

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
DEPARTAMENTO DE MICROBIOLOGIA

Stella de Paula Lopes

Caracterização baseada em genômica evidencia o potencial de bactérias endofíticas  
para promoção do crescimento de plantas

Belo Horizonte - MG  
2020

Stella de Paula Lopes

Caracterização baseada em genômica evidencia o potencial de bactérias endofíticas  
para promoção do crescimento de plantas

Projeto/Monografia apresentado  
Ao departamento de Microbiologia  
da UFMG,  
como requisito para obtenção  
do título de especialista  
versão final

Orientador (a): Prof<sup>a</sup>. Dr<sup>a</sup>. Ubiana de Cássia Mourão Silva

043 Lopes, Stella de Paula.

Caracterização baseada em genômica e análises fenotípicas evidenciam o potencial de bactérias endofíticas para promoção do crescimento de plantas [manuscrito] / Stella de Paula Lopes. – 2020.

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

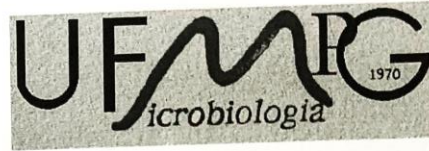
Orientadora: Profa. Dra. Ubiana de Cássia Mourão Silva.

Monografia apresentada ao Curso de Especialização em Microbiologia a Universidade Federal de Minas Gerais – UFMG, como requisito para a obtenção do título de Especialista em Microbiologia.

1. Microbiologia. 2. Endófitos. 3. Bactérias. 4. Genoma de Planta. 5. Agricultura Sustentável. 6. Contenção de Riscos Biológicos. I. Silva, Ubiana de Cássia Mourão. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 579

UFMG



**Universidade Federal de Minas Gerais  
Instituto de Ciências Biológicas  
Programa de Pós-Graduação em Microbiologia**

Às 13:00 horas do dia 13 de fevereiro de 2020 reuniu-se, no Instituto de Ciências Biológicas da UFMG a Banca Debatedora constituída pela Dra. Patrícia Costa Lima da Silva (Departamento de Genética/ICB/UFMG), e a Dra. Ubiana de Cássia Mourão Silva - Orientadora, para avaliar a Monografia intitulada "Caracterização baseada em genômica e análises fenotípicas evidenciam o potencial de bactérias endofíticas para promoção do crescimento de plantas", da aluna Stella de Paula Lopes. Após a apresentação oral pública seguida de uma arguição, a aluna foi APROVADA, considerando as sugestões feitas pela Banca Debatedora. Nada mais havendo a tratar, lavrou-se a presente ata, que será assinada pelos membros da Banca. Belo Horizonte, 13 de fevereiro de 2020.

Dra. Patrícia Costa Lima da Silva

*Patrícia Costa Lima da Silva*

Profa. Ubiana de Cássia Mourão Silva - Orientadora

*Ubiana de Cássia Mourão Silva*

*Prof. Susana Johann*

Profa. Susana Johann

Coordenadora do Curso de Especialização em Microbiologia  
ICB/UFMG

## AGRADECIMENTOS

Primeiramente, obrigada a Deus por ter me guiado nesta jornada.

Obrigada à Professora Vera Lúcia por ter me confiado este projeto e por todas suas contribuições ao longo do seu desenvolvimento

Obrigada a minha orientadora, Professora Ubiana, pelo direcionamento, dedicação, suporte e atenção.

Obrigada Professora Danielle da Glória por ter me apresentado e incentivado a participar no curso de Especialização.

Obrigada a todos os professores da Especialização que contribuíram durante minha formação e aos meus colegas de turma pelo apoio mútuo durante o curso.

Obrigada a Dai, Msc. João Vieira, Prof<sup>a</sup>. Dr<sup>a</sup>. Ubiana pelas contribuições neste trabalho nos experimentos fenotípicos.

Obrigada ao meu amigo Raffi, quem generosamente me ajudou na análise inicial dos dados.

Obrigada à Universidade por conceder um curso de grande excelência e por ter me concedido a bolsa de estudos.

Obrigada as minhas colegas de laboratório Maione e Patrícia pela compreensão e apoio durante meus trabalhos.

Gostaria de agradecer especialmente a meus pais por todo o apoio, amor e confiança.

## RESUMO

Microrganismos endofíticos têm recebido notoriedade por seu potencial como agentes promotores de desenvolvimento vegetal, mesmo sob condições de estresse. A partir disso existe um interesse de aplicação prática desses microrganismos para agricultura sustentável, cultivos em solos salinos e restauração de áreas degradadas por poluentes orgânicos e metais pesados. Para que essa aplicação seja possível, além da comprovação dos efeitos benéficos às plantas, são necessários estudos de biossegurança para garantir que os endofíticos não são patogênicos ou produtores de toxinas de efeito citotóxico e, conseqüentemente, são seguros para todo o ecossistema. A caracterização dos efeitos benéficos dos microrganismos sobre as plantas pode ser feita por meio de estudos fenotípicos. Entretanto, a análise genômica também pode ser utilizada na caracterização como uma ferramenta complementar para predição de características de interesse. Este estudo buscou analisar o genoma de seis bactérias isoladas de plantas de milho: *Ochrobactrum* sp. CNPMS2088, *Bacillus megaterium* UFMG50, *Klebsiella pneumoniae* UFMG51, *Pantoea ananatis* UFMG54, *Microbacterium* sp. UFMG61 e *Pseudomonas putida* UFMG81 que foram investigados com intuito de explorar o potencial genômico e compará-lo ao potencial previamente analisado em experimentos *in vitro* e *in vivo*. Foram investigadas funções relacionadas à: promoção do crescimento de plantas; capacidade de colonizar os tecidos vegetais; potencial de controle de agentes fitopatogênicos; versatilidade fisiológica e metabólica, característica a qual pode favorecer adaptabilidade dessas bactérias às variações ambientais; virulência e resistência a fármacos, importantes para avaliação da biossegurança. Os resultados da análise genômica reforçam o potencial de uso das bactérias estudadas como bioinoculantes já observado anteriormente em experimentos *in vivo* e *in vitro*. Em ambos estudos, as bactérias se destacaram em características como: solubilização de fósforo, produção de fito-hormônios e produção de sideróforos. Considerando a análise genômica, *K. pneumoniae* UFMG51 foi a bactéria que mais se destacou por possuir o maior número de CDS associadas a características de interesse a promoção de crescimento de plantas como fixação de nitrogênio e metabolismo de P, além de demonstrar potencial para biorremediação devido ao total de CDS associados a degradação de compostos aromáticos e resistência a metais pesados. Dessa

maneira, a análise genômica permitiu explorar o potencial das bactérias para funções de interesse que ainda não foram experimentadas em laboratório, o que pode servir de base para orientar futuros trabalhos para comprovação fenotípica dessas características de interesse para bioinoculantes.

Palavras-chave: endofítico, bactérias, bioinoculantes, Bactérias Promotoras de Crescimento em Plantas.

## ABSTRACT

Endophytic microorganisms have received notoriety for their potential as promoters of plant development, even under stress conditions. Because of it, there is an increasing interest in the practical applications of these microorganisms for sustainable agriculture, cultivation in saline soils and restoration of areas degraded by organic pollutants and heavy metals. For it to be possible, in addition to proving the beneficial effects on plants, biosafety studies are necessary to ensure that endophytes are not pathogenic or produce toxins with a cytotoxic effect, which therefore, would be safe for the entire ecosystem. The characterization of the beneficial effects of microorganisms on plants can be done through phenotypic studies. However, genomic analysis can also be applied in the characterization as a complementary tool for searching for potential features of interest. This study sought to analyze the genome of six bacteria isolated from corn plants: *Ochrobactrum* sp. CNPMS2088, *Bacillus megaterium* UFMG50, *Klebsiella pneumoniae* UFMG51, *Pantoea ananatis* UFMG54, *Microbacterium* sp. UFMG61 and *Pseudomonas putida* UFMG81. These bacteria were investigated in order to explore the genomic potential and compare it to the previous results from *in vitro* and *in vivo* experiments. The genomics functions investigated are related to: plant growth promotion; ability to colonize plant tissues; potential to control phytopathogenic agents; physiological and metabolic versatility, which is a characteristic that can favor adaptability of these bacteria to environmental variations; virulence and drug resistance, which is an important assessment for biosafety. The results of the genomic analysis reinforce the potential of using the bacteria studied as bioinoculants already observed in the previous *in vivo* and *in vitro* experiments. The bacteria stood out in the genomic and phenotype studies in characteristics such as: phosphorus solubilization, production of phytohormones and production of siderophores. Considering the genomic analysis, *K. pneumoniae* UFMG51 was the bacteria which had the largest number of CDS associated with features of interest in plant growth promotion such as nitrogen fixation and P metabolism, in addition to demonstrating potential for bioremediation due to the total number of CDS associated to the degradation of aromatic compounds and resistance to heavy metals. Hence, genomic analysis allowed to explore the potential of bacteria for a larger variety of features of interest which have not yet been tested in the laboratory. This type of analysis is an important tool to guide future work to design phenotypic tests to proof



whether these microorganisms are effective in characteristics of interest to be applied as bioinoculants.

Keywords: endophytic, bacteria, bioinoculants, Plant Growth Promoting Bacteria.

## LISTA DE ABREVIATURAS E SIGLAS

BPCP	Bactérias promotoras do crescimento de plantas
CDS	Sequência codificante, do inglês <i>Coding Sequences</i>
P	Fósforo
PGPB	Bactérias promotoras do crescimento de plantas, do inglês <i>Plant Growth Promoting Bacteria</i>

## SUMÁRIO

1. INTRODUÇÃO	10
2. OBJETIVOS	11
2.1 Objetivo geral	11
3. REFERENCIAL TEÓRICO	11
3.1 Microrganismos endofíticos	11
3.2 Importância das bactérias endofíticas	13
3.2.1 <i>Regulação de fito-hormônios</i>	14
3.2.2 <i>Aumento na absorção de nutrientes</i>	15
3.3 Aplicações biotecnológicas de endofíticos	16
3.3.1 <i>Bioinoculantes para agricultura</i>	16
3.3.2 <i>Biorremediação de solos contaminadas</i>	18
3.4 Genômica de procaríotos	18
4. MATERIAL E MÉTODOS	20
4.1 Referencial teórico	20
4.2 Proveniência das bactérias	20
4.3. Sequenciamento genômico, montagem, anotação e busca por características de interesse	21
4.4. Identificação taxonômica	21
5.1. Visão geral da montagem dos genomas e taxonomia das bactérias	22
5.2. Perfil genômico funcional das bactérias promotoras do crescimento de plantas	25
5.2.1. <i>Promoção do crescimento vegetal</i>	27
5.2.3 <i>Biocontrole</i>	32
5.2.4 <i>Versatilidade metabólica</i>	34
5.2.5 <i>Virulência e resistência</i>	36
5.3 <i>Promoção de crescimento e colonização vegetal</i>	38
6. CONCLUSÃO	39
7. REFERÊNCIAS BIBLIOGRÁFICAS	41
8. FIGURAS E TABELAS SUPLEMENTARES	54

## 1. INTRODUÇÃO

Interações entre microrganismos endofíticos e plantas superiores são encontradas de forma ubíqua na natureza. Nas interações mutualistas, os endófitos conferem vantagens nutricionais e adaptativas às plantas em diversos ambientes com estresses bióticos e abióticos em troca de proteção e fotossintatos. Sendo assim, diariamente, pesquisas sobre microrganismos endofíticos têm se multiplicado para compreender a influência desses microrganismos sobre o aumento da absorção de nutrientes e sua influência sobre o crescimento vegetal.

Os microrganismos endofíticos são fungos ou bactérias que estão alocados nas regiões intercelulares das plantas nos mais diversos tecidos. As bactérias endofíticas são também conhecidas na literatura como bactérias promotoras de crescimento de plantas (do inglês *Plant Grow Promoting Bacteria*, PGPB). Tal denominação provém dos efeitos diretos e indiretos exercidos sobre as plantas que colonizam. Dentre os efeitos diretos estão: maior fixação de nitrogênio e solubilização de P, produção de sideróforos e de fito-hormônios. Entre os efeitos indiretos, estão a ação de biocontrole de pragas por meio de metabólitos produzidos.

A agricultura convencional devido ao uso descontrolado de fertilizantes minerais, é associada a problemas de poluição atmosférica, eutrofização e diminuição da fertilidade do solo. Além disso, substâncias agroquímicas de diversas classes têm sido pauta de discussão acerca de seus efeitos negativos sobre ecossistemas e a saúde humana. Por isso, pesquisadores têm analisado o emprego de microrganismos endofíticos, especialmente os rizosféricos, como potenciais biofertilizantes e biopesticidas. Organismos endofíticos, podem impactar positivamente as plantas e elevar a fertilidade do solo, ao contrário dos fertilizantes químicos que elevam o rendimento das plantas até um limiar e que quando usados de forma inadequada comprometem a fertilidade do solo.

Além disso, microrganismos que interagem com as plantas também podem ser fundamentais a processos de biorremediação e adaptação a ambientes desafiantes à

sua fisiologia, como em ambientes de alta salinidade, estresse hídrico e de temperaturas não ótimas. Isto porque os microrganismos são capazes de imobilizar, inativar e sequestrar poluentes orgânicos e metais pesados, o que além de contribuir para a descontaminação do meio, também diminui os efeitos tóxicos sobre a planta. Por sua vez, a contribuição dos microrganismos para o crescimento de plantas em ambientes desafiantes é resultado de sua capacidade de modular o próprio metabolismo e muitas vezes o das plantas de forma a garantir o equilíbrio fisiológico. Essas características são relevantes para a restauração de ambientes contaminados e também são promissoras a expansão da agricultura em áreas não cultiváveis devido a salinidade, temperatura e condições hídricas.

O desenvolvimento de bioinoculantes para as plantas requer inicialmente a seleção de microrganismos com potencial de promoção de crescimento. Sendo assim, neste estudo, seis bactérias isoladas de planta de milho previamente selecionadas por testes fenotípicos tiveram seus genomas analisados com intuito de investigar o potencial para promoção de crescimento e prever características gênicas que possam direta ou indiretamente estar relacionadas a esse potencial.

## **2. OBJETIVOS**

### **2.1 Objetivo geral**

Identificar o potencial para a promoção do crescimento de plantas de bactérias endofíticas isoladas de milho (*Zea mays*) utilizando análise de genômica

## **3. REFERENCIAL TEÓRICO**

### **3.1 Microrganismos endofíticos**

Os microrganismos endofíticos são fungos ou bactérias que estão alocados nas regiões intercelulares das plantas superiores, Gimnospermas e Angiospermas, nos mais diversos tecidos sem causar prejuízos (PIMENTEL *et al.*, 2011). As relações entre organismos endofíticos e plantas são encontradas de forma ubíqua no meio ambiente, mas a natureza da interação ecológica é diversificada, incluindo comensalismo, mutualismo e antagonismo (HALLMANN *et al.*, 1997). O tipo de

interação ecológica estabelecida depende de diversos fatores bióticos e abióticos que variam de acordo com o genótipo da planta e dos microrganismos, condições ambientais e a dinâmica da interação no bioma da planta (HARDOIM *et al.*, 2015). Entre os efeitos positivos descritos na literatura associados aos organismos endofíticos estão diversos mecanismos que influenciam na promoção do desenvolvimento vegetal. Alguns dos seus impactos são: aumento da absorção de nutrientes, produção de fito-hormônios, aumento da tolerância ao estresse biótico e abiótico e aumento da resistência à diversos fitopatógenos (GOUDA *et al.*, 2016; SANTOYO *et al.*, 2016).

Os microrganismos endofíticos podem ser divididos dois subgrupos. O primeiro subgrupo são os endofíticos obrigatórios, os quais dependem do metabolismo da planta para sobrevivência. O segundo subgrupo são os endofíticos facultativos, que passam parte do seu ciclo de vida fora no solo e atmosfera no entorno das plantas (GOUDA *et al.*, 2016; HARDOIM, VAN OVERBEEK e ELSAS, 2008). Entre as bactérias, os filos mais descritos são Actinobacteria, Proteobacteria e Firmicutes; que incluem tanto bactérias gram-positivas, quanto gram-negativas dos gêneros *Achromobacter*, *Acinetobacter*, *Agrobacterium*, *Bacillus*, *Brevibacterium*, *Microbacterium*, *Pseudomonas*, *Xanthomonas* (GOLINSKA *et al.*, 2015).

A colonização das plantas por microrganismos endofíticos depende de fatores tais como: tipo de tecido vegetal, genoma da planta, microorganismo envolvido e fatores ambientais bióticos e abióticos (SUN *et al.*, 2013). O processo de colonização pode ocorrer de duas formas: horizontal e vertical. O processo horizontal ocorre quando um fragmento de folha colonizado ao cair sobre solo libera esporos de fungos e bactérias no ambiente. A partir disso, os microrganismos podem ser atraídas à região da rizosfera por exsudatos da raiz e rizodepósitos ou penetrar nos pelos radiculares. Outros mecanismos de colonização horizontal consistem na quimiotaxia à região aérea pela produção de exsudatos de galhos e folhas ou penetração através de estômatos e feridas no tecido vegetal. O processo de colonização vertical ocorre através das sementes (HARDOIM *et al.*, 2015). Além disso, para as bactérias colonizarem o ambiente da rizosfera antes de chegarem aos tecido internos da planta, elas precisam ser resistentes à luz UV, baixa disponibilidade de nutrientes e

dissecação, condições presentes da superfície aérea vegetal; penetração e colonização de flores e frutos (COMPANT, CLEMENT e SESSITSCH, 2010).

A identificação de genes relacionados a atividades de promoção de crescimento das plantas é de suma importância no estudo de microrganismos endofíticos. Genes de microrganismos endofíticos podem estar associados a diversos fatores de crescimento de plantas, por exemplo: grupos de genes pirroloquinolina quinona (*pqq*) são associados à solubilização de fósforo no solo às plantas (RODRIGUEZ, et al., 2006); genes nitrogenase (*nif*) são importantes às funções de fixação de nitrogênio pelas bactérias endofíticas (GLICK, 2012). Esses genes dentre outros auxiliam na compreensão papel ecológico do microrganismo nas plantas, além de também permitir o desenvolvimento de estudos de biofertilizantes seguros (GUPTA, et al., 2014).

Apesar do importante papel desempenhado por bactérias e fungos endofíticos às plantas, o mecanismo da interação benéfica entre microrganismo-planta e a evolução dessa relação ecológica ainda é pouco compreendida (COMPANT *et al.*, 2016). O enfoque deste trabalho foi sobre as bactérias endofíticas.

### **3.2 Importância das bactérias endofíticas**

As bactérias endofíticas podem atuar sobre o crescimento das plantas e por isso são amplamente conhecidas como BPCP (tradução do inglês para bactérias promotoras de crescimento – *Plant Growth Promoting Bacteria*, *PGPB*). Este efeito está associado a mecanismos diretos e indiretos (Alguns exemplos de mecanismos de ação direta compreendem: facilitação de aquisição de nutrientes como fósforo, nitrogênio e ferro; modulação de fito-hormônios por meio da produção de auxina, citocinina e giberilina ou inibição da síntese do hormônio etileno pela planta. Dentre os mecanismos de ação indireta está a mitigação de efeitos nocivos causados às plantas durante um episódio de infecção fúngica ou bacteriana (WEYENS *et al.*, 2013).

### 3.2.1 Regulação de fito-hormônios

A produção e modulação de fito-hormônios por bactérias endofíticas em plantas promove o crescimento vegetal e a resistência a condições de estresse (VAISHNAV *et al.*, 2019).

Estudos observaram que a bactéria endofítica *Bacillus amyloliquefaciens* RWL-1 é capaz de produzir ácido abscísico (ABA) e melhora a resistência de plantas de arroz sob estresse salino (SHAHZAD *et al.*, 2017). Nesse experimento a inoculação de RWL-1 contribuiu também para o aumento da produção de aminoácidos essenciais pela planta, além de ter induzido a produção de ácido salicílico, hormônio essencial para o aumento da tolerância de plantas de arroz ao estresse salino. Outro exemplo de produção de fito-hormônios por endobactérias foi observado em *Sphingomonas* sp. LK11 capazes de produzir giberilinas, o que contribuiu com um maior crescimento de tomateiros e uma maior concentração de clorofila nas plantas (KHAN, MAZID e QUDDUSI, 2014).

A tecnologia do DNA recombinante também já foi utilizada em bactérias endofíticas com a introdução de genes que as possibilitem de oferecer maiores benefícios às plantas. Por exemplo, bactérias *Sinorhizobium meliloti* receberam genes de grande taxa de transcrição *ipt* (isopenteniltransferase fosfato adenosina), relevantes ao processo de biossíntese de citoquinina. Ao inocular essas bactérias transformadas em alfafas sob estresse hídrico, as plantas demonstraram maior tolerância ao estresse, além de sofrerem menos efeitos negativos oriundos da condição (XU, LI e LUO, 2012).

As endobactérias também exercem um importante papel sobre o doseamento de etileno nas plantas. O etileno é um mediador de respostas a estresse nas plantas, sendo que uma menor quantidade desse hormônio é importante para o desenvolvimento e crescimento das plantas, enquanto situações de estresse estimulam a produção e acúmulo de quantidades maiores desse hormônio que resulta na inibição do crescimento das plantas (GUPTA, SETH e SHARMA, 2016). Apesar disso, bactérias endofíticas conseguem modular a produção de etileno pelas plantas por produzirem enzimas que atuam sobre uma molécula precursora do etileno, o ácido 1-aminociclopropano-1-carboxílico, amplamente conhecido como ACC devido a abreviação do nome em inglês de "*1-Aminocyclopropane-1-carboxylic acid*". A partir



da enzima ACC desaminase das endobactérias, o ACC é clivado em amônia e  $\alpha$ -acetobutirato, o que o impede de ser convertido em etileno (KUMARI *et al.*, 2016). Além disso, a redução da produção de etileno nas plantas também funciona como uma estratégia de colonização das microrganismos endofíticos, uma vez que uma maior concentração de etileno nas plantas reduz a eficiência de colonização das endobactérias na planta hospedeira (ALI, CHARLES e GLICK, 2014).

Bactérias endofíticas também são capazes de influenciar os níveis de auxinas nas plantas. Em um estudo que buscava isolar e analisar o efeito de bactérias endofíticas produtoras de AIA (ácido indolacético), um hormônio da classe da auxina, sobre plantas de arroz verificou que ao aplicar bactérias isoladas que produziam AIA, as plantas alcançaram uma altura 6,7% maior, uma diferença de peso seco dos grãos de arroz de 99,54% comparado ao controle negativo e maior peso seco das raízes em comparação às plantas que não receberam inóculos. Nesse estudo foi observado que diferentes organismos endofíticos produtores de AIA podem conferir diferentes efeitos sobre as plantas. As bactérias associadas ao ganho de altura das plantas foram: *Pseudomonas fragi*, *Bacillus cereus* e *Rhizobium*; ao aumento do peso dos grãos secos foram: *Bacillus aerius*, *Pseudomonas fragi* e *Bacillus cereus*; ao ganho de peso das raízes secas: *Bacillus amyloliquenfaciens* (SUSILOWATIA *et al.*, 2018).

### 3.2.2 Aumento na absorção de nutrientes

#### **Nitrogênio**

Bactérias endofíticas são mais eficientes na fixação de nitrogênio em relação às rizosféricas. Tal fato se dá devido ao ambiente com baixas concentrações de oxigênio no interior dos tecidos vegetais, condição essencial para os processos de fixação de nitrogênio (VAISHNAV *et al.*, 2019). Em leguminosas, a fixação de nitrogênio nas plantas ocorre em nódulos formados nas raízes, onde microrganismos endofíticos reduzem o nitrogênio atmosférico a amônia através da enzima nitrogenase. Estudos metagenômicos reforçam esse papel da microbiota endofítica uma vez que genes relacionados a fixação de nitrogênio foram encontrados no genoma desses microrganismos (SESSITSCH *et al.*, 2012).

#### **Fósforo**

Apesar da maior parte do fósforo presente no solo não estar disponível às plantas, esse elemento tem um papel fundamental na produção de biomassa vegetal. A maior parte do fósforo no solo está sob a forma mineral não solúvel, que é inacessível para a absorção das plantas. Dessa maneira, as plantas utilizam de relações simbióticas com microrganismos que estão presentes tanto no solo ao redor das raízes quanto no tecido vegetal interno para a solubilização de fósforo. Esse processo depende da liberação de metabólitos dos microrganismos, como ácidos orgânicos, os quais atuarão quelando os cátions ligados aos fosfatos, tornando esses em uma forma solúvel e disponível para absorção das plantas (LACAVA e AZEVEDO, 2013; KHAN, *et al.*, 2009).

## **Ferro**

O ferro é um micronutriente essencial às plantas mas que na maioria das vezes encontra-se na sua forma insolúvel no solo como hidróxido de ferro. Esse elemento é um importante cofator de diversas reações enzimáticas das plantas. Apesar disso, microrganismos do solo são capazes de quelar o ferro através de sideróforos e torná-lo solúvel para ser absorvido pelas plantas (CHOUDHARY *et al.*, 2015). Estudos moleculares são relevantes na demonstração da importância dos microrganismos endofíticos para a absorção de ferro pelas plantas. Em um estudo em que se comparava a acumulação de ferro em plantas de arroz, sendo que um grupo recebeu uma linhagem de mutante de endofíticos *Streptomyces sp.* GMKU 3100, o qual não produzia sideróforos, enquanto o outro grupo recebeu a linhagem selvagem, pôde-se constatar o importante papel dos endofíticos produtores de sideróforos na absorção de ferro e os benefícios às plantas (RUNGIN *et al.*, 2012). A produção de sideróforos por endofíticos está associada também a um mecanismo de controle biológico de fitopatógenos, uma vez que diminui a disponibilidade de ferro a esses (CHOUDHARY *et al.*, 2015).

### **3.3 Aplicações biotecnológicas de endofíticos**

#### **3.3.1 Bioinoculantes para agricultura**

A exploração do potencial benéfico dos microrganismos endofíticos às plantas têm sido abordado como uma ferramenta sustentável para a agricultura. Sendo assim, há produtos desenvolvidos conhecidos como bioinoculantes em que linhagens de bactérias, fungos filamentosos e leveduras que possuem atividade comprovada no crescimento e desenvolvimento vegetal são inoculados no solo ou via semente, desempenhando um papel importante tanto na saúde do solo quanto para as plantas (Tallapragada e Seshagiri, 2017).

Para o desenvolvimento de bioinoculantes, é necessário que os microrganismos sejam benéficos às plantas e seguros ao ecossistema. Testes em laboratório podem ser limitados para a observação do potencial bioinoculante de determinadas linhagens por não permitirem a completa expressão das características de promoção de crescimento que seriam expressas na natureza. Durante o desenvolvimento de bioinoculantes, é necessário selecionar microrganismos que sejam resistentes a diversos tipos de estresse como: acidez, alcalinidade e temperaturas altas e baixas. Para melhorar a resistência aos diferentes tipos de estresse, os microrganismos podem sofrer manipulação genética ou serem selecionados a partir de adaptações fenotípicas ao inserir o estresse no ambiente e selecionar as colônias que se adaptaram bem às condições (PAIKRAY e MALIK, 2012).

Os produtos bioinoculantes prometem promover o desenvolvimento vegetal e contribuir com a saúde do solo baseado nas propriedades dos microrganismos de: fixação de nitrogênio, quelação de ferro, solubilização de fosfatos, inibição de microrganismos não benéficos, aumento da resistência à pestes, decomposição da matéria vegetal que contribui diretamente com o aumento de matéria orgânica no solo. Todas essas propriedades atuarão diretamente na melhora da qualidade do solo e crescimento vegetal, o que conseqüentemente impacta na quantidade e qualidade das colheitas e diminui a demanda pelo uso de pesticidas e fertilizantes químicos (BHARDWAJ et al., 2014).

Em um bioinoculante patenteado nos Estados Unidos, os microrganismos utilizados foram: *Pseudomonas fluorescens*, *Pseudomonas striata*, *Bacillus polymyxa*, *Bacillus subtilis*, *Azospirillum*, *Azotobacter*, *Rhizobium*, *Trichoderma herzianum* e *Trichoderma viride*, *Saccharomyces cerevisiae* e *Lactobacillus*. O bioinoculante

apresenta  $10^{10}$  unidades formadoras de colônias inicialmente e é adequado a qualquer tipo de lavoura (PAIKRAY e MALIK, 2012).

Para a implementação do uso de bioinoculantes é importante averiguar sua segurança biológica. Estudos de biossegurança devem analisar os potenciais impactos de metabólitos dos bioinoculantes liberados no ambiente, a viabilidade dos microrganismos e seus efeitos sobre os diferentes níveis tróficos. Dessa maneira, os microrganismos utilizados como biofertilizantes, além de promoverem o crescimento das plantas, deve também ser seguro ao ambiente, plantas e animais (VÍLCHEZ *et al.*, 2016).

### 3.3.2 Biorremediação de solos contaminadas

Outra possibilidade de aplicação biotecnológica para endofíticos é a biorremediação. Esses microrganismos são capazes de degradar diferentes classes de poluentes orgânicos, mineralizando-os e assim, contribuindo inclusive de forma a aumentar a fertilidade do solo. Além disso, a mineralização dos compostos poluentes no tecido vegetal a partir dos organismos endofíticos reduz o efeito da toxicidade sobre as plantas e melhora o metabolismo vegetal mesmo sob condições desfavoráveis (AFZAL *et al.*, 2013; IJAZ *et al.*, 2016). Plantas quando em associação com endofíticos tendem a absorver mais contaminantes do solo, tanto orgânicos quanto minerais, melhorando o desempenho de projetos de fitorremediação sem comprometer a saúde da planta. Poluentes orgânicos podem ser absorvidos por plantas e degradados por endofíticos sem que ocorra o acúmulo nos tecidos vegetais. Em solos contaminados com metais, os endofíticos diminuem a toxicidades desses metais às plantas através, por exemplo, do sistema de sequestro de metais e aumentam a translocação dos metais para tecidos aéreos das plantas (IJAZ *et al.*, 2016).

## 3.4 Genômica de procariotos

O sequenciamento de genomas tem se tornado uma prática corriqueira em laboratórios e fundamental para a classificação e estudo de microrganismos. O sequenciamento permite aplicações práticas como modelagem metabólica,

bioforense, epidemiologia de doenças metabólicas e biovigilância. Além disso, a genômica inovou os estudos de taxonomia e sistemática dos procariotos. Anteriormente, espécies de bactérias eram identificadas a partir de sua morfologia, testes bioquímicos quanto às fontes de carbono e nitrogênio utilizadas e através da coloração de Gram. A partir do desenvolvimento da hibridização do DNA essa técnica passou a ser a principal fonte de classificação de procariotos por mais de cinquenta anos. Porém, com a tecnologia do rápido sequenciamento do DNA, passou-se a utilizar como padrão filogenético de bactérias comparações do gene de RNA ribossomal 16s e outros marcadores filogenéticos, regiões conservadas dentro de espécies. A partir disso, o estudo da genômica quantifica a variação dos genomas, comparando o tamanho de regiões conservadas em relação ao tamanho do genoma entre microrganismos (LAND *et al.*, 2015).

A análise de dados na genômica consiste de dois processos: identificação de genes e associação de função ao gene. A identificação de genes busca identificar as seções codificantes do genoma. Caso um gene transcreva uma proteína, a análise genômica deve determinar a sequência codificante do gene, onde a transcrição e a tradução da proteína se inicia e termina. Para isso, existem programas que oferecem uma plataforma para predizer seções do genoma de procariotos que possuem genes codificantes de proteínas. Alguns dos programas mais populares para essas funções são: prodigal, glimmer e geneMark (SETUBAL, ALMEIDA e WATTAM, 2018). Uma vez que a sequência codificante de proteínas é encontrada, há outros programas de utilizados como ferramentas para prever a função da proteína codificada por aquele gene. Essa associação geralmente é feita por meio de comparação por similaridade de sequências dispostas em bancos de dados e que já têm sua função conhecida. Entretanto, em alguns casos, os programas utilizados para essa função podem ter dificuldade em associar a sequência de genes a uma função, designando o produto como uma proteína hipotética. Isso pode ocorrer em dois cenários: quando uma sequência de genes é similar a outras sequências no banco de dados, mas não possuem uma função conhecida; quando a sequência de nucleotídeos não é suficientemente similar a outras sequências registradas no banco de dados. Alguns exemplos de plataformas usualmente utilizadas para essa função são: IMG, NCBI, RAST (SETUBAL, ALMEIDA e WATTAM, 2018; AZIZ *et al.*, 2008) e Prokka (SEEMANN, 2014).

O sequenciamento e análise comparativa da genômica permitiram ampliar o conhecimento do potencial genético, a ecologia e a evolução dos microrganismos. Esse conhecimento permite compreender o mecanismo molecular dos microrganismos endofíticos em diversos quesitos, desde sua colonização nas plantas, até os tipos de interação sinérgicas com os hospedeiros. Dentre alguns dos genes conhecidos de interação de endofíticos com plantas hospedeiras estão: genes de produção de fito-hormônios e de atividade antimicrobiana. A maior parte desses genes de interação são compartilhados com outras bactérias. A partir disso, a análise genômica permite elucidar as adaptações dos endofíticos nas plantas, como também compreender a interação desses microrganismos às plantas e explorar seu potencial para aplicações biotecnológicas (FRANK, 2011).

## **4. MATERIAL E MÉTODOS**

### **4.1 Referencial teórico**

Para a construção do referencial teórico foi utilizada a plataforma do Google Acadêmico. Na barra de busca desse website foram pesquisados termos referentes à temática de organismos endofíticos na promoção do desenvolvimento vegetal e suas aplicações na agricultura. Todos os termos utilizados na busca estão em inglês para permitir um maior acesso à literatura acadêmica. Alguns dos termos pesquisados foram: *endophytics*, *PGPB*, *Plant-Growth-Promoting-Bacteria*, *plant-growth*, *bioinoculants*.

### **4.2 Proveniência das bactérias**

As bactérias analisadas neste estudo foram isoladas do tecido interno de raízes, folhas ou seiva de milho nos trabalhos de Vieira (2015) e Abreu et al (2017). Essas bactérias foram pré-selecionadas *in vitro* quanto ao potencial de solubilizar e mineralizar fósforo, produção de fitohormônios, produção de sideróforos, formação de biofilme, colonização radicular e produção de EPS, atividade surfactante e promoção do crescimento de milho em casa de vegetação nos trabalhos de Vieira (2015), Silva (2017) e Silva (2018).

### **4.3. Sequenciamento genômico, montagem, anotação e busca por características de interesse**

Para a extração do DNA, foi utilizado o kit de purificação de DNA genômico Promega™ (Madison, USA), no qual seguindo recomendações do kit as culturas de bactérias foram incubadas em caldo TSA por 24 h a temperatura de 28°C sob agitação de 180 rpm. Os genomas foram sequenciados através do protocolo de pareamento de sequências finais na plataforma Illumina MiSeq™. Os softwares A5 e CAP3 foram utilizados para excluir sequências de baixa qualidade, corrigir erros entre os fragmentos gerados e melhorar a qualidade do alongamento da sequência a partir da montagem dos fragmentos. As sequências foram formadas e analisadas usando o protocolo A5 (Coil, et al., 2014). A anotação das sequências foi feita utilizando Prokaryotic Genome Annotation System (PROKKA) (Seemann, 2014) e algoritmos de RAST (RAST-TK) (Brettin et al., 2015). As anotações genômicas foram submetidas na plataforma online de referências genômicas RAST (Rapid Annotation using Subsystem Technology) versão 2.0 para identificar as possíveis funções dos genes dos microrganismos analisados (SILVA, 2018). As anotações feitas no por Prokka e no RAST foram utilizadas para a construção de uma tabela comparativa entre as bactérias para verificar a presença ou ausência de genes e proteínas codificantes no genoma relevantes para a promoção do crescimento de plantas, colonização dos tecidos vegetais, biocontrole, versatilidade metabólica, virulência e resistência. Para a definição dessas categorias e dos genes ou CDS baseou-se nos seguintes trabalhos Silva (2019), Suarez, *et al.*(2019), Gupta, *et al.*(2014), Nelkner, *et al.*(2019), Pinski, *et al.*(2019), além categorização feita pela plataforma Rast.

### **4.4. Identificação taxonômica**

Para a análise taxonômica foi feita atribuição de taxonomia para genes marcadores filogenéticos moleculares: genes do rRNA 16 S e 23 S, girase B (*gyrB*) e da subunidade  $\beta$  da RNA polimerase (*rpoB*) utilizando o alinhamento múltiplo do Blast contra o banco de dados do Genbank e através da análise do perfil de proteínas das culturas bacterianas via espectrometria de massa de tempo de voo de dessorção / ionização assistida por matriz (MALDI-TOF). Uma pequena porção de células bacterianas cultivadas em meio TSA foi transferida para poços da placa e aço

inoxidável do MALDI por esfregaço (Bruker, Alemanha). Posteriormente, as amostras bacterianas foram cobertas com 1µL de ácido fórmico (P.A). Após secar à temperatura ambiente, cobriu-se o esfregaço com 1µL da matriz HCCA (ácido a-ciano-4-hidroxicinâmico, Sigma-Aldrich, Poland). Depois que a matriz foi seca a temperatura ambiente, a placa foi inserida no espectrômetro para medição automatizada e interpretação de dados. Os espectros de massa foram processados com o pacote de software Bruker Daltonik MALDI Biotype (Bruker, Alemanha). Os resultados foram obtidos a partir das 20 melhores correspondências de identificação, juntamente com valores de confiança que variam de 0,00 a 3,00. De acordo com os critérios recomendados pelo fabricante, valores abaixo de 1,70 não permitem uma identificação confiável; valores entre 1,70 e 1,99 permitem a identificação ao nível do gênero; entre 2,00 e 2,29 significa identificação altamente provável no nível do gênero e identificação provável ao nível da espécie; e um valor superior a 2,30 (2,30-3,00) indica identificação altamente provável ao nível da espécie.

## 5. RESULTADOS E DISCUSSÃO

### 5.1. Visão geral da montagem dos genomas e taxonomia das bactérias

**Tabela 1:** Características gerais dos genomas mediante sequenciamento Illumina Miseq, montagem usando o protocolo A5 e anotação no Rast.

Característica genômica	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81	CNPMS2088
Tamanho (Mb)	6,08	6,53	4,98	3,72	6,07	5,53
Conteúdo de GC (%)	37,5	56,8	53,3	68,1	64	53,8
Número de Contigs	54	3	26	35	74	74
Número de CDS	6539	6464	5057	3652	5470	5583

Os resultados expostos na Tabela 1 comparam: cobertura, tamanho, conteúdo GC (guaninas e citosinas), número de contigs, número de CDS. O tamanho dos genomas variou entre 6,53 Mb em UFMG51 e 3,72 Mb em UFMG61. Quanto ao número total de sequências codificantes de proteínas, o maior número foi observado em UFMG50, segundo maior genoma, e o menor número em UFMG61, que também possui o menor tamanho de genoma. Como o genoma bacteriano é caracterizado pela



alta concentração de genes, isso facilita a identificação de CDS, o que também está correlacionado ao tamanho do genoma (Donkor, 2013). Dessa maneira, ao comparar o tamanho do genoma de cada uma das bactérias analisadas, era esperado que UFMG 50 apresentasse uma das maiores quantidades de CDS, enquanto UFMG61 apresentasse a menor. Em relação ao conteúdo de bases GC no genoma, UFMG61 teve o maior (68,1%) e UFMG50 o menor (37,5%). Existe uma variação do conteúdo de GC entre espécies de bactérias, porém essa variação pode também estar atrelada a pressões seletivas negativas (purificadoras) do ambiente (Bohlin *et al.*, 2017). Em relação ao número de contigs, as bactérias UFMG81 e CNPMS2088 apresentaram os maiores números, 74, enquanto UFMG51 teve apenas três contigs. Apesar da diferença na quantidade de contigs formados nas montagens, a relação entre tamanho do genoma e número de contigs demonstra que CNPMS2088 apresenta fragmentos com uma média de tamanho de 74,7 kb e UFMG81 de 82,03 kb, o que representa uma quantidade significativa de genes por contig e, conseqüentemente, atesta a qualidade das montagens, considerando que o tamanho médio de genes bacterianos é de 1 kb e a densidade de região codificadora é de 85% (Smits, 2019).

**Tabela 2:** Identificação de bactérias utilizando a abordagem multivlocus de genes marcadores moleculares e espectrometria de massa.

Bacteria	Gene	Identity %			Maldi Tof	Score
UFMG50	16S	<i>Bacillus aryabhatai</i> – 100%	<i>Bacillus megaterium</i> – 99%	<i>Bacillus flexus</i> -99%		
UFMG50	23S	<i>Bacillus megaterium</i> – 100%	<i>Bacillus flexus</i> -98%			
UFMG50	gyrB	<i>Bacillus megaterium</i> – 99%			<i>B. megaterium</i>	2.1
UFMG50	rpoB	<i>Bacillus megaterium</i> – 99%				
UFMG51	16S	<i>Klebsiella pneumoniae</i> -100%	<i>Klebsiella variicola</i> -99%			
UFMG51	23S	<i>Klebsiella pneumoniae</i> – 99%	<i>Klebsiella variicola</i> -99%	<i>Enterobacter aerogenes</i> -99%	<i>K. pneumoniae</i>	2.3
UFMG51	gyrB	<i>Klebsiella pneumoniae</i> – 99%	<i>Klebsiella variicola</i> -99%			
UFMG51	rpoB	<i>Klebsiella variicola</i> -99%	<i>Klebsiella pneumoniae</i> – 99%	<i>Enterobacter aerogenes</i> -98%		
UFMG54	16S	<i>Pantoea ananatis</i> -100%	<i>Pantoea vagans</i> -99%	<i>Pantoea agglomerans</i> – 99%		
UFMG54	23S	<i>Pantoea ananatis</i> -99%	<i>Pantoea vagans</i> -98%	<i>Pantoea stewartii</i> – 99%		
UFMG54	gyrB	<i>Pantoea ananatis</i> -99%			<i>P. ananatis</i>	2.0
UFMG54	rpoB	<i>Pantoea ananatis</i> -99%				
UFMG61	16S	<i>M. hydrocarbonoxydans</i> -99%	<i>Microbacterium foliorum</i> -99%	<i>M. paraoxydans</i> -97%		
UFMG61	23S	<i>Microbacterium foliorum</i> -99%	<i>Microbacterium pygmaeum</i> -97%	<i>M. paraoxydans</i> -97%		
UFMG61	gyrB	<i>M. phyllosphaerae</i> -91%			<i>Microbacterium sp.</i>	1,9
UFMG61	rpoB	<i>Microbacterium</i> -95%				
UFMG81	16S	<i>Pseudomonas putida</i> – 99 %	<i>P. entomophila</i> – 99 %	<i>Pseudomonas monteilii</i> – 99 %		
UFMG81	23S	<i>Pseudomonas putida</i> – 99 %	<i>P. entomophila</i> – 99 %	<i>Pseudomonas monteilii</i> – 99 %		
UFMG81	gyrB	<i>Pseudomonas putida</i> – 92 %	<i>Pseudomonas entomophila</i> – 93 %		<i>P. putida</i>	2.0
UFMG81	rpoB	<i>Pseudomonas putida</i> – 97 %	<i>Pseudomonas entomophila</i> – 97 %			
CNPMS2088	16S	<i>O. pseudogrignonense</i> – 100%	<i>Ochrobactrum pituitosum</i> -100 %	<i>O. rhizosphaerae</i> -100%		
CNPMS2088	23S					
CNPMS2088	gyrB	<i>O. pseudogrignonense</i> – 98%	<i>Ochrobactrum pituitosum</i> -98%		<i>not identified</i>	–
CNPMS2088	rpoB	<i>O. pseudogrignonense</i> – 99%	<i>Ochrobactrum pituitosum</i> -98%			

A análise de sequências dos marcadores filogenéticos moleculares, rRNA 16S e 23S, girase B (*gyrB*) e subunidade  $\beta$  da RNA polimerase (*rpoB*), em conjunto com a análise do perfil das proteínas ribossômicas via espectrometria de massa (Maldi-ToF) permitiu a identificação do gênero de todas as seis bactérias e a espécie de quatro delas (Tabela 2).

O perfil de proteínas ribossômicas foi determinante para a resolução taxonômica dos isolados UFMG50, UFMG51, UFMG54 e UFMG81. Para UFMG50 foi observada alta identidade filogenética (superior a 99%) com sequências de três espécies do gênero *Bacillus* sp., não permitindo a definição da espécie. No entanto, a obtenção do valor de confiança de 2,1 na espectrometria de massa na atribuição taxonômica do isolado pôde determinar a espécie como *Bacillus megaterium*. UFMG51 apresentou alta identidade com sequências de duas espécies do gênero *Klebsiella* sp. e uma de *Enterobacter* sp. O valor de 2,3 na espectrometria de massa permitiu definir a espécie como *K. pneumoniae*. Já UFMG54 por meio da identificação taxonômica filogenética, teve como possíveis resultados três espécies do gênero *Pantoea* sp. O alto valor de confiança (2,0) na espectrometria de massa definiu a espécie como *P. ananatis*. Para UFMG61, sete possíveis atribuições taxonômicas foram observadas com o uso dos marcadores filogenética, todas espécies do gênero *Microbacterium* sp. Mas neste caso, o valor de confiança obtido na espectrometria de massa (1,9) não permitiu a definição da espécie. Para tal, é preciso que ele seja superior a 2,0. Sendo assim, UFMG61 foi identificada somente no nível de gênero como *Microbacterium* sp. UFMG81 apresentou identidade superior a 92% entre três espécies do gênero *Pseudomonas* sp. O valor de confiança da espectrometria de 2,0 permitiu definir a espécie como *P. putida*. Por fim, CNPMS2088 apresentou identidade para apenas três dos quatro genes utilizados. O gene que não apresentou identidade foi 23S. Os outros genes, porém apresentaram identidade superior a 98% para duas espécies do gênero *Ochrobactrum* sp. A análise por espectrometria não detectou organismo próximo por isso essa bactéria foi identificada apenas ao nível de gênero (*Ochrobactrum* sp.) baseado nos resultados de identificação dos genes marcadores moleculares.

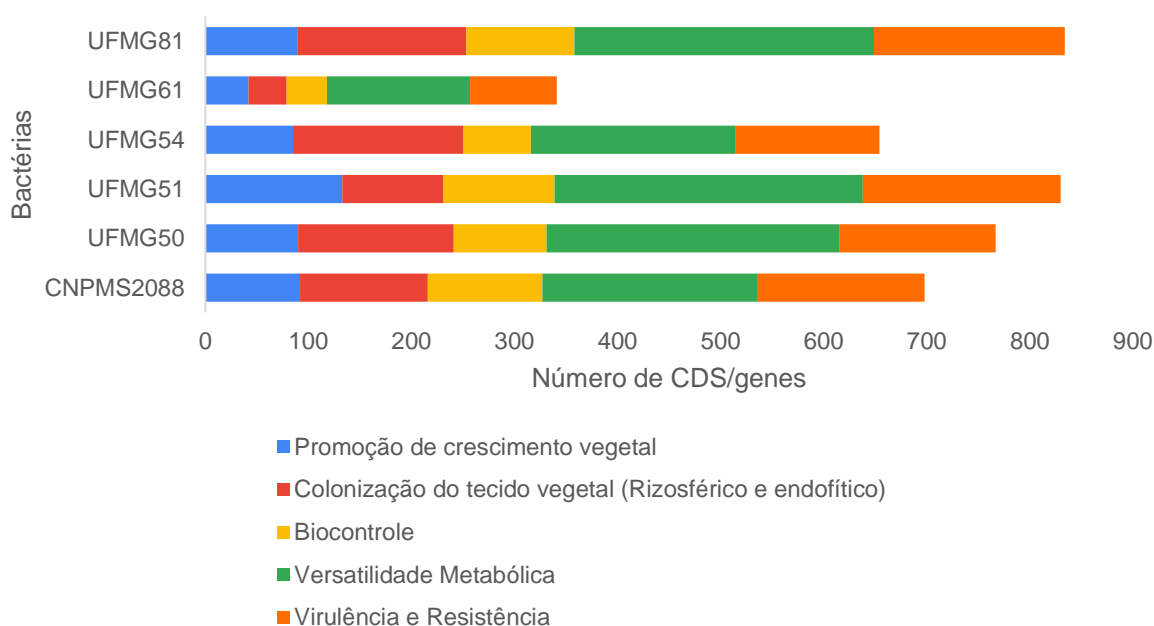
A presença de fagos, pró-fagos, elementos transponíveis e plasmídeos merece destaque por contribuir para a adaptação das bactérias à planta hospedeira e ao ambiente, além de demonstrar o potencial de plasticidade gênica das bactérias, o que está diretamente associada à dinâmica de transferência gênica e biossegurança (GILLIS *et al.*, 2018). Análise na plataforma RAST demonstrou que dentre as bactérias estudadas, aquela com mais CDS associadas a esses elementos móveis foi *Ochrobactrum* sp. CNPMS2088, com 25, enquanto *Microbacterium* sp. UFMG61 não teve CDS para esses elementos (Tabela 3).

**Tabela 3:** Contagem pela plataforma RAST de CDS relacionadas à fagos, pró-fagos, elementos transponíveis e plasmídeos.

	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Fagos, pró-fagos, elementos transponíveis, plasmídeos	25	5	14	8	0	1

## 5.2. Perfil genômico funcional das bactérias promotoras do crescimento de plantas

**Figura 1:** Comparação da distribuição de CDS por categoria (promoção de crescimento de plantas, colonização vegetal, biocontrole, versatilidade metabólica, virulência e resistência).



O genoma das bactérias analisadas neste estudo foi avaliado dentro de cinco grandes categorias: promoção do crescimento de plantas, colonização do tecido vegetal (rizosférica e endofítica), biocontrole, versatilidade metabólica, virulência e resistência (Figura 1). Dentro de cada uma dessas categorias foram listadas CDS e/ou genes relacionados às funções associadas a partir da categorização feita pela plataforma Rast e em artigos publicados (Suarez et al., 2019; Gupta et al., 2014; Nelkner et al., 2019; Pinski et al., 2019; Silva, 2019). Além da pesquisa na anotação feita via Rast, genes e proteínas de interesse foram buscados nas anotações de genômicas geradas pelo Prokka, por ser uma plataforma de anotação reconhecida por sua acuidade (Seemann, 2014). Para obter a soma das CDS por categoria, foi considerada prioritariamente a presença de determinadas proteínas codificantes na anotação genômica. No entanto, devido às diferenças nas anotações de nomenclatura de proteínas, também foram buscados os genes associados à transcrição das proteínas de interesse. A contabilização considerou apenas o gene ou a proteína de interesse nas anotações Prokka para evitar contagens repetidas. A contagem final de CDS nas categorias para cada bactéria considera tanto as proteínas ortólogas quanto a duplicação gênica (parálogos).

A Figura 1 é apresenta de forma comparativa a distribuição total de CDS nas categorias funcionais das diferentes bactérias analisadas. *Klebsiella pneumoniae* UFMG51 e *Pseudomonas putida* UFMG81 se destacaram por possuírem o maior número total de CDS somadas para todas as categorias associadas à promoção de crescimento, 830 e 834 respectivamente. Já *Microbacterium* sp. UFMG61 teve o menor total de CDS, 341. As outras bactérias apresentaram os seguintes totais: *Bacillus megaterium* UFMG50 743 CDS, *Ochrobactrum* CNPMS2088 601 CDS, *Pantoea ananatis* UFMG54 579 CDS. Em todas as bactérias, a categoria que reuniu maior número de CDS no genoma foi referente à versatilidade metabólica. Exceto em UFMG51, as outras cinco bactérias contabilizaram um menor número de CDS na categoria de biocontrole. Em *K. pneumoniae* UFMG51, a categoria com menor número de CDS foi colonização do tecido vegetal (rizosférica e endofítica)(Figura 1).

A partir da presença de CDS para cada bactéria nas tabelas suplementares(S10-S14), foram gerados gráficos de Venn com a finalidade de permitir uma melhor visualização da distribuição das sequências codificantes entre as

bactérias considerando apenas a CDS de interesse e não o número de cópias presentes no genoma como disposto nas tabelas suplementares (Figuras S1-S9). A construção dos gráficos para apenas algumas categorias e subcategorias da análise genômica deu-se com a finalidade de auxiliar na correlação destes resultados com testes prévios fenotípicos realizados por Vieira (2015), Silva (2017), Silva (2018) para verificar se há uma correção entre o genótipo e o fenótipo.

### 5.2.1. Promoção do crescimento vegetal

A categoria de promoção do crescimento vegetal é dividida em quatro subcategorias: fixação de nitrogênio, metabolismo de fosfato, produção de ácidos orgânicos e fito-hormônios (Tabela S10).

Todas as bactérias deste estudo apresentam CDS relacionadas à fixação de Nitrogênio (Tabela S10), incluindo genes e/ou proteínas do complexo nitrogenase e de regulação da fixação de nitrogênio. Dentre os genes e proteínas associados ao complexo nitrogenase identificados dentre as bactérias estão o gene *hupB*, o qual codifica a proteína Periplasmic [NiFeSe] hydrogenase large subunit, e a proteína Nitrogenase-stabilizing/protective protein NifW, codificada pelo gene *nifW*. Em relação à regulação da fixação de nitrogênio, alguns dos genes e proteínas identificados dentre as bactérias analisadas foram: gene *fixN*, o qual codifica a proteína Cytochrome c oxidase subunit 1 homolog, e proteína Nitrogen fixation regulation protein FixK, codificada pelo gene *fixK* (SUAREZ *et al.*, 2019). *K. pneumoniae* UFMG51 se destaca com o maior número de genes e proteínas identificadas relacionadas a essa função, com um total de 33 (Tabela S10). A fixação de nitrogênio por bactérias é um dos principais mecanismos apresentados pelas bactérias com comprovado aumento do crescimento vegetal. *K. pneumoniae* UFMG51, além de apresentar o maior número de CDS para fixação de nitrogênio, apresenta 19 CDS exclusivas em relação às outras bactérias (Tabela S10). A atividade de fixação de nitrogênio por isolados dessa espécie já foi reportada, como no trabalhos de Schmitz, Klopprogge, Grabbe (2002), focado na regulação desse processo nessa espécie.

Microrganismos endofíticos aumentam a disponibilidade de P no solo através da solubilização do fosfato mineral ou da mineralização de fosfato orgânico. Ambos mecanismos são fundamentais à promoção de crescimento vegetal e já se mostraram

como fatores limitantes de desenvolvimento em casos de ausência (Chhabra e Dowling, 2017). Todas as bactérias apresentaram CDS para metabolismo de fósforo. Dentre essas CDS, algumas estão diretamente relacionadas à solubilização e aquisição de fosfato, como por exemplo o gene *pstA*, relacionado à transcrição de Phosphate transport system permease protein PstA, encontrado no genoma de quatro das seis bactérias deste estudo (Tabela S1). Em relação aos testes fenotípicos para as bactérias do estudo, ao analisar a solubilização de P-Ca e P-Fe, *K. pneumoniae* UFMG51 e *P. ananatis* UFMG54 obtiveram os melhores resultados (Tabela S8) (Vieira, 2015 ; Silva, 2017;Silva, 2018). Aliado a isso, na análise genotípica, UFMG51 foi a bactéria que apresentou maior número de CDS para metabolismo de P, com um total de 75 e UFMG54 a quinta maior com 59 CDS. No entanto, ao contabilizar os diferentes tipos de CDS presentes nas bactérias, UFMG51, UFMG54 apresentam alguns dos maiores totais, 51 e 48 respectivamente (Figura S1). Ambas as bactérias não possuem CDS exclusivas (Figura S1). Dentre as CDS pesquisadas para essa subcategoria, quatro são compartilhados entre essas duas bactérias: Guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase (EC 3.6.1.40), Major phosphate-irrepressible acid phosphatase, Putative transcriptional regulator PhnF (Figura S1, Tabela S1).

Fitatos são uma das principais fontes de P orgânico. No entanto, encontram-se em uma forma indisponível para utilização pelas planta. Microrganismos são capazes de produzir fitases, enzimas capazes de mineralizar esse composto, e consequentemente, torná-lo disponível para uma maior captação de P pelas plantas (Eida *et al.*, 2013). Os resultados dos experimentos *in vitro* (Tabela S8) sobre a mineralização de P orgânico corroboram com análise genômica. *Microbacterium sp.* UFMG61, a única bactéria que apresentou resultados negativos nos testes *in vitro* foi também a que apresentou o menor número de CDS relacionadas ao metabolismo de fosfato (Figura S1 e Tabela S1). Além disso, dentre as CDS relacionadas às enzimas fosfatases, *Microbacterium* UFMG61 apresentou o menor número, um total de sete, enquanto *B. megaterium* UFMG50, uma das bactérias a apresentar o fenótipo de mineralização do fitato, apresentou o maior número, possui 37 CDS relacionadas a enzimas fosfatases.

A produção de ácidos orgânicos pelas bactérias endofíticas exerce um papel primordial na solubilização de fosfato a partir de compostos insolúveis para a absorção da planta (Richardson *et al.*, 2009).Com relação a presença de CDS relacionadas

produção de ácidos orgânicos, somente uma é compartilhada por todas as bactérias, a proteína L-lactato desidrogenase (Figura S2 e Tabela S2). *K. pneumoniae* UFMG51 e *P. ananatis* UFMG54 destacaram-se entre as demais com os maior número de tipos de CDS, 12 e 10 (Figura S2), além de quantidade 18 e 13 respectivamente (Tabela S2). Essas duas bactérias compartilham a proteína Low-affinity gluconate transporter, uma CDS em cada bactéria. Vale ressaltar as bactérias UFMG51 e UFMG 54 juntamente com *P. putida* UFMG81 também se destacaram quanto ao maior número de CDS derivadas de PQQ (pirroloquinolina quinona) e ácido glucônico (Tabela S1 e S2), importantes para a via de solubilização de fosfato ligado ao cálcio (Alam et al., 2002; An e Moe, 2016). Nos testes *in vitro* (Tabela S8), foi observada correlação negativa e significativa entre a solubilização de P-Ca e redução do pH (Silva, 2018). No entanto, existe uma relação de correlação entre o aumento da disponibilização de P a partir da liberação de ácidos orgânicos e uma conseqüente redução de pH (Song et al., 2008).

Nos ensaios de promoção do crescimento de milho feitos previamente a esse trabalho, *B. megaterium* UFMG50 aumentou significativamente a quantidade de P liberada no solo dos potes contendo como fonte de P o P-Araxá (Tabela S9) e também foi a bactéria que mais promoveu o crescimento da planta para esta fonte de P, embora não tenha apresentado o mesmo desempenho nos experimentos *in vitro* (Tabela S8). Provavelmente, a presença de fatores ambientais pode ter desencadeado expressões de características de promoção de crescimento no ambiente diferentes das observadas *in vitro* para essa bactéria (Hautefort e Hinton, 2000).

A capacidade de bactérias endofíticas produzirem fito-hormônios, em especial AIA, já foi descrita na literatura como um dos mecanismos de promoção de crescimento vegetal oferecidos por esses microrganismos (Ali, Charles e Glick, 2017). Nos experimentos *in vitro* (Tabela S8) apresentaram capacidade para a produção do fito-hormônio AIA (auxina) as bactérias *P. ananatis* UFMG54, *Microbacterium* sp. UFMG61 e *P. putida* UFMG81. Na análise genômica, todas as bactérias apresentaram CDS relacionadas relacionadas à produção de fito-hormônios. No entanto, a distribuição dessas CDS entre as bactérias não apresentou diferenças de total ou de distribuição significativas para associá-las aos resultados obtidos *in vitro* (Figura S3 e Tabela S3). As três bactérias que se destacaram no teste fenotípico para produção de AIA (Tabela S8) não apresentam CDS compartilhadas apenas entre si ou exclusivas



(Figura S3). Um dos fatores que pode justificar a falta de correlação direta entre a expressão do fenótipo *in vitro* e o potencial genômico dos microrganismos está mais relacionado ao processo de regulação da expressão gênica. A expressão gênica em bactérias endofíticas é regulada por diversos fatores da planta hospedeira, como exemplo a liberação de exsudatos (MARK *et al.*, 2005).

### 5.2.2 Colonização do tecido vegetal (rizosférica e endofítica)

Na categoria colonização do tecido vegetal (rizosférica e endofítica), foram investigadas sete subcategorias: quimiotaxia e motilidade; detoxificação de espécies reativas de oxigênio; *quorum sensing*, biofilme e adesão; produção de exopolissacarídeos; degradação de polímeros vegetais; enzimas piridoxal; atividade de exoprotease. Analisar a capacidade de colonização do tecido vegetal pelos microrganismos endofíticos é essencial para prever sua distribuição nos tecidos internos das plantas (Santos, *et al.*, 2018). A bactéria que apresentou o maior número de CDS nessa categoria foi *P. putida* UFMG81, com 165, a que apresentou menor número foi *Microbacterium* sp. UFMG61, com 37 (Tabela S11).

A formação de biofilme em bactérias promotoras do crescimento confere certas vantagens ao processo de colonização do tecido vegetal. Algumas dessas vantagens são: inibição de organismos competidores, captação de nutrientes, resposta rápida e adaptação a mudanças de condições ambientais (SENEVIRATNE *et al.* 2010). Uma das bactérias que apresentou maior produção de biofilme nos estudos fenotípicos prévios de Vieira (2015), Silva (2017), Silva (2018), *K. pneumoniae* UFMG51, apresentou também o maior número total de CDS relacionadas à formação de biofilmes, adesão e *quorum sensing* um total de 22. A outra bactéria a se destacar nos experimentos fenotípicos para essa subcategoria, *B. megaterium* UFMG50, apresentou o quarto maior total, 14 CDS (Tabela S5). Ao analisar os tipos de CDS presentes nessas bactérias, elas não compartilham nenhuma CDS. Em relação ao ao número de CDS distintas em cada uma dessas bactérias UFMG50 apresenta um total de seis sem nenhuma exclusiva, enquanto UFMG51 apresenta 16 sendo cinco exclusivas: gene *bcsZ* (beta-1,4-glucanase), Biofilm PGA outer membrane secretin PgaA, Biofilm PGA synthesis auxiliary protein PgaD, Biofilm PGA synthesis deacetylase PgaB (EC

3.-), Biofilm PGA synthesis N-glycosyltransferase PgaC (EC 2.4.-.-) (Figura S5, Tabela S5).

Plantas quando sob estresses abióticos, podem produzir metabolitos secundários, tais como espécies reativas de oxigênio, para mitigar os efeitos do estresse sobre o tecido vegetal. No entanto, alguns desses metabólitos podem ser sintetizados por microrganismos endofíticos. Há estudos inclusive que creditam a resistência das plantas a diversos tipos de estresse à produção desses metabólitos pelos microrganismos endofíticos (Ogbe, Finnie e Van Staden, 2020). As seis bactérias apresentaram apenas uma proteína em comum na subcategoria de detoxificação de espécies reativas de oxigênio: Superoxide dismutase [Mn] (EC 1.15.1.1) (Tabela S11). A bactéria a apresentar maior número de CDS/ genes para essa subcategoria foi *P. putida* UFMG81, com 53, enquanto *Microbacterium* sp. UFMG61 apresentou o menor total, cinco.

Motilidade e quimiotaxia são essenciais para a interação entre planta e endofíticos. Microrganismos endofíticos são atraídos por quimiotaxia às raízes através liberação de exsudatos pelas plantas e movimentam-se a essas regiões (Santos, et al., 2018). Na subcategoria de quimiotaxia e mobilidade há CDS compartilhadas pelas seis bactérias concomitantemente. A bactéria a se destacar nesta subcategoria foi *P. ananatis* UFMG 54, com um total de 107 CDS/genes, enquanto *Microbacterium* sp. UFMG 61 apresentou o menor número para a subcategoria, 20 (Tabela S11).

*Quorum sensing*, biofilme e adesão são características de colonização intimamente interligadas. O mecanismo de sinalização celular envolvido em *quorum sensing* pode desencadear a formação de biofilme pelas bactérias, o que conseqüentemente, gera a adesão de uma comunidade de bactérias sobre uma superfície (Munir, et al., 2020). Na subcategoria *Quorum sensing*, biofilme e adesão, duas das proteínas pesquisadas está presente em todas as bactérias: Alkanal monooxygenase alpha chain e Homoserine/homoserine lactone efflux protein. A caracterização *in vitro* de fenótipos relacionados a formação de biofilme das bactérias feita em estudos prévios (Vieira, 2015; Silva, 2017; Silva, 2018) (Tabela S1) corrobora com alguns dos resultados da pesquisa genômica para essa subcategoria. As bactérias que apresentaram maior produção de biofilme, *K. pneumoniae* UFMG51 e *B. megaterium* UFMG50, foram também as que apresentaram alguns dos maiores

totais de CDS relacionadas à formação de biofilmes, adesão e *quorum sensing* encontradas, apresentando um total de 22 (maior número) e 14 (quarto maior número), respectivamente (Tabela S4). Quanto à quantidade de CDS distintas, UFMG50 possui seis, enquanto UFMG 51 possui 16 (Figura S4). Ao observar as CDS/genes encontrados associados a subcategoria de exopolissacarídeos porém, há um perfil discrepante. A avaliação do potencial para produção de exopolissacarídeos deve ser levada em consideração por ser um componente essencial à formação de biofilmes (Sutherland, 1999). Enquanto *B. megaterium* UFMG50 apresentou o maior número de CDS, dez de seis tipos distintos (Tabela S11), ela não se destacou no experimento de produção de EPS *in vitro* (Tabela S8). *K. pneumoniae* UFMG51, a qual não apresentou nenhuma das CDS pesquisadas, nos ensaios *in vitro* teve a maior massa de EPS pelo método de precipitação alcoólica (Vieira, 2015; Silva, 2017; Silva, 2018). Possivelmente, como as montagens dos genomas analisados não estão completas, algumas características podem estar subestimadas. As bactérias que se destacaram nos ensaios de adesão a raiz, *P. ananatis* UFMG54, seguida de *B. megaterium* UFMG50, também não apresentaram o maiores números de genes/CDS relacionados à adesão, biofilmes e *quorum sensing*. Entretanto, ambas, além de UFMG 51, apresentam o gene *lapA*, codificante de uma proteína de adesão de superfície (Tabela S4). Apesar dos exsudatos vegetais atraírem microrganismos, aqueles com capacidade de adesão a superfície da raiz têm maior sucesso na entrada e colonização do interior da planta (PINSKI *et al.*, 2019).

### 5.2.3 Biocontrole

Microrganismos endofíticos podem atuar como agentes de biocontrole à fitopatógenos através da competição direta pela colonização da planta com outros microrganismos do mesmo nicho, assim também como através da atividade antagonista pela produção de metabólitos bioativos (Tewari *et al.*, 2019). A categoria biocontrole focou na produção de metabólitos bioativos pelos microrganismos estudados e foi dividida em cinco subcategorias: quitinase; aquisição de ferro; compostos voláteis; antibióticos; bacteriocinas. Nessa categoria, a bactéria que apresentou o maior número de CDS foi *Ochrobactrum* sp. CNPMS2088 com 111 CDS/genes (Tabela S12).

A produção de sideróforos por microrganismos endofíticos confere vantagens ao processo de colonização do tecido vegetal por ampliar a interação entre plantas e endofíticos (Compant *et al.* 2005), além de desempenhar uma vantagem competitiva a outros microrganismos, inclusive fitopatógenos, por diminuir a disponibilidade de ferro (Miethke e Marahiel, 2007). Experimentos *in vitro* demonstram a produção de sideróforos em ordem decrescente pelas bactérias: *B. megaterium* UFMG50, *P. putida* UFMG81 e *K. pneumoniae* UFMG51 (Tabela S8) (Vieira, 2015; Silva, 2017). Na análise genômica a bactéria a apresentar o maior número de CDS foi *K. pneumoniae* UFMG51 com 77 (Tabela S5). Quando analisada a diversidade de CDS relacionadas à aquisição de ferro, as três bactérias dos resultados do estudo *in vitro* não apresentam CDS compartilhadas apenas entre elas. Entretanto, cada uma apresenta um número significativo de CDS exclusivas que podem estar associadas à maior produção de sideróforos: UFMG 50 possui 17, UFMG 51 possui 15 e UFMG 81 três (Figura S5 e Tabela S5).

Microrganismos endofíticos são uma excelente fonte de substâncias com propriedades bioativas. A partir de seus metabólitos, já foram isoladas substâncias com potencial antibiótico, antifúngico e inclusive antiviral, as quais são importantes para a sobrevivência dos endofíticos em meio a competição com outros microrganismos, mas também influenciam na resistência das plantas a certos fitopatógenos, além de também apresentarem potencial de aplicação na indústria farmacêutica (Christina, Christopher e Bhore, 2013). Na análise genômica, as bactérias apresentaram CDS para quitinases, antibióticos e apenas *Microbacterium* sp. UFMG61 não apresentou CDS para bacteriocinas, dentre as pesquisadas (Tabela S12). Em estudos prévios *P. putida* UFMG81 foi capaz de antagonizar *B. subtilis* nos experimentos *in vitro* e apresentou o número maior de CDS em relação às outras bactérias para produção de antibióticos, 21 (Tabela S8 e S12) (Viana, 2015; Silva, 2017). Já *B. megaterium* UFMG50 antagonizou patógenos fúngicos *in vitro* (Silva, 2018) e apresentou o maior total de CDS relacionadas à produção de bacteriocinas, 12 (Tabela S8 e S12) (Viana, 2015; Silva, 2017). Dentre as CDS exclusivas de UFMG50 está a bacitracina (Tabela S12). Embora não seja comum o relato de bacteriocinas antagonizando fungos, a bacitracina é uma molécula que já foi considerada ativa a fungos como *Neurospora crassa* (Venkateswerlu, 1981).

#### 5.2.4 Versatilidade metabólica

Versatilidade metabólica é essencial para a aquisição de energia frente as mais diversas condições ambientais e competitivas para microrganismos endofíticos. A produção de metabólitos pode diferir entre uma mesma espécie colonizando o tecido rizosférico em relação ao tecido foliar em uma mesma planta hospedeira. Sendo assim, bactérias com grande versatilidade metabólica podem apresentar vantagens colonizadoras e competitivas frente à diferentes condições abióticas, como também bióticas (Brader *et al.*, 2014). A categoria versatilidade metabólica agrupou seis subcategorias, sendo elas: compostos aromáticos; metilotrofia; metabolismo secundário; metabolismo de enxofre; tolerância a estresse; lipopolissacarídeos.

Na pesquisa genômica para essa categoria, destacaram-se as bactérias *K. pneumoniae* UFMG51, com 299 CDS/genes, e *P. putida* UFMG81, com 291 (Tabela S13). Além disso, essas mesmas bactérias apresentam a maior diversidade de CDS/genes relacionadas a essa categoria, com 218, sendo 29 exclusivas, e 190, com 25 exclusivas, respectivamente. As seis bactérias compartilham entre si 33 genes e CDS em comum (Figura S6). A versatilidade metabólica é uma característica importante por favorecer a resistência a condições de estresse ambientais e nutricionais. Sendo assim, um maior acúmulo de CDS associados a essa característica pode favorecer ao *fitness* de sobrevivência e sucesso ecológico dessas bactérias nos mais diferentes ambientes, como o solo rizosférico. Esse habitat além de diverso quanto a parâmetros físico-químicos, apresenta uma grande diversidade microbiana devido a disponibilidade de nutrientes, criando um ambiente desafiador de grande competição para as bactérias ali presentes. Por isso, solos rizosféricos são uma importante fonte de bactérias endofíticas para a planta e ainda uma região para a prospecção de bactérias candidatas ao desenvolvimento de inoculantes/BPCP (MCNEAR JR., 2013)

Biossurfactantes melhoram a solubilidade de moléculas hidrofílicas reduzindo a tensão superficial e interfacial, o que permite uma maior interação entre água e óleo por exemplo (BANAT *et al.*, 2010). A atividade surfactante está ligada a diversas outras funções, dentre elas: aumento da disponibilidade de nutrientes à planta, biocontrole e solubilização de compostos aromáticos, importante nos processos de biorremediação uma vez que acelera a sua degradação desses compostos (SACHDEV e

CAMEOTRA, 2013). Experimento *in vitro* foi conduzido para avaliar a capacidade de produção de oléculas tensoativas pelas bactérias do estudo (Silva, 2017)(Tabela S8). AS bactérias *K. pneumoniae* UFMG51, *P. ananatis* UFMG54 e *Microbacterium* sp. UFMG61 se destacaram na produção de moléculas com altas atividades emulsificantes (E24 superior a 55%) já as *P. putida* UFMG81 e *B. megaterium* UFMG50 na produção de moléculas com capacidade de reduzir a tensão superficial do meio (Silva, 2017) (Tabela S8). Na análise genômica considerando a subcategoria de metabolismo de compostos aromáticos, as bactérias UFMG51 e UFMG 81 se destacaram quanto ao número total de CDS para essa categoria, 72 e 83 respectivamente (Tabela S13). Além disso, essas também foram as bactérias a apresentarem maior diversidade de CDS em seus genomas. UFMG 51 teve o maior número de CDS distintas, 58, sendo que 17 são exclusivas. UFMG81 teve o segundo maior número de CDS distintas, 51, sendo 15 exclusivas (Tabela S7). UFMG51 e UFMG81 compartilham entre si sete CDS: 1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase (EC 1.3.1.25), Anthranilate 1,2-dioxygenase large subunit, Benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10), Benzoate 1,2-dioxygenase beta subunit (EC 1.14.12.10), Benzoate 1,2-dioxygenase, ferredoxin reductase component, Catechol 1,2-dioxygenase (EC 1.13.11.1), Muconolactone isomerase (EC 5.3.3.4) (Tabela S7). Diversos estudos têm mostrado que bactérias degradadoras de compostos aromáticos também apresentam mecanismos de otimizar o contato e absorção desses compostos, com destaque para a produção de biossurfactantes (PATOWARY *et al.*, 2017).

Metilotrofia é uma característica metabólica de algumas bactérias que utilizam metanol e no caso dos endofíticos, promovem o crescimento vegetal, maior tolerância a estresse e produção de fito-hormônios (Shahzad *et al.*, 2019). A análise genômica buscou CDS e genes relacionados a essa subcategoria e todas as bactérias do estudo possuem genes e proteínas relacionados a essa característica (Tabela S13). A bactéria a apresentar o maior total de CDS foi *K. pneumoniae* UFMG51, com 15. Todas as bactérias compartilham seis CDS em comum (Tabela S13).

Microrganismos endofíticos contribuem com o aumento da resistência a estresses às plantas. Em casos de estresses abióticos, por exemplo, endofíticos podem atuar de forma a estimular uma resposta mais rápida das plantas ou também produzindo metabólitos que contribuem para a mitigação dos efeitos do estresse sobre

a planta hospedeira (Redman *et al.*, 1999; Schulz *et al.*, 2002). Foram pesquisados genes/proteínas relacionados à tolerância a estresse para as seis bactérias. Essas CDS estão relacionadas a: choque térmico para altas e baixas temperaturas, estresse osmótico, estresse periplásmico e detoxificação. As bactérias a apresentarem maior total de CDS para essa subcategoria foram: *B. megaterium* UFMG50 com 96, *K. pneumoniae* UFMG51 com 95 e *P. putida* UFMG81 com 93. Dentre as CDS, as seis bactérias compartilham 14 em comum: sensor protein KdpD; pyrroline-5-carboxilate reductase; gene *proB*, codificane da proteína glutamate-5-kinase; gene *kdpA*, codificante da proteína transportadora de potássio ATPase cadeia A; gene *kdpB*, codificante de transportador de potássio ATPase cadeia B; gene *kdpC*, codificante de transportador de potássio ATPase cadeia C; gene *htpX*, codificante da proteína Protease HtpX homolog; heat protein schock 15; gene *grpE*, codificante do cofator HSP-70 (SUAREZ *et al.*, 2019); gamma-glutamyl phosphate reductase; gene *dnaK*, codificante da proteína heat shock protein 70; chaperone protein DnaJ; chaperone protein ClpB; carbono starvation protein A (Tabela S13).

### 5.2.5 Virulência e resistência

Microrganismos endofíticos podem possuir CDS ligadas a fatores de virulência identificados em bactérias patogênicas. No entanto, é importante ressaltar que a regulação e expressão gênica dessas CDS em endofíticos ocorre de forma distinta, o que distingue sua ecologia de patogênicos (Xu *et al.*, 2014). A categoria virulência e resistência deste estudo se agrupou quatro subcategorias: virulência; sistemas de secreção; resistência a metais pesados; resistência a compostos tóxicos e às drogas. A bactéria que apresentou o maior número de CDS para a categoria foi *Klebsiella pneumoniae* UFMG51 com 192, a que apresentou menor número foi *Microbacterium* sp. UFMG61 com 84 (Tabela S14).

A bactéria *K. pneumoniae* UFMG51, a qual pertence a uma espécie de importância médica, era esperada que apresentasse o maior número CDS relacionadas às subcategorias de virulência e resistência a drogas, mas não foi o caso neste estudo. Na subcategoria virulência, ela apresentou o segundo maior número de CDS, 19, enquanto *P. putida* UFMG81 teve o maior total, 24 CDS. E na subcategoria de resistência a drogas UFMG 51 apresentou o segundo maior número de CDS 85,

enquanto *B. megaterium* UFMG50, apresentou o maior número de CDS para essa subcategoria 96 (Tabela S7). Entretanto, ao analisar a diversidade de CDS presentes, UFMG51 apresenta o maior número de CDS distintas relacionadas à virulência e resistência a drogas e substâncias tóxicas, um total de 70, sendo que 13 são exclusivas dessa bactéria (Figura S9). As 13 CDS exclusivas de *K. pneumoniae* UFMG 51 englobando as subcategories de virulência e resistência a drogas e substâncias tóxicas são: Alpha-hemolysin translocation ATP-binding; Beta-lactamase SHV-1; Chloramphenicol acetyltransferase 3; Chromate resistance protein ChrB; Copper homeostasis protein CutF precursor; Fosfomicin resistance protein FosA; Membrane protein, suppressor for copper-sensitivity ScsB; Membrane protein, suppressor for copper-sensitivity ScsD; Metal-dependent hydrolases of the beta-lactamase superfamily III; Multiple antibiotic resistance protein MarA; Secreted protein, suppressor for copper-sensitivity ScsC; Sensor protein of zinc sigma-54-dependent two-component system; Suppression of copper sensitivity: putative copper binding protein ScsA (Tabela S7). De fato, há estudos que comparam o genoma e patogenicidade de estirpes de *K. pneumoniae* de importância médica e endofíticas e apontam diferenças no seu genoma, apesar da presença de fatores de virulência e resistência a drogas em ambos, além de demonstrar em experimentos que a estirpe endofítica possui efeitos de patogenicidade mais atenuados quando inoculados em animais em relação a de importância médica (Fouts *et al.*, 2008).

Sistema secretor em bactérias é amplamente associado como uma ferramenta de virulência em microrganismos patogênicos. Entretanto, sequências codificantes para essa característica também estão presentes em microrganismos endofíticos e apresentam um papel crucial na colonização do tecido vegetal sem trazer prejuízos aos hospedeiro (Piromyou *et al.*, 2015). Todas as bactérias deste estudo apresentam CDS relacionadas a algum sistema de secretor. As bactérias foram analisadas para a presença de proteínas e genes relacionados ao sistema de secretor, os quais incluíam proteínas do sistema *tat*, *sec*, sistemas secretor dos tipos I, II, III e IV. A bactéria com mais CDS para essa subcategoria foi *P. putida* UFMG81 com 51, enquanto *B. megaterium* UFMG50 apresentou o menor número, 16. Estudos *in vivo* sobre colonização radicular destas bactérias demonstrou que UFMG 81 teve a menor taxa de adesão à raiz, mas o segundo maior crescimento no tecido da raiz, enquanto UFMG 50 teve o segundo melhor desempenho de adesão à raiz, mas a menor taxa



de crescimento no tecido da raiz (Tabela 8) (Silva, 2017). Todas as bactérias compartilham 10 CDS em comum: ATP synthase subunit alpha; ATP-dependent zinc metalloprotease FtsH ; cell division ATP-binding protein FtsE; DNA translocase FtsK; Lipoprotein signal peptidase; protein translocase subunit SecA; protein translocase subunit SecE; protein translocase subunit SecY; Signal recognition particle protein; Signal recognition particle receptor FtsY (Tabela S14).

Bactérias endofíticas possuem sistemas de resistência a metais pesados os quais atuam na detoxificação e translocação desses metais. A capacidade das PGPB de acumulação de metais pesados torna-as excelentes aliadas à biorremediação. Através desses mecanismos de resistência a metais pesados, bactérias endofíticas são capazes de promover o crescimento de plantas mesmo sob essa condição de estresse (Jan *et al.*, 2019). As bactérias foram analisadas quanto à presença CDS para resistência a diversos metais pesados, tais como: arsênio, cromo, cobre, cobalto, mercúrio, níquel, molibdênio e zinco. Dentre as sequências pesquisadas, cinco são compartilhadas entre todos os microrganismos: gene *arsC* (Arsenate reductase); gene *actP* (Copper-transporting P-type ATPase); Magnesium and cobalt efflux protein CorC; gene *modA* (Molybdate-binding periplasmic protein precursor); gene *rbsC* (Ribose transport system permease protein RbsC) (Tabela S14). O maior total de CDS para essa subcategoria foi de 42 no genoma da *K. pneumoniae* UFMG51, enquanto o menor foi de 17 em *Microbacterium* sp. UFMG61.

### 5.3 Promoção de crescimento e colonização vegetal

Nos ensaios de promoção do crescimento de milho, *B. megaterium* UFMG50 aumentou significativamente a quantidade de P liberada no solo dos potes contendo como fonte de P, o P-Araxá (Tabela 2) e também foi a bactéria que mais promoveu o crescimento da planta para esta fonte de P, embora não tenha apresentado o mesmo desempenho nos experimentos *in vitro* (Silva,2017; Silva, 2018; Vieira, 2018). Provavelmente, a presença de fatores ambientais pode ter desencadeado expressões de características de promoção ao crescimento no ambiente diferentes das observadas *in vitro* para essa bactéria (Hautefort e Hinton, 2000). No entanto, na análise genômica esta bactéria apresentou a terceira maior soma de CDS presentes nas categorias de Colonização da planta e Promoção de crescimento da planta (241), menor apenas

que o apresentado por *P. putida* UFMG81 (253) e *P. ananatis* UFMG54 (250) (Tabelas S10 e S11). Entretanto, *B. megaterium* UFMG 50 apresentou o maior número de CDS exclusivas (45) dentre as bactérias analisadas o que evidencia que seu alto potencial genético para colonização e promoção do crescimento das plantas pode estar associado às CDS exclusivas apresentadas dentro dessas duas categorias (Figura S9). Além disso, UFMG50 foi uma das bactérias a se destacar com maior número de CDS em resistência a estresse, o que pode também ter corroborado para sua melhor desenvoltura nos testes *in vivo* sob condições controladas em relação às outras bactérias do estudo (Tabela S9 e S13) (Vieira, 2015; Silva, 2017; Silva, 2018). Trabalhos anteriores também endossam *B. megaterium* como um potencial bioinoculante após em experimentos *in vivo* a bactéria demonstrar-se eficaz na promoção de crescimento de plantas quando comparada ao controle sem nenhum inócuo (Grunennvault *et al.*, 2018).

## 6. CONCLUSÃO

A investigação dos genomas das bactérias endofíticas avaliadas neste estudo permitiu explorar o potencial genético dessas bactérias e correlacionar aos resultados obtidos em estudos prévios para diferentes funções investigadas, como solubilização de fosfato, produção de fito-hormônio, sideróforos e biosurfactantes, formação de biofilme, capacidade de adesão e atividade antagonista. A análise genômica também propiciou uma visão mais ampla do potencial das bactérias para expressar outras características relacionadas à promoção de crescimento de plantas que não foram investigadas em seu fenótipo, mas que podem vir a ser expressas no ambiente como o potencial de biorremediação já que todas apresentam CDS relacionadas a resistência a metais pesados e degradação de compostos aromáticos. Dessa maneira, as análises genéticas reforçam os resultados previamente obtidos nos testes *in vivo* e *in vitro* de que as seis bactérias deste estudo possuem potencial para serem utilizadas como bioinoculantes de forma efetiva e segura para potencializar o crescimento e desenvolvimento vegetal. No entanto, ao considerar apenas a análise genômica, *K. pneumoniae* UFMG51 foi a bactéria que mais se destacou em diversas categorias e subcategorias por conter alguns dos maiores números de CDS associadas a fixação de nitrogênio, metabolismo de P, ácidos orgânicos, biofilme,

sideróforos, resistência a estresse, resistência a metais pesados, metabolismo de compostos aromáticos e versatilidade metabólica.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

DE ABREU, C.S. et al. Maize endophytic bacteria as mineral phosphate solubilizers. **Genetics and Molecular Research**, v. 16, n. 1. doi: 10.4238/gmr16019294, 2017.

AFZAL, M. et al. Ecology of Alkane-Degrading Bacteria and Their Interaction with the Plant. In: D., B. F. J. **Molecular Microbial Ecology of the Rhizosphere**. Hoboken: Wiley, 2013. Cap. 92, p. 975-989.

ALAM, S.; KHALIL, A.; AYUB, N.; RASHID, M. In vitro solubilization of inorganic phosphate by Phosphate Solubilizing Microorganisms (PSM) from Maize Rhizosphere. **International Journal of Agriculture & Biology**, v.4, p. 454-458, 2002.

ALI, S.; CHARLES, T. C.; GLICK, B. R. Amelioration of high salinity stress damage by plant growth-promoting bacterial endophytes that contain ACC deaminase. **Plant Physiology and Biochemistry**, v. 80, p. 160-167, 2014.

ALI, S; CHARLES, T. C.; GLICK, B. R. Endophytic phytohormones and their role in plant growth promotion. In: DOTY, S. L. **Functional Importance of Plant Microbiome**. Seattle: Springer, 2017. Cap. 6, p. 89-105.

AMAYA-GÓMEZ, C. V; HIRSCH, A. M.; SOTO, M. J. Biofilm formation assessment in *Sinorhizobium meliloti* reveals interlinked control with surface motility. **BMC Microbiology**, v. 15, p. 58, 2015. London: BioMed Central.

AN, R.; MOE, L. A. Regulation of Pyrroloquinoline Quinone-Dependent Glucose Dehydrogenase Activity in the Model Rhizosphere-Dwelling Bacterium *Pseudomonas putida* KT2440. **Applied and environmental microbiology**, v. 82, n. 16, p. 4955–4964, 2016.

ASAF, S. et al. Complete genome sequencing and analysis of endophytic *Sphingomonas* sp. LK11 and its potential in plant growth. **3 Biotech**, v. 8, n. 9, p. 389, 2018.

ASSUMPÇÃO, L. C. et al. Diversidade e potencial biotecnológico da comunidade bacteriana endofítica de sementes de soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 5, p. 503-510, Maio 2009. ISSN 1678-3921.

AZEVEDO, J. L.; MACCHERONI JÚNIOR, W.; PEREIRA, J. O. Endophytic microorganisms: a review on insect control and recent advances on tropical plants. **Electronic Journal of Biotechnology**, v. 3, p. 40-65, 2000.

AZIZ, R. K. et al. The RAST Server: rapid annotations using subsystems technology. **BMC Genomics**, v. 9, n. 75, 2008.

BANAT, I. M. et al. Microbial biosurfactants production, applications. **Applied Microbiology and Biotechnology**, v. 87, p. 427–444, 2010.

BHARDWAJ, D. et al. Biofertilizers function as key player in sustainable agriculture by improving soil fertility, plant tolerance and crop productivity. **Microbial cell factories**, v. 13, p. 66, 2014.

BOHLIN, J. *et al.* The nucleotide composition of microbial genomes indicates differential patterns of selection on core and accessory genomes. **BMC Genomics**, v. 18, p.151 2017.

BRADER, G., COMPANT, S., MITTER, B., TROGNITZ, F., & SESSITSCH, A. Metabolic potential of endophytic bacteria. **Current opinion in biotechnology**, v. 27, n.100, p. 30–37, 2014.

BRETTIN, T. et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. **Scientific reports**, v. 5, n. 8365, 2015.

CAMERON, D. R.; COOPER, D. G.; NEUFELD, R. J. The mannoprotein of *Saccharomyces cerevisiae* is an effective bioemulsifier. **Applied Environmental Microbiology**, v. 54, n. 6, p. 1420-1425, 1988.

CHANDRA, P.; SINGH, E. Applications and Mechanisms of Plant Growth-Stimulatin Rhizobacteria. In: CHOUDHARY, D. K.; VARMA, A.; TUTEJA, N. **Plant-Microbe Interaction: An Approach to Sustainable Agriculture**. Singapura: Springer, 2016. Cap. 3, p. 37-62.

CHHABRA, S.; DOWLING, D. Endophyte-Promoted Nutrient Acquisition: Phosphorus and Iron. In: DOTY, S. L. **Functional Importance of Plant Microbiome**. Seattle: Springer, 2017. Cap. 3, p. 21-42.

CHOUDHARY, D. K. et al. Bacterial-Mediated Tolerance and Resistance to Plants Under Abiotic and Biotic Stresses. **Journal of Plant Growth Regulation**, v. 35, n. 1, 2015.

CHRISTINA, A.; CHRISTAPHER, V.; BHOORE, S. J. Endophytic bacteria as a source of novel antibiotics: An overview. **Pharmacognosy reviews**, v.7, n.13, p. 11–16, 2013.

COIL, D. et al. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. **Bioinformatics**, v. 31, n. 4, p. 587-589, 2015.

COMPANT, S. et al. Editorial special issue: soil, plants and endophytes. **Plant and Soil**, v. 405, n. 1-2, p. 1-11, 2016.

COMPANT, S.; CLEMENT, C.; SESSITSCH, A. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. **Soil Biology and Biochemistry**, v. 42, n. 5, p. 669-678, 2010.

COMPANT, S.; DUFFY, B.; NOWAK, J.; CLÉMENT, C.; AIT BARKA, E. Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. **Applied Environmental Microbiology**, v. 71, n. 9, p. 4951-4959, 2005.

DONKOR E. S. Sequencing of bacterial genomes: principles and insights into pathogenesis and development of antibiotics. **Genes**, v. 4, n. 4, p. 556–572, 2013.

DUTTA, D. et al. Endophytes: Exploitation as a Tool in Plant Protection. **Brazilian Archives of Biology and Technology**, v. 57, n. 5, p. 621-629, 2014.

FATHALLH EIDA, M.; NAGAOKA, T.; WASAKI, J.; KOUNO, K. Phytate degradation by fungi and bacteria that inhabit sawdust and coffee residue composts. **Microbes and environments**, v.28, n.1, p. 71–80, 2013.

FOUTS, D. E.; TYLER, H. L.; DEBOY, R. T. Complete Genome Sequence of the N<sub>2</sub>-Fixing Broad Host Range Endophyte *Klebsiella pneumoniae* 342 and Virulence Predictions Verified in Mice. **PLoS Genetics**, v. 4, n. 7, 2008.

FRANK, A. C. The Genomes of Endophytic Bacteria. In: PIRTTILÄ, A.; FRANK, A. **Endophytes of Forest Trees. Forestry Sciences**. Dordrecht: Springer, v. 80, 2011. p. 107-136.

GILLIS, A., FAYAD, N., MAKART, L., BOLOTIN, A., SOROKIN, A., KALLASSY, M., & MAHILLON, J. Role of plasmid plasticity and mobile genetic elements in the entomopathogen *Bacillus thuringiensis* serovar *israelensis*. **FEMS Microbiology Reviews**, v.42, n. 6, p. 829–856, 2018.

GLICK, B. R. Plant growth-promoting bacteria: mechanisms and applications. **Scientifica** v. 2012, 2012.

GOLINSKA, P. et al. Endophytic actinobacteria of medicinal plants: diversity and bioactivity. **Antonie van Leeuwenhoek**, v. 108, n. 2, p. 267-289, 2015.

GORDON, S. A.; WEBER, R. P. Colorimetric estimation of indolacetic acid. **Plant Physiology**, v. 26, p.192-195, 1951.

GOUDA, S. et al. Endophytes: A Treasure House of Bioactive Compounds of Medicinal Importance. **Frontiers in microbiology**, v. 7, n. 1538, 2016.

GRUNENVALDT, R. L.; DEGENHARDT-GOLDBACH, J.; TOMASI, J. de C.; SANTOS, G. D. dos; VICENTE, V. A.; DESCHAMPS, C. *Bacillus megaterium*: an endophytic bacteria from callus of *Ilex paraguariensis* with growth promotion activities. **Bioteconología vegetal**, v.18, n. 1, p. 3-13, 2018.

GUPTA, A. et al. Whole genome sequencing and analysis of plant growth promoting bacteria isolated from the rhizosphere of plantation crops coconut, cocoa and arecanut. **PLoS ONE**, v. 9, n. 8, 2014.

GUPTA, S.; SETH, R.; SHARMA, A. Plant Growth-Promoting Rhizobacteria Play a Role as Phytostimulators for Sustainable Agriculture. In: CHOUDHARY, D. K.; VARMA, A.; TUTEJA, N. **Plant-Microbe Interaction: An Approach to Sustainable Agriculture**. Singapore: Springer, 2016. p. 475-493.

HALLMANN, J. et al. Bacterial endophytes in agricultural crops. **Canadian Journal of Microbiology**, v. 43, n. 10, p. 895-914, 1997.

HALLMANN, J. et al. Bacterial endophytes in agricultural crops. **Canadian Journal of Microbiology**, v. 43, p. 895-914.

HARDOIM, P. R. et al. The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of Microbial Endophytes. **Microbiology and molecular biology reviews : MMBR**, v. 79, n. 3, p. 293–320, 2015.

HARDOIM, P. R.; VAN OVERBEEK, L. S.; ELSAS, J. D. Properties of bacterial endophytes and their proposed role in plant growth. **Trends in Microbiology**, v. 16, n. 10, p. 463-471, 2008.

HAUTEFORT, I.; HINTON, J. C. Measurement of bacterial gene expression in vivo. **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 355, n.1397, p. 601–611, 2000.

IJAZ, A. et al. Phytoremediation: recent advances in plant-endophytic synergistic interactions. **Plant and Soil**, v. 405, p. 179-195, 2016.

JAN, R.; KHAN, M. A.; ASAF, S.; LUBNA; LEE, I. J.; KIM, K. M. Metal Resistant Endophytic Bacteria Reduces Cadmium, Nickel Toxicity, and Enhances Expression of Metal Stress Related Genes with Improved Growth of *Oryza Sativa*, via Regulating Its Antioxidant Machinery and Endogenous Hormones. **Plants (Basel)**, v. 23, n. 8, p. 363, 2019.

KHAN, A. A. et al. Phosphorus solubilizing bacteria: occurrence, mechanisms and their role in crop production. **Journal of Agriculture and Biological Science**, v.1, p. 48–58, 2009.

KHAN, T. A.; MAZID, M.; QUDDUSI, S. Role of Organic and Inorganic Chemicals in Plant-Stress Mitigation. In: GAUR, R.; SHARMA, P. **Approaches to Plant Stress and their Management**. New Delhi: Springer, 2014. p. 39-52.



KUMARI, S. et al. Induced drought tolerance through wild and mutant bacterial strain *Pseudomonas simiae* in mung bean (*Vigna radiata* L.). **World Journal of Microbiology and Biotechnology**, v. 32, n. 4, 2016.

LACAVA, P. T.; AZEVEDO, J. L. Endophytic Bacteria: A Biotechnological Potential in Agrobiological System. In: MAHESHWARI, D.; SARAF, M.; AERON, A. **Bacteria in Agrobiological System: Crop Productivity**. Berlin: Springer, 2013. p. 1-44.

LAND, M. et al. Insights from 20 years of bacterial genome sequencing. **Functional and Integrative Genomics**, v. 15, n. 2, p. 141-161, 2015.

MAHESHWARI, R. What is an endophytic fungus? **Current Science**, v. 90, n. 10, p. 1309, 2006.

MANDAL, A. et al. Plant–Microbe Interaction for the Removal of Heavy Metal from Contaminated Site. In: CHOUDHARY, D. K.; VARMA, A.; TUTEJA, N. **Plant-Microbe Interaction: An Approach to Sustainable Agriculture**. Singapura: Springer, 2016. Cap. 11, p. 227-247.

MAPELLI, F. et al. Potential for plant growth promotion of rhizobacteria associated with *Salicornia* growing in Tunisian hypersaline soils. **Biomed Research International**, v. 2013, n. 248078, 2013.

MARK, G. L.; DOW, J. M.; KIELY, P. D.; HIGGINS, H.; HAYNES, J.; BAYSSE, C.; ABBAS, A.; FOLEY, T.; FRANKS, A.; MORRISSEY, J.; O'GARA, F. Transcriptome profiling of bacterial responses to root exudates identifies genes involved in microbe-plant interactions. **Proceedings of the National Academy of Sciences of the United States of America**, v.102, n. 48, p. 17454–17459, 2005.

MCNEAR JR., D. H. The Rhizosphere - Roots, Soil and Everything In Between. **Nature Education Knowledge**, v. 4, n.3, 2013.

MENDES, G., VASSILEV, N. B., BONDUKI, V. H., DA SILVA, I. R., RIBEIRO, J. I., JR, & COSTA, M. D. Inhibition of *Aspergillus niger* phosphate solubilization by fluoride released from rock phosphate. **Applied and environmental microbiology**, v. 79, n.16, p. 4906–4913, 2013.

MIETHKE, M.; MARAHIEL, M. A. Siderophore-Based Iron Acquisition and Pathogen Control. **Microbiology and Molecular Biology Reviews**, v. 71, n. 3, p. 413-451, 2007.

MUNIR, Samman et al . Quorum Sensing Interfering Strategies and Their Implications in the Management of Biofilm-Associated Bacterial Infections. **Brazilian archive of biology and technology**, Curitiba , v. 63, e20190555, 2020 . Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1516-89132020000101001&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-89132020000101001&lng=en&nrm=iso)>. access on 16 Nov. 2020. Epub Aug 31, 2020. <https://doi.org/10.1590/1678-4324-2020190555>.

MURPHY, J. & RILEY, J. P. A modified single solution method for the determination of phosphate in natural waters. **Analytica Chimica Acta**, v. 27, p. 31-36, 1962.

NADARAJAH, K. K. Rhizosphere Interactions: Life Below Ground. In: CHOUDHARY, D. K.; VARMA, A.; TUTEJA, N. **Plant-Microbe Interaction: An Approach to Sustainable Agriculture**. Singapura: Springer, 2016. Cap. 1, p. 3-24.

NASSAR, A. H.; EL-TARABILY, K. A.; SIVASITHAMPARAM, K. Promotion of Plant Growth by an Auxin-Producing Isolate of the Yeast *Williopsis saturnus* Endophytic in Maize (*Zea mays* L.) Roots. **Biology and Fertility of Soils**, v. 42, n. 2, p. 97–108, 2005.

NAUTIYAL, C. S. An efficient microbiological growth medium for screening phosphate solubilizing microorganism. **FEMS Microbiology Letters**, v. 170, p. 265–270, 1999.

NELKNER, J. et al. Genetic Potential of the Biocontrol Agent *Pseudomonas brassicacearum* (Formerly *P. trivialis*) 3Re2-7 Unraveled by Genome Sequencing and Mining, Comparative Genomics and Transcriptomics. **Genes**, v. 10, n. 8, 2019.

NEWMAN, L. A.; REYNOLDS, C. M. Bacteria and phytoremediation: new uses for endophytic bacteria in plants. **Trends in Biotechnology**, v. 23, p. 6-8, 2005.

OGBE, A. A.; FINNIE, J., F.; VAN STADEN, J. The role of endophytes in secondary metabolites accumulation in medicinal plants under abiotic stress. **South African Journal of Botany**, v. 134, p. 126-134, 2020.

OLIVEIRA, C. A.; ALVES, V. M. C.; MARRIEL, I. E.; GOMES, E. A.; SCOTTI, M. R.; CARNEIRO, N. P.; GUIMARÃES, C. T.; SCHAFFERT, R. E.; SÁ, N. M. H. Phosphate solubilizing microorganisms isolated from rhizosphere of maize cultivated in an oxisol of the Brazilian Cerrado Biome. **Soil Biology and Biochemistry**, v. 41, n. 9, p. 1782-1787, 2009.

PATOWARY, K.; Patowary, R.; Kalita, M. C.; Deka, S. Characterization of Biosurfactant Produced during Degradation of Hydrocarbons Using Crude Oil As Sole Source of Carbon. **Frontiers in Microbiology**, v.22, n.8, p. 279, 2017.

PAIKRAY, S. P.; MALIK, V. S. **Novel Formulation Of Microbial Consortium Based Bonoculant For Widespread Use Inagriculture Practices**. US 2012/001580.6 A1, 19 Janeiro 2012.

PINSKI, A.; BETEKHTIN, A.; HUPERT-KOCUREK, K.; MUR, L.; HASTEROK, R. Defining the Genetic Basis of Plant-Endophytic Bacteria Interactions. **International journal of molecular sciences**, v. 20, n.8, p.1947, 2019.

PIMENTEL, M. R. et al. The use of endophytes to obtain bioactive compounds and their application in biotransformation process. **Biotechnology Research International**, v. 2011, 2011.

PIROMYOU, P.; SONGWATTANA, P.; GREETATORN, T.; OKUBO, T.; KAKIZAKI, K. C.; PRAKAMHANG, J.; TITTABUTR, P.; BOONKERD, N.; TEAUMROONG, N.; MINAMISAWA, K. The Type III Secretion System (T3SS) is a Determinant for Rice-Endophyte Colonization by Non-Photosynthetic Bradyrhizobium. **Microbes and environments**, v. 30, n. 4, p. 291–300, 2015.

REDMAN, R. S.; FREEMAN, S.; CLIFTON, D. R.; MORREL, J.; BROWN, G.; RODRIGUEZ, R. J. Biochemical analysis of plant protection afforded by a non-pathogenic endophytic mutant of *Colletotrichum magna*. **Plant Physiology**, v. 119, p. 795–804, 1999.

RICHARDSON, A. E.; HADOBAS, P. A. Soil isolates of *Pseudomonas* spp. that utilize inositol phosphates. **Canadian Journal of Microbiology**, v. 43, p. 509–516, 1997.

RICHARDSON, A.; BAREA, J. M.; MCNEILL, A.; PRIGENT-COMBARET, C. Acquisition of phosphorous and nitrogen in the rhizosphere and plant growth promotion by microorganisms. **Plant Soil**, v. 321, p. 305-339, 2009.

RODRIGUEZ, H.; FRAGA, R.; GOZALEZ, T.; BASHAN, Y. Genetics of phosphate solubilization and its potential applications for improving plant growth-promoting bacteria. **Plant Soil**, v. 287, p. 15–21, 2006.

RODRIGUEZ, R.; REDMAN, R. More than 400 million years of evolution and some plants still can't make it on their own: plant stress tolerance via fungal symbiosis. **Journal of Experimental Botany**, v. 59, p. 1109-1114, 2008.

RUNGIN, S. et al. Plant growth enhancing effects by a siderophore-producing endophytic streptomycete isolated from a Thai jasmine rice plant (*Oryza sativa* L. cv. KDML105). **Antonie Van Leeuwenhoek**, v. 102, n. 3, p. 463-472, 2012.

SACHDEV, D. P.; CAMEOTRA, S. S. Biosurfactants in agriculture. **Applied microbiology and biotechnology**, v.97, n.3, p.1005–1016, 2013.

SANTOS, M. L. et al. Benefits Associated with the Interaction of Endophytic Bacteria and Plants. **Brazilian archives of biology and technology**, Curitiba , v. 61, e18160431, 2018. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1516-89132018000100200&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-89132018000100200&lng=en&nrm=iso)>. Acesso em 16 de novembro de 2020. Epub Nov 08, 2018. <https://doi.org/10.1590/1678-4324-2018160431>.

SANTOYO, G. et al. Plant growth-promoting bacterial endophytes. **Microbiological Research**, v. 183, p. 92-99, 2016.

SCAVINO A.F., PEDRAZA R.O. The Role of Siderophores in Plant Growth-Promoting Bacteria. In: Maheshwari D., Saraf M., Aeron A. (eds) **Bacteria in Agrobiolgy**: Crop Productivity. Berlin: Springer, 2013, p. 265-285.

SCHMITZ, R.A.; KLOPPROGGE, K.; GRABBE, R. Regulation of nitrogen fixation in *Klebsiella pneumoniae* and *Azotobacter vinelandii*: NifL, transducing two environmental signals to the nif transcriptional activator NifA. **Journal of Molecular Microbiology and Biotechnology**, v.4, n.3, p. 235-242, 2002.

SCHULZ, B.; BOYLE, C.; DRAEGER, S.; RÖMMERT, A. K.; KROHN, K. Endophytic fungi: a source of novel biologically active secondary metabolites. **Mycological Research**, v. 109, n. 9, p. 996–1004, 2002.

SCHWYN, B.; NEILANDS, J. B. Universal chemical assay for the detection and determination of siderophores. **Analytical Biochemistry**, v. 160, n. 1, p. 47–56, 1987. Disponível em: <http://www.sciencedirect.com/science/article/pii/0003269787906129>. Acesso em: 3 de janeiro de 2020.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, v. 30, n. 14, p. 2068-2069, 2014.

SENEVIRATNE, G. *et al.* Importance of Biofilm Formation in Plant Growth Promoting Rhizobacterial Action. In: MAHESHWARI, D., K. **Plant Growth and Health Promoting Bacteria**. Berlin: Springer, 2010. Cap. 4, p. 81-95.

SESSITSCH, A. *et al.* Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. **Molecular Plant-Microbe Interactions**, v. 25, n. 1, p. 28-36, 2012.

SETUBAL, J. C.; ALMEIDA, N. F.; WATTAM, A. R. Comparative Genomics for Prokaryotes. In: SETUBAL, J. C.; ALMEIDA, N. F.; WATTAM, A. R. **Comparative Genomics. Methods in Molecular Biology**. Nova Iorque: Humana Press, v. 1704, p. 55-78, 2018.

SHAHZAD, R. *et al.* Inoculation of abscisic acid-producing endophytic bacteria enhances salinity stress tolerance in *Oryza sativa*. **Environmental and Experimental Botany**, v. 136, p. 68-77, 2017.

Shahzad, R. *et al.* Metabolic and proteomic alteration in phytohormone-producing endophytic *Bacillus amyloliquefaciens* RWL-1 during methanol utilization. **Metabolomics**, v. 15, n. 16, 2019.

SHARMA, A. *et al.* Halophilic Bacteria: Potential Bioinoculants for Sustainable Agriculture and Environment Management Under Salt Stress. In: CHOUDHARY, D. K.; VARMA, A.; TUTEJA, N. **Plant-Microbe Interaction: An Approach to Sustainable Agriculture**. Singapura: Springer, 2016. Cap. 14, p. 297-325.

SILVA, D. R. C. **Bacterial selection for indirect mechanisms of plant growth promotion and maize roots colonization**. Monografia – Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 120p, 2017.

SILVA, Ubiana de Cássia. **Diversidade e mecanismos de promoção de crescimento de planta por micro-organismos associados ao milho cultivado em solo adubado com diferentes fontes de fósforo utilizando abordagem dependente e independente de cultivo**. Orientadora: Vera Lúcia dos Santos. 2018. 83 f. Tese (Doutorado em Microbiologia) - Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2018.

SMITS, T.H.M. The importance of genome sequence quality to microbial comparative genomics. **BMC Genomics**, v. 20, p. 662, 2019.

SONG, O. R.; LEE, S. J.; LEE, Y. S.; LEE, S. C.; KIM, K. K.; CHOI, Y. L. Solubilization of insoluble inorganic phosphate by Burkholderia cepacia DA23 isolated from cultivated soil. **Brazilian journal of microbiology: publication of the Brazilian Society for Microbiology**, v. 39, n.1, p.151–156, 2008.

SUAREZ, C. et al. Complete Genome Sequence of the Plant Growth-Promoting Bacterium Hartmannibacter diazotrophicus Strain E19T. **International Journal of Genomics**, v. 2019, 2019.

SUN, H. et al. Isolation, characterization, and antimicrobial activity of endophytic bacteria from Polygonum cuspidatum. **African Journal of Microbiology Research**, v. 7, n. 16, p. 1496-1504, 2013.

SUN, H. et al. Isolation, characterization, and antimicrobial activity of endophytic bacteria from Polygonum cuspidatum. **African Journal of Microbiology Research**, v. 7, n. 16, p. 1496-1504, 2013.

SUSILOWATIA, D. N. et al. **Indole-3-Acetic Acid Producing Bacteria and Its Application on the Growth of Rice**. AIP Conference Proceedings 2002, 020016. [S.l.]: [s.n.]. 2018.

SUSILOWATIA, D. N. et al. Indole-3-Acetic Acid Producing Bacteria and Its Application on the Growth of Rice. **AIP Conference Proceedings**, v. 2002, n. 1, 2018. ISSN 020016.

SUTHERLAND, I. W. Biofilm Exopolysaccharides. In: WINGENDER, J.; NEU, T. R.; FLEMMING, H. C. (eds) **Microbial Extracellular Polymeric Substances**. Berlin: Springer, 1999. Cap. 4, p.73-92.

SZIDERICS, A. H. et al. Bacterial endophytes contribute to abiotic stress adaptation in pepper plants (*Capsicum annuum* L.). **Canadian Journal of Microbiology**, v. 53, n. 11, p. 1195-1202, 2007.

TALLAPRAGADA P., SESHAGIRI S. Application of Bioinoculants for Sustainable Agriculture. In: KUMAR, V.; KUMAR, M.; SHARMA, S.; PRASAD, R. **Probiotics and Plant Health**. Singapore: Springer, 2017. p. 473-495.

TEWARI, S.; SHRIVAS, V. L.; HARIPRASAD, P.; SHARMA, S. Harnessing Endophytes as Biocontrol Agents. In: ANSARI R., MAHMOOD I. (eds) **Plant Health Under Biotic Stress**. Singapore: Springer, 2019. Cap. 10, p. 189-218.

TOMER, S.; SUYAL, D. C.; GOEL, R. Biofertilizers: A Timely Approach for Sustainable Agriculture. In: CHOUDHARY, D. K.; VARMA, A.; TUTEJA, N. **Plant-Microbe Interaction: An Approach to Sustainable Agriculture**. Singapura: Springer, 2016. Cap. 17, p. 375-395.

VAISHNAV, A. et al. Endophytic Bacteria in Plant Salt Stress Tolerance: Current and Future Prospects. **Journal of Plant Growth Regulation**, v. 38, n. 2, p. 650-668, 2019.

VASSILEV, N.; MENDES, G.; COSTA, M.; VASSILEVA, M. Biotechnological tools for enhancing microbial solubilization of insoluble inorganic phosphates. **Geomicrobiology Journal**, v. 31, p. 751–763, 2014.

VENKATESWERLU, G. J. Antifungal activity of bacitracin and its interaction with metals. **Journal of Biosciences**, v. 3, n.1, p. 1-5, 1981.

VIEIRA, João Antônio Côrtes. **Bactérias endofíticas de milho e seu potencial como produtoras de crescimento vegetal e agentes de controle biológico**. Orientadora:

Vera Lúcia dos Santos. 2015. Dissertação (Mestrado em Microbiologia) – Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

VÍLCHEZ, J. I. et al. Biosafety Test for Plant Growth-Promoting Bacteria: Proposed Environmental and Human Safety Index (EHSI) Protocol. **Frontiers in microbiology**, v. 6, p. 1514, 2016.

WEYENS, N. et al. Plant-associated bacteria and their role in the success or failure of metal phytoextraction projects: first observations of a field-related experiemnt. **Microbial Biotechnology**, v. 6, n. 3, p. 288-299, 2013.

XU, J.; LI, X.; LUO, L. Effects of Engineered *Sinorhizobium meliloti* on Cytokinin Synthesis and Tolerance of Alfalfa to Extreme Drought Stress. **Applied and Environmental Microbiology**, v. 78, n. 22, p. 8056-8061, 2012.

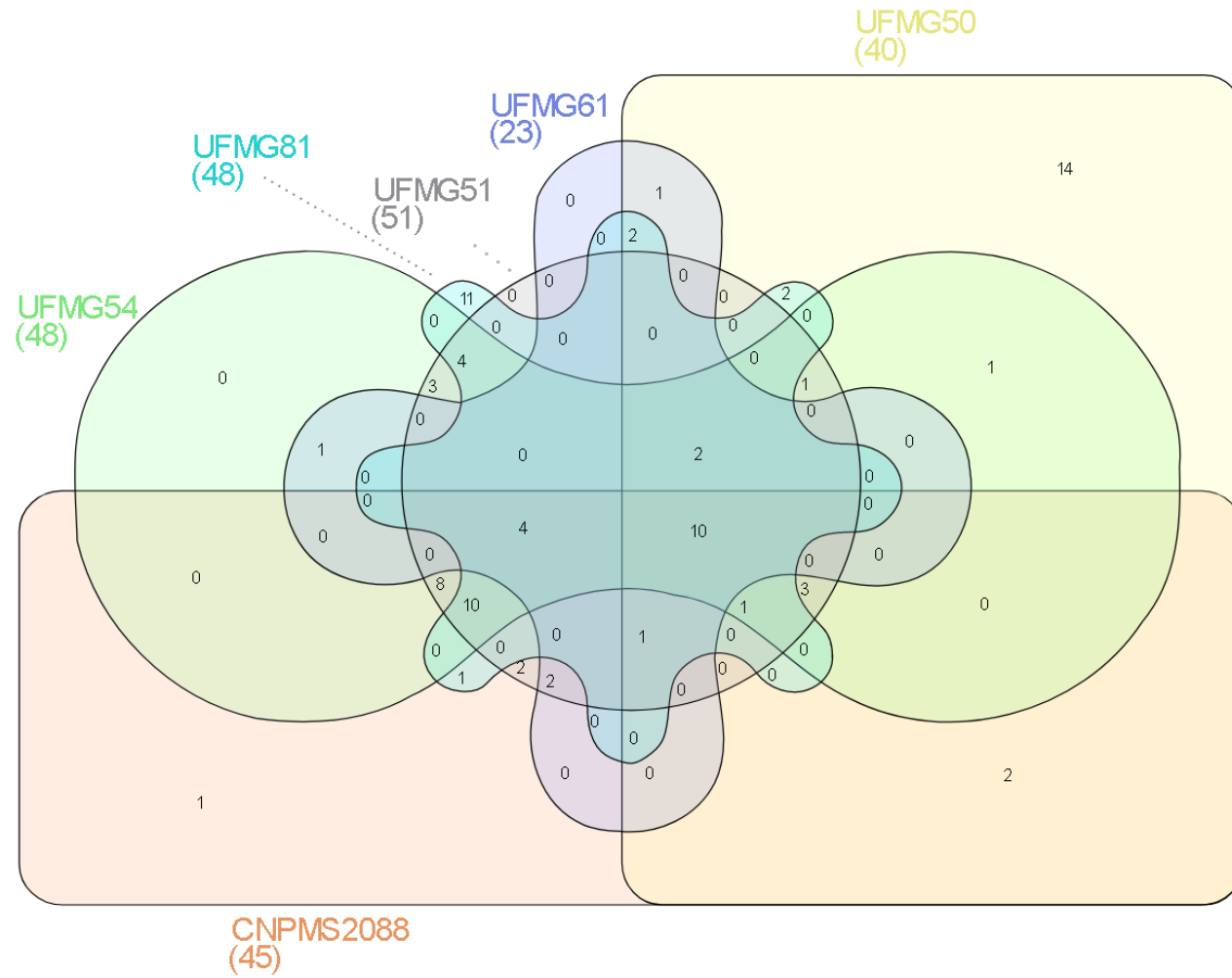
XU, X. H.; SU, Z. Z.; WANG, C. *et al.* The rice endophyte *Harpophora oryzae* genome reveals evolution from a pathogen to a mutualistic endophyte. **Scientific Reports**, v. 4, n. 5783, 2014.



## 8. FIGURAS E TABELAS SUPLEMENTARES

Os seguintes diagramas de Venn e tabelas foram criados a partir da anotação dos genomas das bactérias para subcategorias relacionadas aos resultados de maior destaque dos experimentos *in vivo* e *in vitro* de Vieira (2015), Silva (2017) e Silva, (2018) a fim de analisar uma possível correlação dos resultados fenotípicos ao genoma, além de outras categorias e subcategorias que também merecem maior destaque nas discussões.

**Figura S1:** Gráfico Venn da distribuição de CDS relacionadas ao metabolismo de P para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).



**Tabela S1:** Distribuição de CDS e genes entre as bactérias do estudo relacionados ao metabolismo de fosfato.

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Phosphate Metabolism	70	68	75	59	32	66
Phosphate Metabolism (1/6)	2-aminoethylphosphonate:pyruvate aminotransferase (EC 2.6.1.37)	0	2	0	0	0	0
	alkaline phosphatase	10	0	10	0	0	0
	Alkaline phosphatase (EC 3.1.3.1)	1	3	2	1	0	0
	Alkaline phosphatase 3	0	1	0	0	0	0
	Alkaline phosphatase 4	0	1	0	1	0	0
	Alkaline phosphatase D	0	1	0	0	0	0
	Alkaline phosphatase like protein	1	3	0	0	0	0
	Alkaline phosphatase synthesis sensor protein	1	3	0	0	0	0
	Alkylphosphonate utilization operon protein PhnA	0	1	0	0	0	0
	Cobalt-dependent inorganic pyrophosphatase	0	1	0	0	0	0
	Guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase (EC 3.6.1.40)	0	0	1	1	0	0
	Major phosphate-irrepressible acid phosphatase	0	0	2	1	0	0
	Manganese-dependent inorganic pyrophosphatase	0	1	0	0	0	0
	Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)	0	1	0	0	0	0
	PhnB protein	0	1	0	0	0	0
	Gene <i>phnC</i> (Phosphate-import ATP-binding protein PhnC)	1	0	1	0	3	0
Gene <i>phnG</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnG)	1	0	1	1	0	0	

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate metabolism (2/6)	Gene <i>phnH</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnH)	1	0	1	1	0	0
	Gene <i>phnI</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnI)	1	0	1	1	0	0
	Gene <i>phnK</i> (Putative phosphonates utilization ATP-binding protein PhnK)	1	0	1	1	0	0
	Gene <i>phnL</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnL)	1	0	1	1	0	0
	PhnO protein	0	1	0	0	0	0
	Gene <i>phnV</i> (Putative 2-aminoethylphosphonate transport system permease protein PhnV)	3	0	1	0	1	0
	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	0	1	0	0	0	0
	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	0	1	0	0	0	0
	Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	1	1	1	1	1	1
	Phosphoenolpyruvate phosphomutase (EC 5.4.2.9)	0	1	0	0	0	0
	Phosphonopyruvate decarboxylase (EC 4.1.1.82)	0	1	0	0	0	0
	Polyphosphate glucokinase (EC 2.7.1.63)	0	0	0	1	1	0
	Probable low-affinity inorganic phosphate transporter	0	1	0	0	1	0
	Gene <i>pstS</i> (Phosphate-binding protein PstS)	1	0	1	2	0	2
	Putative transcriptional regulator PhnF	0	0	1	1	0	0

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate Metabolism (3/6)	Pyrophosphate-energized proton pump (EC 3.6.1.1)	1	0	0	0	0	0
	Sensor protein PhoQ	1	0	1	0	0	0
	Alkaline phosphatase L	0	0	0	0	0	1
	Alkaline phosphatase synthesis transcriptional	1	6	1	0	1	1
	Exopolyphosphatase	0	1	1	1	1	1
	Extracytoplasmic function (ECF) sigma factor VreI	0	0	0	0	0	1
	Hemolysin activation/secretion protein associated with VreARI signalling system	0	0	0	0	0	1
	Inorganic pyrophosphatase	2	1	2	2	1	1
	Inorganic pyrophosphatase (EC 3.6.1.1)	2	0	2	2	1	1
	NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)	1	0	1	1	0	1
	Outer membrane TonB-dependent transducer VreA of trans-envelope signaling system	0	0	0	0	0	1
	Gene <i>pgpA</i> (Phosphatidylglycerophosphatase A)	0	0	1	1	0	1
	Gene <i>phnD</i> (Phosphate-import protein PhnD precursor)	2	0	1	1	3	1
	Gene <i>phnN</i> (Ribose 1,5-bisphosphate phosphokinase PhnN)	1	0	1	1	0	1
	Gene <i>phnU</i> (2-aminoethylphosphonate transport system permease PhnU)	1	0	0	0	0	1
	Gene <i>phoR</i> (Alkaline phosphatase synthesis sensor protein PhoR)	1	3	1	1	0	1
	Phosphate import ATP-binding protein PstB	1	0	1	1	0	1
Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	1	2	2	1	1	1	

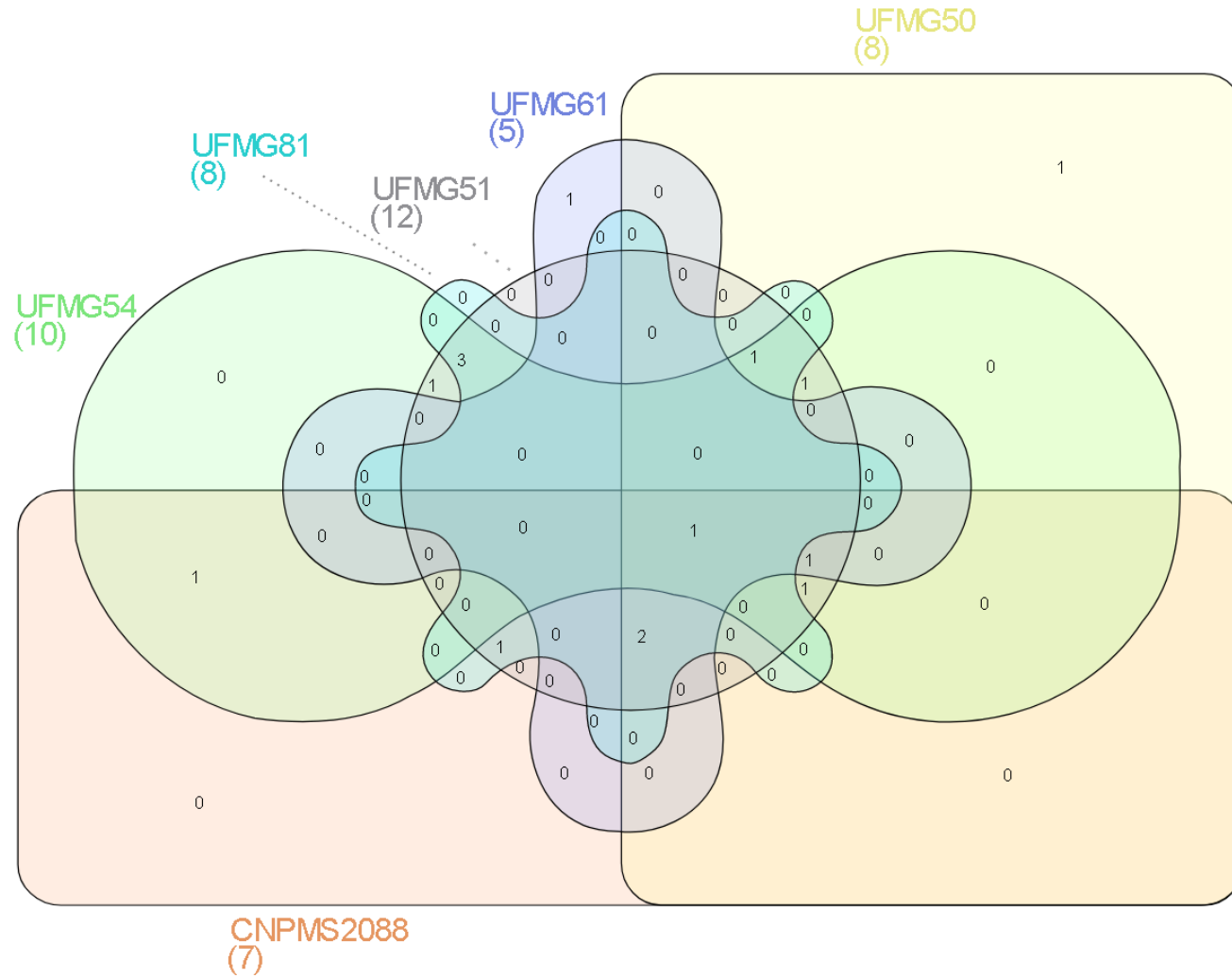
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate Metabolism (4/6)	Phosphate transport system regulatory protein PhoU	1	1	1	1	1	1
	Phosphate-binding DING protein (related to PstS)	0	0	0	0	0	1
	Phosphonate ABC transporter phosphate-binding periplasmic component (TC 3.A.1.9.1)	0	4	0	0	0	1
	Phosphonate uptake and metabolism regulator, LysR-family	0	0	0	0	0	1
	Phosphonoacetaldehyde hydrolase (EC 3.11.1.1)	0	1	0	0	0	1
	Polyphosphate kinase (EC 2.7.4.1)	1	1	2	1	1	1
	Gene <i>pqqB</i> (Pyrroloquinoline quinone biosynthesis protein B)	1	0	1	1	0	1
	Gene <i>pqqD</i> (Pyrroloquinoline quinone biosynthesis protein D)	0	0	1	1	0	1
	Predicted ATPase related to phosphate starvation-inducible protein PhoH	0	1	0	0	1	1
	Gene <i>pstB</i> (Phosphate import ATP-binding protein PstB)	1	0	1	1	0	1
	Putative outer membrane TonB-dependent receptor associated with haemagglutinin family outer membrane protein	0	0	0	0	0	1
	Putative PpiC-type peptidyl-prolyl cis-trans isomerase precursor associated with VreARI signaling system	0	0	0	0	0	1
	Pyrroloquinoline-quinone synthase	0	0	1	1	0	1
	RecA/RadA recombinase	0	0	0	0	0	1
	Ribose 1,5-bisphosphate phosphokinase PhnN	1	0	1	1	0	1

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate Metabolism (5/6)	Sigma factor regulator VreR (cytoplasmic membrane-localized) of trans-envelope signaling system	0	0	0	0	0	1
	Sodium-dependent phosphate transporter	2	0	2	1	0	1
	Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1)	0	0	1	1	0	1
	Gene <i>yliI</i> (Soluble aldose sugar dehydrogenase YliI )	1	0	1	1	0	1
	Exopolyphosphatase (EC 3.6.1.11)	2	1	4	1	1	2
	Filamentous haemagglutinin family outer membrane protein associated with VreARI signalling system	0	0	0	0	0	2
	Low-affinity inorganic phosphate transporter 1	0	1	1	2	1	2
	Phosphate starvation-inducible protein PhoH, predicted ATPase	2	2	4	4	2	2
	Phosphate-import permease protein PhnE	2	0	2	2	4	2
	Phosphate regulon transcriptional regulatory protein PhoB (SphR)	1	4	2	1	1	2
	Phosphonoacetaldehyde hydrolase	0	1	0	0	1	2
	Polyphosphate kinase 2 (EC 2.7.4.1)	1	0	1	1	0	3
	Gene <i>gph</i> (Phosphoglycolate phosphatase)	4	3	1	1	1	4
	Gene <i>phnJ</i> (Alpha-D-ribose 1-methylphosphonate 5-phosphate C-P lyase)	1	0	1	1	0	0
	Gene <i>phnM</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate diphosphatase)	2	0	1	1	0	0
	Gene <i>phnS</i> (2-aminoethylphosphonate binding protein)	1	0	1	1	0	0
	Gene <i>phoA</i> (Alkaline phosphatase H precursor)	1	1	1	1	0	0

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate Metabolism (6/6)	Gene <i>phoD</i> (Alkaline phosphatase D precursor)	0	1	0	0	0	0
	Gene <i>phoP</i> (Alkaline phosphatase synthesis transcriptional regulatory protein PhoP)	2	5	1	2	1	2
	Gene <i>php</i> (Phosphotriesterase-related protein)	0	1	1	1	0	0
	Gene <i>pstA</i> (Phosphate transport system permease protein PstA)	1	1	1	2	0	0
	Gene <i>ubiG</i> (Ubiquinone biosynthesis O-methyltransferase)	3	0	1	2	2	5



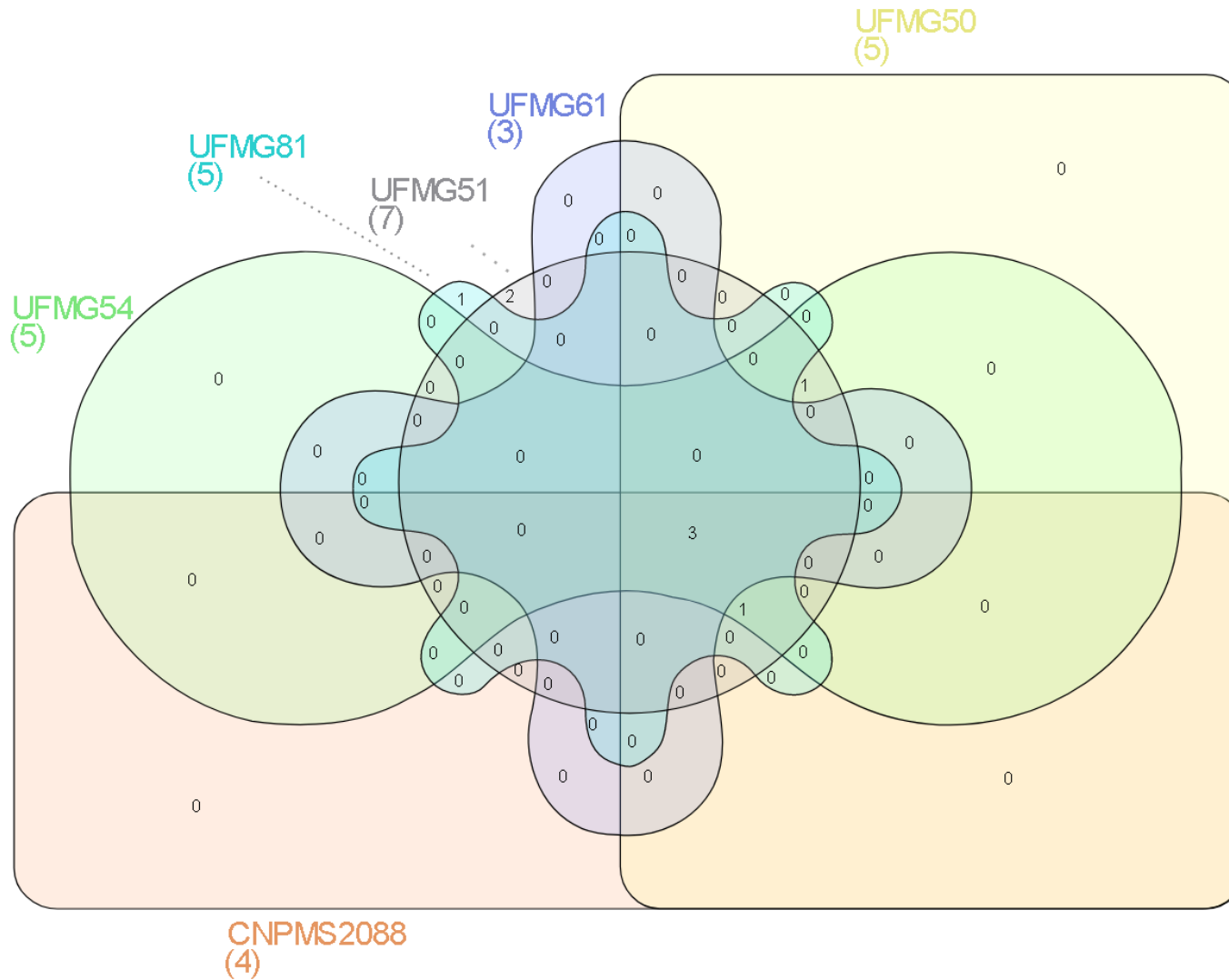
**Figura S2:** Gráfico Venn da distribuição de CDS relacionadas à produção de ácidos orgânicos para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).



**Tabela S2:** Distribuição de CDS e genes entre as bactérias do estudo relacionados à produção de ácidos orgânicos.

SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Organic Acids	9	10	18	13	5	8
Organic acids	(S)-sulfolactate dehydrogenase	0	0	0	0	1	0
	2-ketogluconate reductase	3	0	2	0	0	1
	2,5-diketo-D-gluconic acid reductase A	1	1	1	1	1	0
	2,5-diketo-D-gluconic acid reductase B	1	1	2	0	1	1
	Coenzyme PQQ synthesis protein B	0	0	1	1	0	1
	Coenzyme PQQ synthesis protein D	0	0	1	1	0	1
	D-beta-hydroxybutyrate dehydrogenase	1	2	1	0	1	1
	D-lactate dehydrogenase	0	0	2	2	0	1
	High-affinity gluconate transporter	0	2	1	1	0	1
	L-lactate dehydrogenase	1	1	2	1	1	1
	L-lactate permease	0	1	1	1	0	0
	Low-affinity gluconate transporter	0	0	1	1	0	0
	Malate-2H(+)/Na(+)-lactate antiporter	0	1	0	0	0	0
	Oxalate decarboxylase OxdD	1	0	0	1	0	0
	Putative L-lactate dehydrogenase operon	1	1	3	3	0	0

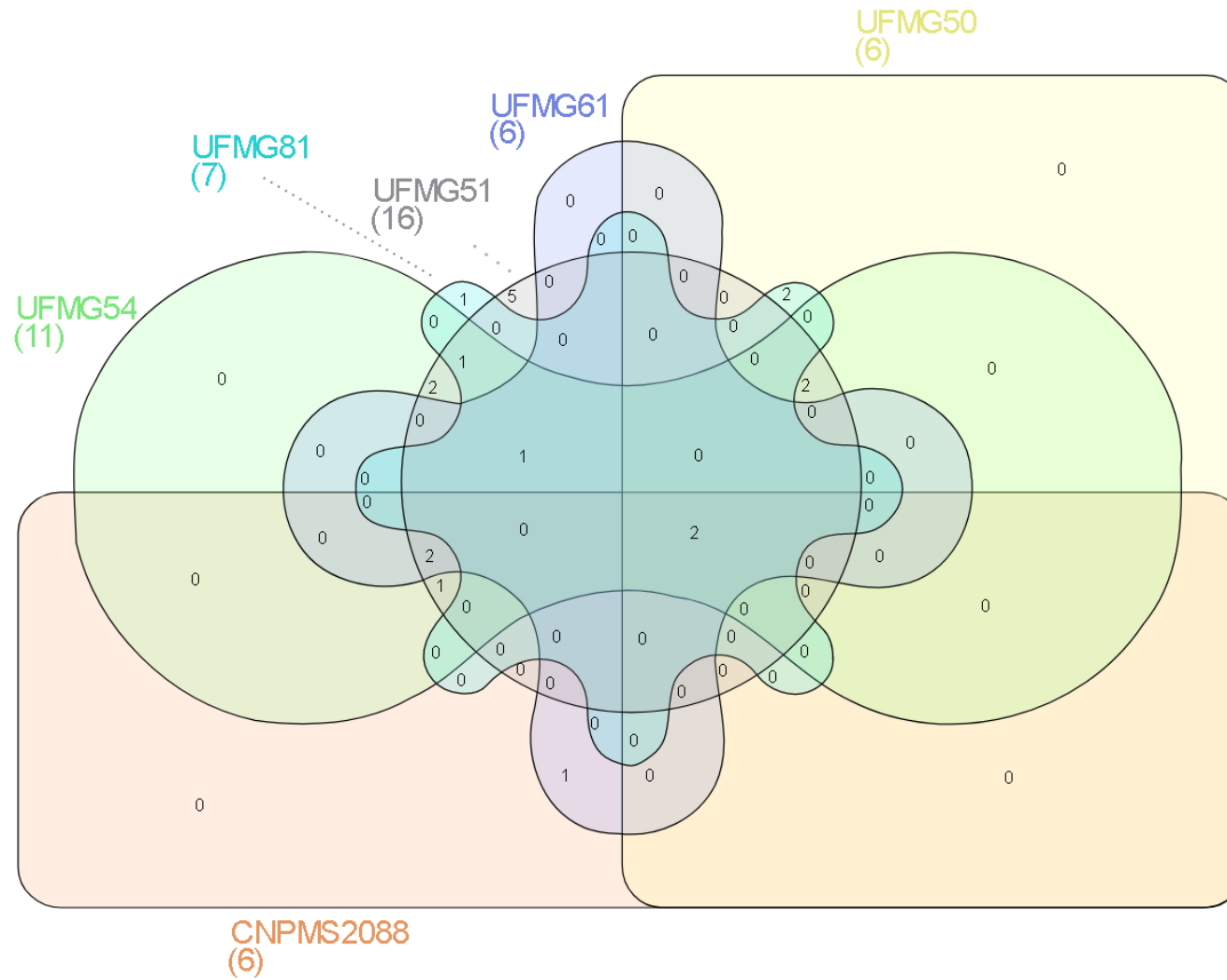
**Figura S3:** Gráfico Venn da distribuição de CDS relacionadas a produção de fito-hormônios para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).



**Tabela S3:** Distribuição de CDS e genes entre as bactérias do estudo relacionados a fito-hormônios.

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	<i>Total Phytohormone</i>	4	6	7	6	3	5
Phytohormone	Anthranilate phosphoribosyltransferase	1	2	1	2	1	1
	Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28)	0	0	0	0	0	1
	Indole-3-acetyl-aspartic acid hydrolase	0	0	1	0	0	0
	Indole-3-pyruvate decarboxylase	0	1	1	1	0	0
	Monoamine oxidase (1.4.3.4)	0	0	1	0	0	0
	Phosphoribosylanthranilate isomerase (EC 5.3.1.24)	1	1	1	1	0	1
	Tryptophan synthase alpha chain	1	1	1	1	1	1
	Tryptophan synthase beta chain	1	1	1	1	1	1

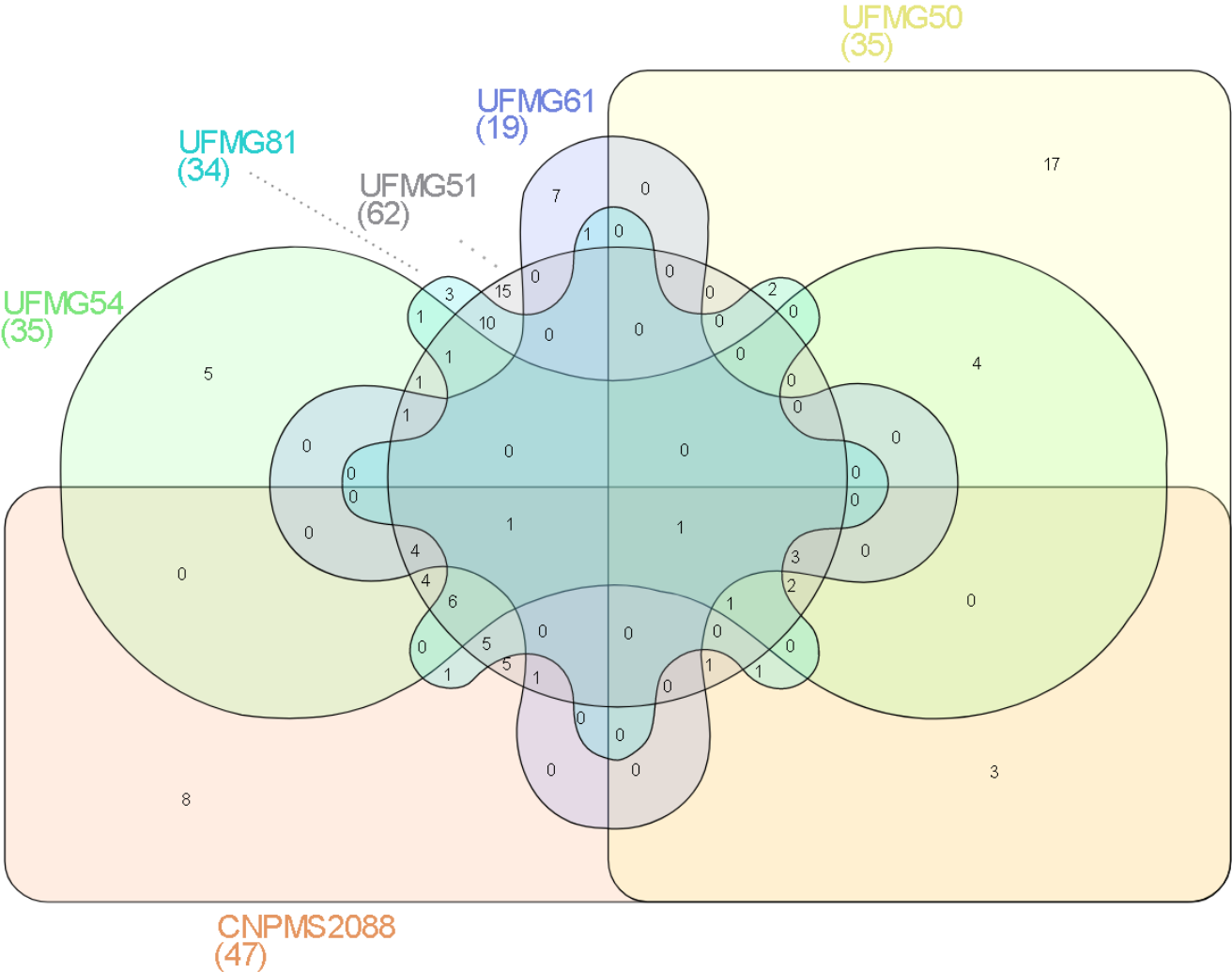
**Figura S4:** Gráfico Venn da distribuição de CDS relacionadas à *quórum sensing*, biofilme e adesão para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).



**Tabela S4:** Distribuição de CDS e genes entre as bactérias do estudo relacionados à *quórum sensing*, biofilme e adesão.

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total <i>Quorum sensing/ Biofilm/ Adhesion</i>	17	14	22	16	8	12
<i>Quorum sensing/ Biofilm/ Adhesion</i>	Alkanal monooxygenase alpha chain	2	5	1	1	2	2
	Autoinducer 2 sensor kinase/phosphatase LuxQ	0	3	0	0	0	1
	Gene bcsA (cellulose synthase catalytic subunit [UDP-forming])	0	0	2	1	1	1
	Gene bcsZ (beta-1,4-glucanase)	0	0	1	0	0	0
	Biofilm PGA outer membrane secretin PgaA	0	0	1	0	0	0
	Biofilm PGA synthesis auxiliary protein PgaD	0	0	1	0	0	0
	Biofilm PGA synthesis deacetylase PgaB (EC 3.-)	0	0	1	0	0	0
	Biofilm PGA synthesis N-glycosyltransferase PgaC (EC 2.4.-.-)	0	0	1	0	0	0
	Gene gumD (polysaccharide biosynthesis glycosyltransferase)	0	0	0	0	0	1
	Homoserine/homoserine lactone efflux protein	4	3	3	3	2	5
	Gene lapA (surface adhesion protein)	0	1	1	1	0	0
	Gene lsrB (Autoinducer 2-binding protein LsrB precursor)	4	0	2	2	1	0
	Gene lsrD (Autoinducer 2 import system permease protein LsrD)	3	0	2	2	1	0
	Gene lsrG (Autoinducer 2-degrading protein LsrG)	0	0	1	1	0	1
	N-acyl homoserine lactonase	1	0	0	0	1	0
	Gene pssA (CDP-diacylglycerol-serine, O-phosphatidyltransferase, acidic exopolysaccharide biosynthesis)	0	0	1	1	0	0
	Putative quorum-quenching lactonase YtnP	0	1	1	2	0	0
	Quorum-sensing regulator protein G	0	0	1	1	0	0
	Transcriptional regulator LsrR	3	0	2	1	0	0
	Gene wapA (Cell wall-associated polypeptide CWBP200)	0	1	0	0	0	1

**Figura S5:** Gráfico Venn da distribuição de CDS relacionadas à aquisição de ferro para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).





**Tabela S5:** Distribuição de CDS e genes entre as bactérias do estudo relacionados à aquisição de ferro.

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Iron Uptake	74	53	77	48	31	72
Iron uptake (1/6)	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EC 1.3.1.28) [enterobactin] siderophore	0	0	1	0	0	0
	2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58) [enterobactin] siderophore	0	0	1	0	0	0
	4'-phosphopantetheinyl transferase (EC 2.7.8.-) [enterobactin] siderophore	0	0	1	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, ATPase component	0	1	0	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, periplasmic iron-binding component	0	1	0	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, permease 2 component	0	1	0	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, permease component	0	3	0	0	0	0
	ABC-type hemin transport system, ATPase component	1	0	0	0	0	0
	ABC-type siderophore export system, fused ATPase and permease components	0	0	0	1	0	0
	Aerobactin siderophore receptor lutA	0	0	1	1	0	1
	Aerobactin synthase (EC 6.3.2.39), aerobactin biosynthesis protein lucC	0	0	0	1	0	0
	Apo-aryl carrier domain of EntB	0	0	1	0	0	0
	Bacterioferritin	1	0	1	1	0	3
	Bacterioferritin-associated ferredoxin	0	0	1	1	0	1
	Biopolymer transport protein ExbB	2	0	2	3	0	5
	Biopolymer transport protein ExbD	2	0	2	3	0	4
Catechololate siderophore receptor Fiu	0	0	1	0	0	0	

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (2/6)	Electron transfer flavoprotein, beta subunit	1	1	0	0	0	0
	Enterobactin esterase	0	0	1	0	0	0
	Enterobactin exporter EntS	2	0	1	1	2	1
	Enterobactin synthase component B	0	0	1	0	0	0
	Enterobactin synthase component D	0	0	1	0	0	1
	Enterobactin synthase component E	0	0	1	0	0	0
	Enterobactin synthase component F	0	0	1	0	0	0
	Enterobactin synthetase component F, serine activating enzyme (EC 2.7.7.-)	0	0	1	0	0	0
	Enterochelin esterase	0	0	1	0	0	1
	Fe(3+) dicitrate transport protein FecA	1	0	1	0	0	3
	Fe(3+)-citrate-binding protein YfmC	1	1	1	1	0	0
	Fe(3+)-pyochelin receptor	1	0	0	0	0	1
	Ferric aerobactin receptor	1	0	1	1	0	1
	Ferric enterobactin receptor	0	0	1	0	0	1
	Ferric enterobactin transport ATP-binding	1	0	2	2	4	0
	Ferric enterobactin transport ATP-binding protein FepC (TC 3.A.1.14.2)	1	0	1	0	0	0
	Ferric enterobactin transport system permease	5	0	2	0	7	0
	Ferric enterobactin transport system permease protein FepD (TC 3.A.1.14.2)	0	0	1	0	0	0
	Ferric enterobactin transport system permease protein FepG (TC 3.A.1.14.2)	1	0	1	0	0	0
	Ferric enterobactin-binding periplasmic protein FepB (TC 3.A.1.14.2)	1	0	1	0	0	0
	Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), ATP-binding protein FhuC	1	0	1	1	0	0
	Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), periplasmic substrate binding protein FhuD	1	0	1	1	0	0
	Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), permease component FhuB	2	0	1	1	0	0

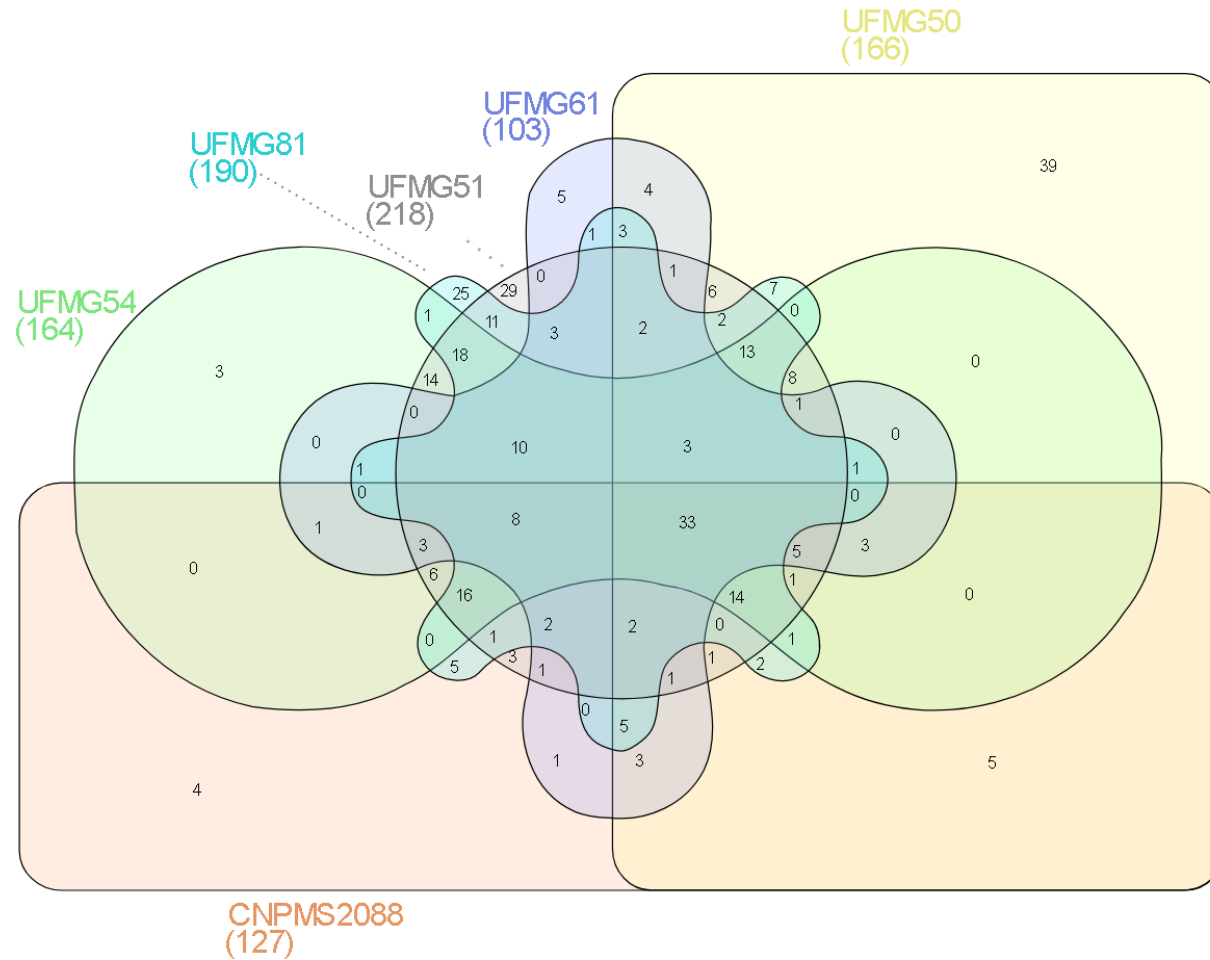
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (3/6)	Ferric iron ABC transporter, ATP-binding protein	2	0	1	0	0	3
	Ferric iron ABC transporter, iron-binding protein	2	0	1	0	0	2
	Ferric iron ABC transporter, permease protein	5	0	2	0	0	2
	Ferric reductase (1.6.99.14)	0	0	1	0	0	0
	Ferric uptake regulation protein	3	1	1	1	0	1
	Ferrichrome receptor FcuA	1	0	1	0	0	3
	Ferrichrome transport ATP-binding protein FhuC (TC 3.A.1.14.3)	1	2	0	0	0	0
	Ferrienterobactin receptor	0	0	2	0	0	0
	Ferrienterobactin-binding periplasmic protein	1	0	1	0	0	0
	Ferrous iron transport periplasmic protein EfeO, contains peptidase-M75 domain and (frequently) cupredoxin-like domain	2	0	1	1	1	0
	Ferrous iron transport permease EfeU	1	2	1	1	1	0
	Ferrous iron transport peroxidase EfeB	1	0	1	1	1	0
	Ferrous-iron efflux pump FieF	1	1	2	1	1	1
	Gene <i>efeU</i> (Ferrous iron uptake protein)	1	1	1	1	1	1
	Gene <i>fhuA</i> (Ferric hydroxamate uptake)	5	0	3	4	0	4
	Gene <i>fhuB</i> (Iron(III)-hydroxamate import system permease protein FhuB)	2	0	1	1	0	0
	Gene <i>fhuC</i> (Iron(3+)-hydroxamate import ATP-binding protein FhuC)	1	0	3	3	1	0
	Gene <i>fhuD</i> (Iron(III)-hydroxamate-binding protein FhuD)	2	2	1	1	0	0
	FIG006045: Sigma factor, ECF subfamily	0	0	0	0	0	6
	Flavin reductase (EC 1.5.1.30)	0	1	0	0	0	0
Heme ABC type transporter HtsABC, heme-binding protein	0	2	0	0	0	0	
Heme ABC type transporter HtsABC, permease protein HtsB	0	1	0	0	0	0	

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (4/6)	Heme ABC type transporter HtsABC, permease protein HtsC	0	2	0	0	0	0
	Heme efflux system ATPase HrtA	0	1	0	0	0	0
	Heme efflux system permease HrtB	0	1	0	0	0	0
	Heme oxygenase (EC 1.14.99.3)	0	0	0	0	2	0
	Heme oxygenase HemO, associated with heme uptake	2	0	0	0	0	0
	Hemin ABC transporter, permease protein	1	0	0	0	0	0
	Hemin import ATP-binding protein HmuV	1	1	1	0	0	1
	Hemin transport protein HmuS	1	0	1	0	0	0
	Iron compound ABC uptake transporter permease protein	0	2	0	0	0	0
	Iron compound ABC uptake transporter substrate-binding protein	1	1	0	0	0	0
	Iron siderophore receptor protein	0	0	0	0	0	6
	Iron siderophore sensor protein	0	0	0	0	0	8
	Iron-chelator utilization protein	2	0	1	2	0	1
	Iron-dependent repressor IdeR/DtxR	0	0	0	0	1	0
	Isochorismatase (EC 3.3.2.1) [enterobactin] siderophore	0	0	1	0	0	0
	Isochorismatase family protein YecD	1	4	2	2	1	1
	Isochorismate synthase (EC 5.4.4.2) [enterobactin] siderophore	0	0	1	0	0	0
	Isochorismate synthase EntC	0	0	1	0	0	0
	L-lysine 6-monooxygenase [NADPH] (EC 1.14.13.59), aerobactin biosynthesis protein lucD	0	0	0	1	0	0
	N(2)-citryl-N(6)-acetyl-N(6)-hydroxylysine synthase (EC 6.3.2.38), aerobactin biosynthesis protein lucA	0	0	0	1	0	0
N6-hydroxylysine O-acetyltransferase (EC 2.3.1.102), aerobactin biosynthesis protein lucB	0	0	0	1	0	0	

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (5/6)	Outer membrane (iron.B12.siderophore.hemin) receptor	0	0	0	0	0	1
	Outer membrane receptor for ferric enterobactin and colicins B, D	0	0	2	0	0	0
	Outer membrane receptor proteins, mostly Fe transport	1	0	0	0	0	0
	Gene pchB (Isochorismate lyase)	0	0	1	0	0	0
	Periplasmic hemin-binding protein	2	0	0	0	0	0
	Petrobactin ABC transporter, ATP-binding protein	0	0	0	0	1	0
	Petrobactin ABC transporter, periplasmic binding protein	0	0	0	0	1	0
	Petrobactin ABC transporter, permease protein I	0	0	0	0	1	0
	Petrobactin ABC transporter, permease protein II	0	0	0	0	1	0
	Possible H <sup>+</sup> -antiporter clustered with aerobactin genes	0	0	0	1	0	0
	Predicted dye-decolorizing peroxidase (DyP), encapsulated subgroup	0	0	0	0	1	1
	Predicted dye-decolorizing peroxidase (DyP), YfeX-like subgroup	0	0	2	0	0	1
	Predicted outer membrane lipoprotein YfeY	0	0	2	0	0	0
	Proofreading thioesterase in enterobactin biosynthesis EntH	0	0	1	0	0	0
	Putative heme iron utilization protein	1	0	0	0	0	0
	Putative siderophore biosynthesis protein SbnA	0	0	1	1	2	1
	Putative TonB-dependent receptor BfrD	0	0	0	0	0	3
	Putative TonB-dependent receptor HI_1217	1	0	0	0	0	0
	Putative TonB-dependent receptor NMB0964	0	0	0	0	0	1
Gene pvdA (ornithine monooxygenase)	0	0	0	0	0	1	

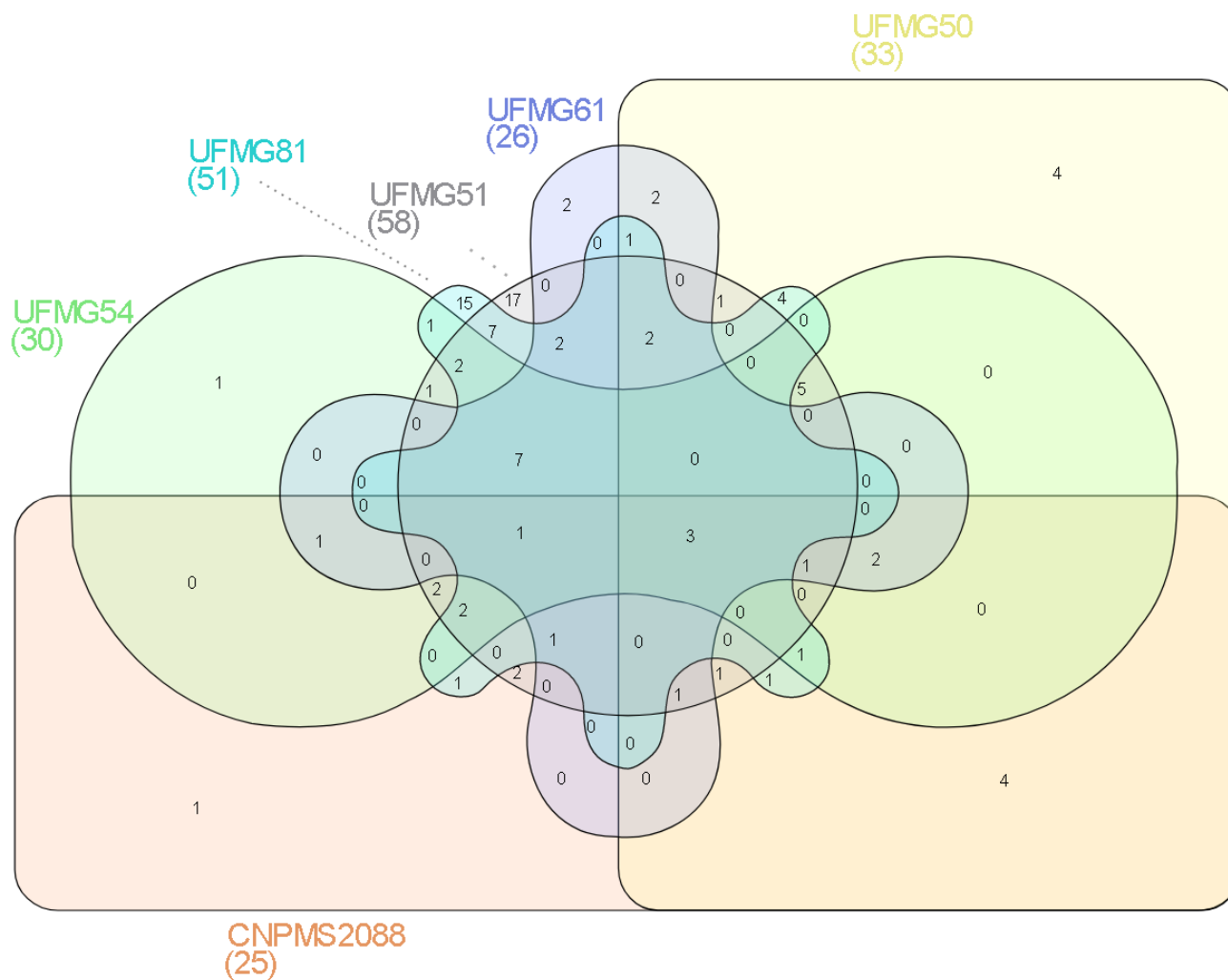
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (6/6)	Sensor histidine kinase colocalized with HrtAB transporter	0	2	0	0	0	0
	Siderophore biosynthesis diaminobutyrate--2-oxoglutarate aminotransferase (EC 2.6.1.76)	0	3	0	0	0	0
	Siderophore biosynthesis L-2,4-diaminobutyrate decarboxylase	0	1	0	1	0	0
	Siderophore biosynthesis protein, monooxygenase	0	1	0	2	0	0
	Siderophore related permease	0	1	0	0	0	0
	Siderophore synthetase component, ligase	0	1	0	1	0	0
	Siderophore synthetase large component, acetyltransferase	0	1	0	0	0	0
	Siderophore synthetase small component, acetyltransferase	0	1	0	1	0	0
	Siderophore transport protein	0	1	0	0	0	0
	Sortase A, LPXTG specific	0	0	0	0	1	0
	TonB-dependent heme receptor A	2	0	0	0	0	0
	TonB-dependent hemin , ferrichrome receptor	0	0	1	0	0	0
	Two-component response regulator colocalized with HrtAB transporter	0	2	0	0	0	0
	Two-component response regulator SA14-24	0	1	0	0	0	0
	Two-component sensor kinase SA14-24	0	4	0	0	0	0
Uncharacterized iron compound ABC uptake transporter, ATP-binding protein	1	1	0	0	0	0	

**Figura S6:** Gráfico Venn da distribuição de CDS relacionadas à categoria versatilidade metabólica para as seis bactérias do estudo. Adaptado de Heberle et al. (2015).



\*A tabela com as CDS e genes de versatilidade metabólica é a Tabela S13 a partir da página 119.

**Figura S7:** Gráfico Venn da distribuição de CDS relacionadas a metabolismo de compostos aromáticos para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).





**Tabela S5:** Distribuição de CDS e genes entre as bactérias do estudo relacionados ao metabolismo de compostos aromáticos.

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total <i>Metabolism of Aromatic Compounds</i>	35	53	72	40	38	83
Metabolism of Aromatic Compounds (1/5)	(S)-2-haloacid dehalogenase 4A	1	0	0	1	1	0
	1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase (EC 1.3.1.25)	0	0	1	0	0	1
	2,3-dihydroxybiphenyl 1,2-dioxygenase (EC 1.13.11.39)	0	0	0	0	0	1
	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-)	0	0	1	0	0	0
	2-hydroxymuconic semialdehyde hydrolase (EC 3.7.1.9)	0	0	1	0	0	0
	2-oxo-hepta-3-ene-1,7-dioic acid hydratase (EC 4.2.-.-)	0	0	1	0	0	0
	3,4-dihydroxyphenylacetate 2,3-dioxygenase (EC 1.13.11.15)	0	0	1	0	0	0
	3-carboxy-cis,cis-muconate cycloisomerase (EC 5.5.1.2)	0	0	2	1	1	1
	3-dehydroquinate dehydratase I (EC 4.2.1.10)	0	1	1	1	0	0
	3-dehydroquinate dehydratase II (EC 4.2.1.10)	1	2	1	1	1	2
	3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)	0	0	0	0	0	1
	3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3)	0	0	0	0	0	1
	3-ketoacyl-CoA thiolase (EC 2.3.1.16)	0	0	0	0	0	5
	3-oxoadipate CoA-transferase subunit A (EC 2.8.3.6)	0	0	1	1	1	1
	3-oxoadipate CoA-transferase subunit B (EC 2.8.3.6)	0	0	1	1	1	1
4-carboxymuconolactone decarboxylase (EC 4.1.1.44)	0	0	5	2	2	3	

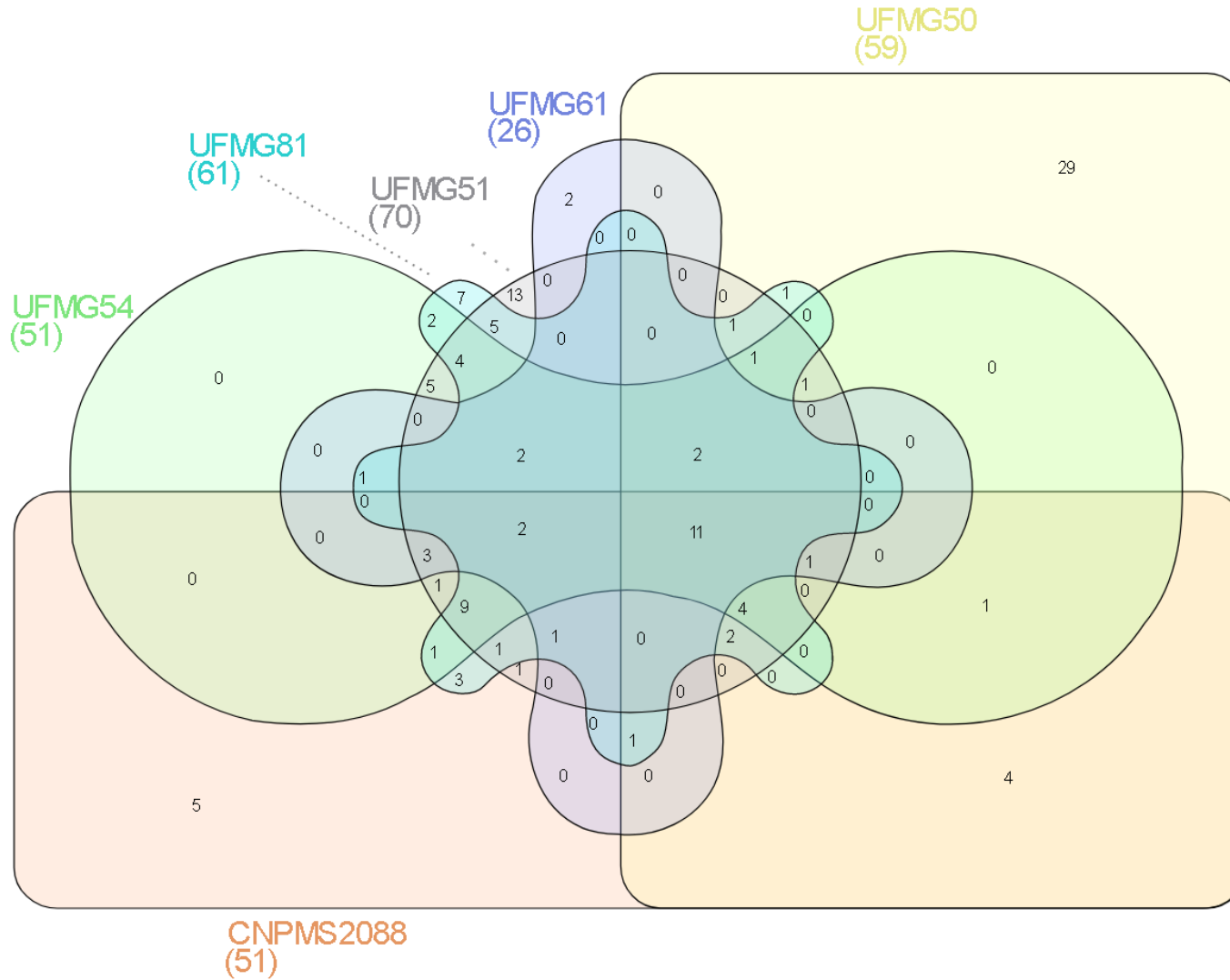
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Metabolism of Aromatic Compounds (2/5)	4-hydroxy-2-oxovalerate aldolase (EC 4.1.3.39)	0	2	0	0	0	1
	4-hydroxybenzoate transporter	0	0	0	2	0	1
	4-hydroxybenzoyl-CoA reductase subunit beta	0	0	0	1	0	0
	4-hydroxybenzoyl-CoA reductase subunit gamma	1	0	0	0	0	2
	4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3)	0	2	0	0	0	0
	4-hydroxyphenylacetate 3-monooxygenase, reductase component (EC 1.6.8.-)	1	0	1	0	0	0
	4-hydroxyphenylacetate symporter, major facilitator superfamily (MFS)	0	0	1	0	0	0
	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	0	1	0	0	0	2
	5-carboxymethyl-2-hydroxymuconate delta-isomerase (EC 5.3.3.10)	0	0	1	0	0	0
	5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.60)	0	0	1	0	0	0
	5-carboxymethyl-2-oxo-hex-3-ene-1,7-dioate decarboxylase (EC 4.1.1.68)	0	0	1	0	0	0
	Acetaldehyde dehydrogenase (EC 1.2.1.10)	0	1	1	0	0	0
	Acetaldehyde dehydrogenase, acetylating, (EC 1.2.1.10) in gene cluster for degradation of phenols, cresols, catechol	0	0	1	0	0	0
	Aldehyde dehydrogenase (EC 1.2.1.3), PaaZ	0	0	0	0	0	1
	Anthranilate 1,2-dioxygenase large subunit	0	0	1	0	0	2
	Aromatic-amino-acid aminotransferase (EC 2.6.1.57)	0	1	0	0	0	2
	benABC operon transcriptional activator BenR	0	0	0	0	0	1

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Metabolism of Aromatic Compounds (3/5)	Benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10)	0	0	1	0	0	1
	Benzoate 1,2-dioxygenase beta subunit (EC 1.14.12.10)	0	0	1	0	0	1
	Benzoate 1,2-dioxygenase, ferredoxin reductase component	0	0	1	0	0	1
	Benzoate MFS transporter BenK	0	1	0	0	1	1
	Benzoate transport protein	2	0	1	1	1	3
	Benzoylformate decarboxylase (EC 4.1.1.7)	0	0	1	0	0	0
	Beta-ketoadipate enol-lactone hydrolase (EC 3.1.1.24)	0	0	1	1	0	2
	biphenyl-2,3-diol 1,2-dioxygenase III-related protein	0	1	0	0	0	0
	Gene <i>bphA3</i> (Biphenyl dioxygenase ferredoxin subunit)	0	0	0	0	1	0
	Carboxymethylenebutenolidase	0	0	2	1	0	0
	Gene <i>catD</i> (3-oxoadipate enol-lactonase 2)	1	2	1	1	1	4
	Catechol 1,2-dioxygenase (EC 1.13.11.1)	0	0	1	0	0	1
	Catechol 2,3-dioxygenase (EC 1.13.11.2)	1	1	0	1	1	0
	Delta(3)-cis-delta(2)-trans-enoyl-CoA isomerase (EC 5.3.3.8)	0	0	0	0	0	1
	Dicarboxylic acid transporter PcaT	0	0	0	0	0	1
	Enoyl-CoA hydratase (EC 4.2.1.17)	0	0	0	0	0	4
	enoyl-CoA hydratase, R-specific	0	0	0	0	0	1
	Fluoroacetate dehalogenase	1	0	1	0	0	0
	Fumarylacetoacetase (EC 3.7.1.2)	0	1	0	0	0	0
	Fumarylacetoacetate hydrolase family protein	1	3	0	1	5	0
Haloacetate dehalogenase H-2	1	0	0	0	0	0	
Haloalkane dehalogenase	1	1	0	1	0	1	
Homogentisate 1,2-dioxygenase (EC 1.13.11.5)	0	1	0	0	0	1	

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Metabolismo of Aromatic Compounds (4/5)	Homoprotocatechuate degradative operon repressor	0	0	1	0	0	0
	Hydroxyaromatic non-oxidative decarboxylase protein B (EC 4.1.1.-)	0	1	1	1	0	0
	Hydroxyaromatic non-oxidative decarboxylase protein C (EC 4.1.1.-)	0	2	1	1	0	0
	Hydroxyaromatic non-oxidative decarboxylase protein D (EC 4.1.1.-)	0	2	1	1	0	0
	Isonitrile hydratase	1	0	3	1	0	0
	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	0	0	0	0	0	8
	Maleylacetoacetate isomerase (EC 5.2.1.2)	0	0	0	0	0	1
	Maleylpyruvate isomerase	0	0	0	0	0	2
	mandelate racemase/muconate lactonizing enzyme family protein	0	0	0	0	1	0
	Gene <i>mhpC</i> (2-hydroxy-6-oxononadienedioate/2-hydroxy-6- oxononatrienedioate hydrolase)	0	0	1	0	0	0
	Monoamine oxidase (1.4.3.4)	0	0	1	0	0	0
	Muconate cycloisomerase (EC 5.5.1.1)	0	0	2	0	1	1
	Muconolactone isomerase (EC 5.3.3.4)	0	0	1	0	0	1
	N-acetylglucosamine repressor	3	3	3	7	6	0
	NAD(P) transhydrogenase subunit alpha	1	0	1	1	0	1
	Gene <i>nagK</i> (Fumarylpyruvate hydrolase)	2	1	1	0	1	0
	Gene <i>ndoA</i> (Naphthalene 1,2-dioxygenase system ferredoxin subunit)	0	1	0	0	1	0
	Gene <i>nhhA</i> (High-molecular weight cobalt-containing nitrile hydratase subunit alpha)	0	0	1	0	0	0
	Nitrile hydratase subunit beta	0	0	1	0	0	0
	Nitrilotriacetate monooxygenase component B (EC 1.14.13.-)	1	1	1	0	0	0
Gene <i>paaJ</i> (3-oxoadipyl-CoA/3-oxo-5,6-dehydrosuberil-CoA thiolase)	1	0	1	0	2	1	
Pca regulon regulatory protein	4	0	2	1	0	1	
Pca regulon regulatory protein PcaR	0	0	2	1	0	1	

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Metabolism of Aromatic Compounds (5/5)	Pentachlorophenol 4-monooxygenase	1	1	0	0	0	0
	Phenylacetaldehyde dehydrogenase (EC 1.2.1.39)	0	0	1	0	0	0
	P-hydroxybenzoate hydroxylase (EC 1.14.13.2)	0	0	1	0	1	1
	Protocatechuate 3,4-dioxygenase alpha chain (EC 1.13.11.3)	0	0	1	1	1	1
	Protocatechuate 3,4-dioxygenase beta chain	0	0	1	1	1	1
	Protocatechuate 3,4-dioxygenase beta chain (EC 1.13.11.3)	0	0	1	1	1	1
	putative 4-hydroxybenzoyl-CoA thioesterase	0	0	0	0	0	1
	Putative hydrolase MhqD	2	1	0	0	0	0
	Gene <i>bcp</i> (Putative peroxiredoxin <i>bcp</i> )	1	1	1	2	1	1
	Putative ring-cleaving dioxygenase MhqA	1	2	0	0	0	0
	Putative ring-cleaving dioxygenase MhqO	2	1	0	0	0	0
	Salicylate hydroxylase (EC 1.14.13.1)	0	1	0	0	1	0
	Shikimate/quininate 5-dehydrogenase I beta (EC 1.1.1.282)	0	2	1	1	0	0
	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)	0	2	1	0	1	1
	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B (EC 2.8.3.5)	0	2	1	0	1	1
	Succinyl-CoA:coenzyme A transferase	1	1	0	0	0	1
	Gene <i>ttgR</i> (Toluene efflux pump <i>ttgABC</i> operon repressor)	0	0	0	0	0	1
	Transcriptional activator of 4-hydroxyphenylacetate 3-monooxygenase operon, XylS/AraC family	0	0	1	0	0	0
	Transcriptional regulator, IclR family	0	5	0	0	0	0
Gene <i>xilF</i> (2-hydroxymuconate semialdehyde hydrolase)	1	0	1	1	0	0	

**Figura S8:** Gráfico Venn da distribuição de CDS relacionadas às subcategorias virulência e resistência a drogas e substâncias tóxicas para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).



**Tabela S7:** Distribuição de CDS e genes entre as bactérias do estudo relacionados a virulência e resistência a drogas e compostos tóxicos.

SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Virulência and Drug and Toxic Compounds Resistance	45	40	61	49	27	56
Virulence	Alpha-hemolysin translocation ATP-binding	0	0	1	0	0	0
	Bifunctional hemolysin/adenylate cyclase	0	0	0	0	0	2
	Hemolysin	0	0	0	0	0	1
	Hemolysin A	0	1	0	0	0	0
	Hemolysin C	2	3	1	2	2	1
	Hemolysin expression-modulating protein Hha	0	0	1	1	0	0
	Hemolysin transporter protein ShIB	0	0	3	1	0	2
	Inner membrane protein translocase component YidC, short form Oxal-like	0	2	0	0	0	0
	Internalin, putative	0	1	0	0	0	0
	L-aspartate oxidase (EC 1.4.3.16)	1	1	2	1	1	1
	LSU ribosomal protein L35p	1	1	1	1	1	1
	Protein YidD	0	1	0	0	0	0
	Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)	1	1	1	1	1	1
	Quinolinate synthetase (EC 2.5.1.72)	1	1	1	1	1	1
	RNA-binding protein Jag	0	1	0	0	0	0
	Translation elongation factor G	1	1	2	1	1	2
	Translation elongation factor Tu	2	1	3	5	2	4
	Translation initiation factor 3	1	1	1	1	1	1
	Type-1 fimbrial protein, A chain	0	0	2	0	0	1
	Virulence sensor histidine kinase PhoQ	0	0	0	1	0	2
Virulence sensor protein BvgS	1	0	0	0	0	4	
Drug and Toxic Compounds Resistance (1/6)	Gene <i>acrB</i> (Acriflavine resistance protein B)	1	0	3	4	0	1
	Acriflavin resistance protein	0	3	0	0	0	0
	Gene <i>acrR</i> (Potential <i>acrAB</i> operon repressor)	1	0	2	1	2	4

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (2/6)	Alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17)	0	2	0	0	0	0
	Aminoglycoside N6'-acetyltransferase (EC 2.3.1.82)	0	2	0	0	0	0
	Arsenic efflux pump protein	0	4	0	0	0	0
	Arsenical resistance operon repressor	0	3	0	0	0	0
	Gene <i>bcr</i> (Sulfonamide resistance protein)	3	2	3	2	2	1
	Beta-lactamase (EC 3.5.2.6)	1	1	1	3	0	1
	Beta-lactamase 3	0	1	0	0	0	0
	Beta-lactamase class A	0	1	0	0	0	0
	Gene <i>hcpA</i> (Beta-lactamase HcpA precursor)	0	0	4	0	0	5
	Beta-lactamase hydrolase-like protein	1	1	0	1	0	0
	Beta-lactamase OXA-18	1	0	0	0	0	0
	Beta-lactamase SHV-1	0	0	1	0	0	0
	Blue copper oxidase CueO precursor	1	0	1	1	0	0
	Gene <i>bmr3</i> (Multidrug-efflux transporter 3)	0	6	1	1	3	1
	Cadmium efflux system accessory protein	0	1	0	0	0	0
	Cd(II)/Pb(II)-responsive transcriptional regulator	0	0	0	0	0	1
	Chloramphenicol acetyltransferase	3	0	0	0	0	1
	Chloramphenicol acetyltransferase 3	0	0	1	0	0	0
	Choloylglycine hydrolase (EC 3.5.1.24)	3	1	0	0	0	0
	Chromate resistance protein ChrB	0	0	1	0	0	0
	Chromate transport protein ChrA	1	0	1	0	0	1
	Cobalt-zinc-cadmium resistance protein	2	2	1	0	0	1
	Cobalt-zinc-cadmium resistance protein CzcD	0	1	0	0	0	0
	CopG protein	0	0	0	0	0	1
	Copper chaperone	1	0	0	0	0	0
	Copper homeostasis protein CutE	1	0	1	1	0	1
	Copper homeostasis protein CutF precursor	0	0	1	0	0	0



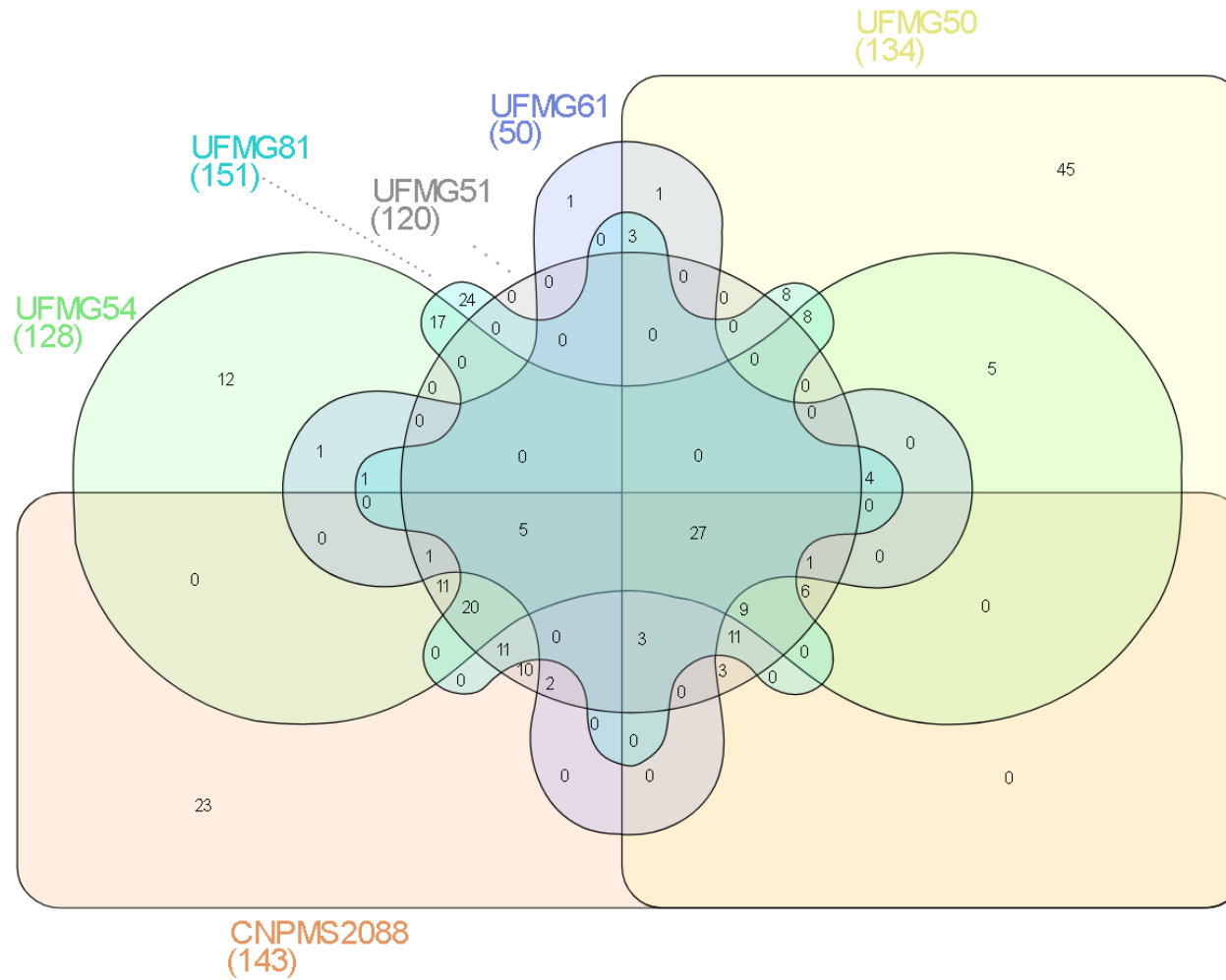
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (3/6)	Copper resistance protein B	0	0	0	0	0	1
	Copper resistance protein CopC	1	0	1	1	1	0
	Copper resistance protein CopD	1	0	1	1	1	0
	Copper resistance protein D	0	2	0	0	0	0
	Copper sensory histidine kinase CusS	0	0	1	0	0	1
	Copper tolerance protein	0	1	0	0	0	1
	Copper-sensing two-component system response regulator CusR	0	0	1	0	0	1
	Copper-translocating P-type ATPase (EC 3.6.3.4)	4	4	3	2	5	3
	Gene <i>cphA</i> (Beta-lactamase precursor)	1	0	0	0	0	1
	Cu(I)-responsive transcriptional regulator	0	0	0	0	1	0
	Cytochrome c heme lyase subunit CcmF	1	0	1	1	0	1
	Cytochrome c heme lyase subunit CcmH	1	0	1	2	0	1
	Cytoplasmic copper homeostasis protein CutC	1	0	1	1	1	0
	D-cysteine desulfhydrase (EC 4.4.1.15)	0	0	1	1	0	0
	Gene <i>drxA</i> (Daunorubicin/doxorubicin resistance ATP-binding protein DrrA)	0	2	1	1	6	1
	Efflux pump membrane transporter BepE	4	0	2	1	0	1
	Efflux pump periplasmic linker BepF	2	0	1	1	0	1
	Gene <i>emrB</i> (Multidrug resistance protein B)	2	6	6	3	0	4
	Gene <i>emrE</i> (Methyl viologen resistance protein C)	1	0	1	0	0	0
	Fosfomycin resistance protein FosA	0	0	1	0	0	0
	Fosfomycin resistance protein FosB	0	1	0	0	0	0
	Fosmidomycin resistance protein	2	1	1	1	0	1
	Glutathione transferase FosA	0	0	1	0	0	1
	Gene <i>macB</i> (Macrolide export ATP-binding/permease protein MacB)	4	0	1	0	1	3
	Macrolide-specific efflux protein MacA	3	0	0	0	0	0

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (4/6)	Magnesium and cobalt efflux protein CorC	1	0	1	1	1	1
	Gene <i>mdtA</i> (Multidrug transporter MdtA)	3	0	3	2	0	4
	Gene <i>mdtB</i> (Multidrug transporter MdtB)	2	0	1	2	0	3
	Gene <i>mdtK</i> (Multidrug-efflux transporter)	0	0	2	1	1	1
	MdtR transcriptional regulator, MarR family	0	3	0	0	0	0
	Membrane component of multidrug resistance system	0	4	0	0	0	0
	Membrane protein, suppressor for copper-sensitivity ScsB	0	0	1	0	0	0
	Membrane protein, suppressor for copper-sensitivity ScsD	0	0	1	0	0	0
	Metal-dependent hydrolases of the beta-lactamase superfamily I	1	2	1	1	0	1
	Metal-dependent hydrolases of the beta-lactamase superfamily III	0	0	1	0	0	0
	Metallothiol transferase FosB	0	2	0	0	0	0
	Metallothiol transferase FosB 2	0	1	0	0	0	0
	Multi antimicrobial extrusion protein (Na <sup>+</sup> )/drug antiporter), MATE family of MDR efflux pumps	1	2	0	0	0	0
	Multicopper oxidase	0	0	0	0	0	3
	Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM, homolog	1	2	0	0	0	0
	Multidrug efflux pump subunit AcrA	0	0	1	1	0	1
	Multidrug efflux transporter MdtP	0	1	0	0	0	0
	Gene <i>mepA</i> (Multidrug export protein MepA)	2	1	1	1	1	0
	Multidrug resistance protein [function not yet clear]	0	3	0	0	0	0
	Multidrug resistance protein D	0	0	1	1	0	0
	Multidrug resistance protein MdtA	3	0	3	2	0	4
	Multidrug resistance protein MdtE	4	0	0	1	0	1
	Multidrug resistance protein MdtH	0	1	3	0	0	1
	Multidrug resistance protein MdtK	0	0	2	1	1	1

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (5/6)	Multidrug resistance protein MdtL	0	0	1	1	0	1
	Multidrug resistance protein MdtN	2	2	2	0	0	1
	Multidrug resistance protein MexA	0	0	1	1	0	1
	Multidrug resistance transporter, Bcr/CflA family	2	1	2	3	1	2
	Multidrug-efflux transporter, major facilitator superfamily (MFS) (TC 2.A.1)	0	1	0	0	0	0
	Multiple antibiotic resistance protein MarA	0	0	1	0	0	0
	Periplasmic divalent cation tolerance protein CutA	0	0	1	1	0	0
	PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	0	0	0	0	1	0
	Phenazine antibiotic resistance protein EhpR	1	0	0	0	0	0
	Probable transcription regulator protein of MDR efflux pump cluster	1	0	0	0	0	0
	Putative multidrug resistance protein EmrK	0	1	2	1	0	3
	Putative multidrug resistance protein EmrY	0	0	0	1	2	1
	Response regulator of zinc sigma-54-dependent two-component system	0	3	1	1	0	0
	Ribosome protection-type tetracycline resistance related proteins, group 2	0	1	0	0	0	0
	Secreted protein, suppressor for copper-sensitivity ScsC	0	0	1	0	0	0
	Sensor protein of zinc sigma-54-dependent two-component system	0	0	1	0	0	0
	Streptogramin A acetyltransferase	1	1	0	0	0	0
	Streptothricin acetyltransferase, Streptomyces lavendulae type	0	1	0	0	0	0
	Suppression of copper sensitivity: putative copper binding protein ScsA	0	0	1	0	0	0
	TetR family regulatory protein of MDR cluster	0	2	0	0	0	0
Topoisomerase IV subunit A (EC 5.99.1.-)	0	1	0	0	0	0	
Topoisomerase IV subunit B (EC 5.99.1.-)	0	1	0	0	0	0	

Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (6/6)	Transcriptional regulator NfxB	0	0	0	0	0	2
	Transcriptional regulator, MerR family	4	8	0	0	6	4
	Translation elongation factor G	0	1	0	0	0	0
	Vancomycin B-type resistance protein VanW	0	2	0	0	0	0
	Virginiamycin A acetyltransferase	0	0	0	1	0	1
	Zinc transporter ZitB	0	0	1	1	0	0

**Figura S9:** Gráfico Venn da distribuição de CDS relacionadas às categorias promoção de crescimento vegetal e colonização rizosférica e endofítica para as seis bactérias do estudo. Adaptado de Heberle *et al.* (2015).



\*As tabelas referentes à figura S7 são as tabelas S10 e S11 a partir da página 91.

**Tabela S8:** Testes *in vitro* para solubilização de diferentes fontes de P e investigação da produção de diferentes compostos relacionados à promoção de crescimento nas plantas das cinco bactérias investigadas neste trabalho. Estes resultados foram obtidos previamente à esse trabalho (Vieira, 2015<sup>1</sup>; Silva, 2017<sup>2</sup>; Silva, 2018<sup>3</sup>).

Categoria	Testes Fenotípicos	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Solubilização de P <sup>1,3</sup>	P-Ca (mg L <sup>-1</sup> )	19	396	504	209	253
	P-Fe (mg L <sup>-1</sup> )	1,5	8	52	1	2
	P-Araxá (mg L <sup>-1</sup> )	40	190	165	255	160
	P-Araxá (Células Sésseis)	–	38	10	9	10
pH final no meio de sol. de P <sup>*1,3</sup>	pH.P- Ca	5,1	4,6	4,2	4,5	4,6
	pH.P-Fe	3,9	3,6	3,2	3,3	3,6
	pH.P-Araxá	1,9	1,6	3,4	3,1	2,6
Mineralização de P <sup>1</sup>	Halo indicativo da hidrólise do fitato	+	+	+	–	+
Fito-hormônios <sup>1</sup>	Produção de AIA (mg L <sup>-1</sup> )	–	–	103	20	63
Sideróforo <sup>1,2</sup>	Produção de quelantes de Fe (μM)	42	0,45	–	–	1,2
Antagonismo <sup>1,2</sup>	Halo de inibição do crescimento	++ **	–	–	–	+ ***
Biofilme <sup>2</sup>	Formação de biofilme em superfície abiótica (UAF****)	85	90	39	15	25
Colonização radicular <sup>2</sup>	Adesão à raiz (UFC g <sup>-1</sup> )	5,2 E <sup>+05</sup>	3,6 E <sup>+05</sup>	7,2 E <sup>+05</sup>	4,6 E <sup>+05</sup>	2,0 E <sup>+05</sup>
	Crescimento no tecido da raiz (UFC g <sup>-1</sup> )	1,0 E <sup>+09</sup>	2,3 E <sup>+09</sup>	5,6 E <sup>+09</sup>	1,3 E <sup>+09</sup>	3,0 E <sup>+09</sup>
EPS <sup>2</sup>	Precipitação do EPS (g L <sup>-1</sup> )	0,27	0,6	0,5	0,18	0,47
Atividade surfactante <sup>2</sup>	E24 (%)	17,3	65	55	55	18,2
	Redução da tensão superficial	27,3	0,0	0,0	0,0	38,3

\* pH inicial no meio de cultura foi 7,0.

\*\* Inibição do crescimento de dois fungos: *Fusarium verticilloides* e *Stenocarpella macrospora*

\*\*\* Inibição do crescimento da bactéria *Bacillus subtilis*

\*\*\*\* Unidades arbitrárias de fluorescência

**Tabela S9:** Resultados dos experimentos *in vivo* para as seis bactérias analisadas representados em porcentagem de desempenho superior em relação ao experimento controle, sem inócuos, para os cultivos em solo acrescido de fosfato de Araxá (A) e Super Fosfato Triplo (ST). Estes resultados foram obtidos previamente à esse trabalho (Vieira, 2015 ; Silva, 2017 ;Silva, 2018).

<b>Categoria</b>	<b>Parâmetros (%)*</b>	<b>UFMG50</b>	<b>UFMG51</b>	<b>UFMG54</b>	<b>UFMG61</b>	<b>UFMG81</b>	<b>CNPMS2088</b>
	<b>Fontes de P</b>	<b>A   ST</b>	<b>A   ST</b>	<b>A   ST</b>	<b>A  ST</b>	<b>A   ST</b>	<b>A   ST</b>
Promoção de crescimento de plantas – <i>in vivo</i>	Tamanho da planta	11   7	-   4,0	-   5,0	-   2,5	-   0,8	7,0   5,0
	Biomassa da parte aérea	15   -	-   -	-   -	-   -	-   -	-   2,0
	Biomassa radicular	20   13	-   23	-   -	7   1,2	-   27	22   11
	Área foliar	17   6	-   16	-   48	-   19	22   9,0	-   16
Conteúdo de P	Biomassa da planta	14   -	22   6	20   8	1   16	-   -	23   2,0
	Solo	44   -	-   -	23   -	8   -	19   -	-   -

\* Valores de cada parâmetro são expressos em porcentagem de aumento em relação ao controle (planta sem a introdução de bactéria) e são apresentados para a planta cultivada com Araxá | Superfosfato triplo, respectivamente, para cada bactéria. O ( - ) indica que não houve porcentagem de aumento para o parâmetro em relação ao experimento controle, sem inócuos.

**Tabela S10:** Distribuição de CDS entre as bactérias do estudo relacionadas à promoção de crescimento às plantas, soma total e distribuição por subcategorias.

Subcategory	Promoção de crescimento de plantas						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Growth Promoting features	92	90	133	85	42	89
Nitrogen fixation (1/2)	Cysteine desulfurase	0	0	2	1	0	1
	FeMo cofactor biosynthesis protein NifB	0	0	1	0	0	0
	fixN (CodCytochrome c oxidase subunit 1 homolog)	1	0	0	0	0	0
	Homocitrate synthase (EC 2.3.3.14)	0	0	1	0	0	0
	Gene <i>hupA</i> (Periplasmic [NiFeSe] hydrogenase small subunit precursor)	0	2	1	1	0	0
	Gene <i>hupB</i> (Periplasmic [NiFeSe] hydrogenase large subunit)	1	0	1	1	0	2
	Iron-sulfur cluster assembly scaffold protein NifU	0	0	1	0	0	0
	Gene NifS (Cysteine desulfurase NifS)	0	1	0	0	0	0
	Nif-specific regulatory protein	0	0	1	0	0	0
	NifT protein	0	0	1	0	0	0
	NifZ protein	0	0	1	0	0	0
	Nitrate reductase	0	0	1	1	0	1
	Nitrite reductase [NAD(P)H]	0	2	2	1	1	1
	Nitrogen assimilation regulatory protein	2	0	1	1	0	0
	Nitrogen fixation regulation protein FixK	1	0	0	0	0	0
	Nitrogenase (molybdenum-iron) alpha chain (EC 1.18.6.1)	0	0	1	0	0	0
	Nitrogenase (molybdenum-iron) beta chain (EC 1.18.6.1)	0	0	1	0	0	0
	Nitrogenase (molybdenum-iron) reductase and maturation protein NifH	0	0	1	0	0	0



Subcategory	Promoção do crescimento de plantas						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Nitrogen fixation (2/2)	Nitrogenase (molybdenum-iron)-specific transcriptional regulator NifA	0	0	1	0	0	0
	Nitrogenase FeMo-cofactor carrier protein NifX	0	0	2	0	0	0
	Nitrogenase FeMo-cofactor scaffold and assembly protein NifE	0	0	1	0	0	0
	Nitrogenase FeMo-cofactor scaffold and assembly protein NifN	0	0	1	0	0	0
	Nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly protein NifB	0	0	1	0	0	0
	Nitrogenase FeMo-cofactor synthesis molybdenum delivery protein NifQ	0	0	1	0	0	0
	Nitrogenase iron protein	0	0	1	0	0	0
	Nitrogenase molybdenum-iron protein alpha chain	0	0	2	0	0	0
	Nitrogenase molybdenum-iron protein beta chain	0	0	2	0	0	0
	Nitrogenase-stabilizing/protective protein NifW	0	0	1	0	0	0
	RNA polymerase sigma-54 factor 2	1	0	0	0	0	0
	Sensor protein FixL	2	0	0	0	0	1
	Serine acetyltransferase	1	1	3	1	1	3
	Transcriptional regulatory protein FixJ	0	0	0	0	0	1
<b>Total Nitrogen Fixation</b>		<b>9</b>	<b>6</b>	<b>33</b>	<b>7</b>	<b>2</b>	<b>10</b>
Phosphate metabolism (1/7)	2-aminoethylphosphonate:pyruvate aminotransferase (EC 2.6.1.37)	0	2	0	0	0	0
	alkaline phosphatase	10	0	10	0	0	0
	Alkaline phosphatase (EC 3.1.3.1)	1	3	2	1	0	0
	Alkaline phosphatase 3	0	1	0	0	0	0
	Alkaline phosphatase 4	0	1	0	1	0	0

Subcategory	Promoção do crescimento de plantas						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate metabolism (2/7)	Alkaline phosphatase D	0	1	0	0	0	0
	Alkaline phosphatase like protein	1	3	0	0	0	0
	Alkaline phosphatase synthesis sensor protein	1	3	0	0	0	0
	Alkylphosphonate utilization operon protein PhnA	0	1	0	0	0	0
	Cobalt-dependent inorganic pyrophosphatase	0	1	0	0	0	0
	Guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase (EC 3.6.1.40)	0	0	1	1	0	0
	Major phosphate-irrepressible acid phosphatase	0	0	2	1	0	0
	Manganese-dependent inorganic pyrophosphatase	0	1	0	0	0	0
	Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)	0	1	0	0	0	0
	PhnB protein	0	1	0	0	0	0
	Gene <i>phnC</i> (Phosphate-import ATP-binding protein PhnC)	1	0	1	0	3	0
	Gene <i>phnG</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnG)	1	0	1	1	0	0
	Gene <i>phnH</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnH)	1	0	1	1	0	0
	Gene <i>phnI</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnI)	1	0	1	1	0	0
Gene <i>phnK</i> (Putative phosphonates utilization ATP-binding protein PhnK)	1	0	1	1	0	0	

Subcategory	Promoção do crescimento de plantas						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate metabolism (3/7)	Gene <i>phnL</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate synthase subunit PhnL)	1	0	1	1	0	0
	PhnO protein	0	1	0	0	0	0
	Gene <i>phnV</i> (Putative 2-aminoethylphosphonate transport system permease protein PhnV)	3	0	1	0	1	0
	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	0	1	0	0	0	0
	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	0	1	0	0	0	0
	Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	1	1	1	1	1	1
	Phosphoenolpyruvate phosphomutase (EC 5.4.2.9)	0	1	0	0	0	0
	Phosphonopyruvate decarboxylase (EC 4.1.1.82)	0	1	0	0	0	0
	Polyphosphate glucokinase (EC 2.7.1.63)	0	0	0	1	1	0
	Probable low-affinity inorganic phosphate transporter	0	1	0	0	1	0
	Gene <i>pstS</i> (Phosphate-binding protein PstS precursor)	1	0	1	2	0	2
	Putative transcriptional regulator PhnF	0	0	1	1	0	0
	Pyrophosphate-energized proton pump (EC 3.6.1.1)	1	0	0	0	0	0
	Sensor protein PhoQ	1	0	1	0	0	0
	Alkaline phosphatase L	0	0	0	0	0	1

Subcategory	Promoção do crescimento de plantas						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate metabolism (4/7)	Alkaline phosphatase synthesis transcriptional	1	6	1	0	1	1
	Exopolyphosphatase	0	1	1	1	1	1
	Extracytoplasmic function (ECF) sigma factor Vrel	0	0	0	0	0	1
	Hemolysin activation/secretion protein associated with VreARI signalling system	0	0	0	0	0	1
	Inorganic pyrophosphatase	2	1	2	2	1	1
	Inorganic pyrophosphatase (EC 3.6.1.1)	2	0	2	2	1	1
	NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)	1	0	1	1	0	1
	Outer membrane TonB-dependent transducer VreA of trans-envelope signaling system	0	0	0	0	0	1
	Gene <i>pgpA</i> (Phosphatidylglycerophosphatase A)	0	0	1	1	0	1
	Gene <i>phnD</i> (Phosphate-import protein PhnD precursor)	2	0	1	1	3	1
	Gene <i>phnN</i> (Ribose 1,5-bisphosphate phosphokinase PhnN)	1	0	1	1	0	1
	Gene <i>phnU</i> (2-aminoethylphosphonate transport system permease PhnU)	1	0	0	0	0	1
	Gene <i>phoP</i> (Alkaline phosphatase synthesis transcriptional regulatory protein PhoP)	2	5	1	2	1	2
	Gene <i>phoR</i> (Alkaline phosphatase synthesis sensor protein PhoR)	1	3	1	1	0	1
Phosphate import ATP-binding protein PstB	1	0	1	1	0	1	

Subcategory	Promoção do crescimento de plantas						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate metabolism (5/7)	Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	1	2	2	1	1	1
	Phosphate transport system regulatory protein PhoU	1	1	1	1	1	1
	Phosphate-binding DING protein (related to PstS)	0	0	0	0	0	1
	Phosphonate ABC transporter phosphate-binding periplasmic component (TC 3.A.1.9.1)	0	4	0	0	0	1
	Phosphonate uptake and metabolism regulator, LysR-family	0	0	0	0	0	1
	Phosphonoacetaldehyde hydrolase (EC 3.11.1.1)	0	1	0	0	0	1
	Polyphosphate kinase (EC 2.7.4.1)	1	1	2	1	1	1
	Gene <i>pqqB</i> (Pyrroloquinoline quinone biosynthesis protein B)	1	0	1	1	0	1
	Gene <i>pqqD</i> (Pyrroloquinoline quinone biosynthesis protein D)	0	0	1	1	0	1
	Predicted ATPase related to phosphate starvation-inducible protein PhoH	0	1	0	0	1	1
	Gene <i>pstB</i> (Phosphate import ATP-binding protein PstB)	1	0	1	1	0	1
	Putative outer membrane TonB-dependent receptor associated with haemagglutinin family outer membrane protein	0	0	0	0	0	1
	Putative PpiC-type peptidyl-prolyl cis-trans isomerase precursor associated with VreARI signaling system	0	0	0	0	0	1
	Pyrroloquinoline-quinone synthase	0	0	1	1	0	1

Subcategory	Promoção do crescimento de plantas						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate metabolism (6/7)	RecA/RadA recombinase	0	0	0	0	0	1
	Ribose 1,5-bisphosphate phosphokinase PhnN	1	0	1	1	0	1
	Sigma factor regulator VreR (cytoplasmic membrane-localized) of trans-envelope signaling system	0	0	0	0	0	1
	Sodium-dependent phosphate transporter	2	0	2	1	0	1
	Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1)	0	0	1	1	0	1
	Gene <i>yliI</i> (Soluble aldose sugar dehydrogenase YliI)	1	0	1	1	0	1
	Exopolyphosphatase (EC 3.6.1.11)	2	1	4	1	1	2
	Filamentous haemagglutinin family outer membrane protein associated with VreARI signalling system	0	0	0	0	0	2
	Low-affinity inorganic phosphate transporter 1	0	1	1	2	1	2
	Phosphate starvation-inducible protein PhoH, predicted ATPase	2	2	4	4	2	2
	Phosphate-import permease protein PhnE	2	0	2	2	4	2
	Phosphate regulon transcriptional regulatory protein PhoB (SphR)	1	4	2	1	1	2
	Phosphonoacetaldehyde hydrolase	0	1	0	0	1	2
	Polyphosphate kinase 2 (EC 2.7.4.1)	1	0	1	1	0	3
	Gene <i>gph</i> (Phosphoglycolate phosphatase)	4	3	1	1	1	4
	Gene <i>phnJ</i> (Alpha-D-ribose 1-methylphosphonate 5-phosphate C-P lyase)	1	0	1	1	0	0

Promoção do crescimento de plantas							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Phosphate metabolism (7/7)	Gene <i>phnM</i> (Alpha-D-ribose 1-methylphosphonate 5-triphosphate diphosphatase)	2	0	1	1	0	0
	Gene <i>phnS</i> (2-aminoethylphosphonate binding protein)	1	0	1	1	0	0
	Gene <i>phoA</i> (Alkaline phosphatase H precursor)	1	1	1	1	0	0
	Gene <i>phoD</i> (Alkaline phosphatase D precursor)	0	1	0	0	0	0
	Gene <i>php</i> (Phosphotriesterase-related protein)	0	1	1	1	0	0
	Gene <i>pstA</i> (Phosphate transport system permease protein PstA)	1	1	1	2	0	0
	Gene <i>ubiG</i> (Ubiquinone biosynthesis O-methyltransferase)	3	0	1	2	2	5
<b>Total Phosphate Metabolism</b>		<b>70</b>	<b>68</b>	<b>75</b>	<b>59</b>	<b>32</b>	<b>66</b>
Organic acids (1/2)	(S)-sulfolactate dehydrogenase	0	0	0	0	1	0
	2-ketogluconate reductase	3	0	2	0	0	1
	2,5-diketo-D-gluconic acid reductase A	1	1	1	1	1	0
	2,5-diketo-D-gluconic acid reductase B	1	1	2	0	1	1
	Coenzyme PQQ synthesis protein B	0	0	1	1	0	1
	Coenzyme PQQ synthesis protein D	0	0	1	1	0	1
	D-beta-hydroxybutyrate dehydrogenase	1	2	1	0	1	1
	D-lactate dehydrogenase	0	0	2	2	0	1
	High-affinity gluconate transporter	0	2	1	1	0	1
	L-lactate dehydrogenase	1	1	2	1	1	1
	L-lactate permease	0	1	1	1	0	0
Low-affinity gluconate transporter	0	0	1	1	0	0	

Promoção do crescimento de plantas							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Organic acids (2/2)	Malate-2H(+)/Na(+)-lactate antiporter	0	1	0	0	0	0
	Oxalate decarboxylase OxdD	1	0	0	1	0	0
	Putative L-lactate dehydrogenase operon	1	1	3	3	0	0
<b>Total Organic Acids</b>		<b>9</b>	<b>10</b>	<b>18</b>	<b>13</b>	<b>5</b>	<b>8</b>
Phytohormone	Anthranilate phosphoribosyltransferase	1	2	1	2	1	1
	Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28)	0	0	0	0	0	1
	Indole-3-acetyl-aspartic acid hydrolase	0	0	1	0	0	0
	Indole-3-pyruvate decarboxylase	0	1	1	1	0	0
	Monoamine oxidase (1.4.3.4)	0	0	1	0	0	0
	Phosphoribosylanthranilate isomerase (EC 5.3.1.24)	1	1	1	1	0	1
	Tryptophan synthase alpha chain	1	1	1	1	1	1
Tryptophan synthase beta chain	1	1	1	1	1	1	
<b>Total Phytohormone</b>		<b>4</b>	<b>6</b>	<b>7</b>	<b>6</b>	<b>3</b>	<b>5</b>



**Tabela S11:** Distribuição de CDS entre as bactérias do estudo relacionadas à colonização rizosférica e endofítica, soma total e distribuição por subcategorias.

SUBCATEGORY	Colonização rizosférica e endofítica						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total colonização do tecido vegetal	124	151	98	165	37	164
Detoxification of reactive oxygen species (1/4)	Alkyl hydroperoxide reductase AhpD	1	1	1	1	0	2
	Alkyl hydroperoxide reductase subunit C-like protein	0	2	2	1	0	2
	bacteriophytochrome heme oxygenase BphO	0	0	0	0	0	1
	Catalase	1	1	0	1	1	2
	Catalase C	1	0	0	0	0	0
	Catalase HP11	0	1	1	1	0	1
	Catalase-peroxidase	0	0	1	1	0	1
	Catalase-related peroxidase	0	0	1	0	0	1
	Cell wall endopeptidase, family M23/M37	0	1	0	0	0	0
	Ferric uptake regulation protein FUR	0	1	1	1	0	1
	Ferroxidase (EC 1.16.3.1)	0	2	0	0	0	0
	FIG005121: SAM-dependent methyltransferase (EC 2.1.1.-)	0	0	2	1	0	1
	Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)	0	4	0	0	0	0
	Fumarate and nitrate reduction regulatory protein	0	0	1	1	0	1
	Gamma-glutamyltranspeptidase (EC 2.3.2.2)	1	0	2	2	0	2
	Glutamate--cysteine ligase (EC 6.3.2.2)	1	0	1	1	0	1
	Glutaredoxin	0	0	0	0	0	1
	Glutaredoxin 1	0	0	1	1	0	0
Glutaredoxin 3 (Grx2)	1	0	0	0	0	0	

SUBCATEGORY	Colonização rizosférica e endofítica (continuação)						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Detoxification of reactive oxygen species (2/4)	Glutaredoxin 3 (Grx3)	0	0	1	1	0	1
	Glutaredoxin-like protein NrdH, required for reduction of Ribonucleotide reductase class Ib	1	0	1	1	1	0
	Glutathione peroxidase (EC 1.11.1.9)	0	0	2	2	1	3
	Glutathione reductase (EC 1.8.1.7)	1	0	1	1	0	1
	Glutathione S-transferase (EC 2.5.1.18)	3	0	6	3	0	3
	Glutathione S-transferase, omega (EC 2.5.1.18)	2	0	1	1	0	1
	Glutathione S-transferase, theta (EC 2.5.1.18)	0	0	1	0	0	0
	Glutathione S-transferase, unnamed subgroup (EC 2.5.1.18)	1	0	0	0	0	3
	Glutathione S-transferase, zeta (EC 2.5.1.18)	0	0	0	0	0	1
	Glutathione synthetase (EC 6.3.2.3)	1	0	1	1	0	1
	Glutathionylspermidine amidohydrolase (EC 3.5.1.78)	0	0	1	0	0	0
	Glutathionylspermidine synthase (EC 6.3.1.8)	0	0	1	0	0	0
	Heme oxygenase HemO, associated with heme uptake	2	0	0	0	0	1
	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	6	1	2	1	0	1
	Iron-binding ferritin-like antioxidant protein	0	2	0	0	0	0
	Iron-responsive regulator Irr	2	0	0	0	0	0
	Lactoylglutathione lyase (EC 4.4.1.5)	2	1	1	2	0	1
	Manganese catalase	0	1	1	0	0	0
	Manganese superoxide dismutase (EC 1.15.1.1)	0	1	0	0	0	0
	Metallothionein	0	0	0	0	0	1
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	0	1	0	0	0	0	
NAD-dependent protein deacetylase of SIR2 family	0	1	0	0	0	0	

Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Detoxification of reactive oxygen species (3/4)	NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)	0	1	0	0	0	0
	Nicotinamidase (EC 3.5.1.19)	0	2	0	0	0	0
	Nicotinate phosphoribosyltransferase (EC 2.4.2.11)	0	2	0	0	0	0
	Nitrite-sensitive transcriptional repressor NsrR	1	1	1	1	0	0
	NnrS protein involved in response to NO	2	0	0	0	0	1
	Non-specific DNA-binding protein Dps	0	2	0	0	0	0
	Organic hydroperoxide resistance protein	3	2	0	1	0	1
	Organic hydroperoxide resistance transcriptional regulator	1	1	0	2	0	1
	Paraquat-inducible protein A	0	0	1	1	0	4
	Paraquat-inducible protein B	0	0	1	1	0	1
	Peroxide stress regulator PerR, FUR family	0	1	0	0	0	0
	Phytochrome, two-component sensor histidine kinase (EC 2.7.3.-)	3	0	0	0	0	1
	Probable glutathione S-transferase (EC 2.5.1.18), YfcF homolog	0	0	1	1	0	0
	Putative manganese catalase	0	1	0	0	0	0
	Redox-sensitive transcriptional activator SoxR	0	0	1	1	0	1
	Redox-sensitive transcriptional regulator (AT-rich DNA-binding protein)	0	1	0	0	0	0
	Rubredoxin	0	0	0	0	0	1
	Rubredoxin-NAD(+) reductase (EC 1.18.1.1)	0	0	0	0	0	1
	Similar to Hydroxyacylglutathione hydrolase, but in an organism lacking glutathione biosynthesis	0	1	0	0	0	0
	Similarity with glutathionylspermidine synthase (EC 6.3.1.8), group 2	0	1	0	0	0	0

Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Detoxification of reactive oxygen species (4/4)	Superoxide dismutase [Cu-Zn]	1	0	1	1	0	0
	Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1)	1	2	1	1	0	0
	Superoxide dismutase [Fe] (EC 1.15.1.1)	0	1	1	0	0	2
	Superoxide dismutase [Mn/Fe]	0	1	0	0	0	1
	Superoxide dismutase [Mn] (EC 1.15.1.1)	1	1	1	1	2	1
	Superoxide dismutase-like protein YojM	0	2	0	0	0	0
	transcriptional regulator, Crp/Fnr family	0	1	0	0	0	0
	Uncharacterized glutathione S-transferase-like protein	1	0	3	1	0	2
	Uncharacterized monothiol glutaredoxin ycf64-like	1	0	0	0	0	0
	Vegetative catalase	0	2	0	0	0	0
	Zinc uptake regulation protein ZUR	0	1	0	0	0	0
<b>Total detoxif. of reactive oxygen species</b>		<b>42</b>	<b>48</b>	<b>46</b>	<b>37</b>	<b>5</b>	<b>53</b>
Chemotaxis/ Mobility (1/5)	Basal-body rod modification protein FlgD	1	1	0	1	0	1
	Gene <i>bdIA</i> (Chemotaxis regulator BdlA)	0	0	0	3	0	2
	Biopolymer transport protein ExbB	2	0	2	3	0	5
	Biopolymer transport protein ExbD	2	0	2	3	0	4
	Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)	1	4	0	0	0	1
	CFA/I fimbrial subunit B	0	0	0	1	0	0
	CFA/I fimbrial subunit E	0	0	0	3	0	0
	Chaperone protein FimC	0	0	2	0	0	0
	Gene <i>cheB</i> (Chemotaxis response regulator protein-glutamate methyltransferase)	1	1	0	1	0	2
	Chemoreceptor glutamine deamidase CheD	0	1	0	0	0	0
	Chemotaxis protein CheA	0	1	0	1	0	2
Chemotaxis protein CheD	0	1	0	0	0	0	

Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Chemotaxi/ Mobility (2/5)	Chemotaxis protein CheV (EC 2.7.3.-)	0	1	0	1	0	3
	Chemotaxis protein CheY	0	1	0	1	0	3
	Chemotaxis protein methyltransferase	0	0	0	1	0	0
	Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)	0	3	0	0	0	0
	Chemotaxis protein PomA	0	2	0	0	1	1
	Chemotaxis regulator - transmits chemoreceptor signals to flagellar motor components CheY	0	1	0	0	0	0
	Gene cheW (Coupling protein CheW)	0	1	0	1	0	3
	Common pilus major fimbriin subunit EcpA	0	0	1	0	0	0
	Fimbria A protein	0	0	0	2	0	0
	Fimbria adhesin EcpD	0	0	1	0	0	0
	Fimbria adhesin protein	0	0	2	0	0	1
	Flagellar assembly protein FliH	0	1	0	0	0	1
	Flagellar basal-body P-ring formation protein FlgA	1	0	0	0	0	1
	Flagellar basal-body rod modification protein FlgD	1	1	0	0	0	1
	Flagellar basal-body rod protein FlgF	2	1	0	0	0	1
	Flagellar biosynthesis protein FlgN	0	0	0	0	0	1
	Flagellar biosynthesis protein FlhA	1	1	0	1	1	1
	Flagellar biosynthesis protein FlhB	1	1	0	0	0	1
	Flagellar biosynthesis protein FlhF	0	1	0	0	0	1
	Flagellar biosynthesis protein FliL	0	1	0	0	0	0
	Flagellar biosynthesis protein FliO	1	0	0	0	0	1
	Flagellar biosynthesis protein FliQ	1	2	0	0	0	1
	Flagellar biosynthesis protein FliS	0	1	0	0	0	1
	Flagellar biosynthesis protein FliZ	0	1	0	0	0	0
	Flagellar biosynthetic protein FlhB	1	1	0	1	1	2
	Flagellar biosynthetic protein FliQ	1	1	0	2	1	1

Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Chemotaxis/ Mobility (3/5)	Flagellar hook protein FlgE	1	0	0	1	1	1
	Flagellar hook-associated protein 1	1	1	0	1	1	1
	Flagellar hook-associated protein FlgK	1	1	0	0	0	1
	Flagellar hook-associated protein FliD	0	1	0	0	0	0
	Flagellar L-ring protein FlgH	1	0	0	0	0	1
	Flagellar motor rotation protein MotA	1	2	0	0	0	2
	Flagellar motor rotation protein MotB	1	2	0	0	0	2
	Flagellar motor switch protein FliG	1	1	0	1	1	1
	Flagellar motor switch protein FliM	0	1	0	1	1	1
	Flagellar motor switch protein FliN	1	1	0	1	1	1
	Flagellar M-ring protein	1	1	0	1	1	1
	Flagellar M-ring protein FliF	1	1	0	0	0	1
	Flagellar P-ring protein FlgI	1	0	0	0	0	1
	Flagellar protein FliD	0	1	0	0	0	0
	Flagellar regulatory protein FleQ	0	0	0	0	0	1
	Flagellar sensor histidine kinase FleS	0	0	0	0	0	1
	Flagellar synthesis regulator FleN	0	1	0	0	0	1
	Flagellar transcriptional regulator FliC	0	0	0	1	0	0
	Flagellin	1	1	0	4	1	1
	Flagellin protein FlaA	1	1	0	0	0	1
	Flagellum-specific ATP synthase FliI	1	1	0	0	0	1
	Gene flbT (flagellar biosynthesis repressor FliB)	1	1	0	0	0	0
	Gene flgA (flagellar basal body P-ring biosynthesis protein FlgA)	0	0	0	1	0	0
	Gene flgB (Putative proximal rod protein)	1	1	0	1	1	1
	Gene flgC (Putative proximal rod protein)	1	1	0	1	1	1
	Gene flgG(Distal rod protein)	1	3	0	1	0	1
	Gene flgH (Basal body L-ring protein)	1	0	0	1	0	1
	Gene flgI (Basal body P-ring protein)	1	0	0	1	0	1

Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Chemotaxis/ Mobility (4/5)	Gene flgL (flagellar hook-associated protein FlgL)	0	1	0	1	1	1
	Gene flgM (flagellar biosynthesis protein FlgM)	0	0	0	1	0	0
	Gene flgN (Flagella synthesis protein FlgN)	0	0	0	1	0	0
	Gene fliE (flagellar hook-basal body protein FliE)	1	1	0	1	1	1
	Gene fliH (flagellar assembly protein H)	0	0	0	1	0	1
	Gene fliJ (flagellar export protein FliJ)	0	1	0	1	0	1
	Gene fliK (Flagellar hook-length control protein FliK)	0	0	0	1	0	1
	Gene fliO (flagellar biosynthesis protein)	0	0	0	1	0	1
	Gene fliP (Flagellar biosynthetic protein FliP precursor)	1	1	0	2	1	1
	Gene fliR (flagellar biosynthesis protein FliR)	0	1	0	1	0	1
	Gene fliY (Sulfate starvation-induced protein 7)	8	0	7	5	1	4
	Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE	0	1	0	0	0	0
	Methyl-accepting chemotaxis protein 3	0	0	0	1	0	0
	Methyl-accepting chemotaxis protein 4	0	0	0	0	0	1
	Methyl-accepting chemotaxis protein I	0	0	0	7	0	0
	Methyl-accepting chemotaxis protein III	0	0	0	11	0	0
	Methyl-accepting chemotaxis protein IV	0	0	0	6	0	0
	Methyl-accepting chemotaxis protein McpB	0	0	0	0	0	1
	Methyl-accepting chemotaxis protein McpC	0	1	0	0	0	0
	Gene motB (Chemotaxis protein MotB)	1	2	0	3	1	2
Negative regulator of flagellin synthesis	0	1	0	0	0	0	

Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Chemotaxis/ Mobility(5/5)	PAP fimbrial minor pilin protein	0	0	0	1	0	0
	Pilin	0	0	1	0	0	0
	Positive regulator of CheA protein activity (CheW)	0	1	0	0	0	0
	Predicted signal transduction protein	0	0	0	0	0	1
	putative ATP synthase YscN	0	1	0	0	1	0
	Putative fimbrial chaperone Yral	0	0	0	1	0	0
	RNA polymerase sigma factor for flagellar operon	0	1	0	0	0	1
	RNA polymerase sigma factor RpoD	1	1	0	0	0	1
	RNA polymerase sigma-54 factor RpoN	2	2	0	0	0	2
	Gene rpoH (RNA polymerase sigma-32 factor)	2	0	1	1	0	1
	Rrf2 family transcriptional regulator	2	2	0	0	0	0
	S-fimbrial adhesin protein SfaS	0	0	1	0	0	0
	S-fimbrial protein subunit SfaG	0	0	1	0	0	0
	Signal transduction histidine kinase CheA (EC 2.7.3.-)	0	1	0	0	0	0
	Gene trg (Ribose and galactose chemoreceptor protein)	0	0	0	9	0	0
	Gene tsr (Serine chemoreceptor protein)	0	0	0	7	0	0
Type 4 prepilin-like proteins leader	0	1	2	1	1	1	
<b>Total Chemotaxis/ Mobility</b>		<b>55</b>	<b>74</b>	<b>23</b>	<b>107</b>	<b>20</b>	<b>90</b>
Quorum sensing/ Biofilm/ Adhesion (1/2)	Alkanal monooxygenase alpha chain	2	5	1	1	2	2
	Autoinducer 2 sensor kinase/phosphatase LuxQ	0	3	0	0	0	1
	Gene bcsA (cellulose synthase catalytic subunit [UDP-forming])	0	0	2	1	1	1
	Gene bcsZ (beta-1,4-glucanase)	0	0	1	0	0	0
	Biofilm PGA outer membrane secretin PgaA	0	0	1	0	0	0



Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
<i>Quorum sensing/ Biofilm/ Adhesion (2/2)</i>	Biofilm PGA synthesis auxiliary protein PgaD	0	0	1	0	0	0
	Biofilm PGA synthesis deacetylase PgaB (EC 3.-)	0	0	1	0	0	0
	Biofilm PGA synthesis N-glycosyltransferase PgaC (EC 2.4.-.-)	0	0	1	0	0	0
	Gene gumD (polysaccharide biosynthesis glycosyltransferase)	0	0	0	0	0	1
	Homoserine/homoserine lactone efflux protein	4	3	3	3	2	5
	Gene lapA (surface adhesion protein)	0	1	1	1	0	0
	Gene lsrB (Autoinducer 2-binding protein LsrB precursor)	4	0	2	2	1	0
	Gene lsrD (Autoinducer 2 import system permease protein LsrD)	3	0	2	2	1	0
	Gene lsrG (Autoinducer 2-degrading protein LsrG)	0	0	1	1	0	1
	N-acyl homoserine lactonase	1	0	0	0	1	0
	Gene pssA (CDP-diacylglycerol-serine, O-phosphatidyltransferase, acidic exopolysaccharide biosynthesis)	0	0	1	1	0	0
	Putative quorum-quenching lactonase YtnP	0	1	1	2	0	0
	Quorum-sensing regulator protein G	0	0	1	1	0	0
	Transcriptional regulator LsrR	3	0	2	1	0	0
	Gene wapA (Cell wall-associated polypeptide CWBP200)	0	1	0	0	0	1
<b>Total <i>Quorum sensing/ Biofilm/ Adhesion</i></b>		<b>17</b>	<b>14</b>	<b>22</b>	<b>16</b>	<b>8</b>	<b>12</b>

		Colonização rizosférica e endofítica (continuação)					
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Exopolysaccharides	Gene <i>algA</i> (alginate biosynthesis protein AlgA 1)	1	0	0	0	0	1
	Gene <i>algD</i> (nucleotide sugar dehydrogenase)	0	0	0	0	0	1
	Gene <i>algJ</i> (alginate O-acetyltransferase)	0	0	0	0	0	3
	Gene <i>algL</i> (mannuronate-specific alginate lyase)	0	0	0	0	0	1
	Gene <i>algX</i> (alginate O-acetyltransferase)	0	0	0	0	0	1
	Glycosyl transferase, group 1 family protein	0	1	0	0	0	0
	Glycosyl transferase, group 2 family protein	0	3	0	0	0	0
	Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)	0	3	0	0	0	0
	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	0	1	0	0	0	0
	Tyrosine-protein kinase transmembrane modulator EpsC	0	1	0	0	0	0
	Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)	0	1	0	0	0	0
<b>Total Exopolysaccharides</b>		<b>1</b>	<b>10</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>7</b>
Degradation of Plant Polymer	Endoglucanase	0	0	2	0	0	0
	Endoglucanase 5A	0	0	0	0	1	0
	Mannanase dehydratase	0	0	2	1	0	0
	Minor endoglucanase Y	0	0	0	1	0	0
<b>Total Degradation of Plant Polymer</b>		<b>0</b>	<b>0</b>	<b>4</b>	<b>2</b>	<b>1</b>	<b>0</b>
PALP Domain	Gene <i>cysK</i> (Cysteine synthase)	1	2	1	1	0	1
	Gene <i>tdcB</i> (L-threonine dehydratase catabolic TdcB)	3	0	1	2	1	0
	O-acetylserine sulfhydrylase	2	0	0	0	1	0
	Threonine synthase	1	2	1	1	1	1
	Tryptophan synthase beta chain	1	1	1	1	1	1
<b>Total PALP Domain</b>		<b>8</b>	<b>5</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>2</b>

Colonização rizosférica e endofítica (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Exoprotease activity	Gene <i>prsD</i> (type I secretion system permease/ATPase PrsD)	1	0	0	0	0	20
<b>Total Exoprotease activity</b>		<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>2</b>

**Tabela S12:** Distribuição de CDS entre as bactérias do estudo relacionadas a biocontrole, soma total e distribuição por subcategorias.

SUBCATEGORY	Biocontrole						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Biocontrole	111	90	108	66	39	105
Iron uptake (1/7)	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EC 1.3.1.28) [enterobactin] siderophore	0	0	1	0	0	0
	2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58) [enterobactin] siderophore	0	0	1	0	0	0
	4'-phosphopantetheinyl transferase (EC 2.7.8.-) [enterobactin] siderophore	0	0	1	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, ATPase component	0	1	0	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, periplasmic iron-binding component	0	1	0	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, permease 2 component	0	1	0	0	0	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, permease component	0	3	0	0	0	0
	ABC-type hemin transport system, ATPase component	1	0	0	0	0	0
	ABC-type siderophore export system, fused ATPase and permease components	0	0	0	1	0	0
	Aerobactin siderophore receptor lutA	0	0	1	1	0	1
	Aerobactin synthase (EC 6.3.2.39), aerobactin biosynthesis protein lucC	0	0	0	1	0	0
	Apo-aryl carrier domain of EntB	0	0	1	0	0	0
	Bacterioferritin	1	0	1	1	0	3
	Bacterioferritin-associated ferredoxin	0	0	1	1	0	1
	Biopolymer transport protein ExbB	2	0	2	3	0	5
	Biopolymer transport protein ExbD	2	0	2	3	0	4
	Catecholate siderophore receptor Fiu	0	0	1	0	0	0

Biocontrole (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (2/7)	Electron transfer flavoprotein, beta subunit	1	1	0	0	0	0
	Enterobactin esterase	0	0	1	0	0	0
	Enterobactin exporter EntS	2	0	1	1	2	1
	Enterobactin synthase component B	0	0	1	0	0	0
	Enterobactin synthase component D	0	0	1	0	0	1
	Enterobactin synthase component E	0	0	1	0	0	0
	Enterobactin synthase component F	0	0	1	0	0	0
	Enterobactin synthetase component F, serine activating enzyme (EC 2.7.7.-)	0	0	1	0	0	0
	Enterochelin esterase	0	0	1	0	0	1
	Fe(3+) dicitrate transport protein FecA	1	0	1	0	0	3
	Fe(3+)-citrate-binding protein YfmC	1	1	1	1	0	0
	Fe(3+)-pyochelin receptor	1	0	0	0	0	1
	Ferric aerobactin receptor	1	0	1	1	0	1
	Ferric enterobactin receptor	0	0	1	0	0	1
	Ferric enterobactin transport ATP-binding	1	0	2	2	4	0
	Ferric enterobactin transport ATP-binding protein FepC (TC 3.A.1.14.2)	1	0	1	0	0	0
	Ferric enterobactin transport system permease	5	0	2	0	7	0
	Ferric enterobactin transport system permease protein FepD (TC 3.A.1.14.2)	0	0	1	0	0	0
	Ferric enterobactin transport system permease protein FepG (TC 3.A.1.14.2)	1	0	1	0	0	0
	Ferric enterobactin-binding periplasmic protein FepB (TC 3.A.1.14.2)	1	0	1	0	0	0
	Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), ATP-binding protein FhuC	1	0	1	1	0	0
Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), periplasmic substrate binding protein FhuD	1	0	1	1	0	0	

Biocontrole (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (3/7)	Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), permease component FhuB	2	0	1	1	0	0
	Ferric iron ABC transporter, ATP-binding protein	2	0	1	0	0	3
	Ferric iron ABC transporter, iron-binding protein	2	0	1	0	0	2
	Ferric iron ABC transporter, permease protein	5	0	2	0	0	2
	Ferric reductase (1.6.99.14)	0	0	1	0	0	0
	Ferric uptake regulation protein	3	1	1	1	0	1
	Ferrichrome receptor FcuA	1	0	1	0	0	3
	Ferrichrome transport ATP-binding protein FhuC (TC 3.A.1.14.3)	1	2	0	0	0	0
	Ferrienterobactin receptor	0	0	2	0	0	0
	Ferrienterobactin-binding periplasmic protein	1	0	1	0	0	0
	Ferrous iron transport periplasmic protein EfeO, contains peptidase-M75 domain and (frequently) cupredoxin-like domain	2	0	1	1	1	0
	Ferrous iron transport permease EfeU	1	2	1	1	1	0
	Ferrous iron transport peroxidase EfeB	1	0	1	1	1	0
	Ferrous-iron efflux pump FieF	1	1	2	1	1	1
	Gene <i>efeU</i> (Ferrous iron uptake protein)	1	1	1	1	1	1
	Gene <i>fhuA</i> (Ferric hydroxamate uptake)	5	0	3	4	0	4
	Gene <i>fhuB</i> (Iron(III)-hydroxamate import system permease protein FhuB)	2	0	1	1	0	0
	Gene <i>fhuC</i> (Iron(3+)-hydroxamate import ATP-binding protein FhuC)	1	0	3	3	1	0
	Gene <i>fhuD</i> (Iron(III)-hydroxamate-binding protein FhuD)	2	2	1	1	0	0
	FIG006045: Sigma factor, ECF subfamily	0	0	0	0	0	6
	Flavin reductase (EC 1.5.1.30)	0	1	0	0	0	0
Heme ABC type transporter HtsABC, heme-binding protein	0	2	0	0	0	0	

Biocontrole (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (4/7)	Heme ABC type transporter HtsABC, permease protein HtsB	0	1	0	0	0	0
	Heme ABC type transporter HtsABC, permease protein HtsC	0	2	0	0	0	0
	Heme efflux system ATPase HrtA	0	1	0	0	0	0
	Heme efflux system permease HrtB	0	1	0	0	0	0
	Heme oxygenase (EC 1.14.99.3)	0	0	0	0	2	0
	Heme oxygenase HemO, associated with heme uptake	2	0	0	0	0	0
	Hemin ABC transporter, permease protein	1	0	0	0	0	0
	Hemin import ATP-binding protein HmuV	1	1	1	0	0	1
	Hemin transport protein HmuS	1	0	1	0	0	0
	Iron compound ABC uptake transporter permease protein	0	2	0	0	0	0
	Iron compound ABC uptake transporter substrate-binding protein	1	1	0	0	0	0
	Iron siderophore receptor protein	0	0	0	0	0	6
	Iron siderophore sensor protein	0	0	0	0	0	8
	Iron-chelator utilization protein	2	0	1	2	0	1
	Iron-dependent repressor IdeR/DtxR	0	0	0	0	1	0
	Isochorismatase (EC 3.3.2.1) [enterobactin] siderophore	0	0	1	0	0	0
	Isochorismatase family protein YecD	1	4	2	2	1	1
	Isochorismate synthase (EC 5.4.4.2) [enterobactin] siderophore	0	0	1	0	0	0
	Isochorismate synthase EntC	0	0	1	0	0	0
	L-lysine 6-monooxygenase [NADPH] (EC 1.14.13.59), aerobactin biosynthesis protein lucD	0	0	0	1	0	0

Biocontrole (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron Uptake (5/7)	N(2)-citryl-N(6)-acetyl-N(6)-hydroxylysine synthase (EC 6.3.2.38), aerobactin biosynthesis protein lucA	0	0	0	1	0	0
	N6-hydroxylysine O-acetyltransferase (EC 2.3.1.102), aerobactin biosynthesis protein lucB	0	0	0	1	0	0
	Outer membrane (iron.B12.siderophore.hemin) receptor	0	0	0	0	0	1
	Outer membrane receptor for ferric enterobactin and colicins B, D	0	0	2	0	0	0
	Outer membrane receptor proteins, mostly Fe transport	1	0	0	0	0	0
	Gene pchB (Isochorismate lyase)	0	0	1	0	0	0
	Periplasmic hemin-binding protein	2	0	0	0	0	0
	Petrobactin ABC transporter, ATP-binding protein	0	0	0	0	1	0
	Petrobactin ABC transporter, periplasmic binding protein	0	0	0	0	1	0
	Petrobactin ABC transporter, permease protein I	0	0	0	0	1	0
	Petrobactin ABC transporter, permease protein II	0	0	0	0	1	0
	Possible H <sup>+</sup> -antiporter clustered with aerobactin genes	0	0	0	1	0	0
	Predicted dye-decolorizing peroxidase (DyP), encapsulated subgroup	0	0	0	0	1	1
	Predicted dye-decolorizing peroxidase (DyP), YfeX-like subgroup	0	0	2	0	0	1
	Predicted outer membrane lipoprotein YfeY	0	0	2	0	0	0
	Proofreading thioesterase in enterobactin biosynthesis EntH	0	0	1	0	0	0



Biocontrole (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron uptake (6/7)	Putative heme iron utilization protein	1	0	0	0	0	0
	Putative siderophore biosynthesis protein SbnA	0	0	1	1	2	1
	Putative TonB-dependent receptor BfrD	0	0	0	0	0	3
	Putative TonB-dependent receptor HI_1217	1	0	0	0	0	0
	Putative TonB-dependent receptor NMB0964	0	0	0	0	0	1
	Gene pvdA (ornithine monooxygenase)	0	0	0	0	0	1
	Sensor histidine kinase colocalized with HrtAB transporter	0	2	0	0	0	0
	Siderophore biosynthesis diaminobutyrate--2-oxoglutarate aminotransferase (EC 2.6.1.76)	0	3	0	0	0	0
	Siderophore biosynthesis L-2,4-diaminobutyrate decarboxylase	0	1	0	1	0	0
	Siderophore biosynthesis protein, monooxygenase	0	1	0	2	0	0
	Siderophore related permease	0	1	0	0	0	0
	Siderophore synthetase component, ligase	0	1	0	1	0	0
	Siderophore synthetase large component, acetyltransferase	0	1	0	0	0	0
	Siderophore synthetase small component, acetyltransferase	0	1	0	1	0	0
	Siderophore transport protein	0	1	0	0	0	0
	Sortase A, LPXTG specific	0	0	0	0	1	0
	TonB-dependent heme receptor A	2	0	0	0	0	0
	TonB-dependent hemin , ferrichrome receptor	0	0	1	0	0	0
	Two-component response regulator colocalized with HrtAB transporter	0	2	0	0	0	0
	Two-component response regulator SA14-24	0	1	0	0	0	0
Two-component sensor kinase SA14-24	0	4	0	0	0	0	

Biococontrole (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Iron uptake (7/7)	Uncharacterized iron compound ABC uptake transporter, ATP-binding protein	1	1	0	0	0	0
	<b>Total Fe uptake</b>	<b>74</b>	<b>53</b>	<b>77</b>	<b>48</b>	<b>31</b>	<b>72</b>
Chitinase	Gene <i>nagA</i> (N-acetylglucosamine-6-phosphate deacetylase)	1	1	1	1	1	1
	<b>Total Chitinase</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
Volatiles	Acetoin catabolism regulatory protein	2	0	0	0	0	3
	Gene <i>acoA</i> (thiamine pyrophosphate-dependent dehydrogenase E1 component subunit alpha)	4	1	1	0	0	0
	Gene <i>acoB</i> (alpha-ketoacid dehydrogenase subunit beta)	0	1	1	0	0	0
	Gene <i>acoC</i> (acetoin dehydrogenase dihydrolipoyllysine-residue acetyltransferase subunit)	3	0	0	0	1	0
	Alpha-acetolactate decarboxylase	0	1	1	1	0	0
	Hydrogen cyanide synthase subunit HcnA	1	0	1	0	1	1
	Hydrogen cyanide synthase subunit HcnB	3	0	1	0	1	1
	Hydrogen cyanide synthase subunit HcnC	2	1	0	0	1	1
	Gene <i>ilvC</i> (Ketol-acid reductoisomerase)	1	1	1	1	1	1
	Gene <i>ilvH</i> (Acetolactate synthase isozyme 3 small subunit)	1	1	1	1	1	1
Gene <i>ilvI</i> (Acetolactate synthase isozyme 3 large subunit)	1	0	2	1	0	1	
<b>Total Volatiles</b>	<b>18</b>	<b>6</b>	<b>9</b>	<b>4</b>	<b>6</b>	<b>9</b>	
Antibiotics (1/2)	Acriflavin resistance protein	0	3	0	0	0	0
	Arsenic efflux pump protein	0	4	0	0	0	0
	Linear gramicidin synthase subunit B	0	0	0	0	0	3
	Linear gramicidin synthase subunit D	0	0	1	1	1	2
	Macrolide export protein MacA	3	0	3	0	0	2
	Macrolide-specific efflux protein MacA	3	0	0	0	0	0

Biococontrole (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Antibiotics (2/2)	Multi antimicrobial extrusion protein (Na <sup>+</sup> )/drug antiporter), MATE family of MDR efflux pumps	1	2	0	0	0	0
	Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM, homolog	1	2	0	0	0	0
	Multidrug efflux pump subunit AcrA	0	0	1	1	0	1
	Multidrug efflux pump subunit AcrB	1	0	3	4	0	1
	Multidrug export protein AcrE	0	0	1	0	0	1
	Multidrug export protein AcrF	0	0	1	0	0	0
	Multidrug export protein EmrA	2	0	3	3	0	2
	Multidrug export protein EmrB	2	6	6	3	0	4
	Multidrug-efflux transporter, major facilitator superfamily (MFS) (TC 2.A.1)	0	1	0	0	0	0
	Probable transcription regulator protein of MDR efflux pump cluster	1	0	0	0	0	0
	Putative multidrug-efflux transporter Rv1634	2	0	0	0	0	0
	Toluene efflux pump outer membrane protein TtgI	1	0	0	0	0	2
	Tyrocidine synthase 3	0	0	0	0	0	3
	<b>Total antibiotics</b>		<b>17</b>	<b>18</b>	<b>19</b>	<b>12</b>	<b>1</b>
Bacteriocin	ABC transporter ATP-binding protein YvcR	0	2	0	0	0	0
	Antilisterial bacteriocin subtilosin	0	1	0	0	0	0
	Bacitracin export ATP-binding protein BceA	0	3	0	0	0	0
	Bacitracin export permease protein BceB	0	1	0	0	0	0
	Bacteriocin BCN5	0	1	0	0	0	0
	Conserved uncharacterized protein CreA	1	0	1	1	0	1
	Inner membrane protein CreD	0	0	1	0	0	1
	Response regulator LiaR	0	1	0	0	0	0
	Two-component response regulator BceR	0	2	0	0	0	0
	Two-component sensor histidine kinase BceS	0	1	0	0	0	0
<b>Total bacteriocin</b>		<b>1</b>	<b>12</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>2</b>

**Tabela S13:** Distribuição de CDS entre as bactérias do estudo relacionadas à versatilidade metabólica, soma total e distribuição por subcategorias.

SUBCATEGORY	Versatilidade metabólica						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Versatilidade Metabólica	209	284	299	198	139	291
Metabolism of Aromatic Compounds (1/6)	(S)-2-haloacid dehalogenase 4A	1	0	0	1	1	0
	1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase (EC 1.3.1.25)	0	0	1	0	0	1
	2,3-dihydroxybiphenyl 1,2-dioxygenase (EC 1.13.11.39)	0	0	0	0	0	1
	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (EC 5.3.3.-)	0	0	1	0	0	0
	2-hydroxymuconic semialdehyde hydrolase (EC 3.7.1.9)	0	0	1	0	0	0
	2-oxo-hepta-3-ene-1,7-dioic acid hydratase (EC 4.2.-.-)	0	0	1	0	0	0
	3,4-dihydroxyphenylacetate 2,3-dioxygenase (EC 1.13.11.15)	0	0	1	0	0	0
	3-carboxy-cis,cis-muconate cycloisomerase (EC 5.5.1.2)	0	0	2	1	1	1
	3-dehydroquinate dehydratase I (EC 4.2.1.10)	0	1	1	1	0	0
	3-dehydroquinate dehydratase II (EC 4.2.1.10)	1	2	1	1	1	2
	3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)	0	0	0	0	0	1
	3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3)	0	0	0	0	0	1
	3-ketoacyl-CoA thiolase (EC 2.3.1.16)	0	0	0	0	0	5

Subcategory	Versatilidade metabólica (continuação)						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Metabolism of Aromatic Compounds (2/6)	3-oxoadipate CoA-transferase subunit A (EC 2.8.3.6)	0	0	1	1	1	1
	3-oxoadipate CoA-transferase subunit B (EC 2.8.3.6)	0	0	1	1	1	1
	4-carboxymuconolactone decarboxylase (EC 4.1.1.44)	0	0	5	2	2	3
	4-hydroxy-2-oxovalerate aldolase (EC 4.1.3.39)	0	2	0	0	0	1
	4-hydroxybenzoate transporter	0	0	0	2	0	1
	4-hydroxybenzoyl-CoA reductase subunit beta	0	0	0	1	0	0
	4-hydroxybenzoyl-CoA reductase subunit gamma	1	0	0	0	0	2
	4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3)	0	2	0	0	0	0
	4-hydroxyphenylacetate 3-monooxygenase, reductase component (EC 1.6.8.-)	1	0	1	0	0	0
	4-hydroxyphenylacetate symporter, major facilitator superfamily (MFS)	0	0	1	0	0	0
	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	0	1	0	0	0	2
	5-carboxymethyl-2-hydroxymuconate delta-isomerase (EC 5.3.3.10)	0	0	1	0	0	0
	5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.60)	0	0	1	0	0	0
	5-carboxymethyl-2-oxo-hex-3-ene-1,7-dioate decarboxylase (EC 4.1.1.68)	0	0	1	0	0	0
	Acetaldehyde dehydrogenase (EC 1.2.1.10)	0	1	1	0	0	0
	Acetaldehyde dehydrogenase, acetylating, (EC 1.2.1.10) in gene cluster for degradation of phenols, cresols, catechol	0	0	1	0	0	0
	Aldehyde dehydrogenase (EC 1.2.1.3), PaaZ	0	0	0	0	0	1

Versatilidade metabólica (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Metabolism of Aromatic Compounds (3/6)	Anthranilate 1,2-dioxygenase large subunit	0	0	1	0	0	2
	Aromatic-amino-acid aminotransferase (EC 2.6.1.57)	0	1	0	0	0	2
	benABC operon transcriptional activator BenR	0	0	0	0	0	1
	Benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10)	0	0	1	0	0	1
	Benzoate 1,2-dioxygenase beta subunit (EC 1.14.12.10)	0	0	1	0	0	1
	Benzoate 1,2-dioxygenase, ferredoxin reductase component	0	0	1	0	0	1
	Benzoate MFS transporter BenK	0	1	0	0	1	1
	Benzoate transport protein	2	0	1	1	1	3
	Benzoylformate decarboxylase (EC 4.1.1.7)	0	0	1	0	0	0
	Beta-ketoadipate enol-lactone hydrolase (EC 3.1.1.24)	0	0	1	1	0	2
	biphenyl-2,3-diol 1,2-dioxygenase III-related protein	0	1	0	0	0	0
	Gene <i>bphA3</i> (Biphenyl dioxygenase ferredoxin subunit)	0	0	0	0	1	0
	Carboxymethylenebutenolidase	0	0	2	1	0	0
	Gene <i>catD</i> (3-oxoadipate enol-lactonase 2)	1	2	1	1	1	4
	Catechol 1,2-dioxygenase (EC 1.13.11.1)	0	0	1	0	0	1
	Catechol 2,3-dioxygenase (EC 1.13.11.2)	1	1	0	1	1	0
	Delta(3)-cis-delta(2)-trans-enoyl-CoA isomerase (EC 5.3.3.8)	0	0	0	0	0	1
	Dicarboxylic acid transporter PcaT	0	0	0	0	0	1
	Enoyl-CoA hydratase (EC 4.2.1.17)	0	0	0	0	0	4
	Enoyl-CoA hydratase, R-specific	0	0	0	0	0	1
Fluoroacetate dehalogenase	1	0	1	0	0	0	
Fumarylacetoacetase (EC 3.7.1.2)	0	1	0	0	0	0	

Versatilidade metabólica (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Metabolismo of Aromatic Compounds (4/6)	Fumarylacetoacetate hydrolase family protein	1	3	0	1	5	0
	Haloacetate dehalogenase H-2	1	0	0	0	0	0
	Haloalkane dehalogenase	1	1	0	1	0	1
	Homogentisate 1,2-dioxygenase (EC 1.13.11.5)	0	1	0	0	0	1
	Homoprotocatechuate degradative operon repressor	0	0	1	0	0	0
	Hydroxyaromatic non-oxidative decarboxylase protein B (EC 4.1.1.-)	0	1	1	1	0	0
	Hydroxyaromatic non-oxidative decarboxylase protein C (EC 4.1.1.-)	0	2	1	1	0	0
	Hydroxyaromatic non-oxidative decarboxylase protein D (EC 4.1.1.-)	0	2	1	1	0	0
	Isonitrile hydratase	1	0	3	1	0	0
	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	0	0	0	0	0	8
	Maleylacetoacetate isomerase (EC 5.2.1.2)	0	0	0	0	0	1
	Maleylpyruvate isomerase	0	0	0	0	0	2
	Mandelate racemase/muconate lactonizing enzyme family protein	0	0	0	0	1	0
	Gene mhpC (2-hydroxy-6-oxononadienedioate/2-hydroxy-6-oxononatrienedioate hydrolase)	0	0	1	0	0	0
	Monoamine oxidase (1.4.3.4)	0	0	1	0	0	0
	Muconate cycloisomerase (EC 5.5.1.1)	0	0	2	0	1	1
	Muconolactone isomerase (EC 5.3.3.4)	0	0	1	0	0	1
	N-acetylglucosamine repressor	3	3	3	7	6	0
	NAD(P) transhydrogenase subunit alpha	1	0	1	1	0	1
Gene nagK (Fumarylpyruvate hydrolase)	2	1	1	0	1	0	

Versatilidade metabólica (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Aromatic Compounds (5/6)	Gene ndoA (Naphthalene 1,2-dioxygenase system ferredoxin subunit)	0	1	0	0	1	0
	Gene nhhA (High-molecular weight cobalt-containing nitrile hydratase subunit alpha)	0	0	1	0	0	0
	Nitrile hydratase subunit beta	0	0	1	0	0	0
	Nitrilotriacetate monooxygenase component B (EC 1.14.13.-)	1	1	1	0	0	0
	Gene paaJ (3-oxoadipyl-CoA/3-oxo-5,6-dehydrosuberil-CoA thiolase)	1	0	1	0	2	1
	Pca regulon regulatory protein	4	0	2	1	0	1
	Pca regulon regulatory protein PcaR	0	0	2	1	0	1
	Pentachlorophenol 4-monooxygenase	1	1	0	0	0	0
	Phenylacetaldehyde dehydrogenase (EC 1.2.1.39)	0	0	1	0	0	0
	P-hydroxybenzoate hydroxylase (EC 1.14.13.2)	0	0	1	0	1	1
	Protocatechuate 3,4-dioxygenase alpha chain (EC 1.13.11.3)	0	0	1	1	1	1
	Protocatechuate 3,4-dioxygenase beta chain	0	0	1	1	1	1
	Protocatechuate 3,4-dioxygenase beta chain (EC 1.13.11.3)	0	0	1	1	1	1
	putative 4-hydroxybenzoyl-CoA thioesterase	0	0	0	0	0	1
	Putative hydrolase MhqD	2	1	0	0	0	0
	Gene bcp (Putative peroxiredoxin bcp)	1	1	1	2	1	1
	Putative ring-cleaving dioxygenase MhqA	1	2	0	0	0	0
	Putative ring-cleaving dioxygenase MhqO	2	1	0	0	0	0
	Salicylate hydroxylase (EC 1.14.13.1)	0	1	0	0	1	0
	Shikimate/quininate 5-dehydrogenase I beta (EC 1.1.1.282)	0	2	1	1	0	0



Versatilidade metabólica (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Aromatic Compounds (6/6)	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)	0	2	1	0	1	1
	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B (EC 2.8.3.5)	0	2	1	0	1	1
	Succinyl-CoA:coenzyme A transferase	1	1	0	0	0	1
	Gene <i>ttgR</i> (Toluene efflux pump <i>ttgABC</i> operon repressor)	0	0	0	0	0	1
	Transcriptional activator of 4-hydroxyphenylacetate 3-monooxygenase operon, <i>XylS/AraC</i> family	0	0	1	0	0	0
	Transcriptional regulator, <i>IclR</i> family	0	5	0	0	0	0
	Gene <i>xilF</i> (2-hydroxymuconate semialdehyde hydrolase)	1	0	1	1	0	0
<b>Total Aromatic Compounds</b>		<b>35</b>	<b>53</b>	<b>72</b>	<b>40</b>	<b>38</b>	<b>83</b>
Methylotrophy (1/2)	(3S)-malyl-CoA thioesterase	1	0	0	0	0	0
	Gene <i>dhaK</i> (Dihydroxyacetone kinase)	2	1	2	1	1	0
	Dihydroneopterin aldolase	1	1	1	1	1	1
	Gene <i>folK</i> (2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase)	1	1	1	1	1	2
	Malate dehydrogenase	1	2	1	1	1	1
	Gene <i>mdcC</i> (Malonate decarboxylase subunit delta)	0	1	1	0	0	0
	Gene <i>mdh</i> (NAD-dependent methanol dehydrogenase)	1	2	1	1	1	1
	Gene <i>panB</i> (3-methyl-2-oxobutanoate hydroxymethyltransferase)	1	1	1	1	1	2
	Phosphoenolpyruvate carboxylase	0	1	1	1	1	1
	Phosphoserine aminotransferase	1	1	1	1	1	1
Phosphoserine phosphatase	1	0	1	1	1	1	

Versatilidade metabólica (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Methylotrophy (2/2)	Gene <i>pqqB</i> (Pyrroloquinoline quinone biosynthesis protein B)	0	0	1	1	0	1
	Gene <i>pqqD</i> (Pyrroloquinoline quinone biosynthesis protein D)	0	0	1	1	0	1
	Pyrroloquinoline-quinone synthase	0	0	1	1	0	1
	Serine 3-dehydrogenase	1	0	2	0	2	0
	Serine hydroxymethyltransferase 1	0	0	0	0	1	0
<b>Total methylotrophy</b>		<b>12</b>	<b>11</b>	<b>17</b>	<b>12</b>	<b>12</b>	<b>13</b>
Secondary metabolismo (1/2)	1,4-dihydroxy-2-naphthoyl-CoA hydrolase	0	0	1	0	0	1
	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	0	0	1	0	0	0
	2-halobenzoate 1,2-dioxygenase large subunit	0	0	1	0	0	1
	2-halobenzoate 1,2-dioxygenase small subunit	0	0	1	0	0	1
	2-keto-3-deoxy-L-fuconate dehydrogenase	1	0	0	0	0	0
	2-succinylbenzoate--CoA ligase	0	1	1	1	0	0
	Catechol 1,2-dioxygenase	0	0	1	0	0	1
	Deferrochelatase/peroxidase EfeB	1	0	1	1	0	0
	Glutathione S-transferase	2	0	0	0	0	1
	Glutathione S-transferase GST-6.0	1	0	2	1	0	0
	Glutathione S-transferase GstA	0	0	1	0	0	0
	Glutathione S-transferase GstB	2	0	3	1	0	2
	Glutathione S-transferase YfcF	0	0	1	1	0	1
	Glyoxal reductase	1	2	0	0	0	0
	Histone deacetylase-like amidohydrolase	1	0	1	0	0	0
	Hydrogen peroxide-inducible genes activator	3	0	5	3	1	3
	Imidazolonepropionase	4	3	1	1	1	1
	Isochorismate synthase EntC	0	0	1	0	0	0
Lanthionine biosynthesis cyclase LanC	0	1	0	0	0	0	
Lanthionine biosynthesis protein LanB	0	1	0	0	0	0	

Versatilidade metabólica (continuação)							
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Secondary Metabolism (2/2)	Non-heme chloroperoxidase	2	0	1	0	2	2
	Organic hydroperoxide resistance protein OhrB	3	2	0	1	1	0
	Protein-L-isoaspartate O-methyltransferase (EC 2.1.1.77)	0	1	0	0	0	0
	Protocatechuate 3,4-dioxygenase alpha chain	0	0	1	1	1	1
	Protocatechuate 3,4-dioxygenase beta chain	0	0	1	1	1	1
	Putative 2-succinylbenzoate--CoA ligase	0	0	0	0	1	0
	Putative deferrochelataase/peroxidase YfeX	0	0	1	0	1	2
	Putative thiol peroxidase	0	1	0	0	0	0
	Putative thioredoxin-2	0	0	0	0	1	0
	Sorbitol-6-phosphate 2-dehydrogenase	0	0	1	1	0	0
	Thermostable monoacylglycerol lipase	0	1	1	0	0	0
	Thiol peroxidase	0	0	1	1	0	1
	Thioredoxin	0	3	1	1	1	0
	Thioredoxin C-1	2	0	0	0	0	0
	Thioredoxin reductase	1	2	1	1	3	0
	Thioredoxin/glutathione peroxidase BtuE	0	0	1	1	1	1
	Thioredoxin-1	0	0	1	2	1	2
	Thioredoxin-2	0	0	1	1	0	1
	Thioredoxin-like protein YdbP	0	1	0	0	0	0
	Trans-aconitate 2-methyltransferase	1	1	1	1	0	2
Ureidoglycolate lyase	2	3	1	2	6	3	
<b>Total Secondary Metabolism</b>		<b>27</b>	<b>23</b>	<b>36</b>	<b>23</b>	<b>22</b>	<b>28</b>
Sulphur Metabolism (1/5)	3'(2'),5'-bisphosphate nucleotidase (EC 3.1.3.7)	0	0	1	1	0	1
	3'(2'),5'-bisphosphate nucleotidase CysQ	1	0	1	1	0	1
	4Fe-4S ferredoxin, iron-sulfur binding	0	0	0	1	0	0
	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	0	1	0	0	0	0

Subcategory	Versatilidade metabólica (continuação)						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Sulphur Metabolism (2/5)	ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components	0	2	0	0	0	0
	Adenylylsulfate kinase (EC 2.7.1.25)	0	1	1	1	0	1
	Adenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.10)	0	1	0	0	0	1
	Adenylylsulfate reductase alpha-subunit (EC 1.8.99.2)	0	0	0	0	0	1
	Adenylylsulfate reductase beta-subunit (EC 1.8.99.2)	0	0	0	0	0	1
	Alkanesulfonate monooxygenase (EC 1.14.14.5)	0	4	0	0	0	0
	Alkanesulfonates ABC transporter ATP-binding protein	0	2	0	0	0	0
	Alkanesulfonates transport system permease protein	0	2	1	0	0	1
	Alkanesulfonates-binding protein	0	2	1	1	0	2
	Alkyl hydroperoxide reductase subunit C-like protein	0	2	2	1	0	2
	Alpha-galactosidase (EC 3.2.1.22)	0	2	0	0	0	0
	Alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17)	0	2	2	1	0	3
	Arylsulfatase (EC 3.1.6.1)	0	0	1	1	0	1
	Bacterial luciferase family protein YtmO, in cluster with L-cystine ABC transporter	0	1	0	0	0	0
	Beta-galactosidase (EC 3.2.1.23)	0	2	0	0	0	0
	Bifunctional enzyme CysN/CysC	1	0	0	0	0	1
	Gene <i>cysA</i> (Sulfate/thiosulfate import ATP-binding protein CysA)	6	1	2	2	3	3
	Gene <i>cysH</i> (Phosphoadenosine phosphosulfate reductase)	1	1	2	1	0	1
	Gene <i>cysL</i> (CysJI operon transcriptional activator)	3	4	1	1	0	1

Subcategory	CDS	Versatilidade metabólica (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Sulphur Metabolism (3/5)	Cystathionine beta-lyase (EC 4.4.1.8)	0	1	0	0	0	0
	Cystathionine gamma-synthase (EC 2.5.1.48)	0	2	0	0	0	0
	Cysteine desulfuration protein SufE	1	0	1	1	0	1
	Gene <i>cysK</i> (Cysteine synthase)	1	2	1	1	0	1
	Ferredoxin	0	5	1	1	0	1
	Ferredoxin--NADP(+) reductase (EC 1.18.1.2)	0	0	1	1	0	2
	FeS cluster assembly protein SufB	1	2	1	1	1	0
	FeS cluster assembly protein SufD	1	0	1	1	1	0
	FIG028220: hypothetical protein co-occurring with HEAT repeat protein	0	0	0	0	0	1
	FMN reductase (EC 1.5.1.29)	0	2	0	0	0	0
	HEAT repeat-containing protein	0	0	0	0	0	1
	HTH-type transcriptional regulator YtII, LysR family	0	1	0	0	0	0
	L-Cystine ABC transporter, ATP-binding protein TcyC	0	1	0	0	0	0
	L-Cystine ABC transporter, ATP-binding protein TcyN	0	1	0	0	0	0
	L-Cystine ABC transporter, periplasmic cystine-binding protein TcyA	0	1	0	0	0	0
	L-Cystine ABC transporter, periplasmic cystine-binding protein TcyK	0	2	0	0	0	0
	L-Cystine ABC transporter, permease protein TcyB	0	1	0	0	0	0
	L-Cystine ABC transporter, permease protein TcyL	0	1	0	0	0	0
	L-Cystine ABC transporter, permease protein TcyM	0	1	0	0	0	0
	L-cystine uptake protein TcyP	0	1	0	0	0	0
O-acetylserine sulfhydrylase	2	0	0	0	1	0	

Subcategory	CDS	Versatilidade metabólica (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Sulphur Metabolism (4/5)	Organosulfonate utilization protein SsuF	0	0	0	0	0	1
	Oxidoreductase probably involved in sulfite reduction	0	0	0	0	0	1
	Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)	0	1	1	1	0	0
	probable beta-D-galactosidase	0	0	1	0	0	0
	probable dibenzothiophene desulfurization enzyme	0	4	2	1	0	2
	Putative alkanesulfonate metabolism utilization regulator	0	0	0	0	0	1
	Putative arylsulfatase regulatory protein	0	1	1	0	0	0
	Putative sulfate permease	0	1	1	0	0	0
	putative sulfoacetate transporter SauU	1	3	0	0	1	1
	Gene <i>sbp</i> (Sulfate binding protein)	1	1	1	1	0	2
	Serine acetyltransferase	1	1	3	1	1	3
	Gene <i>ssuA</i> (Putative aliphatic sulfonates-binding protein precursor)	4	4	4	2	2	4
	Gene <i>ssuB</i> (Aliphatic sulfonates import ATP-binding protein SsuB)	3	3	4	1	1	3
	Gene <i>ssuC</i> (Putative aliphatic sulfonates transport permease protein SsuC)	8	5	7	3	2	3
	Gene <i>ssuE</i> (NAD(P)H-dependent FAD/FMN reductase)	0	1	1	1	0	1
	Gene <i>sufC</i> (putative ATP-dependent transporter SufC)	1	1	1	1	1	0
	Gene <i>sufS</i> (Cysteine desulfurase)	0	1	2	1	0	1
	Sulfate adenylyltransferase subunit 1 (EC 2.7.7.4)	0	0	1	1	0	1
	Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)	1	0	1	1	0	1
	Sulfate adenylyltransferase, dissimilatory-type (EC 2.7.7.4)	0	1	0	0	0	0

Subcategory	CDS	Versatilidade metabólica (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Sulphur Metabolism (5/5)	Sulfate and thiosulfate binding protein CysP	0	0	2	1	0	1
	Sulfate permease, Pit-type	0	1	0	0	0	0
	Sulfate transport system permease protein CysT	2	1	1	1	0	1
	Sulfate transport system permease protein CysW	4	1	4	2	1	4
	Sulfate transporter CysZ	1	0	1	1	1	1
	Sulfate transporter, CysZ-type	0	0	2	1	0	0
	Sulfate-binding protein Sbp	0	1	1	1	0	1
	Sulfite reductase [ferredoxin]	2	0	0	0	0	2
	Sulfite reductase [NADPH] flavoprotein alpha-component (EC 1.8.1.2)	0	1	1	2	0	0
	Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)	0	2	3	1	0	1
	Sulfonate monooxygenase	0	1	0	0	0	0
	Taurine import ATP-binding protein TauB	4	0	2	1	0	1
	Taurine-binding periplasmic protein	1	1	2	1	0	1
	Taurine--pyruvate aminotransferase	0	1	1	0	1	0
	Thiol peroxidase, Bcp-type (EC 1.11.1.15)	1	1	2	1	1	1
	Thiol peroxidase, Tpx-type (EC 1.11.1.15)	0	1	1	1	0	1
Thioredoxin reductase (EC 1.8.1.9)	3	6	1	1	3	1	
<b>Total Sulphur Metabolism</b>		<b>56</b>	<b>100</b>	<b>76</b>	<b>48</b>	<b>21</b>	<b>69</b>
Stress tolerance (1/6)	4-hydroxy-4-methyl-2-oxoglutarate aldolase	2	5	0	0	1	0
	ABC-type Fe <sup>3+</sup> -siderophore transport system, permease 2 component	0	1	0	0	0	0
	Alkyl hydroperoxide reductase AhpD	1	1	1	1	0	2
	Alkyl hydroperoxide reductase subunit C	1	1	1	1	0	1
	Anti-sigma B factor antagonist RsbV	0	0	0	0	1	1
	Gene <i>ahpD</i> (alkylhydroperoxidase AhpD family core domain protein)	1	1	1	1	0	2
	Gene <i>betA</i> (Choline dehydrogenase)	2	0	1	1	0	2

Subcategory	CDS	Versatilidade metabólica (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Stress Tolerance (2/6)	Gene <i>betB</i> (NAD/NADP-dependent betaine aldehyde dehydrogenase)	3	0	1	2	1	3
	Gene <i>betI</i> (transcriptional regulator BetI)	2	0	2	2	4	3
	Gene <i>gbuC</i> (Glycine betaine/carnitine transport binding protein GbuC precursor)	3	0	0	0	0	4
	Carbon starvation protein A	1	1	1	1	1	2
	Carbon storage regulator	0	1	1	1	1	2
	Chaperone protein ClpB	1	1	3	2	1	2
	Chaperone protein DnaJ	2	1	2	1	1	1
	Gene <i>cspA</i> (Cold shock protein CspA)	6	3	1	0	2	3
	Cold shock-like protein CspE	1	1	1	1	0	0
	Cyanoglobin	0	1	1	0	0	0
	Diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	0	6	3	0	0	1
	Gene <i>dnaK</i> (Heat shock protein 70)	2	1	2	2	1	2
	Flavoheмоprotein (Hemoglobin-like protein) (Flavoheмоglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)	0	4	0	0	0	0
	Gamma-glutamyl phosphate reductase	1	3	1	1	1	1
	Gene <i>gltA</i> (Citrate synthase/ Glutamate synthase [NADPH] large chain)	0	2	2	2	1	3
	Gene <i>gltB</i> (Ferredoxin-dependent glutamate synthase 1/ Glutamate synthase [NADPH] small chain)	1	1	1	0	1	1
	Gene <i>proB</i> (Glutamate 5-kinase)	1	2	1	1	1	1
	Glycine betaine transporter OpuD	0	2	0	0	0	0
	Gene <i>grpE</i> (HSP-70 cofactor)	1	1	1	1	1	1
	Gene <i>gstB</i> (Glutathione S-transferase GST-6.0 EC:2.5.1.18)	3	0	5	2	0	2
GTP-binding protein HflX	0	1	0	0	0	0	
Heat shock protein 15	1	1	1	1	1	1	



Subcategory	CDS	Versatilidade metabólica (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Stress Tolerance (3/6)	Heat shock protein HspQ	1	0	1	1	0	0
	Hemoglobin-like protein HbO	0	1	0	0	1	1
	HflC protein	1	0	1	1	0	1
	HflK protein	1	0	1	1	0	1
	Gene <i>hs/O</i> (Heat shock protein 33)	1	1	1	1	0	1
	Gene <i>hs/U</i> (Unfoldase HslU)	1	0	1	1	0	1
	Gene <i>htpX</i> (Protease HtpX homolog)	1	3	1	2	1	1
	HtrA protease/chaperone protein	2	0	1	1	0	1
	Gene <i>ibpA</i> (16 kDa heat shock protein A)	2	1	2	1	0	1
	Intramembrane protease RasP/YluC, implicated in cell division based on FtsL cleavage	1	0	1	0	0	1
	K(+)/H(+) antiporter NhaP2	0	1	1	1	0	2
	Gene <i>kdpA</i> (Potassium-transporting ATPase A chain)	1	1	1	1	1	1
	Gene <i>kdpB</i> (Potassium-transporting ATPase B chain)	1	1	1	1	1	1
	Gene <i>kdpC</i> (Potassium-transporting ATPase C chain)	1	1	1	1	1	1
	Gene <i>kdpE</i> (KDP operon transcriptional regulatory protein KdpE)	1	0	1	1	1	1
	Gene <i>kup</i> (potassium transport protein Kup)	1	0	1	2	0	1
	Multiple antibiotic resistance protein MarA	0	0	1	0	0	0
	Na(+)/H(+) antiporter subunit A	1	1	0	0	1	1
	Na(+)/H(+) antiporter subunit B	0	1	0	0	1	0
	Na(+)/H(+) antiporter subunit C	1	1	0	0	1	0
	Na(+)/H(+) antiporter subunit D	1	1	0	0	1	1
	Na(+)/H(+) antiporter subunit E	0	1	0	0	0	1
	Na(+)/H(+) antiporter subunit F	1	1	0	0	0	1
Na(+)/H(+) antiporter subunit G	1	1	0	0	1	1	

Subcategory	CDS	Versatilidade metabólica (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Stress Tolerance (4/6)	NG,NG-dimethylarginine dimethylaminohydrolase 1 (EC 3.5.3.18)	1	1	0	0	1	0
	Gene <i>nhaA</i> (Sodium/proton antiporter NhaA)	1	0	1	1	1	1
	Gene <i>nhaK</i> (Sodium, potassium, lithium and rubidium/H(+) antiporter)	0	1	2	1	0	1
	Non-heme chloroperoxidase	2	0	1	0	2	2
	Gene <i>opuAA</i> (Glycine betaine transport ATP-binding protein OpuAA)	0	2	0	0	0	0
	Gene <i>opuAB</i> (Glycine betaine transport system permease protein OpuAB)	1	2	0	0	1	5
	Gene <i>opuAC</i> (Glycine betaine-binding protein OpuAC precursor)	0	2	0	0	1	1
	Ornithine aminotransferase (EC 2.6.1.13)	0	1	0	0	1	0
	Gene <i>otsA</i> (Glycine betaine-binding protein OpuAC precursor)	1	0	1	2	1	0
	Gene <i>otsB</i> (Trehalose-6-phosphate phosphatase)	1	0	1	1	1	0
	Outer membrane protein H precursor	0	0	1	1	0	1
	Outer membrane stress sensor protease DegQ, serine protease	0	0	1	1	0	0
	Outer membrane stress sensor protease DegS	0	0	1	1	0	1
	Phage shock protein A	0	0	1	1	0	0
	Phage shock protein B	0	0	1	1	0	0
	Phage shock protein C	0	0	1	1	0	0
	Phage shock protein D	0	0	1	1	0	0
	Psp operon transcriptional activator	0	0	1	1	0	0
	Putative inner membrane protein YjeT (clustered with HflC)	1	0	1	1	0	0
	Pyrroline-5-carboxylate reductase	1	4	2	1	1	1
Right origin-binding protein	0	0	1	0	0	0	

Subcategory	CDS	Versatilidade metabólica (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Stress Tolerance (5/6)	RNA polymerase sigma factor SigB	0	1	0	0	0	0
	RNA-binding protein Hfq	1	1	1	1	0	1
	RsbR, positive regulator of sigma-B	0	16	0	0	0	0
	RsbS, negative regulator of sigma-B	0	1	0	0	0	0
	Sensor protein KdpD	1	1	1	1	1	1
	Serine phosphatase RsbU, regulator of sigma subunit	0	2	0	1	1	2
	Serine-protein kinase RsbW (EC 2.7.11.1)	0	1	0	0	0	2
	SgrR, sugar-phosphate stress, transcriptional activator of SgrS small RNA	0	0	1	1	0	0
	Sigma factor RpoE negative regulatory protein RseA	0	0	2	1	0	1
	Sigma factor RpoE negative regulatory protein RseB precursor	0	0	2	1	0	1
	Starvation sensing protein RspA	0	0	0	1	0	0
	Starvation sensing protein RspB	0	0	2	0	0	0
	Stringent starvation protein A	0	0	1	1	0	1
	Stringent starvation protein B	0	0	0	0	0	1
	Sugar-phosphate stress protein SgrT (embedded in SgrS)	0	0	1	0	0	0
	Gene sodB (Superoxide dismutase [Mn])	1	1	1	1	2	2
	Transcriptional activator RamA	0	0	1	0	0	0
	Trehalase	0	0	0	1	1	1
	Trk system potassium uptake protein TrkA	1	0	1	1	0	1
	Universal stress protein A	0	0	1	1	0	0
	Universal stress protein B	0	0	1	1	0	0
	Universal stress protein C	0	0	1	0	0	0
	Universal stress protein E	0	0	1	1	0	0
	Universal stress protein F	0	0	1	0	0	0
Universal stress protein family 5	0	0	0	0	0	3	
Universal stress protein family 7	0	0	0	0	0	1	

		Versatilidade metabólica (continuação)					
Subcategory	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Stress Tolerance (6/6)	Universal stress protein G	0	0	3	0	0	0
	Gene <i>uspA</i> (Universal stress protein UspA- Required for resistance to DNA-damaging agents)	0	0	2	1	0	0
	<b>Total Stress Tolerance</b>	<b>77</b>	<b>96</b>	<b>95</b>	<b>71</b>	<b>46</b>	<b>93</b>
Lipopolysaccharides	ADP-heptose--LPS heptosyltransferase 2	0	0	1	1	0	1
	Gene <i>lptB</i> (Lipopolysaccharide export system ATP-binding protein LptB)	4	1	3	2	0	3
	Gene <i>lptG</i> (Lipopolysaccharide export system permease protein LptG)	0	0	1	1	0	1
<b>Total Lipopolysaccharides</b>		<b>4</b>	<b>1</b>	<b>5</b>	<b>4</b>	<b>0</b>	<b>5</b>

**Tabela S14:** Distribuição de CDS entre as bactérias do estudo relacionadas à virulência, soma total e distribuição por subcategorias.

Subcategory	Virulência e Resistência						
	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
	Total Virulência e Resistência	162	152	192	140	84	185
Virulence	Alpha-hemolysin translocation ATP-binding	0	0	1	0	0	0
	Bifunctional hemolysin/adenylate cyclase	0	0	0	0	0	2
	Hemolysin	0	0	0	0	0	1
	Hemolysin A	0	1	0	0	0	0
	Hemolysin C	2	3	1	2	2	1
	Hemolysin expression-modulating protein Hha	0	0	1	1	0	0
	Hemolysin transporter protein SHIB	0	0	3	1	0	2
	Inner membrane protein translocase component YidC, short form Oxal-like	0	2	0	0	0	0
	Internalin, putative	0	1	0	0	0	0
	L-aspartate oxidase (EC 1.4.3.16)	1	1	2	1	1	1
	LSU ribosomal protein L35p	1	1	1	1	1	1
	Protein YidD	0	1	0	0	0	0
	Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)	1	1	1	1	1	1
	Quinolinate synthetase (EC 2.5.1.72)	1	1	1	1	1	1
	RNA-binding protein Jag	0	1	0	0	0	0
	Translation elongation factor G	1	1	2	1	1	2
	Translation elongation factor Tu	2	1	3	5	2	4
	Translation initiation factor 3	1	1	1	1	1	1
	Type-1 fimbrial protein, A chain	0	0	2	0	0	1
	Virulence sensor histidine kinase PhoQ	0	0	0	1	0	2
Virulence sensor protein BvgS	1	0	0	0	0	4	
<b>Total Virulence</b>		<b>11</b>	<b>16</b>	<b>19</b>	<b>16</b>	<b>10</b>	<b>24</b>
Secretory system (1/3)	ATP synthase subunit alpha	1	1	1	1	1	1
	Type II secretion system protein L	0	0	1	0	0	0
	Type IV secretion system protein PtlE	0	0	2	0	0	0

Subcategory	CDS	Virulência (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Secretory system (2/3)	Type IV secretion system protein VirB1	0	0	2	0	0	0
	Type IV secretion system protein VirB10	0	0	2	0	0	0
	Type IV secretion system protein VirB11	1	0	2	0	0	0
	Type IV secretion system protein VirB4	1	0	2	0	0	0
	Type IV secretion system protein VirB9	1	0	2	0	0	0
	ATP-dependent zinc metalloprotease FtsH	2	1	1	1	1	1
	ATP-dependent zinc metalloprotease FtsH 4	0	0	0	0	0	1
	Cell division ATP-binding protein FtsE	1	1	1	2	1	1
	Cell division protein FtsA	1	1	2	1	0	1
	DNA translocase FtsK	2	1	1	1	1	1
	Gene <i>gspD</i> (Putative general secretion pathway protein D)	0	0	0	0	0	1
	Membrane protein insertase YidC	1	0	1	1	1	1
	Modulator of FtsH protease HflC	2	0	1	1	0	1
	Peptidoglycan synthase FtsI	1	0	2	2	0	1
	Protein translocase subunit SecA	1	1	1	1	1	1
	Protein translocase subunit SecE	1	1	1	1	1	1
	Protein translocase subunit SecY	1	1	1	1	1	1
	Protein-export protein SecB	1	0	1	1	0	1
	Gene <i>prsD</i> (Type I secretion system ATP-binding protein PrsD)	1	0	0	0	0	2
	Gene <i>prsE</i> (Type I secretion system membrane fusion protein PrsE)	3	0	0	1	0	4
	Signal peptidase I	1	0	1	1	1	1
	Signal recognition particle protein	1	1	1	1	1	1
	Signal recognition particle receptor FtsY	1	1	1	1	1	1
Gene <i>tatA</i> (Sec-independent protein translocase protein TatA)	1	0	1	2	2	1	
Gene <i>tatB</i> (Sec-independent protein translocase protein TatB)	1	0	1	1	2	1	

Subcategory	CDS	Virulência (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Secretory System (3/3)	Gene <i>tatC</i> (Sec-independent protein translocase protein TatC)	1	0	1	1	2	1
	Gene <i>toIC</i> (type I secretion outer membrane protein, TolC family)	1	0	1	1	0	3
	Type II secretion system protein E	0	0	2	1	0	1
	Type II secretion system protein J	0	0	1	0	0	1
	Lipoprotein signal peptidase	1	1	1	1	1	2
	Type I secretion system ATP-binding protein	1	0	0	0	0	2
	Type II secretion system protein D	0	0	1	0	0	2
	Modulator of FtsH protease HflK	2	0	2	2	1	3
	Type II secretion system protein F	0	1	2	2	0	3
	Sensor protein QseC	2	0	1	3	0	4
	Type I secretion system membrane fusion protein	3	0	0	1	0	4
	Gene <i>gspK</i> (Glucosamine kinase GspK)	1	2	1	1	0	0
	Gene <i>icmF</i> (type VI secretion protein IcmF)	0	1	0	0	0	0
	Gene <i>pipB2</i> (Type III effector pipB2)	0	0	0	1	0	0
	Gene <i>secG</i> (protein-export membrane protein SecG)	1	1	1	1	1	0
<b>Total Secretory System</b>		<b>40</b>	<b>16</b>	<b>46</b>	<b>36</b>	<b>20</b>	<b>51</b>
Heavy Metal Resistance (1/3)	Gene <i>arsC</i> (Arsenate reductase)	2	2	2	2	1	1
	Gene <i>actP</i> (Copper-transporting P-type ATPase)	1	1	1	1	1	2
	Glutaredoxin arsenate reductase (EC:1.20.4.1)	0	1	0	0	1	1
	Arsenate reductase (EC:1.20.4.4)	1	0	2	2	0	1
	Gene <i>cadA</i> (Cadmium, zinc and cobalt-transporting ATPase)	0	1	1	0	0	1
	Chromate transport protein	1	2	0	0	0	0
	Cobalt transport protein CbiM	0	0	1	0	0	0

Subcategory	CDS	Virulência (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Heavy Metal Resistance (2/3)	Gene <i>comR</i> (Copper outer membrane regulator)	2	1	1	2	0	0
	Gene <i>copA</i> (Copper resistance protein A precursor)	1	1	1	1	0	3
	Copper-exporting P-type ATPase A	1	1	1	1	0	1
	Gene <i>copB</i> (copper resistance protein B )	0	0	0	0	2	2
	Gene <i>cusF</i> (Cation efflux system protein CusF precursor)	0	0	1	0	0	0
	Gene <i>czcD</i> (Cadmium, cobalt and zinc/H(+)-K(+)-antiporter)	1	1	0	0	1	0
	Divalent metal cation transporter MntH	1	2	2	1	1	0
	DNA-binding heavy metal response regulator	0	0	0	0	0	1
	Gene <i>dps2</i> (DNA starvation/stationary phase protection protein)	0	1	0	0	0	0
	Heavy metal resistance transcriptional regulator HmrR	0	1	0	0	0	1
	Heavy metal sensor histidine kinase	0	0	0	0	0	2
	Gene <i>hmrR</i> (Copper export regulator)	2	0	0	0	1	1
	Gene <i>katB</i> (Catalase)	0	0	0	0	0	1
	Gene <i>katE</i> (Catalase HP11)	1	1	1	1	0	1
	Gene <i>katG</i> (Catalase/oxidase HPI)	0	0	1	1	0	1
	Magnesium and cobalt efflux protein CorC	2	2	4	3	2	3
	Mercuric reductase	0	0	0	0	1	0
	Mercuric resistance operon regulatory protein	0	1	0	0	0	1
	Gene <i>mntB</i> (Manganese transport system membrane protein MntB)	3	0	3	3	1	0
	Gene <i>modA</i> (Molybdate-binding periplasmic protein precursor)	1	1	2	1	1	1



Subcategory	CDS	Virulência (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Heavy Metal Resistance (3/3)	Gene <i>modB</i> (Molybdenum transport system permease protein ModB)	3	1	1	3	0	1
	Nickel and cobalt resistance protein CnrA	0	0	1	0	0	0
	Gene <i>pcoC</i> (Copper resistance protein C precursor)	0	0	1	1	1	0
	Putative chromate transport protein	0	0	1	0	0	1
	Putative oxidoreductase CzcO	0	1	2	1	1	1
	Gene <i>rbsC</i> (Ribose transport system permease protein RbsC)	9	2	9	9	2	1
	Gene <i>rcnA</i> (nickel/cobalt efflux protein RcnA)	0	0	1	0	0	0
	Sensor kinase CusS	1	0	1	0	0	2
	Transcriptional regulatory protein CusR	1	0	1	0	0	1
<b>Total Heavy Metal Resistance</b>		<b>34</b>	<b>24</b>	<b>42</b>	<b>33</b>	<b>17</b>	<b>32</b>
Drug and Toxic Compounds Resistance (1/6)	Gene <i>acrB</i> (Acriflavine resistance protein B)	1	0	3	4	0	1
	Acriflavin resistance protein	0	3	0	0	0	0
	Gene <i>acrR</i> (Potential <i>acrAB</i> operon repressor)	1	0	2	1	2	4
	Alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17)	0	2	0	0	0	0
	Aminoglycoside N6'-acetyltransferase (EC 2.3.1.82)	0	2	0	0	0	0
	Arsenic efflux pump protein	0	4	0	0	0	0
	Arsenical resistance operon repressor	0	3	0	0	0	0
	Gene <i>bcr</i> (Sulfonamide resistance protein)	3	2	3	2	2	1
	Beta-lactamase (EC 3.5.2.6)	1	1	1	3	0	1
	Beta-lactamase 3	0	1	0	0	0	0
	Beta-lactamase class A	0	1	0	0	0	0
	Beta-lactamase hydrolase-like protein	1	1	0	1	0	0

SUBCATEGORY	CDS	Virulência e Resistência (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (2/6)	Gene <i>hcpA</i> (Beta-lactamase HcpA precursor)	0	0	4	0	0	5
	Beta-lactamase OXA-18	1	0	0	0	0	0
	Beta-lactamase SHV-1	0	0	1	0	0	0
	Blue copper oxidase CueO precursor	1	0	1	1	0	0
	Gene <i>bmr3</i> (Multidrug-efflux transporter 3)	0	6	1	1	3	1
	Cadmium efflux system accessory protein	0	1	0	0	0	0
	Cd(II)/Pb(II)-responsive transcriptional regulator	0	0	0	0	0	1
	Chloramphenicol acetyltransferase	3	0	0	0	0	1
	Chloramphenicol acetyltransferase 3	0	0	1	0	0	0
	Choloylglycine hydrolase (EC 3.5.1.24)	3	1	0	0	0	0
	Chromate resistance protein ChrB	0	0	1	0	0	0
	Chromate transport protein ChrA	1	0	1	0	0	1
	Cobalt-zinc-cadmium resistance protein	2	2	1	0	0	1
	Cobalt-zinc-cadmium resistance protein CzcD	0	1	0	0	0	0
	CopG protein	0	0	0	0	0	1
	Copper chaperone	1	0	0	0	0	0
	Copper homeostasis protein CutE	1	0	1	1	0	1
	Copper homeostasis protein CutF precursor	0	0	1	0	0	0
	Copper resistance protein B	0	0	0	0	0	1
	Copper resistance protein CopC	1	0	1	1	1	0
	Copper resistance protein CopD	1	0	1	1	1	0
	Copper resistance protein D	0	2	0	0	0	0
	Copper sensory histidine kinase CusS	0	0	1	0	0	1
	Copper tolerance protein	0	1	0	0	0	1
	Copper-sensing two-component system response regulator CusR	0	0	1	0	0	1

SUBCATEGORY	CDS	Virulência e Resistência (continuação)					
		CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compound Resistance (3/6)	Copper-translocating P-type ATPase (EC 3.6.3.4)	4	4	3	2	5	3
	Gene <i>cphA</i> (Beta-lactamase precursor)	1	0	0	0	0	1
	Cu(I)-responsive transcriptional regulator	0	0	0	0	1	0
	Cytochrome c heme lyase subunit CcmF	1	0	1	1	0	1
	Cytochrome c heme lyase subunit CcmH	1	0	1	2	0	1
	Cytoplasmic copper homeostasis protein CutC	1	0	1	1	1	0
	D-cysteine desulfhydrase (EC 4.4.1.15)	0	0	1	1	0	0
	Gene <i>drmA</i> (Daunorubicin/doxorubicin resistance ATP-binding protein DrrA)	0	2	1	1	6	1
	Efflux pump membrane transporter BepE	4	0	2	1	0	1
	Efflux pump periplasmic linker BepF	2	0	1	1	0	1
	Gene <i>emrB</i> (Multidrug resistance protein B)	2	6	6	3	0	4
	Gene <i>emrE</i> (Methyl viologen resistance protein C)	1	0	1	0	0	0
	Fosfomycin resistance protein FosA	0	0	1	0	0	0
	Fosfomycin resistance protein FosB	0	1	0	0	0	0
	Fosmidomycin resistance protein	2	1	1	1	0	1
	Glutathione transferase FosA	0	0	1	0	0	1
	Gene <i>macB</i> (Macrolide export ATP-binding/permease protein MacB)	4	0	1	0	1	3
	Macrolide-specific efflux protein MacA	3	0	0	0	0	0
	Magnesium and cobalt efflux protein CorC	1	0	1	1	1	1
	Gene <i>mdtA</i> (Multidrug transporter MdtA)	3	0	3	2	0	4
	Gene <i>mdtB</i> (Multidrug transporter MdtB)	2	0	1	2	0	3
	Gene <i>mdtK</i> (Multidrug-efflux transporter)	0	0	2	1	1	1
	MdtR transcriptional regulator, MarR family	0	3	0	0	0	0
Membrane component of multidrug resistance system	0	4	0	0	0	0	

Virulência e Resistência (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (4/6)	Fosfomycin resistance protein FosA	0	0	1	0	0	0
	Fosfomycin resistance protein FosB	0	1	0	0	0	0
	Fosmidomycin resistance protein	2	1	1	1	0	1
	Glutathione transferase FosA	0	0	1	0	0	1
	Gene <i>macB</i> (Macrolide export ATP-binding/permease protein MacB)	4	0	1	0	1	3
	Macrolide-specific efflux protein MacA	3	0	0	0	0	0
	Magnesium and cobalt efflux protein CorC	1	0	1	1	1	1
	Gene <i>mdtA</i> (Multidrug transporter MdtA)	3	0	3	2	0	4
	Gene <i>mdtB</i> (Multidrug transporter MdtB)	2	0	1	2	0	3
	Gene <i>mdtK</i> (Multidrug-efflux transporter)	0	0	2	1	1	1
	MdtR transcriptional regulator, MarR family	0	3	0	0	0	0
	Membrane component of multidrug resistance system	0	4	0	0	0	0
	Membrane protein, suppressor for copper-sensitivity ScsB	0	0	1	0	0	0
	Membrane protein, suppressor for copper-sensitivity ScsD	0	0	1	0	0	0
	Metal-dependent hydrolases of the beta-lactamase superfamily I	1	2	1	1	0	1
	Metal-dependent hydrolases of the beta-lactamase superfamily III	0	0	1	0	0	0
	Metallothiol transferase FosB	0	2	0	0	0	0
	Metallothiol transferase FosB 2	0	1	0	0	0	0
	Multi antimicrobial extrusion protein (Na <sup>+</sup> )/drug antiporter, MATE family of MDR efflux pumps	1	2	0	0	0	0
	Multicopper oxidase	0	0	0	0	0	3

Virulência e Resistência (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (5/6)	Multidrug and toxin extrusion (MATE) family efflux pump YdhE/NorM, homolog	1	2	0	0	0	0
	Multidrug efflux pump subunit AcrA	0	0	1	1	0	1
	Multidrug efflux transporter MdtP	0	1	0	0	0	0
	Gene <i>mepA</i> (Multidrug export protein MepA)	2	1	1	1	1	0
	Multidrug resistance protein [function not yet clear]	0	3	0	0	0	0
	Multidrug resistance protein D	0	0	1	1	0	0
	Multidrug resistance protein MdtA	3	0	3	2	0	4
	Multidrug resistance protein MdtE	4	0	0	1	0	1
	Multidrug resistance protein MdtH	0	1	3	0	0	1
	Multidrug resistance protein MdtK	0	0	2	1	1	1
	Multidrug resistance protein MdtL	0	0	1	1	0	1
	Multidrug resistance protein MdtN	2	2	2	0	0	1
	Multidrug resistance protein MexA	0	0	1	1	0	1
	Multidrug resistance transporter, Bcr/CfiA family	2	1	2	3	1	2
	Multidrug-efflux transporter, major facilitator superfamily (MFS) (TC 2.A.1)	0	1	0	0	0	0
	Multiple antibiotic resistance protein MarA	0	0	1	0	0	0
	Periplasmic divalent cation tolerance protein CutA	0	0	1	1	0	0
	PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	0	0	0	0	1	0
	Phenazine antibiotic resistance protein EhpR	1	0	0	0	0	0
	Probable transcription regulator protein of MDR efflux pump cluster	1	0	0	0	0	0
	Putative multidrug resistance protein EmrK	0	1	2	1	0	3
	Putative multidrug resistance protein EmrY	0	0	0	1	2	1

Virulência e Resistência (continuação)							
SUBCATEGORY	CDS	CNPMS2088	UFMG50	UFMG51	UFMG54	UFMG61	UFMG81
Drug and Toxic Compounds Resistance (6/6)	Response regulator of zinc sigma-54-dependent two-component system	0	3	1	1	0	0
	Ribosome protection-type tetracycline resistance related proteins, group 2	0	1	0	0	0	0
	Secreted protein, suppressor for copper-sensitivity ScsC	0	0	1	0	0	0
	Sensor protein of zinc sigma-54-dependent two-component system	0	0	1	0	0	0
	Streptogramin A acetyltransferase	1	1	0	0	0	0
	Streptothricin acetyltransferase, Streptomyces lavendulae type	0	1	0	0	0	0
	Suppression of copper sensitivity: putative copper binding protein ScsA	0	0	1	0	0	0
	TetR family regulatory protein of MDR cluster	0	2	0	0	0	0
	Topoisomerase IV subunit A (EC 5.99.1.-)	0	1	0	0	0	0
	Topoisomerase IV subunit B (EC 5.99.1.-)	0	1	0	0	0	0
	Transcriptional regulator NfxB	0	0	0	0	0	2
	Transcriptional regulator, MerR family	4	8	0	0	6	4
	Translation elongation factor G	0	1	0	0	0	0
	Vancomycin B-type resistance protein VanW	0	2	0	0	0	0
	Virginiamycin A acetyltransferase	0	0	0	1	0	1
	Zinc transporter ZitB	0	0	1	1	0	0
<b>Total Drug and Toxic Compound Resistance</b>	<b>77</b>	<b>96</b>	<b>85</b>	<b>55</b>	<b>37</b>	<b>78</b>	