbiol.**, 14(6):1528-1543. <https://doi.org/10.1111/j.1462-2920.2012.02740.x>
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. **J. Mol. Evol.** 16(2):111-120. <https://doi.org/10.1007/BF01731581>
- Kitasato S. (1889). Ueber den Tetanusbacillus. **Ztschr. Hyg. u. Infektionskrankh.**, 7:225-234.
- Koch R. (1882). The etiology of tuberculosis in **Milestones in Microbiology: 1556 to 1940** (Brock, T. D., ed). Washington, USA: ASM Press. 1998, 109-115.
- Koch, R (1876). The etiology of anthrax, based on the life history of Bacillus Anthracis. in **Milestones in Microbiology: 1556 to 1940** (Brock, T. D., ed). Washington, USA: ASM Press. 1998, 89-95.
- Koch, R. (1881). Methods for the study of pathogenic organisms in **Milestones in Microbiology: 1556 to 1940** (Brock, T. D., ed). Washington, USA: ASM Press. 1998, 101-108.
- Könneke M; Bernhard AE; de la Torre JR *et al.* (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. **Nature**, 437(7058):543-546. <https://doi.org/10.1038/nature03911>
- Koren S; Walenz BP; Berlin K *et al.* (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. **Genome Res.**, 27(5): 722-736. <https://doi.org/10.1101/gr.215087.116>
- Krawczyk PS; Lipinski L; Dziembowski A. (2018). PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. **Nucleic Acids Res.**, 46(6):e35. <https://doi.org/10.1093/nar/gkx1321>
- Krieg NR. (1988). Bacterial classification: an overview. **Canad. J. Microbiol.**, 34:536-540.
- Krishnan R; Menon RR; Likhitha *et al.* (2017). *Novosphingobium pokkali* sp nov, a novel rhizosphere-associated bacterium with plant beneficial properties isolated from saline-tolerant pokkali rice. **Res. Microbiol.** 168(2):113-121. <https://doi.org/10.1016/j.resmic.2016.09.001>
- Kubo K; Lloyd KG; Biddle JF. *et al.* (2012). Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. **ISME J.**, 6(10):1949-1965. <https://doi.org/10.1038/ismej.2012.37>
- Kumar D; Mondal AK; Kutum R; Dash D. (2016). Proteogenomics of rare taxonomic phyla: A prospective treasure trove of protein coding genes. **Proteomics**, 16(2):226-240. <https://doi.org/10.1002/pmic.201500263>
- Kumar R; Verma H; Heider S. *et al.* (2017). Comparative genomic analysis reveals habitat-specific genes and regulatory hubs within the genus *Novosphingobium*. **mSystems** 2, e00020-17, <https://doi.org/10.1128/mSystems.00020-17>
- Ladino-Orjuela G; Gomes E; da Silva R; Salt C; Parsons JR. (2016) Metabolic Pathways for Degradation of Aromatic Hydrocarbons by Bacteria. In: **Reviews of Environmental Contamination and Toxicology** Volume 237, Springer, Cham. https://doi.org/10.1007/978-3-319-23573-8_5
- Lagier JC; Edouard S; Pagnier I *et al.* (2015). Current and past strategies for bacterial culture in clinical microbiology. **Clin. Microbiol. Rev.**, 28(1):208-236. <https://doi.org/10.1128/CMR.00110-14>
- Lane DJ. (1991). 16S/23S rRNA sequencing in **Nucleic acid techniques in bacterial systematics**. New York: John Wiley and Sons. P. 115-175.
- Larkin MA; Blackshields G; Brown NP. *et al.* (2007). Clustal W and Clustal X version 2.0. **Bioinformatics**, 23(21): 2947-2948. <https://doi.org/10.1093/bioinformatics/btm404>

- Lasica AM; Ksiazek M; Madej M; Potempa J. (2017). The Type IX Secretion System (T9SS): Highlights and Recent Insights into Its Structure and Function. **Front. Cell. Infect. Microbiol.**, 7:215. <https://doi.org/10.3389/fcimb.2017.00215>
- Lazar CS; Baker BJ; Seitz KW; Teske AP. (2017). Genomic reconstruction of multiple lineages of uncultured benthic archaea suggests distinct biogeochemical roles and ecological niches. **ISME J.**, 11(5):1118-1129.
- Lee JC; Kim SG; Whang KS. (2014a). *Novosphingobium aquiterrae* sp. nov., isolated from ground water. **Int J Syst Evol Microbiol.**, 64(Pt 9):3282-3287. <https://doi.org/10.1099/ijs.0.060749-0>
- Lee LH; Azman, AS; Zainal N. *et al.* (2014). *Novosphingobium malaysiense* sp. nov. isolated from mangrove sediment. 64(pt4):1194–1201. <https://doi.org/10.1099/ijs.0.059014-0>
- Lefort V; Desper R; Gascuel O. (2015). FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. **Mol. Biol. Evol.**, 32(10), 2798–2800. <https://doi.org/10.1093/molbev/msv150>
- Lehtovirta, LE; Prosser JI; Nicol GW. (2009). Soil pH regulates the abundance and diversity of Group 1.1c Crenarchaeota. **FEMS Microbiol Ecol.**, 70(3):367-76. <https://doi.org/10.1111/j.1574-6941.2009.00748.x>
- Lehtovirta-Morley, LE; Ge C; Ross, J *et al.* (2014). Characterisation of terrestrial acidophilic archaeal ammonia oxidisers and their inhibition and stimulation by organic compounds. **FEMS Microbiol. Ecol.**, 89:542-552. <https://doi.org/10.1111/1574-6941.12353>
- Lehtovirta-Morley; LE; Ross J; Hink L *et al.* (2016). Isolation of '*Candidatus* Nitrosocosmicus franklandus', a novel ureolytic soil archaeal ammonia oxidiser with tolerance to high ammonia concentration. **FEMS Microbiol. Ecol.**, 92(5):1-31. <https://doi.org/10.1093/femsec/fiw057>
- Li S; Liu J; Sun K; Yang Z; Ling W. (2020). Degradation of 17 β -estradiol by *Novosphingobium* sp. ES2-1 in aqueous solution contaminated with tetracyclines. **Environ. Pollution**, 260:114063. <https://doi.org/10.1016/j.envpol.2020.114063>
- Li X.; Plésiat P; Nikaido H. (2015). The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. **Clin. Microbiol. Rev.** 28 337–418. <https://doi.org/10.1128/CMR.00117-14>
- Li Y; Ding K., Wen; X. (2016). A novel ammonia-oxidizing archaeon from wastewater treatment plant: Its enrichment, physiological and genomic characteristics. **Scient. Rep.**, 6:23747. <https://doi.org/10.1038/srep23747>
- Locey KJ; Lennon JT. (2016). Scaling laws predict global microbial diversity. **Proc. Natl. Acad. Sci. USA**, 113(21):5970-5. <https://doi.org/10.1073/pnas.1521291113>
- Ludwig, W. e Schleifer, K. W. (1994). Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. **FEMS Microbiol. Rev.**, 15:155-173.
- Ma L; Kim J; Hatzenpichler R. *et al.* (2014). Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in Human Microbiome Project's Most Wanted taxa. **Proc. Natl. Acad. Sci. USA**, 111(27): 9768-9773. <https://doi.org/10.1073/pnas.1404753111>
- Maccallum WG; Hastings TW. (1899). A case of acute endocarditis caused by micrococcus zymogenes (nov. Spec.), with a description of the microorganism. **J. Exp. Med.**, 4(5-6):521-534.
- Mackenzie C; Choudhary M; Larimer FW *et al.* (2001). The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4.1. **Photosynth. Res.**, 70(1):19-41. <https://doi.org/10.1023/A:1013831823701>
- Magrum, LJ; Luehrsén KR; Woese CR. (1978). Are Extreme Halophiles Actually Bacteria? **J. Molec. Evol.**, 11: 1-8
- Manzanares P; Vallés S; Ramòn D; Orejas M. (2007) α -L-rhamnosidases: Old and New Insights. *In: Industrial Enzymes*. Springer, Dordrecht. https://doi.org/10.1007/1-4020-5377-0_8
- Marchesi JR; Sato T; Weightman AJ *et al.* (1998). Design and Evaluation of Useful Bacterium-Specific PCR Primers That Amplify Genes Coding for Bacterial 16S rRNA. **Appl. Environ. Microbiol.**, 64(2):795.
- Martinez-Murcia AJ; Benlloch S; Collins MD. (1992). Phylogenetic interrelationships of members of the genera *Aeromonas* and *Plesiomonas* as determined by 16S ribosomal DNA sequencing: lack of congruence with results of DNA-DNA hybridizations. **Int. J. Syst. Bacteriol.**, 42(3):412-421. <https://doi.org/10.1099/00207713-42-3-412>
- Mathrani IM; Boone DR; Mah RA *et al.* (1988). *Methanohalophilus zhilinae* sp. nov., an alkaliphilic, halophilic, methylotrophic methanogen. **Int. J. Syst. Bacteriol.**, 38(2):139-42.
- Matsuyama H; Kamesaki T; Sasaki R *et al.* (2003). Production of two types of exopolysaccharide by *Novosphingobium rosa*. **J. Biosci. Bioeng.**, 95(2):152-156. [https://doi.org/10.1016/s1389-1723\(03\)80121-4](https://doi.org/10.1016/s1389-1723(03)80121-4)

- Medema MH; Blin K; Cimermancic P; Jager V *et al.* (2011). antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters. **Nucleic Acids Res.**, W339-46. <https://doi.org/10.1093/nar/gkr466>
- Mehetre GT; Paranjpe AS; Dastager SG; Dharne MS. (2016). Complete metagenome sequencing based bacterial diversity and functional insights from basaltic hot spring of Unkeshwar, Maharashtra, India. **Genomics Data**, 7:140–143. <https://doi.org/10.1016/j.gdata.2015.12.031>
- Meier-Kolthoff JP; Göker M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. **Nat. Commun.**, 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>
- Mellmann A; Cloud J; Maier T *et al.* (2008). Evaluation of matrix-assisted laser desorption ionization-time-of-flight mass spectrometry in comparison to 16S rRNA gene sequencing for species identification of nonfermenting bacteria. **J. Clin. Microbiol.**, 46(6):1946-1954. <https://doi.org/10.1128/JCM.00157-08>
- Melo-Nascimento AODS; Treumann C; Neves C *et al.* (2018). Functional characterization of ligninolytic *Klebsiella* spp. strains associated with soil and freshwater. **Arch. Microbiol.**, ;200(8):1267-1278. <https://doi.org/10.1007/s00203-018-1532-0>
- Meng J; Xu J; Qin D *et al.* (2014). Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. **ISME J.**, 8:650-659. <https://doi.org/10.1038/ismej.2013.174>
- Møller MS; Svensson B. (2016). Structural biology of starch-degrading enzymes and their regulation. **Curr. Opin. Struct. Biol.**, 40:33-42. <https://doi.org/10.1016/j.sbi.2016.07.006>
- Morais, D., Pylro, V., Clark, I. M. *et al.* (2016). Responses of microbial community from tropical pristine coastal soil to crude oil contamination. **Peer J.**, 4, e1733. <https://doi.org/10.7717/peerj.1733>
- Moreira MAS; de Souza EC; Moraes CA. (2004). Multidrug efflux systems in Gram-negative bacteria. **Braz. J. Microbiol.**, 35(1-2), 19-28. <https://doi.org/10.1590/S1517-83822004000100003>
- Morgan JLW; Acheson JF; Zimmer J. (2017). Structure of a Type-1 Secretion System ABC Transporter. **Structure**, 25(3):522-529. <https://doi.org/10.1016/j.str.2017.01.010>
- Morris JJ; Johnson ZI; Szul MJ *et al.* (2011). Dependence of the cyanobacterium *Prochlorococcus* on hydrogen peroxide scavenging microbes for growth at the ocean's surface. **PLoS One**, 6:e16805. <https://doi.org/10.1371/journal.pone.0016805>
- Morris JJ; Kirkegaard R; Szul MJ *et al.* (2008). Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by “helper” heterotrophic bacteria. **Appl. Environ Microbiol.**, 74: 4530–4534. <https://doi.org/10.1128/AEM.02479-07>
- Muller EEL; Glaab E; May P *et al.* (2013). Condensing the omics fog of microbial communities. **Trends Microbiol.**, 21(7):325-333. <https://doi.org/10.1016/j.tim.2013.04.009>
- Myers N; Mittermeier RA; Mittermeier CG *et al.* (2000). Biodiversity hotspots for conservation priorities. **Nature**, 403:853-858. <https://doi.org/10.1038/35002501>
- Nesme J; Achouak W; Agathos SN. (2016). Back to the future of soil metagenomics. **Front. Microbiol.**, 7:1-5. <https://doi.org/10.3389/fmicb.2016.00073>
- Ngo HT; Trinh H; Kim JH., *et al.* (2016). *Novosphingobium lotistagni* sp. nov., isolated from a lotus pond. **Int. J. Syst. Evol. Microbiol.**, 66(11):4729-4734. <https://doi.org/10.1099/ijsem.0.001418>
- Nguyen STC; Freund HL; Kasanjian J; Berlemont R. (2018). Function, distribution, and annotation of characterized cellulases, xylanases, and chitinases from CAZy. **Appl. Microbiol. Biotechnol.**, 102(4):1629-1637. <https://doi.org/10.1007/s00253-018-8778-y>
- Nguyen TM; Myung SW; Jang H; Kim J. (2016). Description of *Novosphingobium flavum* sp. nov., isolated from soil. **Int. J. Syst. Evol. Microbiol.**, 66(9):3642-3650. <https://doi.org/10.1099/ijsem.0.001242>
- Niharika N; Moskalikova H; Kaur, J *et al.* (2013). *Novosphingobium barchaimii* sp. nov., isolated from hexachlorocyclohexane-contaminated soil. **Int. J. Syst. Evol. Microbiol.**, 2013 Feb,63(Pt 2):667-72. <https://doi.org/10.1099/ijms.0.039826-0>
- Notomista E; Pennacchio F; Cafaro V *et al.* (2011). The marine isolate *Novosphingobium* sp. PP1Y shows specific adaptation to use the aromatic fraction of fuels as the sole carbon and energy source. **Microb. Ecol.**, 61(3):582-594. <https://doi.org/10.1007/s00248-010-9786-3>
- Oliveira-Filho A. T; Ratter JA. (2002). Vegetation Physiognomies and Woody Flora of the Cerrado Biome. *In: The Cerrados of Brazil – ecology and natural history of a neotropical savanna*. New York: Columbia University Press. Capítulo 6, p.91-120.
- Olsen GJ; Woese CR. (1993). Ribosomal RNA: a key to phylogeny. **FASEB J.**, 7:114-123.
- Olsen GJ; Lane DJ; Giovannoni SJ *et al.* (1986). Microbial ecology and evolution: A ribosomal RNA approach. **Annual Rev. Microbiol.**, 40:337-365.
- Park BJ; Park SJ; Yoon DN *et al.* (2010). Cultivation of autotrophic ammonia-oxidizing archaea from marine sediments in coculture with sulfur-oxidizing bacteria. **Appl. Environ. Microbiol.**, 76(22): 7575-7587. <https://doi.org/10.1128/AEM.01478-10>

- Parks DH; Imelfort M; Skennerton CT; Hugenholtz P; Tyson GW. (2014). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. **Genome Res.**, 25: 1043-1055. <https://doi.org/10.1101/gr.186072.114>
- Parks DH; Rinke C; Chuvochina M *et al.* (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. **Nature Microbiol.**, 2, 1533–1542. <https://doi.org/10.1038/s41564-017-0012-7>
- Parte AC. (2014). LPSN - List of prokaryotic names with standing in nomenclature. **Nuc. Acid. Res.**, 42(D1):613-616. <https://doi.org/10.1093/nar/gkt1111>
- Peixoto J; Silva LP; Krüger, R. H. (2017). Brazilian Cerrado soil reveals an untapped microbial potential for unpretreated polyethylene biodegradation. **J. Hazard. Mater.**, 15,324(Pt B):634-644. <https://doi.org/10.1016/j.jhazmat.2016.11.037>
- Peng Y; Leung HC; Yiu SM; Chin FY. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. **Bioinformatics**, 28(11):1420-1428. <https://doi.org/10.1093/bioinformatics/bts174>
- Pester M; Rattei T; Flechl S *et al.* (2012). *amoA*-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of *amoA* genes from soils of four different geographic regions. **Environ. Microbiol.**, 14(2):525-539. <https://doi.org/10.1111/j.1462-2920.2011.02666.x>
- Petri RJ. (1887). Eine kleine Modification des Koch'schen Plattenverfahrens in **Milestones in Microbiology: 1556 to 1940**. (Brock, T. D., ed), Washington, USA: ASM Press. 1998, 218-219.
- Pine MJ; Barker HA. (1954). Studies on the Methane Bacteria. **J. Bacteriol.**, 68(5): 589–591.
- Pollock T; Armentrout R. (1999). Planktonic/sessile dimorphism of polysaccharide-encapsulated sphingomonads. **J. Ind. Microbiol. Biotech.**, 23, 436–441. <https://doi.org/10.1038/sj.jim.2900710>
- Pukatzki S; Ma AT; Sturtevant D *et al.* (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model system. **Proc. Natl. Acad. Sci. USA**, 103(5):1528-1533. <https://doi.org/10.1073/pnas.0510322103>
- Quirino BF; Pappas GJ; Tagliaferro AC *et al.* (2009). Molecular phylogenetic diversity of bacteria associated with soil of the savanna-like Cerrado vegetation. **Microbiol. Res.**, 164(1):59-70. <https://doi.org/10.1016/j.micres.2006.12.001>
- Raina V; Nayak T; Ray L; Kumari K; Suar M. (2019). Approach for designation and description of novel microbial species. In: **Microbial diversity in the genomic era** (pp. 137–152). India: Elsevier.
- Rappé MS; Connon SA; Vergin KL; Giovannoni, S. J. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. **Nature**, 418:630-633. <https://doi.org/10.1038/nature00917>
- Rasband WS. (1997). ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA, <https://imagej.nih.gov/ij/>
- Rinke C; Schwientek P; Sczyrba A *et al.* (2013). Insights into the phylogeny and coding potential of microbial dark matter. **Nature**, 499(7459):431-437. <https://doi.org/10.1038/nature12352>
- Rodríguez Hernández J; Cerón Cucchi ME; Cravero S *et al.* (2018). The first complete genomic structure of *Butyrivibrio fibrisolvens* and its chromid. **Microb. Genom.**, 4(10):e000216. <https://doi.org/10.1099/mgen.0.000216>
- Rodriguez-Conde S; Molina L; González P; García-Puente A; Segura A. (2016). Degradation of phenanthrene by *Novosphingobium* sp. HS2a improved plant growth in PAHs-contaminated environments. **Appl. Microbiol. Biotechnol.**, 100(24):10627-10636. <https://doi.org/10.1007/s00253-016-7892-y>
- Rosselló-Mora R; Amann, R. (2001). The species concept for prokaryotes. **FEMS Microbiol. Rev.**, 25:39-67.
- Rozzak D; Colwell RR (1987). Survival strategies of bacteria in the natural environment. **Microbiol. Rev.**, 51(3): 365-379.
- Rutebemberwa A; Stevens MJ; Perez MJ *et al.* (2014). *Novosphingobium* and its potential role in chronic obstructive pulmonary diseases: insights from microbiome studies. **PLoS One.**, 9(10):e111150. <https://doi.org/10.1371/journal.pone.0111150>
- Saitou N; Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. **Mol. Biol. Evol.**, 4(4):406-425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
- Sánchez O; Garrido L; Forn I *et al.* (2011). Molecular characterization of activated sludge from a seawater-processing wastewater treatment plant. **Microb. Biotechnol.**, 4(5):628-642. <https://doi.org/10.1111/j.1751-7915.2011.00256.x>
- Sasser, M. (1990). Identification of bacteria by gas chromatography of cellular fatty acids. **MIDI Technical Note** (MIDI, Newark, Delaware),101, 1–7.
- Saxena A; Anand S; Dua A *et al.* (2013). *Novosphingobium lindaniclasticum* sp. nov., a hexachlorocyclohexane (HCH)-degrading bacterium isolated from an HCH dumpsite. **Int. J. Syst. Evol. Microbiol.**, 63(Pt 6):2160-2167. <https://doi.org/10.1099/ijs.0.045443-0>

- Schloss PD; Girard RA; Martin T; Edwards J; Thrash JC. (2016). Status of the archaeal and bacterial census: An update. **mBio**, 7(3): e00201-16. <https://doi.org/10.1128/mBio.00201-16>
- Schloss PD.; Westcott SL; Ryabin T *et al.* (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. **Appl. Environ. Microbiol.**, 75(23):7537-7541. <https://doi.org/10.1128/AEM.01541-09>
- Schulthess B; Brodner K; Bloemberg GV *et al.* (2013). Identification of Gram-Positive Cocci by Use of Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry: Comparison of Different Preparation Methods and Implementation of a Practical Algorithm for Routine Diagnostics. **J. Clin. Microbiol.**, 51(6): 1834–1840. <https://doi.org/10.1128/JCM.02654-12>
- Schwechheimer C; Kuehn MJ. (2015). Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. **Nat. Rev. Microbiol.**, 13(10):605-619. <https://doi.org/10.1038/nrmicro3525>
- Sengupta D; Datta S; Biswas D. (2018). Towards a better production of bacterial exopolysaccharides by controlling genetic as well as physico-chemical parameters. **Appl. Microbiol. Biotechnol.**, 102(4):1587-1598. <https://doi.org/10.1007/s00253-018-8745-7>
- Seufferheld M; Lea CR; Vieira M; Oldfield E; Docampo R. (2004). The H(+)-pyrophosphatase of *Rhodospirillum rubrum* is predominantly located in polyphosphate-rich acidocalcisomes. **J. Biol. Chem.**, 279(49):51193-51202. <https://doi.org/10.1074/jbc.M406099200>
- Seufferheld M; Vieira MC; Ruiz FA; Rodrigues CO; Moreno SN; Docampo R. (2003). Identification of organelles in bacteria similar to acidocalcisomes of unicellular eukaryotes. **J Biol Chem.**, 278(32):29971-29978. <https://doi.org/10.1074/jbc.M304548200>
- Sha S; Zhong J; Chen B; Lin L; Luan T. (2017). *Novosphingobium guangzhouense* sp. nov., with the ability to degrade 1-methylphenanthrene. **Int. J. Syst. Evol. Microbiol.**, 67(2):489-497. <https://doi.org/10.1099/ijsem.0.001669>
- Sharrar AM; Crits-Christoph A; Méheust R *et al.* (2020). Bacterial Secondary Metabolite Biosynthetic Potential in Soil Varies with Phylum, Depth, and Vegetation Type. **mBio**, 11 (3) e00416-20. <https://doi.org/10.1128/mBio.00416-20>
- Sheu SY; Cai CY; Kwon SW; Chen WM. (2020). *Novosphingobium umbonatum* sp. nov., isolated from a freshwater mesocosm. **Int. J. Syst. Evol. Microbiol.**, 70(2):1122-1132. <https://doi.org/10.1099/ijsem.0.003889>
- Sheu SY; Chen ZH; Chen WM. (2016). *Novosphingobium piscinae* sp. nov., isolated from a fish culture pond. **Int. J. Syst. Evol. Microbiol.**, 66(3):1539-1545. <https://doi.org/10.1099/ijsem.0.000914>
- Shokralla S; Spall JL; Gibson JF; Hajibabaei M. (2012). Next-generation sequencing technologies for environmental DNA research. **Mol. Ecol.**, 21, 1794–1805.
- Silva MRSS. (2012). Diversidade de comunidades bacterianas de solo de Cerrado em resposta a diferentes alterações dos ecossistemas. **Tese de doutorado**, Departamento de Ecologia do Instituto de Ciências Biológicas, Universidade de Brasília.
- Simão FA; Waterhouse RM; Ioannidis P; Kriventseva EV; Zdobnov EM. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, 31(19):3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Simon HM; Jahn CE; Bergerud LT; *et al.* (2005). Cultivation of Mesophilic Soil Crenarchaeotes in Enrichment Cultures from Plant Roots. **App. Environ. Microbiol.**, 71(8):4751-4760. <https://doi.org/10.1128/AEM.71.8.4751-4760.2005>
- Singh H; Du J; Yang JE *et al.* (2015). *Novosphingobium aquaticum* sp. nov., isolated from lake water in Suwon, Republic of Korea. **Antonie Van Leeuwenhoek**, 108(4):851-858.
- Singh R; Kumar M; Mittal A; Mehta PK. (2016). Microbial enzymes: industrial progress in 21st century. **3 Biotech.**, 6(2), 174. <https://doi.org/10.1007/s13205-016-0485-8>
- Smalla K; Jechalke S; Top EM. (2015). Plasmid Detection, Characterization, and Ecology. **Microbiol. Spectr.**, 3(1). <https://doi.org/10.1128/microbiolspec.PLAS-0038-2014>
- Sohn JH; Kwon KK; Kang JH; Jung HB; Kim SJ. (2004). *Novosphingobium pentaromativorans* sp. nov., a high-molecular-mass polycyclic aromatic hydrocarbon-degrading bacterium isolated from estuarine sediment. **Int. J. Syst. Evol. Microbiol** 54(Pt 5):1483-1487. <https://doi.org/10.1099/ijms.0.02945-0>.
- Sowers KR; Baron SF; Ferry, J.G. (1984). *Methanosarcina acetivorans* sp. nov., an Acetotrophic Methane-Producing Bacterium Isolated from Marine Sediments. **Appl. Environ. Microbiol.**, 47(5):971-978.
- Stackebrandt E; Goebel BM. (1994). Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. **Int. J. Syst. Evol. Microbiol.**, 44(4):846-849. <https://doi.org/10.1099/00207713-44-4-846>
- Staley BF; de Los Reyes FL; Barlaz MA. (2011). Effect of spatial differences in microbial activity, pH, and substrate levels on methanogenesis initiation in refuse. **Appl. Environ. Microbiol.**, 77(7):2381-2391. <https://doi.org/10.1128/AEM.02349-10>

- Staley BF; de los Reyes FL; Barlaz MA (2012). Comparison of *Bacteria* and *Archaea* communities in municipal solid waste, individual refuse components, and leachate. **FEMS Microbiol. Ecol.**, 79(2):465-473. <https://doi.org/10.1111/j.1574-6941.2011.01239.x>
- Staley JT; Konopka, A. (1985). Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. **Annual Rev. Microbiol.**, 39:321-346. <https://doi.org/10.1146/annurev.mi.39.100185.001541>
- Starikova EV; Tikhonova PO; Prianichnikov NA. *et al.* (2020) Phigaro: high-throughput prophage sequence annotation, **Bioinformatics**, 36(12):3882–3884, <https://doi.org/10.1093/bioinformatics/btaa250>
- Stetter KO; Fiala G; Huber G; Huber R; Seegerer, A. (1990). Hyperthermophilic microorganisms. **FEMS Microbiol. Rev.**, 75:117-124
- Stieglmeier M; Klingl A; Alves RE *et al.* (2014). *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum *Thaumarchaeota*. **Int. J. Syst. Evol. Microbiol.**, 64:2738-2752. <https://doi.org/10.1099/ijs.0.063172-0>
- Sun DL; Jiang X; Wu QL; Zhou NY. (2013). Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. **App. Environ. Microbiol.**, 79(19): 5962-5969. <https://doi.org/10.1128/AEM.01282-13>
- Sun Y; Liu Y; Pan J; Wang F; Li M. (2020). Perspectives on Cultivation Strategies of *Archaea*. **Microb. Ecol.**, 79, 770–784 <https://doi.org/10.1007/s00248-019-01422-7>
- Suzuki Y; Nishijima S; Furuta Y *et al.* (2019). Long-read metagenomic exploration of extrachromosomal mobile genetic elements in the human gut. **Microbiome**, 7(1):119. <https://doi.org/10.1186/s40168-019-0737-z>
- Suzuki S; Horinouchi S; Beppu S. (1988). Growth of a Tryptophanase-producing Thermophile, *Symbiobacterium thermophilum* gen. nov., sp. nov., Is Dependent on Co-culture with a Bacillus sp. **J. Gen. Microbiol.**, 134, 2353-2362.
- Takeuchi M; Hamana K; Hiraishi A. (2001). Proposal of the genus *Sphingomonas sensu stricto* and three new genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, on the basis of phylogenetic and chemotaxonomic analyses. **Intern. J. System. Bacteriol.**, 51(Pt 4):1405-1417. <https://doi.org/10.1099/00207713-51-4-1405>
- Takeuchi M; Sakane T; Yanagi M *et al.* (1995). Taxonomic Study of Bacteria Isolated from Plants: Proposal of *Sphingomonas rosa* sp. nov., *Sphingomonas pruni* sp. nov., *Sphingomonas asaccharolytica* sp. nov., and *Sphingomonas mali* sp. nov. **45(2):334-341**. <https://doi.org/10.1099/00207713-45-2-334>
- Tamaki H. (2019). Cultivation Renaissance in the Post-Metagenomics Era: Combining the New and Old. **Microbes Environ.**, 34(2):117-120. <https://doi.org/10.1264/jsme2.ME3402rh>
- Tamura K; Peterson D; Peterson N; Stecher G; Nei M; Kumar S. (2011). MEGA5: Molecular evolutionary genetics analyses using maximum likelihood, evolutionary distance and maximum parsimony methods. **Mol. Biol. Evol.**, 28: 2731-2739. <https://doi.org/10.1093/molbev/msr121>
- Tanaka Y; Hanada S; Manome A *et al.* (2004). *Catellibacterium nectariphilum* gen. nov., sp. nov., which requires a diffusible compound from a strain related to the genus *Sphingomonas* for vigorous growth. **Int. J. Syst. Evol. Microbiol.** 54:955–959. <https://doi.org/10.1099/ijs.0.02750-0>
- Tarayre C; Nguyen HT; Brognaux A *et al.* (2016). Characterisation of Phosphate Accumulating Organisms and Techniques for Polyphosphate Detection: A Review. **Sensors**, 16(6): 797. <https://doi.org/10.3390/s16060797>
- Tatusov RL; Fedorova ND; Jackson JD *et al.* (2003). The COG database: an updated version includes eukaryotes. **BMC Bioinformatics**, 4:41. <https://doi.org/10.1186/1471-2105-4-41>
- Tatusov RL; Koonin EV; Lipman DJ. (1997). A genomic perspective on protein families. **Science**, ;278(5338):631-637. <https://doi.org/10.1126/science.278.5338.631>
- Tirola MA; Busse HJ; Kämpfer P; Männistö MK. (2005). *Novosphingobium lentum* sp. nov., a psychrotolerant bacterium from a polychlorophenol bioremediation process. **Int J Syst Evol Microbiol.** 55(Pt 2):583-588. <https://doi.org/10.1099/ijs.0.63386-0>
- Tirola MA; Männistö MK; Puhakka JA; Kulomaa MS. (2002). Isolation and characterization of *Novosphingobium* sp. strain MT1, a dominant polychlorophenol-degrading strain in a groundwater bioremediation system. **Appl. Environ. Microbiol.**, 68(1):173-180. <https://doi.org/10.1128/aem.68.1.173-180.2002>
- Timonen S; Bomberg, M. (2009). Archaea in dry soil environments. **Phytochem. Rev.**, 8(3):505-518. <https://doi.org/10.1007/s11101-009-9137-5>
- Tindall BJ; Sikorski J; Smibert RA; Krieg NR. (2007); Phenotypic characterization and the principles of comparative systematics. *In: Methods for General and Molecular Bacteriology*, 3rd edn. pp330–393. Washington, D.C. American Society for Microbiology.

- Tindall BJ. (1990a). A comparative study of the lipid composition of *Halobacterium saccharovorum* from various sources. **Syst. Appl. Microbiol.** 13, 128-130. [https://doi.org/10.1016/S0723-2020\(11\)80158-X](https://doi.org/10.1016/S0723-2020(11)80158-X)
- Tindall BJ. (1990b). Lipid composition of *Halobacterium lacusprofundi*. **FEMS Microbiol. Letts.**, 66, 199-202. [https://doi.org/10.1016/0378-1097\(90\)90282-U](https://doi.org/10.1016/0378-1097(90)90282-U)
- Tørresen OK; Star B; Mier P *et al.* (2019). Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases, **Nuc. Acid. Res.**, 47(21):10994–11006. <https://doi.org/10.1093/nar/gkz841>
- Torsvik V; Øvreas L. (2002). Microbial diversity and function in soil: from genes to ecosystems. **Curr. Opin. Microbiol.**, 5(3):240–245.
- Tourna M; Stieglmeier M; Spang A *et al.* (2011). *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. **Proc. Natl. Acad. Sci. USA**, 108(20):8420-8425. <https://doi.org/10.1073/pnas.1013488108>
- Toyofuku M; Nomura N; Eberl L. (2019). Types and origins of bacterial membrane vesicles. **Nat. Rev. Microbiol.**, 17, 13–24. <https://doi.org/10.1038/s41579-018-0112-2>
- Troncone L. (2011) A study of the biotechnological applications of *Novosphingobium puteolanum* PP1Y. **Tese de doutorado**, Universidade de Nápoli Federico II, Itália.
- Tsao J; Kaneshiro SM; Yu S; Clark DS. (1993). Continuous Culture of *Methanococcus jannaschii*, an Extremely Thermophilic Methanogen. **Biotech. Bioeng.**, 43:258-261. <https://doi.org/10.1002/bit.260430309>
- Tyc O; Song C; Dickschat JS; Vos M; Garbeva P. (2017). The Ecological Role of Volatile and Soluble Secondary Metabolites Produced by Soil Bacteria. **Trends Microbiol.**, 25(4):280-292. <https://doi.org/10.1016/j.tim.2016.12.002>
- Vaksmas A; van Alen TA; Ettwig KF *et al.* (2017). Stratification of Diversity and Activity of Methanogenic and Methanotrophic Microorganisms in a Nitrogen-Fertilized Italian Paddy Soil. **Front. Microbiol.**, 8:2127. <https://doi.org/10.3389/fmicb.2017.02127>
- Vanwonterghem I; Jensen PD; Rabaey K; Tyson GW. (2016). Genome-centric resolution of microbial diversity, metabolism and interactions in anaerobic digestion. **Environ. Microbiol.**, 18(9):3144-3158. <https://doi.org/10.1111/1462-2920.13382>
- Vaz-Moreira I; Nunes OC; Manaia CM. (2011). Diversity and antibiotic resistance patterns of *Sphingomonadaceae* isolates from drinking water. **Appl. Environ. Microbiol.**, 77(16), 5697–5706. <https://doi.org/10.1128/AEM.00579-11>
- Vedler E. (2009) Megaplasmids and the Degradation of Aromatic Compounds by Soil Bacteria. *In: Microbial Megaplasmids. Microbiology Monographs*, vol 11. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-85467-8_2
- Vissers EW; Bodelier PLE; Muyzer G; Laanbroek HJ. (2009). A nested PCR approach for improved recovery of archaeal 16S rRNA gene fragments from freshwater samples. **FEMS Microbiol. Lett.**, 298(2):193-198. <https://doi.org/10.1111/j.1574-6968.2009.01718.x>
- Vives-Peris V; Gómez-Cadenas A; Pérez-Clemente RM. (2018). Salt stress alleviation in citrus plants by plant growth-promoting rhizobacteria *Pseudomonas putida* and *Novosphingobium* sp. **Plant Cell Rep.**, 37, 1557–1569. <https://doi.org/10.1007/s00299-018-2328-z>
- Walderhaug MO; Polarek JW; Voelkner P *et al.* (1992). KdpD and KdpE, proteins that control expression of the kdpABC operon, are members of the two-component sensor-effector class of regulators. **J Bacteriol.**, 174(7):2152-2159. <https://doi.org/10.1128/jb.174.7.2152-2159.1992>
- Wallden K; Rivera-Calzada A; Waksman G. (2010). Type IV secretion systems: versatility and diversity in function. **Cell. Microbiol.**, 12(9), 1203–1212. <https://doi.org/10.1111/j.1462-5822.2010.01499.x>
- Wang H; Fewer DP; Holm L *et al.* (2014). Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. **Proc. Natl. Acad. Sci. USA**, 111(25):9259-9264. <https://doi.org/10.1073/pnas.1401734111>
- Wang J; Wang C; Li J *et al.* (2018). Comparative Genomics of Degradative *Novosphingobium* Strains With Special Reference to Microcystin-Degrading *Novosphingobium* sp. THN1. **Front. Microbiol.**, 9:2238. <https://doi.org/10.3389/fmicb.2018.02238>
- Wang DZ; Kong LF; Li YY; Xie ZX. (2016). Environmental microbial community proteomics: Status, challenges and perspectives. **Int. J. Mol. Sci.**, 17, 1275. <https://doi.org/10.3390/ijms17081275>
- Wang J; Gao F; Liu Z *et al.* (2012). Pathway and molecular mechanisms for malachite green biodegradation in *Exiguobacterium* sp. MG2. **PloS One**, 7(12), e51808. <https://doi.org/10.1371/journal.pone.0051808>
- Wang J; Goh KM; Salem D *et al.* (2019). Genome analysis of a thermophilic exopolysaccharide-producing bacterium - *Geobacillus* sp. WSUCF1. **Sci. Rep.**, 9, 1608. <https://doi.org/10.1038/s41598-018-36983-z>

- Wanyonyi WC; Onyari JM; Shiundu JM; Mulaa FJ. (2017). Biodegradation and detoxification of malachite green dye using novel enzymes from bacillus cereus strain km201428: kinetic and metabolite analysis. **Energy Procedia**, 119:38-51. <https://doi.org/10.1016/j.egypro.2017.07.044>
- Watve M; Shejval V; Sonawane C *et al.* (2000). The 'K' selected oligophilic bacteria, a key to uncultured diversity? **Curr. Sci.**, 78:1535:1542.
- Weber EB; Lehtovirta-Morley LE; Prosser JI; Gubry-Rangin C. (2015). Ammonia oxidation is not required for growth of Group 1.1c soil Thaumarchaeota. **FEMS Microbiol. Ecol.**, 91(3). <https://doi.org/10.1093/femsec/fiv001>
- Weisburg WG; Barns SM; Pelletier DA; Lane DJ. (1991). 16S Ribosomal DNA Amplification for Phylogenetic Study. **J. Bacteriol.**, 173(2): 697-703.
- Whitman WB; Woyke T; Klenk HP *et al.* (2015). Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase III: the genomes of soil and plant-associated and newly described type strains. **Stand. Genomic Sci.**, 10:26. <https://doi.org/10.1186/s40793-015-0017-x>
- Wick RR; Judd LM; Gorrie CL; Holt KE (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. **PloS Comput. Biol.**, 13(6): e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Winkel M; Mitzscherling J; Overduin PP *et al.* (2018). Anaerobic methanotrophic communities thrive in deep submarine permafrost. **Sci. Rep.**, 8(1):1291. <https://doi.org/10.1038/s41598-018-19505-9>
- Winogradsky S. 1890. On the nitrifying organisms in **Milestones in Microbiology: 1556 to 1940** (Brock, T. D., ed). Washington, USA: ASM Press. 1998, 231- 233.
- Woese CR; Fox GE. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. **Proc. Natl. Acad. Sci. USA**, 74(11):5088-5090.
- Woese CR; Kandler O; Wheelis ML. (1990). Towards a natural system of organisms: Proposal for the domains *Archaea*, *Bacteria* and *Eucarya*. **Proc. Natl. Acad. Sci USA**, 87:4576-4579. <https://doi.org/10.1073/pnas.87.12.4576>
- Woese CR; Magrum LJ; Fox, G. E. (1978). Archaeobacteria. **J. Molec. Evol.**, 11(3): 245-251.
- Wu M; Huang H; Li G. *et al.* (2017). The evolutionary life cycle of the polysaccharide biosynthetic gene cluster based on the Sphingomonadaceae. **Sci. Rep.**, 7, 46484. <https://doi.org/10.1038/srep46484>
- Yabuuchi E; Kosako Y; Fujiwara N *et al.* (2002) Emendation of the genus *Sphingomonas* Yabuuchi *et al.* 1990 and junior objective synonymy of the species of three genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, in conjunction with *Blastomonas ursincola*. **Int. J. Syst. Evol. Microbiol.**, 52(Pt 5):1485-1496. <https://doi.org/10.1099/00207713-52-5-1485>
- Yadav V; Yadav PK; Yadav S; Yadav KDS. (2010). α -l-Rhamnosidase: a review. **Process Biochem.**, 45:1226–1235. <https://doi.org/10.1016/j.procbio.2010.05.025>
- Yan Q-X; Hong Q; Han P *et al.* (2007) Isolation and characterization of a carbofuran-degrading strain *Novosphingobium* sp. FND-3. **FEMS Microbiol. Lett.**, 271:207–213. <https://doi.org/10.1111/j.1574-6968.2007.00718.x>
- Yoon SH; Ha SM; Kwon S *et al.* (2017). Introducing EzBioCloud: A taxonomically united database of 16S rRNA and whole genome assemblies. **Int. J. Syst. Evol. Microbiol.**, 67:1613-1617. <https://doi.org/10.1099/ijsem.0.001755>
- Yu T; Liang Q; Niu M; Wang F. (2017). High occurrence of Bathyarchaeota (MCG) in the deep-sea sediments of South China Sea quantified using newly designed PCR primers. **Environ. Microbiol. Rep.**, 9(4):374-382. <https://doi.org/10.1111/1758-2229.12539>
- Yuan J; Lai Q; Zheng T; Shao Z. (2009). *Novosphingobium indicum* sp. nov., a polycyclic aromatic hydrocarbon-degrading bacterium isolated from a deep-sea environment. **Int J Syst Evol Microbiol.**, 59(Pt 8):2084-2088. <https://doi.org/10.1099/ijms.0.002873-0>
- Yuan Z; Tam VH. (2008). Polymyxin B: a new strategy for multidrug-resistant Gram-negative organisms. **Int J Syst Evol Microbiol.**, 58(Pt 5):661-668. <https://doi.org/10.1517/13543784.17.5.661>
- Yücel O; Holert J; Ludwig KC. (2017). A Novel Steroid-Coenzyme A Ligase from *Novosphingobium* sp. Strain Chol11 Is Essential for an Alternative Degradation Pathway for Bile Salts. **Appl. Environ. Microbiol.**, 84(1): pii: e01492-17.
- Yun SH; Lee SY; Choi CW *et al.* (2017). Proteomic characterization of the outer membrane vesicle of the halophilic marine bacterium *Novosphingobium pentaromativorans* US6-1. **J. Microbiol.**, 55(1):56-62. <https://doi.org/10.1007/s12275-017-6581-6>
- Zengler K; Toledo G; Rappé M *et al.* (2002). Cultivating the uncultured. **Proc. Natl. Acad. Sci. USA**, 99(24):15681-15686. <https://doi.org/10.1073/pnas.252630999>
- Zengler K; Walcher M; Clark G *et al.* (2005). High-throughput cultivation of microorganisms using microcapsules. **Methods Enzymol.**, 397:124-30. [https://doi.org/10.1016/S0076-6879\(05\)97007-9](https://doi.org/10.1016/S0076-6879(05)97007-9)
- Zhang H; Yohe T; Huang L *et al.* (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. **Nucleic Acids Res.**, 46(W1):W95-W101. <https://doi.org/10.1093/nar/gky418>

- Zhang L; Gao JS; Kim SG *et al.* (2016). *Novosphingobium oryzae* sp. nov., a potential plant-promoting endophytic bacterium isolated from rice roots. **Int. J. Syst. Evol. Microbiol.**, 66(1):302-307. <https://doi.org/10.1099/ijsem.0.000718>
- Zheng J; Guan Z; Cao S *et al.* (2015). Plasmids are vectors for redundant chromosomal genes in the *Bacillus cereus* group. **BMC Genomics**, 16(1):6. <https://doi.org/10.1186/s12864-014-1206-5>
- Zheng J; Peng D; Ruan L; Sun M. (2013). Evolution and dynamics of megaplasmids with genome sizes larger than 100 kb in the *Bacillus cereus* group. **BMC Evol Biol.**, 13:262. <https://doi.org/10.1186/1471-2148-13-262>
- Zhu L; Wu X; Li O; Qian C; Gao H. (2012). Cloning and characterization of genes involved in nostoxanthin biosynthesis of *Sphingomonas elodea* ATCC 31461. **PLoS One.**, 7(4):e35099. <https://doi.org/10.1371/journal.pone.0035099>
- Zuckerandl E; Pauling L. (1965). Molecules as documents of evolutionary history. **J. Theoretic. Biol.**, 8(2):357-366.

ANEXO I – Publicações realizadas durante o período do doutorado

- Artigos publicados em periódicos:

Belmok, A.; de Cena, J.A.; Kyaw, C.M., Damé-Teixeira N. (2020). The Oral Archaeome: A Scoping Review. *Journal of Dental Research*, 99 (6): 630-643. <https://doi.org/10.1177/0022034520910435>

Belmok, A.; Rodrigues-Oliveira, T.; Lopes, F. A. C.; Miranda, H. S.; Krüger, R. H.; Kyaw, C. M. (2019). Long-term effects of periodical fires on archaeal communities from Brazilian Cerrado soils. *Archaea*. 2019(11):1-11. <https://doi.org/10.1155/2019/6957210>

Dame-Teixeira, N.; de Cena, J. A.; Côrtes, D. A.; **Belmok, A.**; Borges, L. G.; Marconatto, L.; Giongo, A.; Kyaw, C. M. (2019) Presence of archaea in dental caries biofilms. *Archives of Oral Biology*, 110:104606. <https://doi.org/10.1016/j.archoralbio.2019.104606>

Rodrigues-Oliveira, T.; **Belmok, A.**; Vasconcellos, D.; Schuster, B.; Kyaw, C.M. (2017). Archaeal S-Layers: Overview and Current State of the Art. *Frontiers in Microbiology*, v. 8, p. 1-17. <https://doi.org/10.3389/fmicb.2017.02597>

Rodrigues, T.; **Belmok, A.**; Catão, E.; Kyaw, C. M. (2016). Archaea in Natural and Impacted Brazilian Environments. *Archaea*, 2016(1): 1-14. <https://doi.org/10.1155/2016/1259608>

- Capítulos de livro:

Belmok, Aline; Rodrigues-Oliveira, Thiago; Kyaw, Cynthia Maria. Southern blot: detecção de fragmentos de DNA imobilizados em membranas. In: Marlene Teixeira De-Souza; Marcelo de Macedo Brígido; Andréa Queiroz Maranhão. (Org.). *Técnicas Básicas em Biologia Molecular*. 2ed. Brasília: Editora UnB, 2016, v., p. 231-250.

Rodrigues-Oliveira, T. R.; **Belmok, Aline**; Kyaw, Cynthia Maria. Northern blot: detecção de RNA por hibridização em membranas. In: Marlene Teixeira De-Souza; Marcelo de Macedo Brígido; Andréa Queiroz Maranhão. (Org.). *Técnicas Básicas em Biologia Molecular*. 2ed. Brasília: Editora UnB, 2016, v., p. 265-272.