

RALF BRUNO MOURA LOPES

**MONTAGEM E ANÁLISE DO GENOMA PARCIAL DE *Bacillus thuringiensis*
BAC3151 ENDOFÍTICO DAS FOLHAS DO FEIJOEIRO COMUM (*Phaseolus
vulgaris*)**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Microbiologia Agrícola, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2015

Ficha catalográfica preparada pela Biblioteca Central da
Universidade Federal de Viçosa - Campus Viçosa

T

L864m
2015
Lopes, Ralf Bruno Moura, 1990-
Montagem e análise do genoma parcial de *Bacillus thuringiensis* BAC3151 endofítico das folhas do feijoeiro comum (*Phaseolus vulgaris*) / Ralf Bruno Moura Lopes. - Viçosa, MG, 2015.

xii, 67f. : il. (algumas color.) ; 29 cm.

Orientador: Marisa Vieira de Queiroz.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Referências bibliográficas: f. 47-67.

1. Controle biológico. 2. Genoma - Análise. 3. *Bacillus thuringiensis*. 4. *Phaseolus vulgaris*. I. Universidade Federal de Viçosa. Departamento de Microbiologia. Programa de Pós-graduação em Microbiologia Agrícola. II. Título.

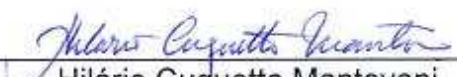
CDD 22. ed. 632.96

RALF BRUNO MOURA LOPES

**MONTAGEM E ANÁLISE DO GENOMA PARCIAL DE *Bacillus thuringiensis*
BAC3151 ENDOFÍTICO DAS FOLHAS DO FEIJOEIRO COMUM (*Phaseolus
vulgaris*)**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Microbiologia Agrícola, para obtenção do título de *Magister Scientiae*.

APROVADA: 25 de fevereiro de 2015.


Hilário Cuquetto Mantovani
(Coorientador)


Maria Cristina Dantas Vanetti
(Coorientadora)


Juliana Lopes Rangel Fietto


Denise Mara Soares Bazzolli
(Presidente da Banca)

Aos meus pais, Ademilson e Rosângela.
Aos meus irmãos, Gustavo e Guilherme.

“Não há vitórias sem lutas,
muito menos conquistas sem batalhas.”
Leonardo Luis Gomes

AGRADECIMENTOS

À Universidade Federal de Viçosa, pela estrutura e pela formação desde a graduação.

Ao Programa de Pós-Graduação em Microbiologia Agrícola, pela formação e pela oportunidade de realização deste trabalho.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

À professora Marisa Vieira de Queiroz, pela orientação, ensinamentos e apoio nos trabalhos de pesquisa.

Ao professor Jeronimo Conceição Ruiz do Centro de Pesquisa René Rachou (CPqRR), pelo auxílio na anotação do genoma.

Aos professores Hilário Cuquetto Mantovani, Maria Cristina Dantas Vanetti, Denise Mara Soares Bazzolli e Juliana Lopes Rangel Fietto, pelos aconselhamentos e sugestões.

Aos demais professores do Departamento de Microbiologia, que contribuíram para a minha formação.

À Lara, Elvira e Grace do CPqRR pelo auxílio na anotação e pelas conversas oportunas.

Aos funcionários do Departamento de Microbiologia, pela ajuda constante.

Aos meus pais, Ademilson e Rosângela, pelo amor, ensinamentos de vida, apoio, dedicação e esforço que me permitiram chegar até aqui.

Ao meu padrinho, José Tarcísio (*in memoriam*), e às minhas tias Lurdes (*in memoriam*) e Lili, pelo amor, incentivo, apoio constantes.

A todos os meus familiares, pelo apoio, amizade e confiança.

Ao amigo Tiago Leite, por tantas conversas oportunas sobre pesquisa e também sobre assuntos diversos.

A todos os amigos do Laboratório de Genética Molecular de Micro-organismos, pela convivência e amizade.

A todas as pessoas que, direta ou indiretamente, colaboraram para a realização deste trabalho.

BIOGRAFIA

Ralf Bruno Moura Lopes, filho de Ademilson Geraldo Lopes e Rosângela de Fátima Moura Lopes, nasceu em Sete Lagoas, Minas Gerais, em 3 de junho de 1990. Em 2008, ingressou na Universidade Federal de Viçosa, graduando-se em Ciências Biológicas em março de 2013. Em abril do mesmo ano, iniciou o curso de Mestrado do Programa de Pós-Graduação em Microbiologia Agrícola na mesma instituição.

SUMÁRIO

LISTA DE FIGURAS	viii
LISTA DE TABELAS	x
RESUMO.....	xi
ABSTRACT	xii
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	3
2.1. Bactérias endofíticas	3
2.2. <i>Bacillus thuringiensis</i>	5
2.2.1. Características gerais.....	5
2.2.2. Distribuição no ambiente	6
2.2.3. Utilização para biocontrole na agricultura.....	7
2.3. Breve histórico da ciência genômica	9
2.4. Montagem de genomas.....	12
2.5. Anotação genômica.....	14
2.5.1. Anotação automática.....	14
2.5.2. Curadoria manual	15
2.6. Projetos genômicos de <i>B. thuringiensis</i>	16
2.7. Mineração de genomas e bioprospecção de produtos de <i>Bacillus</i>	17
3. MATERIAL E MÉTODOS	23
3.1. Isolado bacteriano e condições de crescimento	23
3.2. Extração do DNA e sequenciamento do genoma.....	23
3.3. Tratamento dos dados e montagem.....	23
3.4. Anotação do genoma	23
3.5. Análise filogenômica	24
3.6. Ordenação e orientação dos scaffolds	24
3.7. Rearranjos cromossomais em <i>B. thuringiensis</i>	24
3.8. Compartilhamento de genes entre linhagens	25

4. RESULTADOS E DISCUSSÃO.....	26
4.1. Características gerais do <i>draft</i> do genoma de <i>B. thuringiensis</i> BAC3151	26
4.2. Análise filogenômica	30
4.3. Ordenação e orientação dos scaffolds	31
4.4. Alta similaridade na arquitetura do genoma	35
4.5. Compartilhamento de genes entre linhagens	37
4.6. Genes putativos de <i>B. thuringiensis</i> BAC3151 associados a antagonismo ...	40
5. CONCLUSÕES	46
6. REFERÊNCIAS	47

LISTA DE FIGURAS

Figura 1. Representação esquemática de diferentes tipos de interações entre bactérias endofíticas e plantas, levando à síntese de metabólitos. Os metabólitos produzidos por bactérias endofíticas podem exibir importantes funções para a planta (BRADER et al., 2014).....5

Figura 2. Distribuição filogenética dos projetos genômicos de (A) *Bacteria* e (B) *Archaea* registrados no *Genomes Online Database* (GOLD) (<https://gold.jgi-psf.org/distribution>) até dezembro de 2014 10

Figura 3. Visão geral do *draft* da sequência do genoma de *B. thuringiensis* BAC3151. A) Exibição circular do *draft* do genoma anotado. Os círculos representam, a partir do mais externo: círculo 1, posição das bases do DNA; círculo 2, regiões codificadoras de proteínas transcritas no sentido horário; círculo 3, regiões codificadoras de proteínas transcritas no sentido anti-horário; círculo 4, regiões codificadoras de tRNA (amarelo) e rRNA (vermelho); círculo 5, conteúdo G+C. B) Distribuição do comprimento dos scaffolds. C) Histograma do conteúdo G+C. D) Distribuição do tamanho dos *gaps*.....27

Figura 4. Análise filogenômica de *B. thuringiensis* baseada em 1276 genes *core* das linhagens usadas. A árvore foi construída com PHYLIP (FELSENSTEIN, 1989), utilizando o método de Neighbor-Joining com 1000 replicatas. Os valores de *bootstrap* são mostrados nos nós. A barra da escala mostra cinco substituições de nucleotídeos por 1000 nucleotídeos. O número de acesso das sequências utilizadas é dado entre parênteses.....31

Figura 5. Alinhamento da sequência cromossômica de *B. thuringiensis* MC28 com o *draft* do genoma de *B. thuringiensis* BAC3151 apresentando 36 LCBs após a ordenação e orientação dos scaffolds do *draft*. LCBs homólogos são conectados por linhas, mas somente os maiores LCBs estão representados. As regiões ampliadas e a seta mostram os segmentos invertidos de BAC3151, que são discutidos no texto. As linhas vermelhas verticais da sequência de BAC3151 indicam os limites entre os scaffolds. Embora a visão geral das regiões compartilhadas entre as duas sequências seja clara, alguns rearranjos podem ser ofuscados por um excesso de ordenamento.....32

Figura 6. Alinhamento de cromossomos de diferentes linhagens de *B. thuringiensis* e o *draft* do genoma de BAC3151. Segmentos de mesma cor representam regiões homólogas, sem rearranjos internos, denominadas de blocos localmente colineares (LCBs). Regiões invertidas são claramente apresentadas como blocos abaixo da linha horizontal. Essas regiões são indicadas pelos números 1, 2 e 3, e são discutidas no texto. A escala mostra as coordenadas de cada sequência36

Figura 7. Desenvolvimento do pangenoma, genoma *core* e CDSs únicas de *B. thuringiensis* com o aumento do número de genomas. Modelos de ajuste de quadrados mínimos não-lineares predizem o desenvolvimento do pangenoma (A), *core* genômico (B) e CDSs únicas (C) dos genomas descritos na tabela 3 com

base nas funções $5422,261x^{0,399}$, $1935,165e^{(-x/3,664)}$ + 3367,738 e $1793,351e^{(-x/2,355)}$ + 339,683, respectivamente38

Figura 8. Diagramas de Venn construídos com EDGAR (BLOM et al., 2009) apresentando o compartilhamento de CDSs de BAC3151 com outros genomas. A) 1. BMB171, 2. MC28, 3. *sv. thuringiensis* IS5056, 4. Al Hakam, 5. BAC3151; B) 1. YBT-1518, 2. *sv. chinensis* CT-43, 3. YBT-020, 4. *sv. kurstaki* HD73, 5. BAC3151; C) 1. Bt407, 2. HD-789, 3. *sv. konkukian* 97-27, 4. BAC3151. Regiões sobrepostas representam CDSs comuns entre os genomas analisados. Os valores fora das regiões sobrepostas indicam o número de CDSs em cada genoma sem ortólogas nos outros genomas.....40

LISTA DE TABELAS

Tabela 1. Ferramentas computacionais amplamente utilizadas na anotação de genomas bacterianos	15
Tabela 2. Resumo do processamento dos dados obtidos do sequenciamento e montagem	26
Tabela 3. Propriedades gerais de genomas de <i>Bacillus thuringiensis</i> utilizados neste estudo.....	28
Tabela 4. Ordem e orientação final de todos os scaffolds do <i>draft</i> genômico de <i>B. thuringiensis</i> BAC3151	34
Tabela 5. Classes de genes putativos chave potencialmente envolvidos em atividade antagonista de <i>B. thuringiensis</i> BAC3151	42

RESUMO

LOPES, Ralf Bruno Moura, M.Sc., Universidade Federal de Viçosa, fevereiro de 2015. **Montagem e análise do genoma parcial de *Bacillus thuringiensis* BAC3151 endofítico das folhas do feijoeiro comum (*Phaseolus vulgaris*).** Orientadora: Marisa Vieira de Queiroz. Coorientadores: Hilário Cuquetto Mantovani, Maria Cristina Dantas Vanetti e Denise Mara Soares Bazzolli.

Bacillus thuringiensis é uma espécie comumente associada a solos, insetos e plantas. Diversas linhagens são importantes para diferentes formas de controle biológico na agricultura e em outras áreas. Neste trabalho, foi sequenciado e anotado o genoma de *B. thuringiensis* BAC3151 endofítico do feijoeiro comum (*Phaseolus vulgaris*), um isolado que tem demonstrado atividade antimicrobiana contra bactérias patogênicas, incluindo bactérias fitopatogênicas do feijoeiro, possuindo, portanto, potencial para biocontrole de doenças nessa leguminosa. O genoma de BAC3151 e o de outros *B. thuringiensis* publicados forneceu a oportunidade para realizarmos análises de genômica comparativa entre diferentes linhagens dessa espécie dinâmica. A análise filogenômica utilizando os genes *core* das linhagens avaliadas agrupou BAC3151 com *B. thuringiensis* MC28 isolado de solo. Também foram identificados rearranjos cromossômicos de grandes segmentos em BAC3151 e em outras linhagens, sugerindo que rearranjos podem contribuir para diferenças funcionais dessas linhagens. O genoma *core* das linhagens estudadas consistiu em 3413 sequências codificadoras de proteínas (CDSs) e incluiu as CDSs com funções fisiológicas gerais e de sobrevivência da espécie, enquanto as CDSs do genoma acessório foi associada a plasmídeos e a regiões cromossômicas específicas. Além disso, foram encontrados múltiplos genes potencialmente envolvidos na ação antagonista de BAC3151 contra outros micro-organismos e insetos, incluindo genes inseticidas, policetídeos sintases, bacteriocinas, peptídeos não-ribossômicos sintetases, sideróforos, quitinases, N-acil homoserina lactonase e outros. Este estudo mostrou diversas variações entre linhagens de *B. thuringiensis* em nível genômico e também que *B. thuringiensis* BAC3151 endofítico do feijoeiro é potencialmente útil para o desenvolvimento de estratégias para o controle biológico de doenças nessa importante leguminosa.

ABSTRACT

LOPES, Ralf Bruno Moura, M.Sc., Universidade Federal de Viçosa, February, 2015. **Assembly and analysis of partial genome of *Bacillus thuringiensis* BAC3151 from the leaves of the common bean (*Phaseolus vulgaris*)**. Adviser: Marisa Vieira de Queiroz. Co-advisers: Hilário Cuquetto Mantovani, Maria Cristina Dantas Vanetti and Denise Mara Soares Bazzoli.

Bacillus thuringiensis is a species commonly associated with soil, insects and plants. Several strains are important for different types of biological control in agriculture and other areas. In this work, was sequenced and annotated the genome of endophytic *B. thuringiensis* BAC3151 of common bean (*Phaseolus vulgaris*), an isolate that has demonstrated antimicrobial activity against pathogenic bacteria, including plant pathogenic bacteria of the common bean, having, in this way, potential for biocontrol of diseases in this legume. The genome of BAC3151 and the other *B. thuringiensis* published provided the opportunity to accomplish comparative genomic analysis between different strains of this dynamic species. The phylogenomic analysis using the core genes of the evaluated strains grouped BAC3151 with *B. thuringiensis* MC28 isolated from soil. Also, large chromosomal rearrangements were identified in BAC3151 and other strains, suggesting that rearrangements may contribute to functional differences of these strains. The core genome of the strains studied consisted of 3413 sequences encoding proteins (CDSs) and included the CDSs with general physiological functions and species survival, while the accessory genome CDSs were associated with plasmids and specific chromosomal regions. In addition, were found multiple potential genes involved in the antagonistic action of BAC3151 against other microorganisms and insects, including insecticidal genes, polyketide synthases, bacteriocins, nonribosomal peptide synthetases, siderophores, chitinases, N-acyl homoserine lactonase and others. This study showed several variations among strains of *B. thuringiensis* in genomic level and also that endophytic *B. thuringiensis* BAC3151 of common bean is potentially useful for the development of strategies for biological control in this important legume.

1. INTRODUÇÃO

Bacillus thuringiensis é uma bactéria Gram-positiva formadora de endósporos. A espécie é amplamente distribuída, sendo encontrada em uma variedade de ambientes, como solo, inseto, superfície de folhas e em tecidos vegetais internos. A maioria dos estudos com *B. thuringiensis* tem focado nos efeitos diretos contra insetos praga relevantes para a agricultura, embora diversas novas funções na proteção de plantas contra fitopatógenos vêm sendo descobertas. Por exemplo, a síntese de zwittermicina A, um antibiótico eficiente principalmente contra oomicetos e fungos fitopatogênicos relacionados. Outro exemplo é a produção de N-acil homoserina lactonase, que hidrolisa N-acil homoserina lactona (AHL), uma molécula sinal do sistema de *quorum sensing* bacteriano, suprimindo a virulência de patógenos.

Bactérias endofíticas colonizam o interior da planta hospedeira sem causar danos aparentes. A colonização endofítica bem sucedida por *B. thuringiensis* tem sido relatada em plantas como algodão, soja, feijoeiro, milho, cana-de-açúcar, repolho e lentilha. Costa et al. (2012) reportaram que 32,9% dos isolados de bactérias endofíticas cultiváveis das folhas de diferentes cultivares do feijoeiro comum (*Phaseolus vulgaris*) pertenceram ao filo Firmicutes, incluindo diferentes isolados de *B. thuringiensis*. No entanto, muito ainda permanece desconhecido sobre a associação de *B. thuringiensis* com leguminosas. Estudos sobre essa associação têm mostrado que *B. thuringiensis* inoculado na rizosfera ou em sementes não somente coloniza o interior de leguminosas, mas também aumenta a nodulação das raízes e o crescimento vegetal.

Por outro lado, as doenças microbianas do feijoeiro comum constituem uma das principais causas da redução da produtividade dessa importante cultura. O controle de doenças causadas por fitopatógenos é geralmente feito pela combinação de manejo da cultura, defensivos agrícolas e cultivares resistentes. A aplicação de defensivos agrícolas é o principal método de controle, porém seu uso é controverso em virtude dos efeitos prejudiciais ao ambiente e da seleção de patógenos resistentes. Há também doenças microbianas para as quais não existe ainda nenhum produto de eficiência comprovada. Assim, a utilização microorganismos capazes de proteger o vegetal contra fitopatógenos tem emergido

como uma alternativa mais racional e segura para o manejo de culturas. Neste contexto, bactérias endofíticas podem ser úteis para controle biológico de doenças, uma vez que, muitas vezes, ocupam o mesmo hábitat que micro-organismos fitopatogênicos e não causam danos aparentes à planta hospedeira.

B. thuringiensis BAC3151 endofítico de *P. vulgaris* tem demonstrado atividade antimicrobiana *in vitro* contra bactérias fitopatogênicas do feijoeiro bem como potencial para inibição de *quorum sensing*, referida como *quorum quenching*, de bactérias Gram-negativas. De fato, estudos genômicos têm revelado o gênero *Bacillus* como uma surpreendente fonte potencial de antimicrobianos, como policetídeos, peptídeos não-ribossomais bem como antimicrobianos incomuns e outros metabólitos secundários. Esses estudos também permitem pesquisar os eventos mudam a arquitetura do genoma, como a perda ou a duplicação gênica, rearranjos e aquisição de genes por transferência horizontal. No presente trabalho, o genoma de BAC3151 foi sequenciado para melhor investigar os determinantes genéticos potencialmente envolvidos na síntese de antimicrobianos que poderiam ser úteis para o biocontrole de fitopatógenos, além de analisar diferenças genômicas do isolado e outras linhagens.

2. REVISÃO DE LITERATURA

2.1. Bactérias endofíticas

Bactérias endofíticas são conhecidas há mais de 120 anos (HARDOIM et al., 2008) e correspondem a micro-organismos, cultiváveis ou não, que colonizam o interior de uma planta hospedeira e não causam danos aparentes a ela (HALLMANN et al., 1997). Embora bactérias fixadoras de nitrogênio nos nódulos do feijoeiro comum (*Phaseolus vulgaris*) têm sido bem estudadas e caracterizadas (MARTÍNEZ-ROMERO, 2003; MARTÍNEZ-ROMERO, 2009) e a comunidade bacteriana endofítica das sementes e raízes dessa leguminosa ter sido reportada (LÓPEZ-LÓPEZ, et al., 2010), pouco ainda é conhecido sobre as bactérias endofíticas dos tecidos aéreos de *P. vulgaris*. Costa et al. (2012) descreveram as bactérias endofíticas cultiváveis das folhas das cultivares Vermelhinho, Talismã e Ouro Negro de *P. vulgaris* e encontraram que 36,7% dos isolados perteceram a Proteobacteria, 32,9% a Firmicutes, 29,7% a Actinobacteria e 0,6% a Bacteroidetes. Esses resultados mostraram ainda diferença na estrutura da comunidade bacteriana endofítica de acordo com a cultivar do feijoeiro.

Em geral, a entrada de micro-organismos no tecido vegetal pode ocorrer pelos estômatos, lenticelas, lesões e superfícies emergentes de raízes laterais e de radículas germinativas (HUANG, 1986). No entanto, as principais portas de entrada para bactérias endofíticas parecem ser as lesões que ocorrem naturalmente como resultado do crescimento da planta, além dos pelos radiculares e das junções epidermais (SPRENT e FARIA, 1988). Algumas bactérias também possuem penetração ativa em tecidos do hospedeiro, produzindo enzimas celulolíticas e pectinolíticas (HUREK et al., 1994).

Uma vez dentro do tecido vegetal, elas podem permanecer em um local específico, como o córtex da raiz, ou colonizar a planta de forma sistêmica por meio do transporte através dos elementos condutores ou do apoplasto. A capacidade de dispersão sistêmica foi demonstrada por endofíticos como *Erwinia* sp. em algodão (MISAGHI e DONNDELINGER, 1990), *Pseudomonas aureofaciens* em milho (LAMB et al., 1996) e *Bacillus thuringiensis* em leguminosas (TANUJA et al., 2013). A colonização de tecidos específicos por bactérias parece ser espécie específica. Geralmente, bactérias endofíticas

colonizam o espaço intercelular (QUADT-HALLMANN e KLOEPPER, 1996), sendo poucos os relatos demonstrando a colonização intracelular (FROMMEL et al., 1991; HUREK et al., 1994; QUADT-HALLMANN e KLOEPPER, 1996). A colonização do sistema vascular também foi reportada (GARDNER et al., 1982; LAMB et al., 1996).

Os padrões de colonização de bactérias endofíticas em tecidos de plantas são fortemente dependentes de fatores bióticos e abióticos que interagem entre si (HALLMANN et al., 1997). Um dos fatores bióticos mais importantes é a presença de outros micro-organismos associados ao vegetal, como fungos, vírus e outras bactérias (HALLMANN et al., 1998). Como o espaço e os nutrientes fornecidos pelo hospedeiro são fatores limitantes, existem vários tipos de interação entre os micro-organismos que possuem o mesmo hábitat, incluindo antagonismo, simbiose e mutualismo. Os tecidos internos das plantas fornecem um ambiente mais uniforme e protetor para os micro-organismos do que a superfície vegetal, onde exposições a condições ambientais extremas, como temperatura, potencial osmótico e radiação ultravioleta são fatores limitantes para a sobrevivência (HALLMANN et al., 1997). Um exemplo da influência de fatores abióticos na população de bactérias endofíticas é o que ocorre em sítios contaminados com metais pesados, petróleo, solventes, hidrocarbonetos aromáticos policíclicos ou outros contaminantes orgânicos, onde vegetais crescendo nesses solos são capazes de recrutar micro-organismos que podem degradar os poluentes. Características do solo, como pH, salinidade e textura também afetam as bactérias endofíticas de forma indireta, visto que alteram a comunidade bacteriana da rizosfera, fonte de bactérias endofíticas em potencial (DOTY, 2008; RYAN et al., 2008).

Em relação à estratégia de vida, as bactérias endofíticas podem ser classificadas como endofíticas obrigatórias ou facultativas. As bactérias endofíticas obrigatórias são estritamente dependentes da planta hospedeira para o seu crescimento e sobrevivência. A transmissão dessas bactérias ocorre verticalmente ou por meio da veiculação de um vetor. Já os micro-organismos endofíticos facultativos podem ser caracterizados por alternar entre o interior dos tecidos vegetais e o ambiente. A diversidade microbiana endofítica encontrada nas plantas pode ser explicada principalmente pela capacidade de diversos endofíticos

facultativos entrarem e persistirem no interior da planta (ROSENBLUETH e MARTINEZ-ROMERO, 2006).

Além disso, as bactérias endofíticas podem propiciar muitas vantagens para os vegetais, incluindo a promoção do crescimento (BARKA et al., 2002; TANUJA et al., 2013), indução dos mecanismos de defesa sistêmicos (MISHRA et al., 2006; BAKKER et al., 2007), redução de sintomas de doenças causadas por vários patógenos (COOMBS et al., 2004; KLOPPER et al., 2004;) controle de insetos e nematoides (D'ALESSANDRO et al., 2014), produção de compostos anti-herbivoria (SCOTT, 2001; SULLIVAN et al., 2007), fixação biológica de nitrogênio (MARTÍNEZ et al., 2003; JHA e KUMAR, 2007) e o aumento da tolerância aos estresses ambientais (ANDREOLLI et al., 2013). A figura 1 apresenta alguns dos principais tipos de interações entre bactérias endofíticas e plantas, envolvendo a síntese de metabólitos com diferentes papéis.

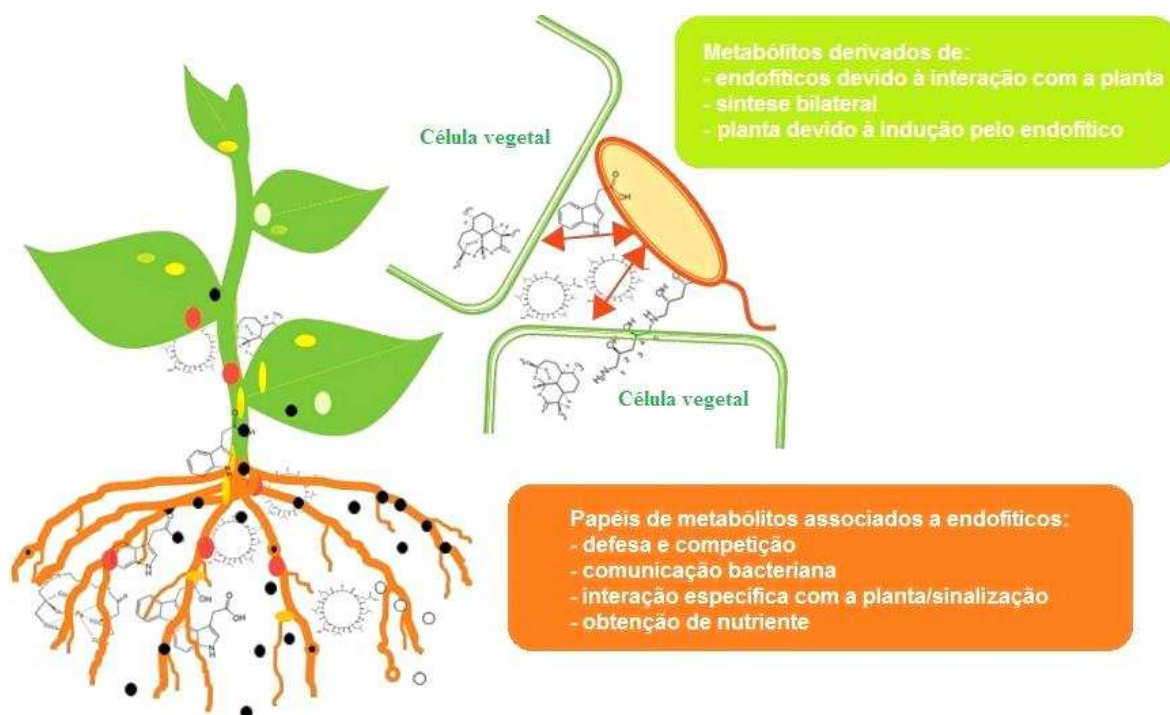


Figura 1. Representação esquemática de diferentes tipos de interações entre bactérias endofíticas e plantas, levando à síntese de metabólitos. Os metabólitos produzidos por bactérias endofíticas podem exibir importantes funções para a planta (BRADER et al., 2014).

2.2. *Bacillus thuringiensis*

2.2.1. Características gerais

Bacillus thuringiensis foi primeiramente isolado de larvas mortas do bicho-da-seda (*Bombyx mori*) pelo bacteriologista Shigetane Ishiwata no Japão em

1901, enquanto ele estudava a doença de soto nesse inseto, e a espécie foi inicialmente nomeada como *Bacillus sotto* (ISHIWATA, 1901). No entanto, em 1911, o microbiologista Ernst Berliner isolou uma linhagem relacionada em larvas da traça da farinha (*Anagasta kuehniella*) no estado alemão da Turíngia e, posteriormente, a nomeou como *B. thuringiensis*, denominação que foi mantida (BERLINER, 1911; BERLINER, 1915).

B. thuringiensis é uma bactéria Gram-positiva, anaeróbia facultativa, mesófila e quimioheterotrófica, crescendo principalmente na faixa de 10 a 40 °C. Possui forma de um bastonete com dimensões de cerca de 1,1 µm de largura e 4,0 µm de comprimento, sendo capaz de se movimentar devido à presença de flagelos peritríquios (dependendo do sorovar) (STAHLY et al., 1991). Quando em condições ambientais desfavoráveis, a divisão celular é interrompida e inicia-se o ciclo de esporulação típico dos bacilos, gerando esporos elipsoidais localizados na região central ou paracentral da célula. Simultaneamente ao processo de esporulação, ocorre produção de inclusões cristalinas, características de cada variedade *B. thuringiensis* e tóxicas para insetos (FEITELSON et al., 1994; IBRAHIM et al., 2010). Essas inclusões cristalinas, também chamadas de proteínas Cry ou δ-endotoxina, são liberadas com o esporo no momento da lise celular (HANNAY e FITZ-JAMES, 1995). Além disso, elas constituem uma característica importante para diferenciar *B. thuringiensis* de espécies similares (FEDERICI, 1999).

2.2.2. Distribuição no ambiente

A espécie *B. thuringiensis* é amplamente distribuída, podendo ser encontrada em quase todas as partes do mundo e associada a vários ambientes, como solo, água, grãos armazenados (BRAVO et al., 1998), insetos mortos e vivos (SUZUKI et al., 2004) e superfície e interior de plantas (BRAVO et al., 1998; MONNERAT et al., 2003; COSTA et al., 2012).

No entanto, não há certeza sobre o hábitat primário de *B. thuringiensis* (MEADOWS, 1993). Uma possibilidade é que a bactéria seja originalmente um entomopatógeno, uma vez que seria improvável a utilização de grande quantidade de energia durante a esporulação para a formação de cristal que não apresentasse um grande valor adaptativo. Como o esporo e o cristal liberados pela

célula tendem a permanecer juntos no ambiente, após a ingestão por uma larva suscetível e a subsequente morte do inseto, pode então ocorrer a germinação do esporo (THOMAS et al., 2000; VILAS-BÔAS et al., 2000 e SUZUKI et al., 2004).

Uma segunda hipótese considera que *B. thuringiensis* seja natural do solo. A germinação de esporos em solo esterilizado é descrita em vários estudos e depende da umidade, disponibilidade de nutrientes, pH, além do próprio tipo de solo (AKIBA, 1986). No entanto, trabalhos posteriores têm mostrado que os esporos de *B. thuringiensis* podem muitas vezes não germinar em solo não esterilizado (VILAS-BÔAS et al., 2000) e que, nesse tipo de solo, a morte de células vegetativas inoculadas é significativamente maior que em solo estéril (THOMAS et al., 2000; FERREIRA et al., 2003). Segundo Arantes et al. (2002), essa menor viabilidade pode ser devida a uma baixa capacidade de *B. thuringiensis* se multiplicar nesse ambiente, possivelmente por não ser um membro original das comunidades microbianas do solo, permanecendo principalmente na forma de esporos.

Outra possibilidade menos discutida, mas apresentada por Smith e Couche (1991), é a espécie ser natural do filoplano de árvores. Os autores isolaram *B. thuringiensis* de diversas espécies arbóreas e em quantidades bastante altas para serem decorrentes somente da disseminação pelo vento ou pela chuva. Ainda segundo os autores, o solo seria apenas um reservatório, enquanto o hábitat primário poderia ser o filoplano de árvores.

2.2.3. Utilização para biocontrole na agricultura

B. thuringiensis é uma espécie que produz uma variedade de compostos liberados para o meio extracelular que são potencialmente úteis para diferentes tipos de biocontrole na agricultura. Esses compostos incluem antibióticos (FICKERS, 2012), quitinases (GOMAA, 2012) e outros antifúngicos (SADFI et al., 2001), bacteriocinas (CHERIF et al., 2003), lactonases (DONG et al., 2004), sideróforos (BODE, 2009) e a característica δ -endotoxina (proteínas Cry), que constitui o produto mais estudado e mais empregado de *B. thuringiensis* no mundo (POLANCZYK e ALVES, 2003).

O primeiro estudo empregando *B. thuringiensis* para biocontrole em campo foi conduzido por Husz (1929), como parte de um programa internacional de

controle da lagarta do milho europeia (*Ostrinia nubilalis*). Embora nessa época não fosse conhecida ainda a causa da atividade entomopatogênica da bactéria, os resultados obtidos foram promissores e a partir de então foi crescente o interesse na espécie. A primeira formulação à base de *B. thuringiensis*, a *Sporeine*, foi produzida na França em 1938. Na década de 1950, iniciou-se a produção em larga escala, com o lançamento do bioinseticida comercial *Thuricide*, seguido por produtos similares em empresas localizadas na Rússia, Alemanha, Estados Unidos e antiga Tchecoslováquia (DE MAAGD et al., 1999). Inicialmente, os produtos eram utilizados somente para o controle de larvas de lepidópteros. Contudo, principalmente a partir da década de 1970, novas subespécies da bactéria se mostraram eficientes contra insetos das ordens Diptera, Coleoptera (WEISER, 1986; EDWARDS et al., 1988), Hymenoptera, Hemiptera e Orthoptera (FEILTEISON, 1994). Também foi relatada toxicidade para espécies de nematoides, protozoários, ácaros e moluscos (FEILTEISON, 1994; WANG et al., 2013; LIU et al., 2014).

A eficácia e a especificidade das linhagens de *B. thuringiensis* e suas toxinas no controle de insetos favorecem a formulação de biopesticidas à base dessa bactéria e mais de cem formulações foram colocadas no mercado mundial desde o primeiro produto lançado, sendo responsáveis por mais de 90% do faturamento total com bioinseticidas (POLANCZYK e ALVES, 2003). A aplicação direta dos cristais e esporos de *B. thuringiensis* representa a principal forma de utilização desses biopesticidas, que são inócuos para o meio ambiente, insetos benéficos e outros animais, inclusive para os seres humanos (CÁRDENAS et al., 2001; BRAR et al., 2006).

Na América do Norte, produtos à base de *B. thuringiensis* são muito utilizados para o controle de pragas florestais, principalmente *Lymantria dispar*, *Choristoneura fumiferana* e *Choristoneura occidentalis*, além dos insetos prejudiciais às culturas agrícolas. Na Austrália, esses bioinseticidas são empregados contra pragas de culturas como algodão, plantas frutíferas, plantas ornamentais e fumo (GLARE e O'CALLAGHAM, 2000; VAN FRANKENHUYZEN, 2000). Eles também são utilizados contra insetos danosos de grandes culturas, florestas e hortaliças em, aproximadamente, um milhão de hectares que abrangem cerca de trinta províncias do centro-sul da China. No Egito, *B. thuringiensis* é

utilizado principalmente no controle de pragas do algodão (SALAMA e MORRIS, 1993). Cuba e México lideram a utilização dos bioinseticidas à base desse micro-organismo na América Latina, especialmente para o manejo de insetos nas culturas de algodão, banana, batata, citros, hortaliças, fumo, milho e pastagens. No Brasil, cerca de trinta pragas de importância agrícola e mosquitos vetores de doenças são controladas *B. thuringiensis* (POLANCZYK e ALVES, 2003).

Apesar da ampla utilização, o custo de bioinseticidas à base de *B. thuringiensis*, na maioria das vezes, é superior ao de inseticidas químicos, os cristais não podem penetrar nas plantas e a persistência da maioria das formulações em campo é baixa (GELERNTER e SCHWAB, 1993; VAN FRANKENHUYSEN, 1993; NAVON 2000). Nesse sentido, pesquisas para minimizar essas limitações estão sendo desenvolvidas, principalmente por meio da prospecção de novos genes *cry*, da produção de plantas expressando esses genes, do desenvolvimento de novas formulações e também da ampliação do conhecimento das interações de *B. thuringiensis* com outros entomopatógenos e com inseticidas químicos (NAVON, 2000).

2.3. Breve histórico da ciência genômica

Com o advento da tecnologia de sequenciamento de DNA baseada em didesoxinucleotídeos desenvolvida por Sanger et al. (1977b), foi viabilizado o progresso dos primeiros projetos de sequenciamento de genomas. Essa metodologia possibilitou que, no mesmo ano, fosse realizado o sequenciamento completo do DNA do bacteriófago Φ X174 (SANGER et al., 1977a). Apesar desse avanço inicial, foi somente após anos de desenvolvimento tecnológico utilizando a metodologia que os primeiros genomas de organismos celulares, as bactérias *Haemophilus influenzae* e *Mycoplasma genitalium*, foram completamente sequenciados (FLEISCHMANN et al., 1995; FRASER et al., 1995). A partir de então, a ciência genômica começou a se desenvolver conjuntamente com a bioinformática, modificando completamente o cenário biológico e caracterizando os mais diversos organismos presentes na árvore da vida.

Principalmente em razão do menor tamanho do genoma, os domínios *Bacteria* e *Archaea* são os que apresentam o maior número de organismos com genoma finalizado e projetos genoma em andamento. Existem, no total, 39183

projetos genoma de bactérias atualmente, sendo que Proteobacteria, Firmicutes e Actinobacteria representam 43,8%, 30,8% e 14,5%, respectivamente, desses projetos (<https://gold.jgi-psf.org/distribution>) (dados até de dezembro de 2014) (Figura 2). A predominância desses filos é explicada principalmente pela prioridade de sequenciamento do genoma de organismos já previamente estudados e caracterizados por grupos de pesquisa ao redor do mundo (KYRPIDES, 2009).

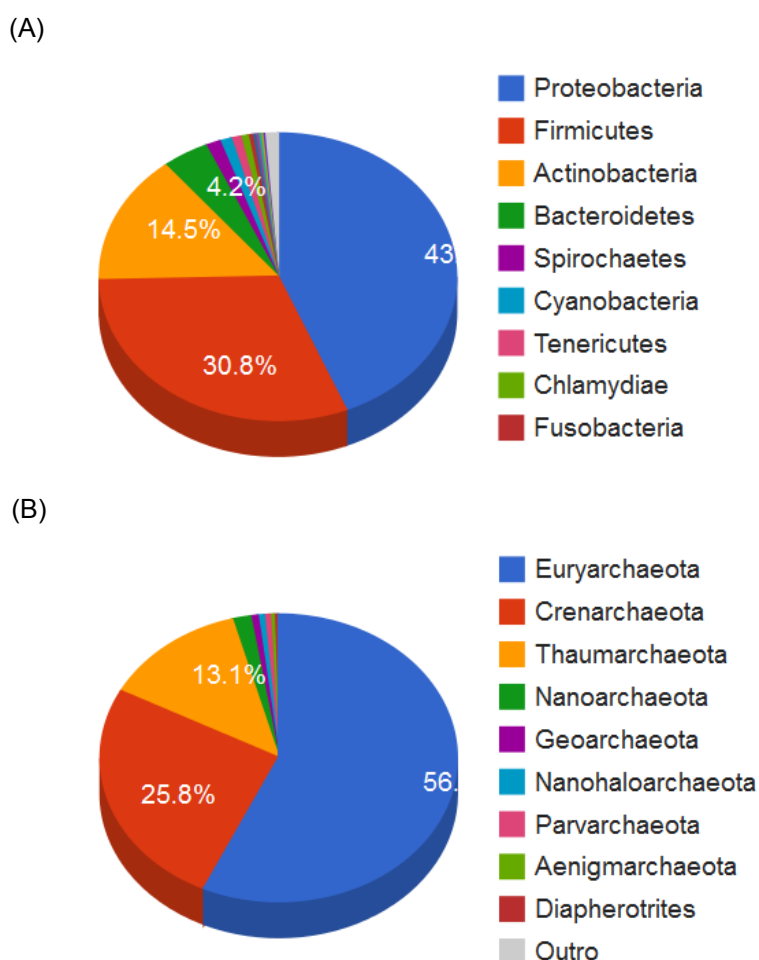


Figura 2. Distribuição filogenética dos projetos genômicos de (A) *Bacteria* e (B) *Archaea* registrados no *Genomes Online Database* (GOLD) (<https://gold.jgi-psf.org/distribution>) até dezembro de 2014.

Desde a publicação do *draft* do genoma humano (VENTER et al., 2001), a capacidade de geração de dados genômicos com menores custos, maior rapidez e maior eficiência tem aumentado ainda mais, especialmente em virtude do desenvolvimento das novas tecnologias de sequenciamento de DNA, os chamados sequenciadores de nova geração. Existem diversas tecnologias

disponíveis no mercado voltadas para o sequenciamento de DNA em larga escala. Essas tecnologias diferem principalmente quanto ao mecanismo empregado no processo de sequenciamento, ao volume de dados gerados, à precisão de bases, à tecnologia de operação, ao tamanho dos fragmentos gerados e ao custo operacional (MACLEAN et al., 2009).

Nos primeiros projetos de sequenciamento genômico com a tecnologia de Sanger, os fragmentos sequenciados eram em número pequeno e com tamanho de até 1.000 pb, cujo processamento era realizado em períodos de semanas ou meses, em extenuante trabalho de laboratório. Contudo, no início deste milênio, a automatização das tecnologias de sequenciamento foi aperfeiçoada até um ponto em que equipamentos totalmente automatizados passaram a gerar sequências de modo paralelizado durante 24 horas por dia e grandes centros de sequenciamento de genomas ao redor do mundo abrigaram dezenas dessas máquinas de sequenciamento. Isso, por sua vez, levou à necessidade de criação de novos algoritmos montadores de genoma, utilizando sequências de, geralmente, 35 a 1000 bases de comprimento, com taxas de erro de sequenciamento variando de 0,5 e 15%, e que ainda podem conter artefatos complexos como repetições (MARDIS, 2008).

O crescente avanço na tecnologia também tem levado ao surgimento de diversos consórcios e projetos de grande escala ao redor do mundo. O Projeto Microbioma Humano foi iniciado com o objetivo de caracterizar toda a comunidade microbiana associada aos diferentes tecidos do organismo humano e os resultados previstos poderão contribuir para um melhor entendimento sobre a fisiologia humana e a predisposição a doenças (TURNBAUGH et al., 2007). Já o projeto *TerraGenome* surgiu com o intuito de sequenciar o metagenoma completo de diferentes amostras de solos ao redor do mundo (<http://www.terragenome.org/>). Por fim, o maior e mais ambicioso projeto de sequenciamento genômico atual, o *Earth Microbiome Project*, iniciado em 2010, tem como objetivo a caracterização de 200.000 amostras provenientes de diferentes ambientes do planeta. Por meio de enfoques metagenômicos, metatranscriptômicos e sequenciamento de amplicons, os pesquisadores prevêem a reconstrução de, aproximadamente, 500.000 genomas microbianos, bem como a construção de um atlas de genes que conterá informações e modelos metabólicos específicos para cada bioma

amostrado (<http://www.earthmicrobiome.org/>).

2.4. Montagem de genomas

A montagem de um genoma consiste em um conjunto de procedimentos em que se busca organizar um grande número de sequências curtas de DNA em um espaço linear com o objetivo de representar a(s) molécula(s) completa(s) de DNA que compõe(m) o genoma da espécie estudada (DOHM et al., 2007).

A enorme quantidade de dados de sequência de DNA gerados pelas tecnologias de sequenciamento de nova geração (NGS) juntamente com os artefatos e erros inerentes a cada tecnologia de sequenciamento, desafiam os projetos de montagem de genomas completos de diferentes espécies. Há menos de uma década, apenas o montador *Newbler* era disponível para esse fim, aplicado à montagem de fragmentos gerados por sequências produzidas pelo sequenciador Roche 454 (*software* proprietário). Apresentada em meados de 2007, a versão híbrida do montador MIRA (CHEVREUX et al., 2004) foi o primeiro montador de uso livre desenvolvido para montar segmentos provenientes do sequenciador 454 e misturas de segmentos do sequenciador 454 e do método de Sanger. No final de 2007, o montador SHARCGS (DOHM et al., 2007) foi desenvolvido para montagem de segmentos curtos, oriundos da tecnologia Illumina, e rapidamente seguido por uma série de outros *softwares*. Atualmente, existem diversas opções de montadores capazes de utilizar diferentes tamanhos de fragmentos, inúmeros formatos de arquivos e que são aplicados a genomas de diferentes complexidades (HUNT et al., 2014).

Os algoritmos de montagem do genoma funcionam tomando todos os segmentos de DNA de uma vez, alinhando-os uns com os outros e tentando identificar as regiões onde dois segmentos de leitura se sobrepõem. Essas sobreposições podem ser incorporadas linearmente no processo de montagem. Quanto mais curtas as sequências, maior a quantidade de sobreposições necessárias para que se possa executar a montagem. Além disso, a cobertura genômica, isto é, o número de vezes que uma determinada região do genoma é coberta por segmentos de leitura, contribui para aumentar a acurácia na região considerada (EKBLÖM e WOLF, 2014).

A utilização de segmentos de leitura com extremidades pareadas, ou seja, com sobreposição de sequência de DNA nas suas duas extremidades, facilita o processo de obtenção de sequências montadas. Quando duas ou mais das sequências obtidas do processo de sequenciamento são emparelhadas (por possuírem extremidades pareadas), são geradas sequências maiores chamadas de *contigs*. Por sua vez, os *contigs* podem ser emparelhados, gerando sequências ainda maiores chamadas de *scaffolds*. Regiões com ausência de sobreposição são usualmente encontradas, caracterizando os *gaps* (BAKER, 2012). A determinação da sequência presente nos *gaps* pode ser obtida por meio da amplificação especificamente dessas regiões com subsequente ressequenciamento, possibilitando alcançar a montagem final do genoma (ALKAN et al., 2011).

A tecnologia NGS gera facilmente dois ou três bilhões de segmentos de leitura de DNA com 100 cópias cada, que podem ser usados na montagem do genoma da espécie (BAKER, 2012). A montagem representa, naturalmente, um desafio de alta complexidade, devido ao pequeno tamanho das sequências iniciais. Durante a montagem de genomas, os fragmentos de leitura geralmente são alinhados com uma sequência genômica conhecida como “referência” para a montagem (HUNT et al., 2014). Na ausência de um genoma de referência, as sequências de leitura devem ser usadas para uma montagem *de novo*. O termo *de novo* vem do latim e significa “desde o princípio”. Refere-se, portanto, aos métodos utilizados para determinar a sequência completa do DNA quando não há nenhuma sequência genômica disponível para uso como referência. As diferentes estratégias de montagem *de novo* têm vantagens e desvantagens em velocidade e precisão quando comparadas entre si. A montagem *de novo* é quase sempre complexa e difícil, particularmente quando o genoma é grande e possui sequências que se repetem muitas vezes, causando falhas na montagem (ALKAN et al., 2011). A decisão para usar a estratégia de montagem *de novo* ou baseada na referência, caso esta última esteja disponível, baseia-se na aplicação biológica, no custo, no esforço para atingir a acurácia necessária e nas considerações de tempo de montagem (EKBLÖM e WOLF, 2014).

Novos avanços no processo de montagem são esperados com a integração de bancos de dados e com a exploração das múltiplas estratégias de

sequenciamento, sempre com o propósito de enfrentar o desafio da montagem de genomas complexos. Como mencionado, os genomas grandes e marcados pela abundância de sequências repetidas ainda constituem um grande desafio para o desenvolvimento de algoritmos de montagem a partir de sequências curtas de DNA. Normalmente, vários montadores são combinados para contornar esse problema. Enquanto isso, a precisão e o comprimento dos fragmentos sequenciados vêm aumentando paulatinamente (ZHANG et al., 2011). O desafio é fazer com que os algoritmos de montagem sejam otimizados para conseguir processar a vasta quantidade de informações dos sequenciamentos NGS que têm suas particularidades para processamento, tanto pelo volume quanto pelos tipos de dados específicos de cada sequenciador (CHAISSON e PEVZNER, 2008).

2.5. Anotação genômica

2.5.1. Anotação automática

A interpretação de dados genômicos brutos envolve a identificação e a anotação de diversos elementos presentes no DNA, como sequências codificadoras de proteínas, tRNAs e rRNAs, pseudogenes e DNA repetitivo (MÉDIGUE e MOSZER, 2007). O interesse da comunidade científica na anotação genômica inclui o esclarecimento dos dados biológicos obtidos e a sua inserção em um contexto para o entendimento de processos biológicos (STEIN, 2001). A interpretação dos dados pode ser realizada nos níveis de nucleotídeo, de aminoácido e de processos biológicos.

Quando a anotação é em nível de nucleotídeos, o intuito principal é a predição de prováveis genes ou *Open Reading Frame* (ORFs) (LUKASHIN e BORODOVSKY, 1998; STEIN, 2001). Os programas que desempenham essa tarefa geralmente apresentam cálculos estatísticos para reconhecer características intrínsecas das regiões codificadoras (STEIN, 2001). A utilização de genes ortólogos de outros organismos também constitui uma estratégia importante para a predição gênica (STEIN, 2001; MÉDIGUE e MOSZER, 2007).

Após a predição de genes, é importante inferir quais produtos eles codificam e em quais processos biológicos irão atuar (STEIN, 2001). Diversos programas são utilizados nesses passos da anotação, não somente para identificar produtos e suas funções, mas também para padronizar nomenclaturas

de maneira universal (REEVES et al., 2009). A anotação automática fornece um mapa final composto por vários elementos genéticos caracterizados e que terá seu conteúdo aprimorado na curadoria manual. A tabela 1.1 apresenta os principais programas utilizados na anotação automática e suas respectivas funções.

Tabela 1. Ferramentas computacionais amplamente utilizadas na anotação de genomas bacterianos.

Utilização	Programa	Referência
Predição de genes	FgenesB	http://www.softberry.com/
	Glimmer	Delcher et al., 1999
Predição de tRNA	tRNA-Scan-SE	Lowe e Eddy, 1997
	ARAGORN	Laslett et al., 2004
Predição de rRNA	RNAmmmer	Lagesen et al., 2007
	Rfam	Griffiths-Jones et al., 2005
Predição de elementos repetitivos	<i>Repeat Scout</i>	Price et al., 2005
Predição de motivos, domínios e famílias proteicas	InterProScan	Hunter et al., 2009
Predição de peptídeo sinal	SignalP	Bendtsen et al., 2004
Predição de promotores	BPROM	http://www.softberry.com/
	PromScan	Studholme e Dixon, 2003
	BDGP	Reese, 2001
Predição de pseudogenes	Consed	Gordon et al., 1998

2.5.2. Curadoria manual

A curadoria manual é uma etapa crucial para a finalização e validação da anotação automática de uma dada sequência genômica (MÉDIGUE e MOSZER, 2007). É nessa etapa que a avaliação humana (do curador) se faz necessária para corrigir e legitimar as informações genômicas preditas anteriormente (STEIN, 2001). Como resultado da curadoria manual, há o aumento da acurácia dos dados que serão disponibilizados para a comunidade científica (REEVES et al., 2009) e, conseqüentemente, a diminuição de uma posterior propagação de erros oriundos da anotação automática (GUNDOGDU et al., 2007).

A realização da curadoria manual ocorre com análises de busca por similaridade em banco de dados proteicos contendo proteínas bem caracterizadas

e conservadas e, assim, é fornecida uma descrição mais detalhada sobre o produto de determinado gene ou família gênica (OVERBEEK et al., 2005). Além disso, pode-se, nesse momento, caracterizar as vias metabólicas, domínios e motivos proteicos conservados presentes no genoma, bem como utilizar uma denominação uniforme para enzimas e outros produtos (RUST et al., 2002).

Alguns grupos que trabalham com genômica de procariotos têm somado esforços para realizar inclusive a reanotação manual de genomas já depositados em bancos de dados (GUNDOGDU et al., 2007; LUO et al., 2009). Esses esforços se baseiam em necessidades como realizar a atualização das informações em diversos genomas que se encontram desatualizadas nos bancos de dados, fornecendo descrições mais detalhadas para seus genes e respectivas proteínas (GUNDOGDU et al., 2007), além de correção ou diminuição de erros oriundos da anotação automática, que pode ser incompleta ou designar falsas descrições (MADUPU et al., 2010), tornando, assim, as informações genômicas disponíveis mais acuradas.

Sabe-se ainda que nos bancos de dados há cerca de 30 a 40% de proteínas sem função conhecida ou proteínas hipotéticas (HANSON et al., 2010). Uma falha nesses bancos é que, em muitos deles, como no caso do *GenBank*, as informações não sofrem atualizações periódicas, ficando obsoletas e induzindo ao erro em novas anotações (SALZBERG, 2007). Muitas das proteínas anotadas como hipotéticas já apresentam descrição disponível atualmente, bem como domínios conservados que possam auxiliar na descrição de suas funções, e a sua atualização vem sendo realizada no processo de curadoria manual, tanto inédita quanto vinculada à reanotação (GUNDOGDU et al., 2007; LUO et al., 2009).

Com efeito, outras abordagens fazem uso dos resultados obtidos pós-curadoria, como as abordagens de genômica comparativa. A comparação genômica entre espécies relacionadas ou entre linhagens de um dado micro-organismo pode, então, ser realizada com maior segurança (RUST et al., 2002).

2.6. Projetos genômicos de *B. thuringiensis*

Atualmente, existem 42 genomas para *B. thuringiensis* depositados no *National Center for Biotechnology Information* (NCBI) (<http://www.ncbi.nlm.nih.gov/genome/browse/>), dos quais 15 correspondem a

genomas fechados e 27 *drafts* de genomas (genomas não completamente montados) (dados até dezembro de 2014).

Os genomas depositados para a espécie variam de 4,93 Mb (para *B. thuringiensis* JM-Mgvxx-63) a 7,09 Mb (para *B. thuringiensis* T01-328), sendo que, nos genomas fechados, 135 sequências de plasmídeos no total foram identificadas e disponibilizadas. Nos *draft* de genomas, possivelmente alguns *contigs* ou *scaffolds* podem corresponder a plasmídeos, levando ao aumento desse valor total com o fechamento de outros genomas, embora determinadas linhagens possam não possuir plasmídeos. As linhagens com genoma menor também possuem um menor número de genes, com *B. thuringiensis* JM-Mgvxx-63 apresentando 4931 genes, enquanto *B. thuringiensis* T01-328 possui 7518 genes. O número de sequências codificadoras de proteínas para essas duas linhagens é 4719 e 7403, respectivamente. Já o conteúdo G+C entre todos os genomas sequenciados não varia muito e permanece entre 34,5 e 35,5%.

O interesse nas linhagens sequenciadas de *B. thuringiensis* concentra-se, principalmente, para o uso em biocontrole, em função de características como atividade inseticida (GUAN et al., 2012; LIU et al., 2013; MURAWSKA et al., 2013; PALMA et al., 2014), nematocida (LIU et al., 2014), moluscicida (WANG et al., 2013) e quitinolítica (MARCON et al., 2014). Outras investigações em algumas das linhagens incluem a utilização para a engenharia genética, em virtude da alta eficiência de transformação (HE et al., 2010; JEONG et al., 2014), o estudo da formação de biofilme (ZWICK et al., 2012) ou mesmo do envolvimento em patologia humana (um caso particular e incomum) (HAN et al., 2006).

2.7. Mineração de genomas e bioprospecção de produtos de *Bacillus*

Os recentes avanços com o sequenciamento de genomas vêm destacando o gênero *Bacillus* como uma inesperada fonte promissora de compostos bioativos. De fato, para alguns deles, como *Bacillus subtilis*, mais de 4% do genoma têm sido encontrados potencialmente envolvidos na síntese de compostos como peptídeos não-ribossômicos, policetídeos e antimicrobianos incomuns (KUNST et al., 1997; ARGUELLES-ARIAS et al., 2009; CHEN et al., 2009).

Peptídeo sintetases não-ribossômicos (NRPSs) são grandes enzimas multimodulares que sintetizam peptídeos não-ribossômicos a partir da ligação

sucessiva de aminoácidos ativados. Além dos domínios principais que catalisam a reação específica para a incorporação de um aminoácido, as NRPSs apresentam domínios mais especializados que realizam modificações nesses monômeros, como epimerização, metilação, oxidação, redução, formilação e heterociclicização. São esses domínios especializados que permitem as NRPSs sintetizarem um grande número de estruturas diversificadas e com uma variedade de funções biológicas que não poderiam ser obtidas pela maquinaria ribossomal (FELNAGLE et al., 2008). A diversidade estrutural é responsável pelas diferentes atividades biológicas, alvos e mecanismos de ação dos peptídeos não-ribossomais, possibilitando a exploração biotecnológica desses peptídeos (FICKERS, 2012).

Desde a caracterização do primeiro peptídeo não-ribossomal no início dos anos de 1970, a gramicidina, produzida por *Bacillus brevis*, muitos outros têm sido isolados e seus *clusters* biossintéticos caracterizados. Com a emergência de poderosas ferramentas moleculares e de bioinformática, a identificação até mesmo de novos metabólitos por meio da mineração de genomas tem se tornado uma realidade (ANSARI et al., 2004; VAN LANEN e SHEN, 2006). Por exemplo, estudos de genomas de *B. thuringiensis* levaram à predição de uma NRPS constituída de sete módulos que poderiam estar envolvidos na síntese de um heptalipeptídeo similar à kurstacina (ABDERRAHMANI et al., 2011). A descoberta de novos compostos naturais com a mineração de genomas é um sinal encorajador, sugerindo que essa metodologia poderá levar ao isolamento de novas moléculas de interesse biotecnológico.

Policetídeo sintases (PKSs) produzem a estrutura central dos policetídeos a partir de monômeros de acil-coenzima A (JENKE-KODAMA e DITTMANN, 2009). PKSs do tipo I são similares a NRPSs, em que diferentes domínios catalíticos são encontrados em uma única cadeia polipeptídica, enquanto PKSs do tipo II apresentam os domínios catalíticos em proteínas mono ou bifuncionais (HERTWECK et al., 2007). Além dos domínios centrais envolvidos na biossíntese, domínios adicionais das PKSs, como cetoreductase, desidratase, metiltransferase, enoilreductase e oxidase são responsáveis pela modificação química do policetídeo em crescimento. PKSs do tipo II frequentemente apresentam um domínio ciclase responsável pela formação de estruturas aromáticas (MEURER et al., 1997). Similarmente aos peptídeos não-ribossomais,

devido ao mecanismo versátil de montagem, os policetídeos exibem notável diversidade tanto em termos de estrutura quanto em atividade biológica. Por exemplo, os genes envolvidos na síntese de difidina, macrolatina e bacileno foram caracterizados em *Bacillus amyloliquefaciens* (CHEN et al., 2007; ARGUELLES-ARIAS et al., 2009). Enquanto a difidina apresenta um amplo espectro de atividade antibacteriana (WILSON et al., 1987; ZIMMERMAN et al., 1987) e tem sido promissora no antagonismo a *Erwinia amylovora*, um fitopatógeno causador do fogo bacteriano em várias rosáceas, como maçã e pêra (ZWEERINK e EDISON, 1987), a macrolatina e o bacileno apresentam um espectro menor, não têm sido associados ao biocontrole na agricultura, mas possuem potencial para uso na medicina (GUSTAFSON et al., 1989; PATEL et al., 1995).

Além de peptídeos não-ribossomais e de policetídeos, diversos *Bacillus* também podem produzir interessantes peptídeos antimicrobianos sintetizados pelos ribossomos, as bacteriocinas. Linhagens de *B. thuringiensis*, por exemplo, sintetizam uma variedade de bacteriocinas e de peptídeos inibitórios semelhantes às bacteriocinas. Os peptídeos antimicrobianos de *B. thuringiensis* são as turicinas, tochicinas e entomocinas, como a turicina B439 de *B. thuringiensis* B439 (AHERN et al., 2003), turicina 17 de *B. thuringiensis* NEB17 (LEE et al., 2009b), turicina H de *B. thuringiensis* SF361 (LEE et al., 2009a), turicina S de *B. thuringiensis* sv. *entomocidus* HD198 (CHEHIMI et al., 2007), tochicina de *B. thuringiensis* sv. *tochigiensis* HD868 (PAIK et al., 1997), entomocina 9 de *B. thuringiensis* sv. *entomocidus* HD9 (CHERIF et al., 2003) e entomocina 110 de *B. thuringiensis* HD110 (CHERIF et al., 2008). *B. thuringiensis* sv. *tochigiensis* BGSC4Y1 e *B. thuringiensis* sv. *monterrey* BGSC4AJ1 podem produzir subtilosina (READ et al., 2009). Outros peptídeos antimicrobianos, como morricina 269, kurstacina 287, queniacina 404 e tolvorticina 524 de linhagens mexicanas também foram reportados (BARBOZA-CORONA et al., 2007; DE LA FUENTE-SALCIDO et al., 2008).

Exemplos de bacteriocinas com potencial para biocontrole na agricultura incluem a ericina S, produzida por *B. subtilis* A1/3, que é ativa contra *Clavibacter michiganensis*, o agente causal do cancro bacteriano do tomate, e o peptídeo Bac 14, produzido por *B. subtilis* 14B isolado de rizosfera, ativo contra *Agrobacterium*

tumefaciens (ABRIOUEL et al., 2011). Outras aplicações de bacteriocinas na agricultura também são possíveis. A promoção do crescimento vegetal pela turicina 17 produzida por *B. thuringiensis* NEB17 tem sido demonstrada (LEE et al., 2009 a,b). Peptídeos com atividade antifúngica também podem ser aplicados no controle de deterioração de plantas e no controle pós-colheita de frutas e legumes, como no caso do peptídeo produzido por *B. amyloliquefaciens* RC-2, ativo contra *Colletotrichum demantium* e outros fungos e bactérias fitopatogênicos (ABRIOUEL et al., 2011).

Em bactérias Gram-negativas, o mecanismo de controle da expressão gênica dependente de densidade populacional, o *quorum sensing*, envolve moléculas sinalizadoras, especialmente N-acil homoserina lactonas (AHLs) (FUQUA et al., 2001), e bactérias fitopatogênicas podem utilizar esse mecanismo na regulação da produção de fatores de virulência (VON BODMAN et al., 2003). No entanto, o *quorum sensing* pode ser inibido por enzimas que degradam AHL, como as lactonases. Dong et al. (2004) mostraram que linhagens de *B. thuringiensis* produtoras de N-acil homoserina lactonase suprimem a virulência dependente de *quorum sensing* do patógeno bacteriano *Pectobacterium carotovorum* (*Erwinia carotovora*). Alternativamente, vegetais geneticamente modificados para expressar *aiiA*, o gene da N-acil homoserina lactonase de *Bacillus*, vem sendo estudados. A superexpressão desse gene em plantas de tabaco, por exemplo, reduziu significativamente a incidência do fogo selvagem causado por *Pseudomonas syringae* pv. *tabaci* (QUIÑONES et al., 2005). Esses resultados mostram o potencial promissor para explorar antagonistas que inibem o mecanismo de comunicação celular bacteriana para o controle e prevenção de doenças.

Além disso, o crescimento e a produção vegetal também podem ser seriamente afetados por estresses abióticos. Entre esses estresses, a deficiência de ferro constitui o principal fator que leva à redução da produtividade, principalmente em solos calcáreos, em que a solubilidade do ferro é extremamente baixa (KOBAYASHI et al., 2005). Um dos mecanismos mais utilizados por micro-organismos para a aquisição de ferro é a produção e a excreção de sideróforos, quelantes de baixo peso molecular que se ligam ao Fe^{3+} com alta afinidade e o transportam para o interior da célula (ZAWADZKA et al.,

2009). Dessa maneira, uma forma pela qual micro-organismos associados a plantas e que são produtores de sideróforos contribuem para a saúde vegetal é por meio da complexação desses compostos com o ferro, tornando-o menos disponível para fitopatógenos (CHEN et al., 2009). Também tem sido demonstrado que, em condições experimentais de solo estéril, as plantas apresentam sintomas de deficiência de ferro e possuem níveis bastante baixos do elemento na raiz, o que evidencia a importância dos micro-organismos para a aquisição do ferro para o vegetal (RADDADI et al., 2007). Para *Bacillus*, a produção de vários sideróforos foi detectada, principalmente esquizoquineno de *Bacillus megaterium* e outros de *B. subtilis*, *Bacillus licheniformis* e do grupo *B. cereus*. Bacilibactina e petrobactina, por exemplo, são produzidas por *B. cereus* (PARK et al., 2005), *Bacillus anthracis* (CENDROWSKI et al., 2004) e *B. thuringiensis* (WILSON et al., 2006).

Outra classe de produtos de *Bacillus* é a das quitinases. A quitina é um polissacarídeo insolúvel formado por cadeias lineares de resíduos de N-acetilglicosamina unidos por ligação β -1,4-glicosídica. Ela é um dos biopolímeros mais abundantes do planeta, atrás apenas da celulose, e é o principal componente da parede celular dos fungos, além de ser encontrada no esqueleto externo de insetos e também em algas, crustáceos e estruturas internas de outros invertebrados (BHATTACHRYA et al., 2007). A cada ano, uma vasta quantidade de resíduos de quitina é liberada no ambiente, o que pode gerar problemas ambientais, e a decomposição desse polímero é importante para manter o balanço de carbono e nitrogênio do ecossistema (HAYES et al., 2008). O catabolismo de quitina envolve a clivagem inicial do polímero por quitinase, produzindo oligossacarídeos e, então, a posterior hidrólise dos monômeros de N-acetilglicosamina por quitobiasas (SUGINTA et al., 2000). Esses derivados têm uma variedade de aplicações agrícolas, industriais e médicas, incluindo uso em cosméticos e atividade antitumoral (BANSODE e BAJEKAL, 2006).

Diversas espécies de *Bacillus*, como *Bacillus alvei*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus cereus* e *B. thuringiensis*, podem produzir quitinases durante o crescimento (ROJAS-AVELIZAPA et al., 1999; GOMAA, 2012). As quitinases são usadas como potencial agente de biocontrole para muitos fungos patogênicos e, dessa forma, podem ser úteis para o controle de doenças de

plantas (TANG et al., 2012), conservação de alimentos e sementes (DE LA VEGA et al., 2006), além de biorremediação (WANG e HUANG, 2001) e preparação de oligossacarídeos e N-acetil-D-glicosamina (WANG et al., 2006). Além disso, o papel potencial de quitinases no biocontrole de insetos, como mosquitos, tem sido sugerido (HAYES et al., 2008). Quitinases podem ser empregadas ainda na área de saúde humana, como na composição de preparações oftálmicas, e em outras áreas biotecnológicas (DAHIYA et al., 2006).

Dessa maneira, muitos produtos de *Bacillus* têm sido descobertos nos últimos anos e, embora tenham ocorrido grandes avanços na sua bioprospecção, novos compostos de interesse biotecnológico ainda são esperados ser encontrados bem como a otimização da sua produção.

3. MATERIAL E MÉTODOS

3.1. Isolado bacteriano e condições de crescimento

Bacillus thuringiensis BAC3151 pertence à bacterioteca do Laboratório de Genética Molecular de Micro-organismos/BIOAGRO da Universidade Federal de Viçosa (UFV). O isolado foi previamente obtido das folhas de *P. vulgaris* (COSTA et al., 2012) e o cultivo realizado em meio TSA (*Tryptone Soy Agar*) 10% (triptona 1,5 g/l, peptona de soja 0,5 g/l, NaCl 1,5 g/l, pH 7,3) a 28°C. O estoque foi mantido a -80°C em criotubos contendo meio TSA 10% e 40% de glicerol.

3.2. Extração do DNA e sequenciamento do genoma

Para o sequenciamento do genoma, o DNA total foi isolado da cultura crescida por 16 horas utilizando o *Wizard genomic DNA purification kit* (Promega), de acordo com as instruções do fabricante para bactéria Gram-positiva. O DNA total purificado (25,278 µg e $A_{260}/A_{280} = 2,05$) foi sequenciado pela MacroGen Inc. (Seoul, Coréia do Sul) utilizando o sistema de sequenciamento HiSeq 2000 (100 bp paired-end) (Illumina). Um total de 21.512.378 *reads* de alta qualidade e 2.172.750.178 bases de alta qualidade foram obtidos, produzindo uma cobertura de aproximadamente 100X.

3.3. Tratamento dos dados e montagem

Os dados brutos obtidos do sequenciamento foram processados usando o pacote Phred-Phrap-Consed (RASKO et al., 2008). O valor mínimo de qualidade usado no programa *base-calling* foi Q20 (que representa acurácia das bases de pelo menos 99%). Os scaffolds foram montados utilizando SOAPdenovo versão 2.04 (LUO et al., 2012).

3.4. Anotação do genoma

O processo de anotação envolveu o uso de vários algoritmos em um procedimento de múltiplos passos. A anotação estrutural foi realizada utilizando os seguintes *softwares*: FgenesB: preditor de gene (<http://www.softberry.org/>); RNAmmer: preditor de rRNA (LAGESEN et al., 2007) e tRNAscan-SE: preditor de tRNA (LOWE e EDDY, 1997). A anotação funcional foi realizada por análises de

similaridade com o BLAST (ALTSCHUL et al., 1990), usando o banco não redundante (NR) e o banco de domínios conservados (CDD), e análise com InterProScan (JONES et al., 2014) (*e-value* <1e-6). A identificação de genes inseticidas também foi feita com *BtToxin_scanner* (YE et al., 2012). Além disso, a anotação manual foi realizada com o Artemis (RUTHERFORD et al., 2000).

3.5. Análise filogenômica

Análise filogenômica de *B. thuringiensis* foi baseada nos genes *core* de BAC3151 e outras 27 linhagens obtidas do NCBI em <ftp://ftp.ncbi.nih.gov/genomes/>. No processo, os genes *core* foram alinhados com o MUSCLE versão 3.8.31 (EDGAR, 2004), as regiões de não correspondência do alinhamento foram mascaradas e removidas posteriormente por GBLOCKS (CASTRESANA, 2000). Em seguida, as regiões restantes de todos os alinhamentos foram concatenadas, formando um alinhamento de grandes sequências, que depois foi usado como entrada para o PHYLIP v. 3.696 (FELSENSTEIN, 1989). A árvore foi construída utilizando o método de Neighbor-Joining com 1000 replicatas.

3.6. Ordenação e orientação dos scaffolds

A determinação da ordem e da orientação dos scaffolds do genoma de *B. thuringiensis* BAC3151 foi feita com o *software* Mauve v. 2.3.1 por meio do algoritmo *Mauve Contig Mover* (DARLING et al., 2010), utilizando a sequência de *B. thuringiensis* MC28 como referência e os parâmetros padrão. É um processo iterativo que só é finalizado quando encontra um resultado repetitivo.

3.7. Rearranjos cromossomais em *B. thuringiensis*

A análise de rearranjos cromossomais em *B. thuringiensis* foi realizada a partir do alinhamento do *draft* do genoma de BAC3151 (com a ordem e a orientação dos scaffolds determinadas) com o cromossomo de outras linhagens da mesma espécie (Tabela 3), usando o algoritmo *Progressive Mauve* (RISSMAN et al., 2009) do Mauve 2.3.1 com os parâmetros padrão.

3.8. Compartilhamento de genes entre linhagens

Efficient Database framework for comparative Genome Analyses using BLAST score Ratios (EDGAR) (BLOM et al., 2009) foi utilizado para analisar o pangenoma e *core* genômico de *B. thuringiensis*, além das sequências codificadoras únicas (*singletons*) de BAC3151. Nesse processo, são considerados os resultados do blast recíproco entre grupos de genomas. As sequências codificadoras de proteínas (CDSs) de um genoma que não possuem *score ratio value* (SRV) (*score* observado/*score* máximo) maior que o ponto de corte em nenhum dos outros genomas analisados são consideradas *singletons*. O SRV do ponto de corte (0,33) foi gerado durante a criação do banco dos genomas usados na análise. No cálculo do genoma *core*, são levadas em consideração somente as CDSs com o melhor *hit* recíproco do blast em todos os genomas utilizados.

4. RESULTADOS E DISCUSSÃO

4.1. Características gerais do *draft* do genoma de *B. thuringiensis* BAC3151

O *draft* da sequência genômica de *B. thuringiensis* BAC3151 possui 5.743.871 pb, distribuídos em 42 scaffolds (com tamanho médio de 136.758 pb e N50 de 552.900 pb), e um conteúdo G+C total de 34,9% (A=1.867.003 bases [32,5%]; T=1.871.097 [32,57%]; C=1.008.063 [17,55%]; G=996.602 [17,35%]; N=1.106 [0,01%]). A tabela 2 mostra o resultado do processamento dos *reads* até a montagem dos scaffolds.

Tabela 2. Resumo do processamento dos dados obtidos do sequenciamento e montagem.

Nº inicial de <i>reads</i>	24.391.936
Nº de <i>reads</i> filtrados (\geq Q20)	21.512.378
Nº de contigs	112
Nº de scaffolds	42

Um total de 5562 sequências codificadoras de proteínas (CDSs) foi predito, sendo que as CDSs compreendem 4.604.274 pb, o que resulta em uma porcentagem codificadora de proteínas de aproximadamente 80,2%. O conteúdo G+C das CDSs é um pouco maior que o global e corresponde a 35,87% (A=1.596.789 bases [34,68%]; T=1.355.335 [29,43%]; C=689.264 [14,97%]; G=962.377 [20,9%]; N=509 [0,01%]). A densidade gênica é 0,968 CDS/kb, sendo o tamanho médio das CDSs igual a 827 pb. O *draft* do genoma também possui 53 genes codificadores de tRNA e 4 de rRNA. Uma visão geral do *draft* genômico de BAC3151 pode ser vista na figura 3 e as principais características deste e dos outros genomas de *B. thuringiensis* utilizados neste trabalho são apresentadas na tabela 3.

Linhagens de *B. thuringiensis* demonstraram ter similar conteúdo G+C, porcentagem codificadora e comprimento médio das CDSs. No entanto, diferenças significantes são observadas no tamanho do genoma e também na composição de genes específicos das linhagens. Elementos genéticos móveis,

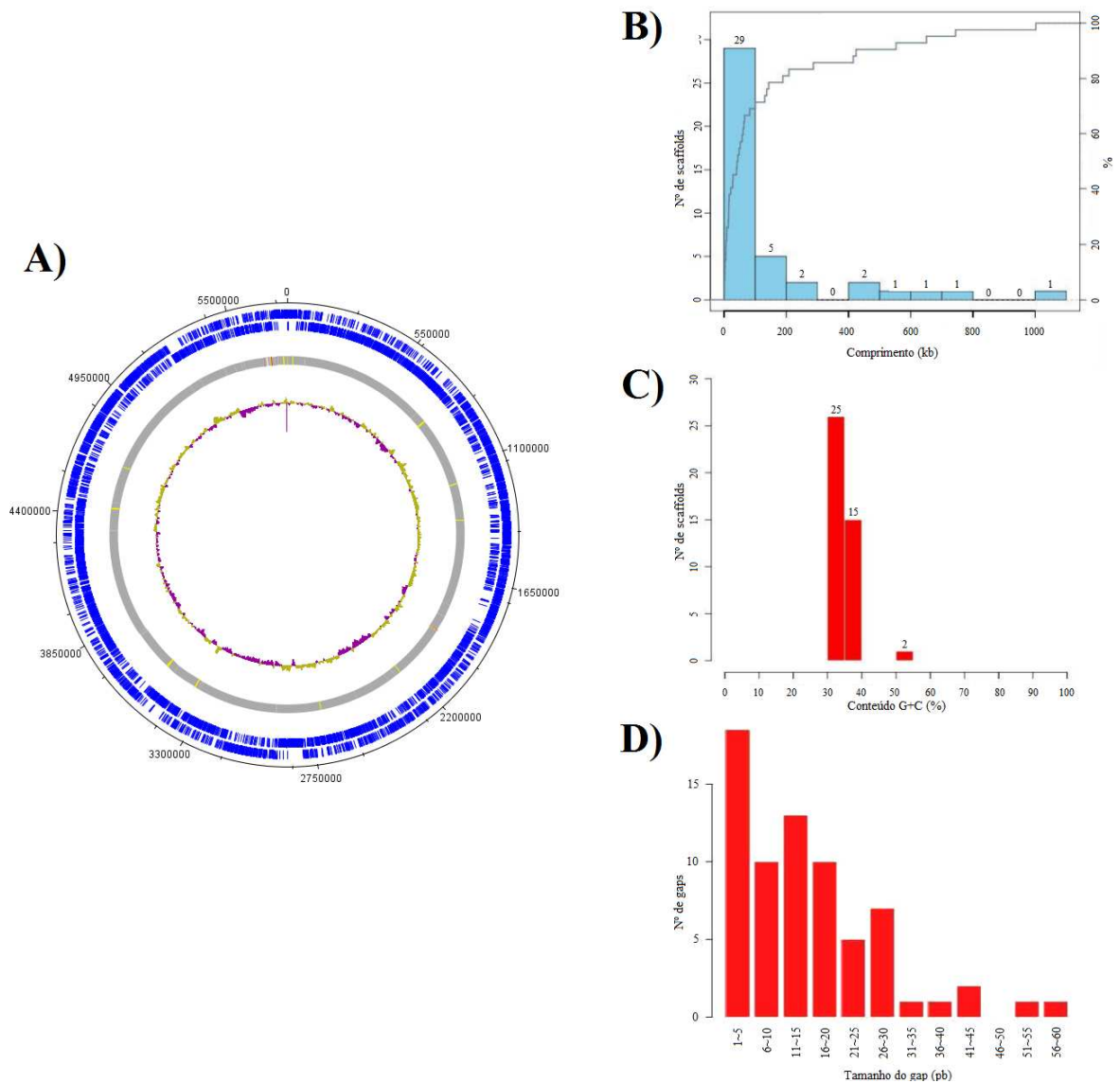


Figura 3. Visão geral do *draft* da sequência do genoma de *B. thuringiensis* BAC3151. A) Exibição circular do *draft* do genoma anotado. Os círculos representam, a partir do mais externo: círculo 1, posição das bases do DNA; círculo 2, regiões codificadoras de proteínas transcritas no sentido horário; círculo 3, regiões codificadoras de proteínas transcritas no sentido anti-horário; círculo 4, regiões codificadoras de tRNA (amarelo) e rRNA (vermelho); círculo 5, conteúdo G+C. B) Distribuição do comprimento dos scaffolds. C) Histograma do conteúdo G+C. D) Distribuição do tamanho dos *gaps*.

como plasmídeos, transposons, fagos, integrons e ilhas genômicas possuem uma contribuição importante para essas diferenças (FROST et al., 2005; FANG et al., 2011; HE et al., 2011). Além disso, 86 proteínas do proteoma predito de *B. thuringiensis* BAC3151 não geraram alinhamento significativo com os bancos utilizados ($e\text{-value} < 1e\text{-6}$) e podem ser codificadas por genes ainda não caracterizados. Esse número é maior que o das outras linhagens de *B. thuringiensis* com genoma disponível.

Tabela 3. Propriedades gerais de genomas de *Bacillus thuringiensis* utilizados neste estudo.

Organismo	Status	Tamanho (Mb)	% GC	Total de genes	Descrição	Número de acesso	Referência
<i>B. thuringiensis</i> BAC3151	Draft	5,74387	34,90	5619	Endofítico com atividade antimicrobiana.	Em processamento	Este estudo
<i>B. thuringiensis</i> MC28	Completo	6,69453	34,92	6843	Isolado de solo. Tóxico para lepidópteros e dípteros.	CP003687.1 CP003688.1 CP003693.1 CP003690.1 CP003689.1 CP003692.1 CP003691.1 CP003694.1	Guan et al. (2012)
<i>B. thuringiensis</i> Bt407	Completo	6,13434	35,02	6590	Linhagem <i>cry</i> ⁻ utilizada para estudos de determinantes genéticos de patogenicidade. É derivada de Bt407 isolado de inseto.	CP003889.1 CP003896.1 CP003893.1 CP003897.1 CP003890.1 CP003894.1 CP003891.1 CP003898.1 CP003895.1 CP003892.1	Sheppard et al. (2013)
<i>B. thuringiensis</i> BMB171	Completo	5,64305	35,19	5513	Linhagem <i>cry</i> ⁻ derivada do sorovar <i>kurstaki</i> YBT-1463 curado do plasmídeo que codifica toxina inseticida.	CP001903.1 CP001904.1	He et al. (2010)
<i>B. thuringiensis</i> HD-789	Completo	6,33463	35,18	6626	Usado para mapeamento genético e análise de polimorfismo de nucleotídeo único (SNP).	CP003763.1 CP003769.1 CP003764.1 CP003766.1 CP003765.1 CP003768.1 CP003767.1	Doggett et al. (2013)
<i>B. thuringiensis</i> YBT-1518	Completo	6,67292	35,29	6877	Linhagem isolada de solo.	CP005935.1 CP005937.1 CP005939.1 CP002486.1 CP005938.1 CP005940.1 CP005936.1	-
<i>B. thuringiensis</i> sv. <i>thuringiensis</i> str. IS5056	Completo	6,77159	34,91	7060	Isolado de solo. Produz cristal altamente tóxico para larva de <i>Trichoplusia ni</i> .	CP004123.1 CP004126.1 CP004127.1 CP004131.1 CP004128.1 CP004136.1 CP004137.1 CP004132.1 CP004135.1 CP004130.1 CP004134.1 CP004129.1 CP004124.1 CP004125.1 CP004133.1	Murawska et al. (2013)
<i>B. thuringiensis</i> sv. <i>kurstaki</i> str. HD73	Completo	5,90857	35,19	6334	Pertence ao sorotipo 3ab e é tóxico para larvas de lepidópteros.	CP004069.1 CP004071.1 CP004072.1 CP004073.1	Liu et al. (2013)

Tabela 3. Continuada.

						<u>CP004070.1</u> <u>CP004075.1</u> <u>CP004076.1</u> <u>CP004074.1</u>	
<i>B. thuringiensis</i> sv. <i>konkukian</i> str. 97-27	Completo	5,31479	35,36	5343	Isolado de tecido humano necrosado. Foi considerado provável patógeno oportunista. É estreitamente relacionado a <i>Bacillus anthracis</i> .	<u>AE017355.1</u> <u>CP000047.1</u>	Han et al. (2006)
<i>B. thuringiensis</i> str. Al Hakam	Completo	5,31303	35,41	4945	Coletado pela Comissão Especial das Nações Unidas em instalação suspeita de armas biológicas no Iraque.	<u>CP000485.1</u> <u>CP000486.1</u>	Challacombe et al. (2007)
<i>B. thuringiensis</i> sv. <i>finitimus</i> YBT-020	Completo	5,68238	35,38	5931	Usado em estudos da associação entre cristais proteicos e esporo.	<u>CP002508.1</u> <u>CP002509.1</u> <u>CP002510.1</u>	Zhu et al. (2011)
<i>B. thuringiensis</i> sv. <i>chinensis</i> CT-43	Completo	6,15115	35,12	6380	Altamente tóxico para lepidópteros e dípteros. Pode produzir níveis altos de turingiensina, que tem atividade inseticida contra ampla gama de insetos.	<u>CP001907.1</u> <u>CP001913.1</u> <u>CP001908.1</u> <u>CP001917.1</u> <u>CP001909.1</u> <u>CP001915.1</u> <u>CP001912.1</u> <u>CP001910.1</u> <u>CP001914.1</u> <u>CP001916.1</u> <u>CP001911.1</u>	He et al. (2011)

4.2. Análise filogenômica

Marcadores convencionais de análise filogenética, como o rDNA 16S, são essencialmente idênticos para linhagens de *B. thuringiensis* (bem como outras do grupo *Bacillus cereus*) (BAVYKIN, et al., 2004) e a relação entre essas linhagens utilizando genes *housekeeping* é frequentemente inconsistente (HELGASON et al., 2004). Por isso, a análise filogenômica foi realizada com 1276 genes *core* de 27 linhagens (Figura 4).

BAC3151 foi agrupado junto com *B. thuringiensis* MC28 isolado de solo e que é altamente tóxico para lepidópteros e dípteros (TAN et al., 2009; TAN et al., 2010). Esse resultado é consistente com o fato de que o solo é uma fonte de bactérias endofíticas em potencial (DOTY, 2008; RYAN et al., 2008). Outras linhagens isoladas de ambientes diferentes também foram agrupadas juntas. Por exemplo, linhagens obtidas de solo, como YBT-1518 e IS5056, agruparam com linhagens isoladas de insetos (ATCC 10792, T01001 [ZWICK et al., 2012] e Bt407), enquanto outras linhagens de solo (BGSC 4CC1 e BGSC 4BA1 [ZWICK et al., 2012]) foram mais similares ao sorovar *konkukian* 97-27, um patógeno humano oportunista (HAN et al., 2006).

Estudos prévios baseados em *multilocus sequence typing* (MLST) (PRIEST et al., 2004), além de outras técnicas como mapeamento genômico (CARLSON et al., 1996), eletroforese em gel de campo pulsado (CARLSON et al., 1994), eletroforese de enzimas *multilocus* (HELGASON et al., 2000), mapeamento de número variável de repetições em tandem (VNTR), BOX-PCR *fingerprinting* (KIM et al., 2002), polimorfismo de comprimento de fragmento amplificado (AFLP) (TICKNOR et al., 2001) e técnicas proteômicas baseadas em espectrometria de massa integradas com métodos estatísticos (DWORZANSKI et al., 2010) relevam a extensiva similaridade entre as linhagens de *B. thuringiensis* bem como outras do grupo *B. cereus*. O agrupamento de linhagens funcionalmente distintas (como patógeno de inseto, patógeno humano oportunista, endofítico com atividade antimicrobiana) está de acordo com estudo prévio da relação filogenética entre linhagens de *B. thuringiensis* baseada no genoma completo das linhagens (ZWICK et al. 2012).

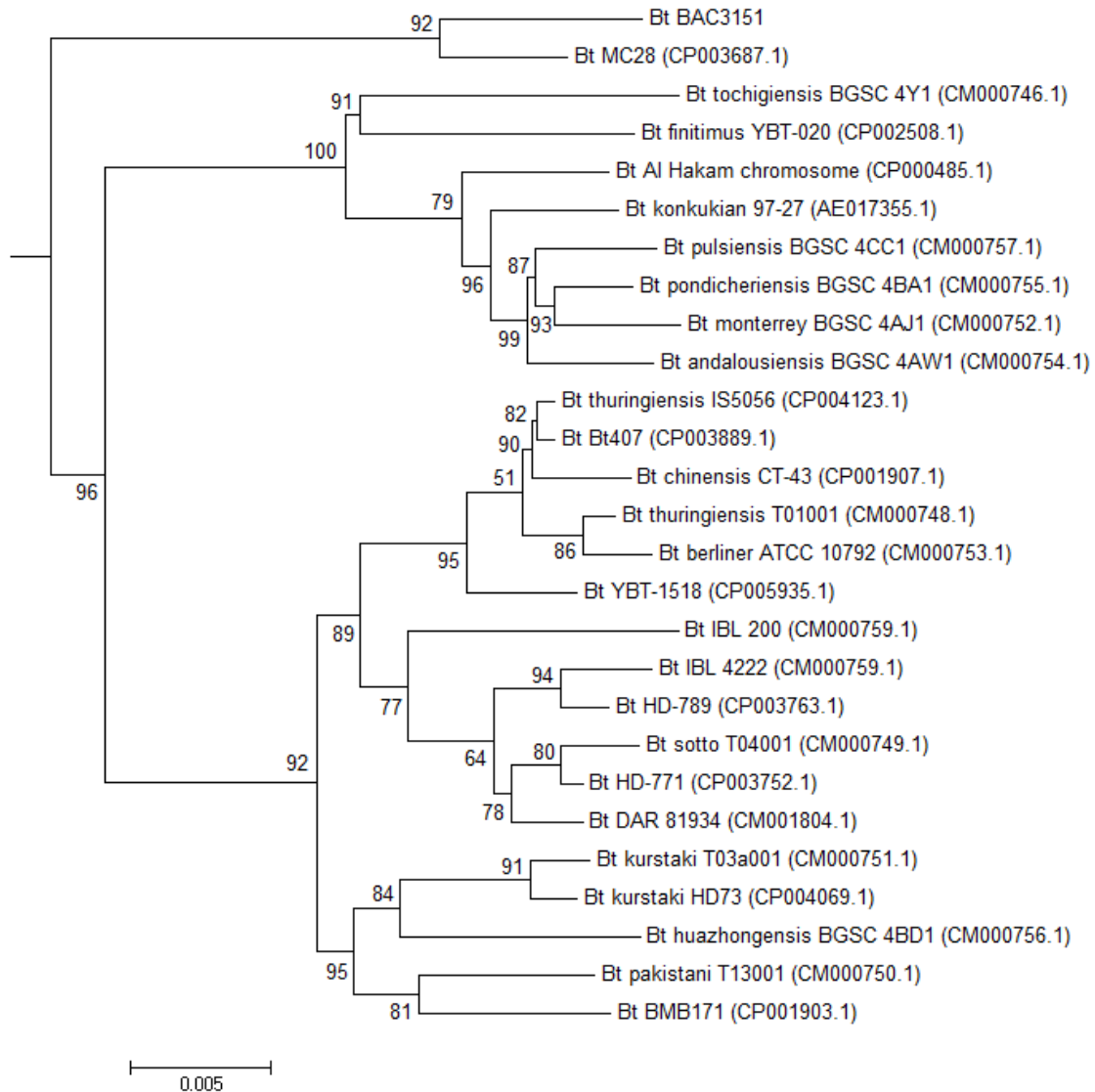


Figura 4. Análise filogenômica de *B. thuringiensis* baseada em 1276 genes *core* das linhagens usadas. A árvore foi construída com PHYLIP (FELSENSTEIN, 1989), utilizando o método de Neighbor-Joining com 1000 replicatas. Os valores de *bootstrap* são mostrados nos nós. A barra da escala mostra cinco substituições de nucleotídeos por 1000 nucleotídeos. O número de acesso das sequências utilizadas é dado entre parênteses.

4.3. Ordenação e orientação dos scaffolds

B. thuringiensis MC28, que foi a linhagem mais próxima a BAC3151 (item 4.2), foi usado como referência para determinar a ordem e a orientação dos scaffolds do nosso isolado.

A figura 5 representa o alinhamento do cromossomo de MC28 com o *draft* do genoma de BAC3151 para esse propósito. As maiores regiões genômicas presumivelmente homólogas e internamente livres de rearranjos, referidas como blocos localmente colineares (LCBs), são representadas por blocos coloridos.

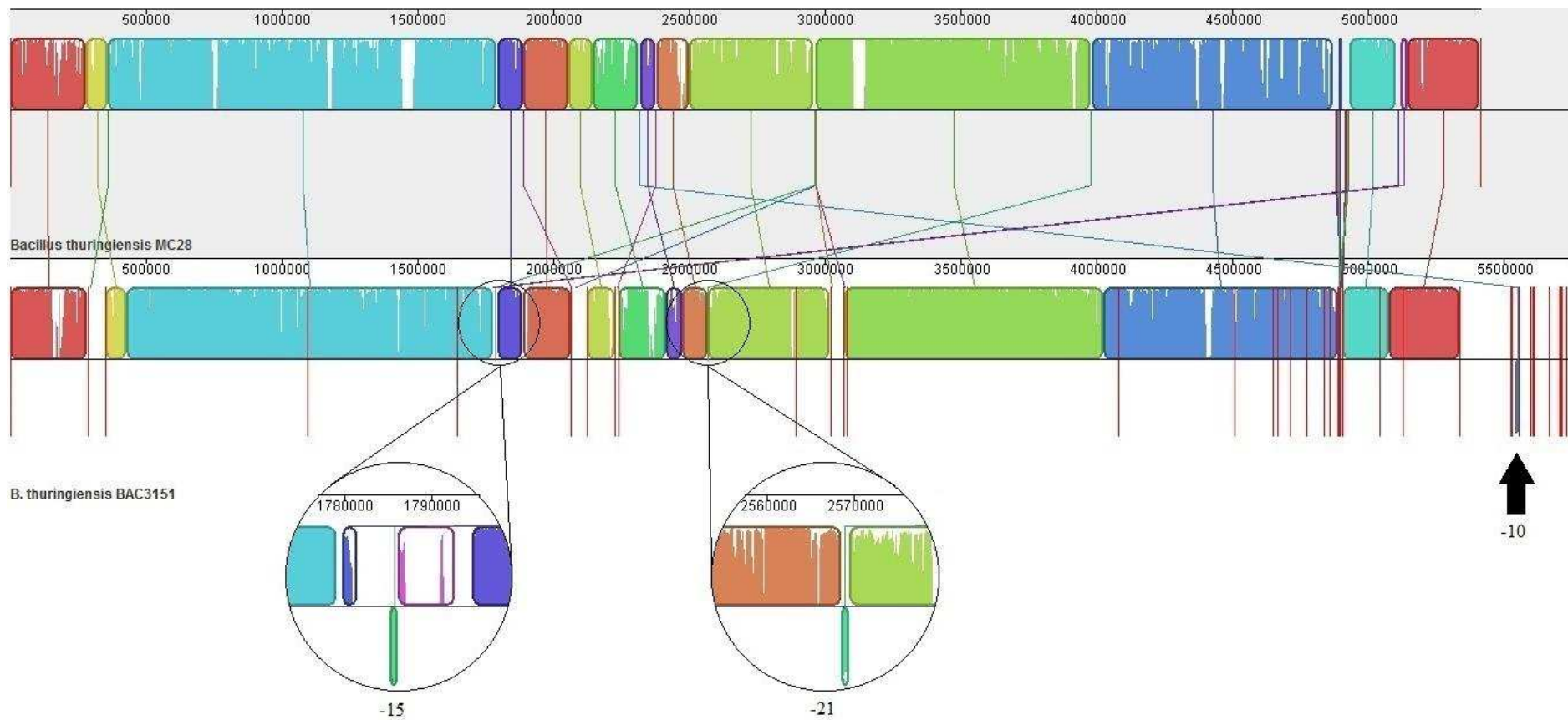


Figura 5. Alinhamento da sequência cromossômica de *B. thuringiensis* MC28 com o *draft* do genoma de *B. thuringiensis* BAC3151 apresentando 36 LCBs após a ordenação e orientação dos scaffolds do *draft*. LCBs homólogos são conectados por linhas, mas somente os maiores LCBs estão representados. As regiões ampliadas e a seta mostram os segmentos invertidos de BAC3151, que são discutidos no texto. As linhas vermelhas verticais da sequência de BAC3151 indicam os limites entre os scaffolds. Embora a visão geral das regiões compartilhadas entre as duas sequências seja clara, alguns rearranjos podem ser ofuscados por um excesso de ordenamento.

Quando um LCB se localiza acima da linha horizontal, a região alinhada possui a mesma orientação que a correspondente na sequência de referência, enquanto que os LCBs abaixo da linha horizontal indicam regiões que alinham na orientação inversa. Segmentos externos aos blocos não possuem homologia detectável entre as sequências alinhadas. A similaridade das regiões genômicas está representada no interior dos LCBs, sendo que a altura do perfil corresponde ao nível de conservação da região. Os limites dos LCBs indicam pontos de rearranjos genômicos, incluindo pontos de ganho ou perda de sequência. Encontrar o arranjo dos scaffolds que minimize o número de LCBs é equivalente a encontrar a sua provável ordem e orientação. Foram encontrados 36 LCBs, com, no mínimo, 317 nucleotídeos homólogos identificados nas regiões colineares.

Além de reordenações de segmentos genômicos, três LCBs invertidos foram identificados em BAC3151. O primeiro (indicado como -15) possui 915 pb e localiza-se no scaffold 16. Essa região apresenta uma sequência codificadora de proteína hipotética e outra parcial de regulador transcricional de bacteriófago. O segundo (indicado como -21) apresenta 981 pb, está localizado no scaffold 3 e possui apenas uma CDS de uma oxidoreductase (aldo/ceto redutase). O terceiro (indicado como -10) é o maior e possui 13178 pb que incluem dez CDSs do scaffold 10, cujos produtos são transposase, β -lactamase, proteínas de modificação de peptídeo e proteínas hipotéticas, além do segmento 5' de uma CDS de proteína hipotética do scaffold 11. A transposase do último segmento invertido pertence à IS605, embora as extremidades da sequência de inserção não tenham sido identificadas, possivelmente em virtude de degeneração do elemento transponível. As regiões genômicas rearranjadas contribuem para diferenciar BAC3151 de MC28, enquanto outras análises realizadas neste trabalho mostram que os dois isolados são mais semelhantes entre si em relação às demais linhagens de *B. thuringiensis* investigadas.

A tabela 4 apresenta a posição final dos scaffolds (numerada em pseudocoordenadas) bem como a orientação *forward* ou *reverse*. Apesar da acurácia, é necessária certa cautela na interpretação dos resultados dos arranjos dos scaffolds, uma vez que muitos rearranjos genômicos verdadeiros podem ocorrer em sequências repetidas, dificultando o alinhamento das sequências. Rissman et al. (2009) relataram uma acurácia entre 90,4% e 99,4% da ordenação

e orientação de sequências de diferentes *drafts* de genoma utilizando o *Mauve Contig Mover*.

A determinação da ordem e da orientação dos scaffolds facilita o fechamento de *draft* de genomas e análises comparativas. A utilização de um genoma de referência para essa finalidade mais similar ao *draft* produz uma ordenação mais acurada e uma cobertura maior das sequências, além de estabelecer a orientação mais provável das mesmas.

Tabela 4. Ordem e orientação final de todos os scaffolds do *draft* genômico de *B. thuringiensis* BAC3151.

Nome	Orientação	Posição inicial	Posição final
scaffold9	<i>reverse</i>	1	287073
scaffold4	<i>forward</i>	287074	351153
scaffold26 ^a	<i>forward</i>	351154	1095058
scaffold24	<i>reverse</i>	1095059	1647958
scaffold16	<i>forward</i>	1647959	2063090
scaffold6 ^a	<i>reverse</i>	2063091	2124415
scaffold32	<i>forward</i>	2124416	2225266
scaffold31	<i>forward</i>	2225267	2240649
scaffold3	<i>forward</i>	2240650	2891218
scaffold19	<i>forward</i>	2891219	3022029
scaffold20 ^a	<i>forward</i>	3022030	3066338
scaffold23 ^a	<i>reverse</i>	3066339	3080235
scaffold5	<i>reverse</i>	3080236	4082243
scaffold13 ^a	<i>forward</i>	4082244	4506425
scaffold25	<i>forward</i>	4506426	4650005
C1123 ^a	<i>reverse</i>	4650006	4664851
C1173 ^a	<i>forward</i>	4664852	4713332
scaffold27	<i>reverse</i>	4713333	4771614
scaffold2	<i>forward</i>	4771615	4837291
scaffold18	<i>forward</i>	4837292	4858510
scaffold1 ^a	<i>forward</i>	4858511	4887195
C1055	<i>reverse</i>	4887196	4889722
C1051	<i>reverse</i>	4889723	4892122
C1083 ^a	<i>reverse</i>	4892123	4897873
C1105 ^a	<i>reverse</i>	4897874	4905973
scaffold28	<i>reverse</i>	4905974	5043094
scaffold7	<i>reverse</i>	5043095	5125671
scaffold21	<i>forward</i>	5125672	5334211
scaffold12	<i>reverse</i>	5334212	5523281
scaffold8	<i>forward</i>	5523282	5526453
scaffold10	<i>forward</i>	5526454	5554308

Tabela 4. Continuada.

scaffold11	<i>forward</i>	5554309	5596095
scaffold17	<i>forward</i>	5596096	5604454
scaffold22	<i>forward</i>	5604455	5611028
scaffold29	<i>forward</i>	5611029	5663233
scaffold30	<i>forward</i>	5663234	5703273
C1031	<i>forward</i>	5703274	5704386
C1035	<i>forward</i>	5704387	5705583
C1037	<i>forward</i>	5705584	5706865
C1079	<i>forward</i>	5706866	5712279
C1127	<i>forward</i>	5712280	5727434
C1131	<i>forward</i>	5727435	5743871

^a Scaffolds com localização alternativa possível.

4.4. Alta similaridade na arquitetura do genoma

Genomas podem sofrer mudanças em pequena e em grande escala durante a evolução. Uma relação colinear global foi encontrada nos genomas utilizados neste estudo, apesar de alguns rearranjos observados (Figura 6). Além de reordenações de segmentos genômicos, BAC3151 e MC28 apresentaram dois LCBs invertidos. O primeiro (indicado pelo número 1 na figura 6) possui um tamanho médio de 27557 pb, e inclui, principalmente, transposase relacionada à IS231, resolvase de transposon, eritromicina esterase (um gene de resistência à eritromicina) (essas três CDSs em BAC3151), além de sequências codificando transportadores de aminoácidos, de magnésio e outros cátions para o interior da célula, diferentes reguladores transcricionais e proteínas hipotéticas. Possivelmente, esse rearranjo seja resultado da presença de elemento genético móvel. O segundo rearranjo (indicado pelo número 2) possui tamanho médio de 19870 pb e é encontrado também nas linhagens IS5056, CT-43, HD73, YBT-1518 e Bt407, porém em uma localização diferente do genoma. Ele inclui, principalmente, sequências codificadoras de hemolisina, proteínas envolvidas na esporulação e na germinação de esporos, além de ativador transcricional da família AraC, que pode regular genes com diferentes funções, desde metabolismo de carbono a respostas a estresses (GALLEGOS et al., 1997; MARTIN e ROSNER, 2001). Interessantemente, *B. thuringiensis* sv. *kurstaki* HD73 apresenta uma região genômica invertida bem maior que a das outras linhagens (indicada pelo número 3) e que inclui oito LCBs com genes putativos diversos.

Rearranjos genômicos podem ser observados em diferentes tipos de organismos (TILLIER e COLLINS, 2000; EICHLER e SANKOFF, 2003). Em

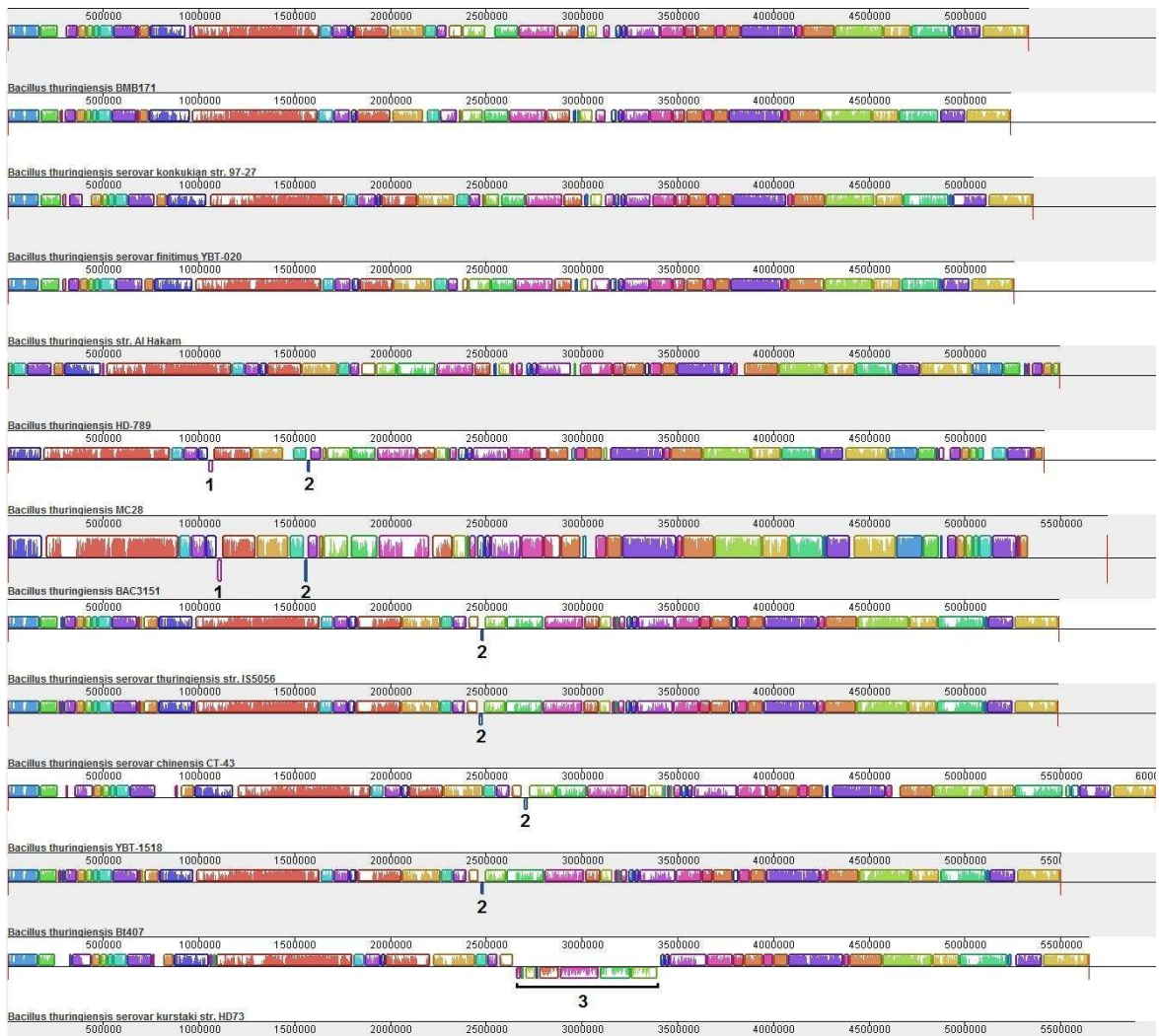


Figura 6. Alinhamento de cromossomos de diferentes linhagens de *B. thuringiensis* e o *draft* do genoma de BAC3151. Segmentos de mesma cor representam regiões homólogas, sem rearranjos internos, denominadas de blocos localmente colineares (LCBs). Regiões invertidas são claramente apresentadas como blocos abaixo da linha horizontal. Essas regiões são indicadas pelos números 1, 2 e 3, e são discutidas no texto. A escala mostra as coordenadas de cada sequência.

Shigella e *Salmonella*, eles têm sido associados a elementos genéticos móveis específicos (DARLING et al., 2010). No gênero *Bacillus*, apesar de Rouli et al. (2014) não terem identificado rearranjos em linhagens de *Bacillus anthracis*, tais alterações genômicas têm sido reportadas em espécies relacionadas. Um segmento invertido de mais de 1,2 milhões de pares de bases foi encontrado no cromossomo de *Bacillus amyloliquefaciens* TA208, presumivelmente devido à infecção fágica (ZHANG et al., 2011). Em *Bacillus subtilis*, foi relatada a influência de rearranjo na expressão gênica (STRAGIER et al., 1989). Em *B. thuringiensis*, embora rearranjos de DNA em plasmídeo têm sido reportados (ZGHAL e JAOUA, 2006), não têm sido encontrados estudos que abordem rearranjos cromossomais

em grande escala para a espécie. Esses rearranjos podem ter implicações importantes na expressão gênica das linhagens e, conseqüentemente, na adaptação a diferentes ambientes e também nas suas potencialidades de uso biotecnológico.

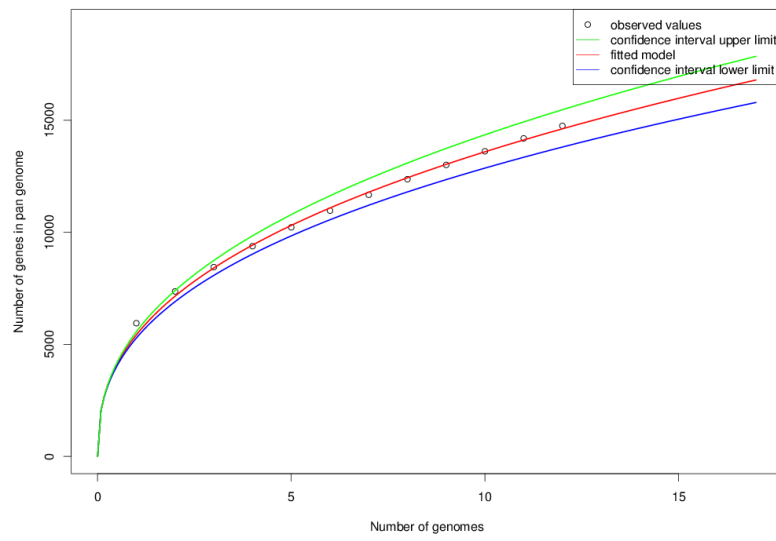
4.5. Compartilhamento de genes entre linhagens

Foi calculado o pangenoma, genoma core e o número médio de singletons de *B. thuringiensis* com base no draft do genoma de BAC3151 mais os 11 genomas completos descritos na tabela 3. O pangenoma das linhagens analisadas foi estimado em 14755 CDSs, apresentando 3413 CDSs core, o que corresponde a aproximadamente 49,3% a 71,1% das CDSs. O número médio de singletons foi de 326 CDSs.

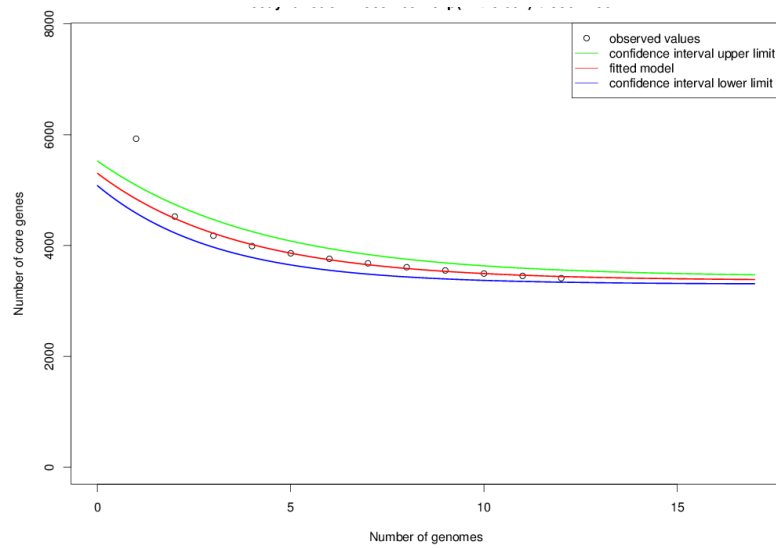
À medida que mais sequências completas de genomas de uma espécie tornam-se disponíveis, o tamanho do pangenoma dessa espécie geralmente aumentará devido a um aumento no número de genes acessórios. Baseado em modelos matemáticos, é predito que novos genes serão descobertos dentro do pangenoma de muitas espécies bacterianas ainda que centenas ou, possivelmente, milhares de sequências genômicas completas tenham sido caracterizadas (MEDINI, 2005). A figura 7A apresenta o aumento do número de CDSs do pangenoma de *B. thuringiensis* com o aumento do número de genomas analisados. Dessa forma, *B. thuringiensis* apresenta um pangenoma “aberto”. Essa observação está de acordo com a hipótese de que um pangenoma aberto é a norma para o gênero *Bacillus*, embora *Bacillus anthracis* seja exceção (FANG et al., 2011).

As razões pelas quais foram usados exclusivamente os 12 genomas em vez de todos os genomas de *B. thuringiensis* disponíveis no NCBI para os cálculos são duas. Primeiro, o pangenoma da espécie é aberto. Segundo, para as análises do genoma core (Figura 7B) e de singletons (Figura 7C), a curva permanecia praticamente a mesma com a adição de novos genomas e, assim, as análises estariam somente levemente viesadas. A análise do genoma core mostrou que ele conserva as CDSs das funções fisiológicas gerais e de sobrevivência da espécie. Em um estudo prévio com somente sete linhagens de *B. thuringiensis* (FANG et al., 2011), a proporção de genes compartilhados entre

(A)



(B)



(C)

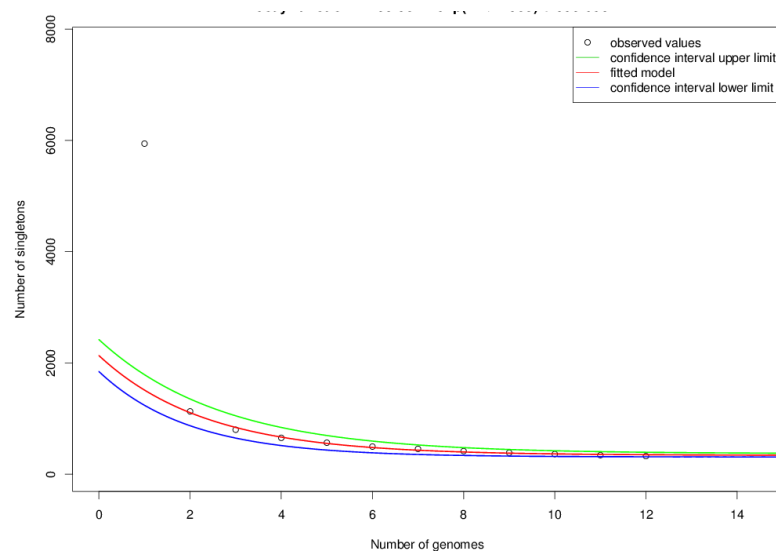


Figura 7. Desenvolvimento do pangenoma, genoma *core* e singletons de *B. thuringiensis* com o aumento do número de genomas. Modelos de ajuste de quadrados mínimos não-lineares

predizem o desenvolvimento do pangenoma (A), *core* genômico (B) e singletons (C) dos genomas com base nas funções $5422,261x^{0,399}$, $1935,165e^{(-x/3,664)}$ + 3367,738 e $1793,351e^{(-x/2,355)}$ + 339,683, respectivamente.

as linhagens variou dentro da faixa observada por nós quando foram utilizadas as 12 linhagens. A figura 8 mostra o compartilhamento de CDSs de BAC3151 com os outros *B. thuringiensis*. O nosso isolado possui mais CDSs em comum com MC28 (4951 genes) e menos com Al Hakam (3965 CDSs). Essa observação reforça a maior proximidade de BAC3151 com MC28 observada na análise filogenética. O menor número de genes compartilhados com Al Hakam pode ser resultado não somente de uma relação menos próxima, mas também porque Al Hakam possui um menor repertório gênico, diminuindo assim a possibilidade de compartilhamento.

Além disso, houve uma grande variação do número de *singletons* entre as linhagens de *B. thuringiensis*, como observado por Fang et al. (2011). A terminologia *singletons* descreve as CDSs encontradas em somente uma das linhagem do nosso estudo; as mesmos CDSs talvez possam ser encontradas em outras linhagens ou espécies. *B. thuringiensis* BAC3151 apresentou 221 *singletons*, cerca de 4,0% do total de CDSs, enquanto, nos outros genomas, esse valor variou de aproximadamente 0,4% a 11,% (para o sorovar *chinensis* CT-43 e para Bt407, respectivamente). A maioria dos *singletons* foi associada a proteínas hipotéticas, conforme observado em outras linhagens e espécies (SMITS et al., 2010; JOSEPH et al., 2012). O grande número de proteínas hipotéticas encontradas em várias linhagens pode ser resultado de genes provenientes da transferência horizontal, incluindo de organismos distantes filogeneticamente, para os quais a caracterização molecular não tem sido feita ainda. De fato, *B. thuringiensis* pode facilmente adquirir material genético de outras bactérias, como relatado anteriormente (GONZALEZ et al., 1982; HAACK et al.,1996; VILAS-BÔAS et al.,1998). A aquisição de genes externos pode contribuir para a melhor a adaptação de subpopulações de bactérias a nichos específicos (LAING et al., 2010). Outros *singletons* encontrados estão envolvidos principalmente em regulação gênica, síntese de componentes celulares, resposta a estímulo e processamento da informação genética (replicação e reparo).

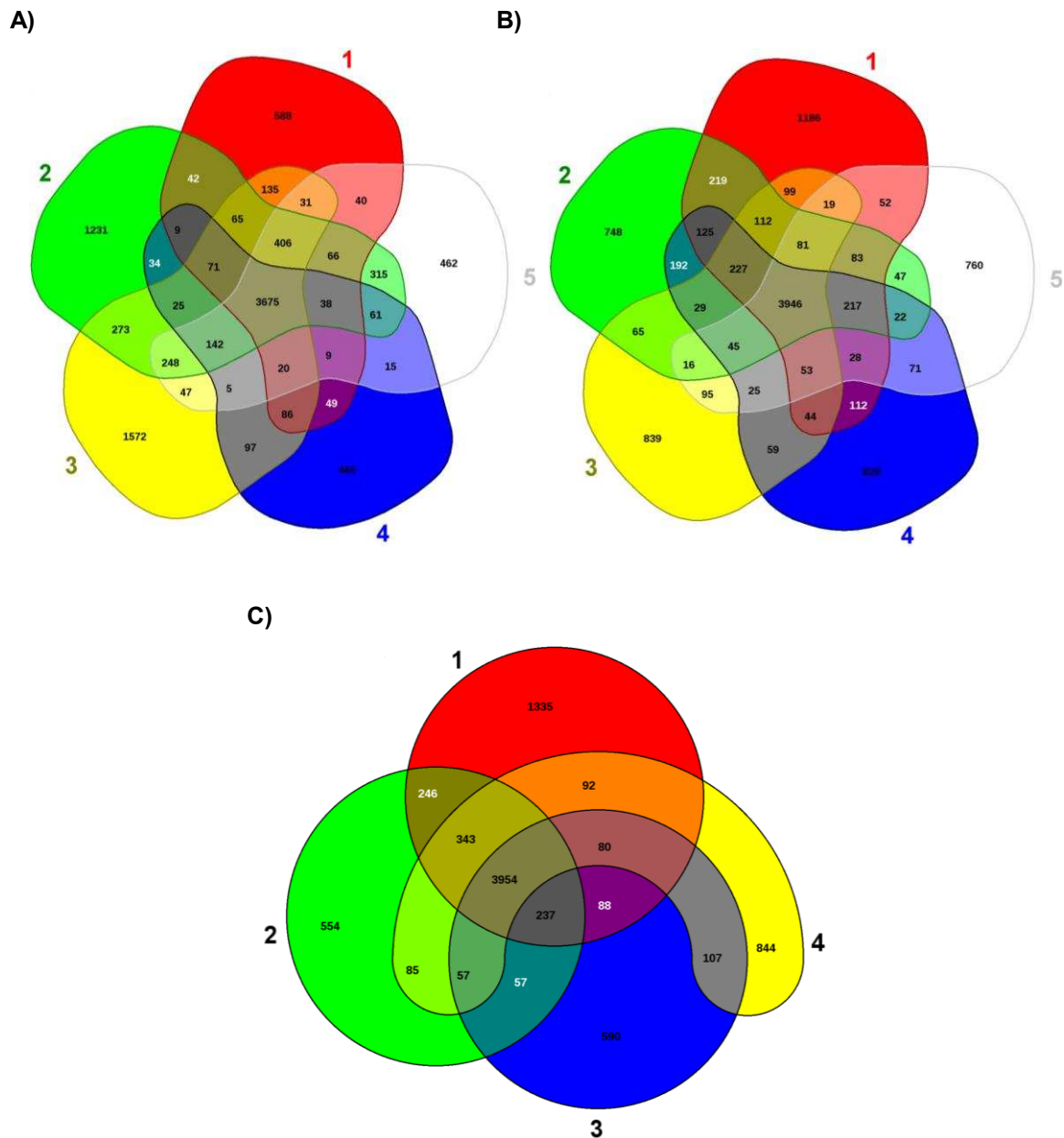


Figura 8. Diagramas de Venn construídos com EDGAR (BLOM et al., 2009) apresentando o compartilhamento de CDSs de BAC3151 com outros genomas. A) 1. BMB171, 2. MC28, 3. sv. *thuringiensis* IS5056, 4. Al Hakam, 5. BAC3151; B) 1. YBT-1518, 2. sv. *chinensis* CT-43, 3. YBT-020, 4. sv. *kurstaki* HD73, 5. BAC3151; C) 1. Bt407, 2. HD-789, 3. sv. *konkukian* 97-27, 4. BAC3151. Regiões sobrepostas representam CDSs comuns entre os genomas analisados. Os valores fora das regiões sobrepostas indicam o número de CDSs em cada genoma sem ortólogos nos outros genomas.

4.6. Genes putativos de *B. thuringiensis* BAC3151 associados a antagonismo

Os micro-organismos competem por fontes limitadas de recursos dentro da comunidade e os mecanismos de competição incluem efeitos antagonistas. *B. thuringiensis* BAC3151 tem demonstrado atividade antimicrobiana *in vitro* contra bactérias Gram-negativas e Gram-positivas, incluindo *Curtobacterium*

flaccumfaciens pv. *flaccumfaciens*, responsável pela murcha bacteriana do feijoeiro (dados não publicados). Um repertório de genes putativos chave do isolado potencialmente envolvidos em antagonismo foi encontrado e é apresentado na tabela 5, que inclui genes inseticidas, policetídeo sintases, bacteriocinas, peptídeo sintetases não-ribossomais, proteínas de biossíntese de sideróforos, quitinases, N-acil homoserina lactonase e outros. Esses genes corroboram a atividade antimicrobiana observada e ainda apontam para possibilidades de novos testes que poderão ampliar as potencialidades de uso de BAC3151 para controle biológico de doenças de plantas.

Dois genes inseticidas foram encontrados em BAC3151, um (C1127_orf00006) com 95% de similaridade com um gene de proteína mosquitocida de *B. thuringiensis* IBL 4222 e outro (scaffold24_orf00006) com 94,7% de similaridade com uma δ -endotoxina inseticida de *B. thuringiensis* MC28. O número de genes inseticidas varia entre as linhagens (FANG et al., 2011), que podem produzir um ou mais cristais proteicos parasporais. As proteínas inseticidas individuais que constituem os cristais têm um espectro de atividade tóxica restrita a poucas espécies de uma ordem particular de insetos, principalmente lepidópteros, dípteros, coleópteros e himenópteros (RHO et al., 2007). Atividade contra nematoides e outros invertebrados também tem sido descrita (WEI et al., 2003; KOTZE et al., 2005; RHO et al., 2007). A combinação de toxinas de uma dada linhagem de *B. thuringiensis* define o espectro de atividade dessa linhagem. Entretanto, há muitas linhagens produtoras de cristais sem atividade tóxica conhecida (RHO et al., 2007). Análises adicionais poderão confirmar a existência de cristais parasporais em BAC3151 e a sua toxicidade, pois avanços no desenvolvimento de novos biopesticidas baseados em *B. thuringiensis* e a manipulação de seus genes dependem da variedade de linhagens disponíveis e seus correspondentes genes inseticidas.

B. thuringiensis BAC3151 também possui o gene *aiiA* (codificador de N-acil homoserina lactonase) que, presumivelmente, é o responsável pela observada inibição do *quorum sensing* de bactérias Gram-negativas pelo isolado (LOPES, 2013). Muitas bactérias fitopatogênicas Gram-negativas utilizam o *quorum sensing* como mecanismo de regulação de diferentes atividades biológicas, incluindo a produção de fatores de virulência (VON BODMAN et al., 2003). O *aiiA*

Tabela 5. Classes de genes putativos chave potencialmente envolvidos em atividade antagonista de *B. thuringiensis* BAC3151

Classe	Tamanho (pb)	Subject	Bit score	Cobertura	E-value	Similaridade
Policetídeos (sintases)						
scaffold9_orf00043	294	[GenBank:WP_001215030.1]	606	0,996599	1,00E-171	1
scaffold16_orf00424	47	[GenBank:WP_000142746.1]	94,7	0,978723	1,00E-17	1
scaffold32_orf00058	120	[GenBank:YP_008780959.1]	250	0,991667	1,00E-64	1
Bacteriocinas (síntese/maturação)						
scaffold3_orf00655	68	[GenBank:YP_008781619.1]	136	0,985294	2,00E-30	1
scaffold5_orf00527	915	[GenBank:WP_000754144.1]	1800	0,998907	0	0,971585
scaffold21_orf00106	1435	[GenBank:WP_002037942.1]	2831	0,987456	0	0,989422
scaffold26_orf00189	513	[GenBank:WP_016090937.1]	1068	0,998051	0	0,998051
scaffold26_orf00190	649	[GenBank:WP_000191878.1]	1320	0,998459	0	0,978428
scaffold29_orf00015	942	[GenBank:WP_016090759.1]	1760	0,998938	0	0,953291
Peptídeos não-ribossomais (sintetases)						
scaffold24_orf00132	74	[GenBank:YP_008785493.1]	154	0,986486	9,00E-36	1
scaffold28_orf00020	1193	[GenBank:WP_000761471.1]	2465	0,999162	0	1
Sideróforos (síntese)						
scaffold19_orf00032	253	[GenBank:YP_008781664.1]	491	0,996047	1,00E-136	0,948617
scaffold24_orf00456	612	[GenBank:WP_000861773.1]	1272	0,998366	0	1
scaffold24_orf00453	327	[GenBank:YP_008785115.1]	680	0,996942	0	1
Quitinases						
scaffold1_orf00008	235	[GenBank:YP_006831774.1]	473	0,995745	1,00E-131	1
scaffold19_orf00019	360	[GenBank:YP_008781652.1]	714	0,997222	0	0,977778
scaffold28_orf00015	688	[GenBank:WP_002050263.1]	1316	0,986919	0	0,954412
N-acil homoserina lactonase						
scaffold3_orf00391	250	[GenBank:YP_006829423.1]	512	0,996	1,00E-143	1
Outros (síntese de antibióticos)						
scaffold24_orf00119	206	[GenBank:WP_002038636.1]	432	0,995146	1,00E-119	1
scaffold26_orf00002	311	[GenBank:YP_006827114.1]	390	0,649518	1,00E-106	0,995074
scaffold32_orf00043	492	[GenBank:YP_006828972.1]	399	0,457317	1,00E-108	0,884956
scaffold32_orf00063	108	[GenBank:WP_002038952.1]	219	0,990741	3,00E-55	1
Genes inseticidas						
C1127_orf00006	1584	[GenBank: EEN00236.1]	1024	0,985	0	0,95
scaffold24_orf00006	292	[GenBank: YP_006828518.1]	547	0,996575	1,00E-153	0,94863

tem sido encontrado em diferentes sorovares de *B. thuringiensis* (LEE et al., 2002) e a inibição da comunicação química de bactérias fitopatogênicas pode constituir uma estratégia eficiente para o controle de doenças de plantas. Os estudos indicam que o uso de inibidores de *quorum sensing* para esse propósito é promissor, ampliando as atuais formas de prevenção de bacterioses (DONG et al., 2004).

Além disso, genes envolvidos na biossíntese e maturação de bacteriocinas, principalmente lantibióticos, estão presentes. Embora os lantibióticos sejam mais bem descritos para *B. subtilis*, linhagens de *B. thuringiensis*, como *tochigiensis* BGSC4Y1 e *monterrey* BGSC4AJ1, também possuem potencial para produzi-los (READ et al., 2009). Bacteriocinas de *Bacillus* são cada vez mais importantes devido ao seu espectro de inibição por vezes mais amplo que o de muitas bactérias do ácido láctico e que pode incluir bactérias Gram-negativas, leveduras e outros fungos, além de bactérias Gram-positivas (ABRIOUEL et al., 2011).

Outra classe de genes de BAC3151 potencialmente envolvidos em biocontrole é a das quitinases. Três genes putativos de quitinase (genes *chi*) foram identificados no genoma. Em *B. thuringiensis*, até sete cópias têm sido reportadas (MARCON et al., 2014). *Bacillus* spp. capazes de degradar quitina em solo (KUZU et al., 2012), filoplano (SMITT e COUCHE, 1991) e em associação com insetos (SNEH et al., 1993) têm sido descritos. Interessantemente, a atividade quitinolítica de *B. thuringiensis* pode apresentar efeito sinérgico sobre a toxicidade inseticida. Isso pode ser aplicado para o desenvolvimento de melhores produtos inseticidas pela combinação de ambos os fatores em uma linhagem ou em formulações (WIWAT et al., 2000; LIU et al., 2002).

Genes responsáveis pela biossíntese de sideróforos também são encontrados no genoma do isolado. Esses genes incluem *rhbC*, envolvido na síntese de rizobactina, e *asbA* e *asbE*, envolvidos na síntese de petrobactina. Embora a rizobactina seja mais comumente associada a *Rhizobium*, *rhbC* é frequentemente encontrado nos genomas sequenciados de *B. thuringiensis*, como na linhagem Bt407 e no sorovar *kurstaki* YBT-1520. Por sua vez, a petrobactina é mais característica de espécies de *Bacillus*, como *B. thuringiensis* (WILSON et al., 2006), *B. cereus* (PARK et al., 2005) e *Bacillus anthracis* (CENDROWSKI et al., 2004). Sideróforos de *B. thuringiensis* podem ser especialmente importantes para

o biocontrole na agricultura, uma vez que *B. thuringiensis* é naturalmente associado a plantas e a competição pelo ferro pode levar à inibição de fitopatógenos, além de aumentar suprimento do elemento para o vegetal. Os fitopatógenos geralmente não são capazes de produzir sistemas comparáveis de aquisição de ferro (RADDADI et al., 2007).

Além desses, genes de policetídeos sintases e peptídeo sintetases não-ribossomais foram encontrados. Devido ao mecanismo versátil de síntese dos policetídeos e dos peptídeos não-ribossomais, eles podem exibir estruturas diversificadas e com uma variedade de funções biológicas, incluindo atividade antimicrobiana. Existem poucas informações na literatura sobre esses produtos de *B. thuringiensis*. No entanto, o crescente aumento de genomas disponíveis para a espécie oferece a oportunidade para a mineração de genes potencialmente envolvidos na sua síntese e cujas aplicabilidades poderão ser exploradas. Policetídeos e peptídeos não-ribossomais são, na verdade, difíceis de serem obtidos, o que torna a expressão heteróloga uma alternativa para o uso. Sistemas de produção heteróloga também podem facilitar o estudo das propriedades catalíticas das policetídeos sintases e das peptídeo sintetases não-ribossomais. Exemplos de hospedeiros heterólogos para produção dessas proteínas incluem *Escherichia coli*, *B. subtilis* e *Streptomyces coelicolor* (RUDE e KHOSLA, 2004).

Outros genes, que presumivelmente participam da síntese de compostos antibióticos, foram encontrados no genoma de *B. thuringiensis* BAC3151 (brevemente descritos na tabela 5). Esses genes muitas vezes fazem parte de *clusters* envolvidos em um complexo metabolismo secundário característico de linhagens produtoras de diversos antimicrobianos. Neste trabalho, foram abordados os genes putativos chave dos múltiplos mecanismos que podem ser responsáveis por biocontrole. Embora a fração do genoma de BAC3151 atribuída à síntese de antimicrobianos aparentemente não seja tão grande quanto a de alguns *Bacillus* spp. (KUNST et al., 1997; CHEN et al., 2009), pesquisas futuras deverão analisar a funcionalidade dos genes encontrados, a conservação das vias de síntese e as propriedades dos eventuais produtos bem como a sua aplicabilidade na agricultura ou mesmo em outras áreas. Dessa forma, o isolado

pode representar uma fonte de metabólitos secundários e enzimas importantes para controlar micro-organismos patogênicos.

5. CONCLUSÕES

Neste trabalho, foi verificado que *B. thuringiensis* BAC3151 endofítico do feijoeiro comum foi mais similar a *B. thuringiensis* MC28 isolado de solo. Rearranjos cromossômicos específicos identificados em BAC3151 e outros rearranjos em outras linhagens de *B. thuringiensis* foram encontrados e podem ter implicações importantes na funcionalidade dessas linhagens. As análises genômicas também revelaram genes de BAC3151 sem similaridade significativa com outros dos bancos de dados públicos bem como genes não encontrados nas outras linhagens analisadas e que serão melhor investigados futuramente. Além disso, genes putativos envolvidos em múltiplas formas de antagonismo estão presentes. Dessa maneira, os resultados apresentados por nós contribuirão para o melhor entendimento da organização do genoma de linhagens de *B. thuringiensis* e para o desenvolvimento de estratégias para o controle biológico de doenças do feijoeiro.

6. REFERÊNCIAS

ABDERRAHMANI, A.; TAPI, A.; NATECHE, F.; CHOLLET, M.; LECLERE, V.; WATHELET, B.; HACENE, H.; JACQUES, P. Bioinformatics and molecular approaches to detect NRPS genes involved in the biosynthesis of kurstakin from *Bacillus thuringiensis*. **Appl Microbiol Biotechnol**, v. 92, p. 571-581, 2011.

ABRIOUEL, H.; FRANZ, C. M.; BEN OMAR, N.; GALVEZ, A. Diversity and applications of *Bacillus* bacteriocins. **FEMS Microbiol Rev**, v. 35, p. 201-232, 2011.

AHERN, M.; VERSCHUEREN, S.; VAN SINDEREN, D. Isolation and characterisation of a novel bacteriocin produced by *Bacillus thuringiensis* strain B439. **FEMS Microbiol Lett**, v. 220, p. 127-131, 2003.

AKIBA, Y. Microbial ecology of *Bacillus thuringiensis*. **Appl Ent Zool**, v. 21, p. 76-80, 1986.

ALKAN, C.; SAJJADIAN, S.; EICHLER, E. E. Limitations of next-generation genome sequence assembly. **Nat Methods**, v. 8, p. 61-65, 2011.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. **J Mol Biol**, v. 215, p. 403-410, 1990.

ANDREOLLI, M.; LAMPIS, S.; POLI, M.; GULLNER, G.; BIRO, B.; VALLINI, G. Endophytic *Burkholderia fungorum* DBT1 can improve phytoremediation efficiency of polycyclic aromatic hydrocarbons. **Chemosphere**, v. 92, p. 688-694, 2013.

ANSARI, M. Z.; YADAV, G.; GOKHALE, R. S.; MOHANTY, D. NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. **Nucleic Acids Res**, v. 32, p. W405-413, 2004.

ARANTES, O. M. N.; VILAS-BÔAS, L. A.; VILAS-BÔAS, G. F. L. T. *Bacillus thuringiensis*: estratégias no controle biológico. In: SERAFINI, L. A.; BARROS, N. M.; AZEVEDO, J. L. (org). **Biotecnologia na agricultura e agroindústria**, Guaíba: Agropecuária, v. 2, p. 269-293, 2002.

ARGUELLES-ARIAS, A.; ONGENA, M.; HALIMI, B.; LARA, Y.; BRANS, A.; JORIS, B.; FICKERS, P. *Bacillus amyloliquefaciens* GA1 as a source of potent antibiotics and other secondary metabolites for biocontrol of plant pathogens. **Microb Cell Fact**, v. 8, p. 63, 2009.

ARONSON, A. I.; BECKMAN, W.; DUNN, P. *Bacillus thuringiensis* and related insect pathogens. **Microbiol Rev**, v. 50, p. 1-24, 1986.

BAKER, M. *De novo* genome assembly: what every biologist should know. **Nat Meth**, v. 9, p. 333-337, 2012.

BAKKER, P. A. H. M.; PIETERSE, C. M. J.; VAN LOON, L. C. Induced systemic resistance by fluorescent *Pseudomonas* spp. **Phytopathology**, v. 97, p. 239-243, 2007.

BANSODE, V. B.; BAJEKAL, S. S. Characterization of chitinases from microorganisms isolated from Lonar Lake. **Ind J Biotech**, v. 5, p. 357-363, 2006.

BARBOZA-CORONA, J. E.; VAZQUEZ-ACOSTA, H.; BIDESHI, D. K.; SALCEDO-HERNANDEZ, R. Bacteriocin-like inhibitor substances produced by Mexican strains of *Bacillus thuringiensis*. **Arch Microbiol**, v. 187, p. 117-126, 2007.

BARKA, E. A.; GOGNIES, S.; NOWAK, J.; AUDRAN, J. C.; BELARBI, A. Inhibitory effect of endophyte bacteria on *Botrytis cinerea* and its influence to promote the grapevine growth. **Biol Control**, v. 24, p. 135-142, 2002.

BAVYKIN, S. G.; LYSOV, Y. P.; ZAKHARIEV, V.; KELLY, J. J.; JACKMAN, J.; STAHL, D. A.; CHERNI, A. Use of 16S rRNA, 23S rRNA, and *gyrB* gene sequence analysis to determine phylogenetic relationships of *Bacillus cereus* group microorganisms. **J Clin Microbiol**, v. 42, p. 3711-3730, 2004.

BENDTSEN, J. D.; NIELSEN, H.; VON HEIJNE, G.; BRUNAK, S. Improved prediction of signal peptides: SignalP 3.0. **J Mol Biol**, v. 340, p. 783-795, 2004.

BERLINER, E. Eber die schlaffsucht der Mehlmotterraupe. (*Ephesis kuehniella* Zell.) undihren Erreger *Bacillus thuringiensis*, n. sp. **Z. ang. Entomology**, v. 2, p. 29-56, 1915.

BERLINER, E. Über die schlaffsucht der Mehlmotterraupe. **Z. Gesante Getreidewesen**, v. 3, p. 63-70, 1911.

BHATTACHRYA, D.; NAGPURE, A.; GUPTA, R. K. Bacterial chitinase: properties and potential. **Critic Rev Biotechnol**, v. 27, p. 21-28, 2007.

BLOM, J.; ALBAUM, S. P.; DOPPEMEIER, D.; PUHLER, A.; VORHOLTER, F. J.; ZAKRZEWSKI, M.; GOESMANN, A. EDGAR: a software framework for the comparative analysis of prokaryotic genomes. **BMC Bioinformatics**, v. 10, p. 154, 2009.

BODE, H.B. Entomopathogenic bacteria as a source of secondary metabolites. **Curr Opin Chem Biol**, v. 13, p. 2224–230, 2009.

BRADER, G.; COMPANT, S.; MITTER, B.; TROGNITZ, F.; SESSITSCH, A. Metabolic potential of endophytic bacteria. **Curr Opin Biotechnol**, v. 27, p. 30-37, 2014.

BRAR, S. K.; VERMA, M.; TYAGI, R. D.; VALÉRO, J. R. Recent advances in downstream processing and formulations of *Bacillus thuringiensis* based biopesticides. **Process Biochem**, v. 41, p. 323-342, 2006.

BRAVO, A.; SARABIA, S.; LOPEZ, L.; ONTIVEROS, H.; ABARCA, C.;

ORTIZ, A.; ORTIZ, M.; LAURA, L.; VILLALOBOS, F. J.; PEÑA, G.; NUNES-VALDEZ, M. E.; SOBERÓN, M.; QUINTERO, R. Characterization of *cry* genes in a Mexican *Bacillus thuringiensis* strain collection. **Appl Environ Microbiol**, v. 64, p. 4965-4972, 1998.

CÁRDENAS, M. I.; GALÁN-WONG, L.; FERRÉ-MANZANERO, J.; PEREYRA-ALFÉREZ, B. Selección de toxinas *cry* contra *Thichoplusia ni*. **Cienc Uanl**, v. 4, p. 51-62, 2001.

CARLSON, C. R.; CAUGANT, D. A.; KOLSTO, A. B. Genotypic diversity among *Bacillus cereus* and *Bacillus thuringiensis* strains. **Appl Environ Microbiol**, v. 60, p. 1719-1725, 1994.

CARLSON, C. R.; JOHANSEN, T.; KOLSTO, A. B. The chromosome map of *Bacillus thuringiensis* subsp. *canadensis* HD224 is highly similar to that of the *Bacillus cereus* type strain ATCC 14579. **FEMS Microbiol Lett**, v. 141, p. 163-167, 1996.

CASTRESANA J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. **Mol Biol Evol**, v. 17, p. 540-552, 2000.

CENDROWSKI, S.; MACARTHUR, W.; HANNA, P. *Bacillus anthracis* requires siderophore biosynthesis for growth in macrophages and mouse virulence. **Mol Microbiol**, v. 51, p. 407-417, 2004.

CHAISSON, M.; PEVZNER, P. Short read fragment assembly of bacterial genomes. **Genome Res**, v. 18, p. 324, 2008.

CHALLACOMBE, J. F.; ALTHERR, M. R.; XIE, G.; BHOTIKA, S. S.; BROWN, N.; BRUCE, D.; CAMPBELL, C. S.; CAMPBELL, M. L.; CHEN, J.; CHERTKOV, O.; CLELAND, C.; DIMITRIJEVIC, M.; DOGGETT, N. A.; FAWCETT, J. J.; GLAVINA, T.; GOODWIN, L. A.; GREEN, L. D.; HAN, C. S.; HILL, K. K.; HITCHCOCK, P.; JACKSON, P. J.; KEIM, P.; KEWALRAMANI, A. R.; LONGMIRE, J.; LUCAS, S.; MALFATTI, S.; MARTINEZ, D.; MCMURRY, K.; MEINCKE, L. J.; MISRA, M.; MOSEMAN, B. L.; MUNDT, M.; MUNK, A. C.; OKINAKA, R. T.; PARSON-QUINTANA, B.; REILLY, L. P.; RICHARDSON, P.; ROBINSON, D. L.; SAUNDERS, E.; TAPIA, R.; TESMER, J. G.; THAYER, N.; THOMPSON, L. S.; TICE, H.; TICKNOR, L. O.; WILLS, P. L.; GILNA, P.; BRETTIN, T. S. The complete genome sequence of *Bacillus thuringiensis* Al Hakam. **J Bacteriol**, v. 189, p. 3680-3681, 2007.

CHEHIMI, S.; DELALANDE, F.; SABLE, S.; HAJLAOUI, M. R.; VAN DORSSELAER, A.; LIMAM, F.; PONS, A. M. Purification and partial amino acid sequence of thuricin S, a new anti-*Listeria* bacteriocin from *Bacillus thuringiensis*. **Can J Microbiol**, v. 53, p. 284-290, 2007.

CHEN, X. H.; KOUMOUTSI, A.; SCHOLZ, R.; EISENREICH, A.; SCHNEIDER, K.; HEINEMEYER, I.; MORGENSTERN, B.; VOSS, B.; HESS, W. R.; REVA, O.; JUNGE, H.; VOIGT, B.; JUNGBLUT, P. R.; VATER, J.; SUSSMUTH, R.; LIESEGANG, H.; STRITTMATTER, A.; GOTTSCHALK, G.; BORRISS, R. Comparative analysis of the complete genome sequence of the

plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. **Nat Biotechnol**, v. 25, p. 1007-1014, 2007.

CHEN, X. H.; KOUMOUTSI, A.; SCHOLZ, R.; SCHNEIDER, K.; VATER, J.; SUSSMUTH, R.; PIEL, J.; BORRIS, R. Genome analysis of *Bacillus amyloliquefaciens* FZB42 reveals its potential for biocontrol of plant pathogens. **J Biotechnol**, v. 140, p. 27-37, 2009.

CHERIF, A.; CHEHIMI, S.; LIMEM, F.; HANSEN, B. M.; HENDRIKSEN, N. B.; DAFFONCHIO, D.; BOUDABOUS, A. Detection and characterization of the novel bacteriocin entomocin 9, and safety evaluation of its producer, *Bacillus thuringiensis* ssp. *entomocidus* HD9. **J Appl Microbiol**, v. 95, p. 990-1000, 2003.

CHERIF, A.; REZGUI, W.; RADDADI, N.; DAFFONCHIO, D.; BOUDABOUS, A. Characterization and partial purification of entomocin 110, a newly identified bacteriocin from *Bacillus thuringiensis* subsp. *entomocidus* HD110. **Microbiol Res**, v. 163, p. 684-692, 2008.

CHEVREUX, B.; PFISTERER, T.; DRESCHER, B.; DRIESEL, A. J.; MULLER, W. E.; WETTER, T.; SUHAI, S. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. **Genome Res**, v. 14, p. 1147-1159, 2004.

COOMBS, J. T.; MICHELSEN, P. P.; FRANCO, C. M. M. Evaluation of endophytic actinobacteria as antagonists of *Gaeumannomyces graminis* var. *tritici* in wheat. **Biol Control**, v. 29, p. 359-366, 2004.

COSTA, L. E. O.; QUEIROZ, M. V.; BORGES, A. C.; MORAES, C. A.; ARAÚJO, E. F. Isolation and characterization of endophytic bacteria isolated from the leaves of the common bean (*Phaseolus vulgaris*). **Braz J Microbiol**, v. 43, p. 1562-1575, 2012.

CRICKMORE, N.; BAUM, J.; BRAVO, A.; LERECLUS, D.; NARVA, K.; SAMPSON, K.; SCHNEPF, E.; SUN, M.; ZEIGLER, D. R. ***Bacillus thuringiensis* toxin nomenclature**. Disponível em: <http://www.btnomenclature.info/>. Acesso em: 20 de dezembro de 2014.

DAHIYA, N.; TEWARI, R.; HOONDAL, G. S. Biotechnological aspects of chitinolytic enzymes: a review. **Appl Microbiol Biotechnol**, v. 71, p. 773-782, 2006.

D'ALESSANDRO, M.; ERB, M.; TON, J.; BRANDENBURG, A.; KARLEN, D.; ZOPFI, J.; TURLINGS, T. C. Volatiles produced by soil-borne endophytic bacteria increase plant pathogen resistance and affect tritrophic interactions. **Plant Cell Environ**, v. 37, p. 813-826, 2014.

DARLING, A. C.; MAU, B.; BLATTNER, F. R.; PERNA, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. **Genome Res**, v. 14, p. 1394-1403, 2004.

DE LA FUENTE-SALCIDO, N.; GUADALUPE ALANIS-GUZMAN, M.; BIDESHI, D. K.; SALCEDO-HERNANDEZ, R.; BAUTISTA-JUSTO, M.; BARBOZA-CORONA, J. E. Enhanced synthesis and antimicrobial activities of bacteriocins produced by Mexican strains of *Bacillus thuringiensis*. **Arch Microbiol**, v. 190, p. 633-640, 2008.

DE LA VEGA, L. M.; BARBOZA-CORONA, J. E.; AGUILAR-USCANGA, M. G.; RAMIREZ-LEPE, M. Purification and characterization of an exochitinase from *Bacillus thuringiensis* subsp. *aizawai* and its action against phytopathogenic fungi. **Can J Microbiol**, v. 52, p. 651-657, 2006.

DE MAAGD, R. A.; BOSCH, D.; STIEKEMA, W. *Bacillus thuringiensis* toxin-mediated insect resistance in plants. **Trends Plant Sci**, v. 4, p. 9-13, 1999.

DELCHER, A. L.; HARMON, D.; KASIF, S.; WHITE, O.; SALZBERG, S. L. Improved microbial gene identification with GLIMMER. **Nucleic Acids Res**, v. 27, p. 4636-4641, 1999.

DIDELOT, X.; FALUSH, D. Inference of bacterial microevolution using multilocus sequence data. **Genetics**, v. 175, p. 1251-1266, 2007.

DOGGETT, N. A.; STUBBEN, C. J.; CHERTKOV, O.; BRUCE, D. C.; DETTER, J. C.; JOHNSON, S. L.; HAN, C. S. Complete genome sequence of *Bacillus thuringiensis* serovar *israelensis* strain HD-789. **Genome Announc**, v. 1, p. 1-2, 2013.

DOHM, J. C.; LOTTAZ, C.; BORODINA, T.; HIMMELBAUER, H. SHARCGS, a fast and highly accurate short-read assembly algorithm for de novo genomic sequencing. **Genome Res**, v. 17, p. 1697-1706, 2007.

DONG, Y. H.; ZHANG, X. F.; XU, J. L.; ZHANG, L. H. Insecticidal *Bacillus thuringiensis* silences *Erwinia carotovora* virulence by a new form of microbial antagonism, signal interference. **Appl Environ Microbiol**, v. 70, p. 954-960, 2004.

DOTY, S. L. Enhancing phytoremediation through the use of transgenics and endophytes. **New Phytol**, v. 179, p. 318-333, 2008.

DWORZANSKI, J. P.; DICKINSON, D. N.; DESHPANDE, S. V.; SNYDER, A. P.; ECKENRODE, B. A. Discrimination and phylogenomic classification of *Bacillus anthracis-cereus-thuringiensis* strains based on LC-MS/MS analysis of whole cell protein digests. **Anal Chem**, v. 82, p. 145-155, 2010.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Res**, v. 32, p. 1792-1797, 2004.

EICHLER, E. E.; SANKOFF, D. Structural dynamics of eukaryotic chromosome evolution. **Science**, v. 301, p. 793-797, 2003.

EKBLOM, R.; WOLF, J. B. A field guide to whole-genome sequencing, assembly and annotation. **Evol Appl**, v. 7, p. 1026-1042, 2014.

FANG, Y.; LI, Z.; LIU, J.; SHU, C.; WANG, X.; ZHANG, X.; YU, X.; ZHAO, D.; LIU, G.; HU, S.; ZHANG, J.; AL-MSSALLEM, I.; YU, J. A pangenomic study of *Bacillus thuringiensis*. **J Genet Genomics**, v. 38, p. 567-576, 2011.

FEDERICI, B. A. *Bacillus thuringiensis* in Biological Control. In: BELLOWS, T.; FISHER, T. W. (eds). **Handbook of Biological Control**. Califórnia: Universidade da Califórnia, p. 575-593, 1999.

FEITELSON, J. S. Novel pesticidal delta-endotoxins from *Bacillus thuringiensis*. In: **Proceedings of XXVII Annual Meeting of the Society for Invertebrate Pathology**, Montpellier, France, p. 184, 1994.

FELNAGLE, E. A.; JACKSON, E. E.; CHAN, Y. A.; PODEVELS, A. M.; BERTI, A. D.; MCMAHON, M. D.; THOMAS, M. G. Nonribosomal peptide synthetases involved in the production of medically relevant natural products. **Mol Pharm**, v. 5, p. 191-211, 2008.

FELSENSTEIN, J. PHYLIP: Phylogeny Inference Package (version3.2). **Cladistics**, v. 5, p. 164–166, 1989.

FERREIRA, L. H. P. L.; SUZUKI, M. T.; ITANO, E. N.; ONO, M. A.; ARANTES, O. M. N. Ecological aspects of *Bacillus thuringiensis* in an Oxisol. **Sci Agric**, v. 60, p.19-22, 2003.

FICKERS, P. Antibiotic compounds from *Bacillus*: Why are they so amazing? **Am J Biochem Biotechnol**, v. 8, p. 38-43, 2012.

FLEISCHMANN, R. D.; ADAMS, M. D.; WHITE, O.; CLAYTON, R. A.; KIRKNESS, E. F.; KERLAVAGE, A. R.; BULT, C. J.; TOMB, J. F.; DOUGHERTY, B. A.; MERRICK, J. M.; ET AL. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. **Science**, v. 269, p. 496-512, 1995.

FRASER, C. M.; GOCAYNE, J. D.; WHITE, O.; ADAMS, M. D.; CLAYTON, R. A.; FLEISCHMANN, R. D.; BULT, C. J.; KERLAVAGE, A. R.; SUTTON, G.; KELLEY, J. M.; FRITCHMAN, R. D.; WEIDMAN, J. F.; SMALL, K. V.; SANDUSKY, M.; FUHRMANN, J.; NGUYEN, D.; UTTERBACK, T. R.; SAUDEK, D. M.; PHILLIPS, C. A.; MERRICK, J. M.; TOMB, J. F.; DOUGHERTY, B. A.; BOTT, K. F.; HU, P. C.; LUCIER, T. S.; PETERSON, S. N.; SMITH, H. O.; HUTCHISON, C. A., 3RD; VENTER, J. C. The minimal gene complement of *Mycoplasma genitalium*. **Science**, v. 270, p. 397-403, 1995.

FROMMEL, M. I.; NOWAK, J.; LAZAROVITS, G. Growth enhancement and developmental modifications of *in vitro* grown potato (*Solanum tuberosum* spp. *tuberosum*) as affected by a nonfluorescent *Pseudomonas* sp. **Plant Physiol**, v. 96, p. 928-936, 1991.

FROST, L. S.; LEPLAE, R.; SUMMERS, A. O.; TOUSSAINT, A. Mobile genetic elements: the agents of open source evolution. **Nat Rev Microbiol**, v. 3, p. 722-732, 2005.

FUQUA, C.; PARSEK, M. R.; GREENBERG, E. P. Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. **Annu Rev Genet**, v. 35, p. 439-468, 2001.

GALLEGOS, M. T.; SCHLEIF, R.; BAIROCH, A.; HOFMANN, K.; RAMOS, J. L. Arac/XylS family of transcriptional regulators. **Microbiol Mol Biol Rev**, v. 61, p. 393-410, 1997.

GARDNER, J. M.; FELDMAN, A. W.; ZABLOTOWICZ, R. M. Identity and behavior of xylem-residing bacteria in rough lemon roots of florida citrus trees. **Appl Environ Microbiol**, v. 43, p. 1335-1342, 1982.

GELERNTER, W.; SCHWAB, G. E. Transgenic bacteria, viruses, algae and other microorganisms as *Bacillus thuringiensis* toxin delivery systems. In: ENTWISTLE, P. F.; CORY, J. S.; BAILEY, M. J.; HIGGS, S. (ed). **Bacillus thuringiensis, an environmental biopesticide: theory and practice**, Chichester: John Wiley & Sons, p. 89-104, 1993.

GISH, W.; STATES, D. J. Identification of protein coding regions by database similarity search. **Nat Genet**, v. 3, p. 266-272, 1993.

GLARE, T. R.; O'CALLAGHAN, M. *Bacillus thuringiensis*: biology, ecology and safety. In: ENTWISTLE, P. F.; CORY, J. S.; BAILEY, M. J.; HIGGS, S. (ed). **Bacillus thuringiensis, an environmental biopesticide: theory and practice**, Chichester: John Wiley, p. 350, 2000.

GOMAA, E. Z. Chitinase Production by *Bacillus thuringiensis* and *Bacillus licheniformis*: their potential in antifungal biocontrol. **J Microbiol**, v. 50, p. 103-111, 2012.

GONZALEZ, J. M., JR.; BROWN, B. J.; CARLTON, B. C. Transfer of *Bacillus thuringiensis* plasmids coding for delta-endotoxin among strains of *B. thuringiensis* and *B. cereus*. **Proc Natl Acad Sci USA**, v. 79, p. 6951-6955, 1982.

GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: a graphical tool for sequence finishing. **Genome Res**, v. 8, p. 195-202, 1998.

GRIFFITHS-JONES, S.; MOXON, S.; MARSHALL, M.; KHANNA, A.; EDDY, S. R.; BATEMAN, A. Rfam: annotating non-coding RNAs in complete genomes. **Nucleic Acids Res**, v. 33, p. D121-124, 2005.

GUAN, P.; AI, P.; DAI, X.; ZHANG, J.; XU, L.; ZHU, J.; LI, Q.; DENG, Q.; LI, S.; WANG, S.; LIU, H.; WANG, L.; LI, P.; ZHENG, A. Complete genome sequence of *Bacillus thuringiensis* serovar *sichuansis* strain MC28. **J Bacteriol**, v. 194, p. 6975, 2012.

GUNDOGDU, O.; BENTLEY, S. D.; HOLDEN, M. T.; PARKHILL, J.; DORRELL, N.; WREN, B. W. Re-annotation and re-analysis of the *Campylobacter jejuni* NCTC11168 genome sequence. **BMC Genomics**, v. 8, p. 162, 2007.

GUSTAFSON, K. R.; CARDELLINA, J. H., 2ND; FULLER, R. W.; WEISLOW, O. S.; KISER, R. F.; SNADER, K. M.; PATTERSON, G. M.; BOYD, M. R. AIDS-antiviral sulfolipids from cyanobacteria (blue-green algae). **J Natl Cancer Inst**, v. 81, p. 1254-1258, 1989.

HAACK, B. J.; ANDREWS, R. E.; LOYNACHAN, T. E. Tn916-mediated genetic exchange in soil. **Soil Biol Biochem**, v. 28, p. 765-771, 1996.

HALLMANN, J.; QUADT-HALLMANN, A.; MAHAFFEE, W. F.; KLOEPPER, J. Bacterial endophytes in agricultural crops. **Can J Microbiol**, v. 43, p. 895-914, 1997.

HALLMANN, J.; QUADT-HALLMANN, A.; RODRIGUEZ-KABANA, R.; KLOEPPER, J. W., Interactions between *Meloidogyne incognita* and endophytic bacteria in cotton and cucumber. **Soil Biol Biochem**, v. 30, p. 925-937, 1998.

HAN, C. S.; XIE, G.; CHALLACOMBE, J. F.; ALTHERR, M. R.; BHOTIKA, S. S.; BROWN, N.; BRUCE, D.; CAMPBELL, C. S.; CAMPBELL, M. L.; CHEN, J.; CHERTKOV, O.; CLELAND, C.; DIMITRIJEVIC, M.; DOGGETT, N. A.; FAWCETT, J. J.; GLAVINA, T.; GOODWIN, L. A.; GREEN, L. D.; HILL, K. K.; HITCHCOCK, P.; JACKSON, P. J.; KEIM, P.; KEWALRAMANI, A. R.; LONGMIRE, J.; LUCAS, S.; Malfatti, S.; MCMURRY, K.; MEINCKE, L. J.; MISRA, M.; MOSEMAN, B. L.; MUNDT, M.; MUNK, A. C.; OKINAKA, R. T.; PARSON-QUINTANA, B.; REILLY, L. P.; RICHARDSON, P.; ROBINSON, D. L.; RUBIN, E.; SAUNDERS, E.; TAPIA, R.; TESMER, J. G.; THAYER, N.; THOMPSON, L. S.; TICE, H.; TICKNOR, L. O.; WILLS, P. L.; BRETTIN, T. S.; GILNA, P. Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. **J Bacteriol**, v. 188, p. 3382-3390, 2006.

HANNAY, C. L.; FITZ-JAMES, P. C. The protein-crystals of *Bacillus thuringiensis berliner*. **Can J Microbiol**, v. 1, p. 694-710, 1995.

HANSON, A. D.; PRIBAT, A.; WALLER, J. C.; DE CRECY-LAGARD, V. 'Unknown' proteins and 'orphan' enzymes: the missing half of the engineering parts list-and how to find it. **Biochem J**, v. 425, p. 1-11, 2010.

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

HAYES, M.; CARNEY, B.; SLATER, J.; BRUCK, W. Mining marine shellfish wastes for bioactive molecules: chitin and chitosan - Part A: extraction methods. **Biotechnol J**, v. 3, p. 871-877, 2008.

HE, J.; SHAO, X.; ZHENG, H.; LI, M.; WANG, J.; ZHANG, Q.; LI, L.; LIU, Z.; SUN, M.; WANG, S.; YU, Z. Complete genome sequence of *Bacillus thuringiensis* mutant strain BMB171. **J Bacteriol**, v. 192, p. 4074-4075, 2010.

HE, J.; WANG, J.; YIN, W.; SHAO, X.; ZHENG, H.; LI, M.; ZHAO, Y.; SUN, M.; WANG, S.; YU, Z. Complete genome sequence of *Bacillus thuringiensis* subsp. *chinensis* strain CT-43. **J Bacteriol**, v. 193, p. 3407-3408, 2011.

HELGASON, E.; OKSTAD, O. A.; CAUGANT, D. A.; JOHANSEN, H. A.; FOUET, A.; MOCK, M.; HEGNA, I.; KOLSTO, A. B. *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* - one species on the basis of genetic evidence. **Appl Environ Microbiol**, v. 66, p. 2627-2630, 2000.

HELGASON, E.; TOURASSE, N. J.; MEISAL, R.; CAUGANT, D. A.; KOLSTO, A. B. Multilocus sequence typing scheme for bacteria of the *Bacillus cereus* group. **Appl Environ Microbiol**, v. 70, p. 191-201, 2004.

HENDRIKSEN, N. B.; HANSEN, B. M. Long-term survival and germination of *Bacillus thuringiensis* var. *kurstaki* in a field trial. **Can J Microbiol**, v. 48, p. 256-261, 2002.

HERTWECK, C.; LUZHETSKYY, A.; REBETS, Y.; BECHTHOLD, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. **Nat Prod Rep**, v. 24, p. 162-190, 2007.

HUANG, J. Ultrastructure of bacterial penetration in plants. **Annu Rev Phytopathol**, v. 24, p. 141-157, 1986.

HUNT, M.; NEWBOLD, C.; BERRIMAN, M.; OTTO, T. D. A comprehensive evaluation of assembly scaffolding tools. **Genome Biol**, v. 15, p. R42, 2014.

HUNTER, S.; APWEILER, R.; ATTWOOD, T. K.; BAIROCH, A.; BATEMAN, A.; BINNS, D.; BORK, P.; DAS, U.; DAUGHERTY, L.; DUQUENNE, L.; FINN, R. D.; GOUGH, J.; HAFT, D.; HULO, N.; KAHN, D.; KELLY, E.; LAUGRAUD, A.; LETUNIC, I.; LONSDALE, D.; LOPEZ, R.; MADERA, M.; MASLEN, J.; MCANULLA, C.; MCDOWALL, J.; MISTRY, J.; MITCHELL, A.; MULDER, N.; NATALE, D.; ORENGO, C.; QUINN, A. F.; SELENGUT, J. D.; SIGRIST, C. J.; THIMMA, M.; THOMAS, P. D.; VALENTIN, F.; WILSON, D.; WU, C. H.; YEATS, C. InterPro: the integrative protein signature database. **Nucleic Acids Res**, v. 37, p. D211-215, 2009.

HUREK, T.; REINHOLD-HUREK, B.; VAN MONTAGU, M.; KELLENBERGER, E. Root colonization and systemic spreading of *Azoarcus* sp. strain BH72 in grasses. **J Bacteriol**, v. 176, 1913-1923, 1994.

HUSZ, B. On the use of *Bacillus thuringiensis* in the fight against the corn borer. **Instit Corn Borer Invest Sci Repub**, v.2, p. 99-110, 1929.

IBRAHIM, M. A.; GRIKO, N.; JUNKER, M.; BULLA, L. A. *Bacillus thuringiensis*: a genomics and proteomics perspective. **Bioeng Bugs**, v. 1, p. 31-50, 2010.

ISHIWATA, S. On a kind of severe flacherie (sotto disease). **Danihon Sanshi Kaiho**, v. 114, p. 1-5, 1901.

JENKE-KODAMA, H.; DITTMANN, E. Evolution of metabolic diversity: insights from microbial polyketide synthases. **Phytochemistry**, v. 70, p. 1858-1866, 2009.

JEONG, H.; PARK, S. H.; CHOI, S. K. Genome sequence of the acrySTALLIFEROUS *Bacillus thuringiensis* serovar *israelensis* strain 4Q7, widely used as a recombination host. **Genome Announc**, v. 2, p. e00231-14, 2014.

JHA, P. N.; KUMAR, A. Endophytic colonization of *Typha australis* by a plant growth-promoting bacterium *Klebsiella oxytoca* strain GR-3. **J Appl Microbiol**, v. 103, p. 1311-1320, 2007.

JONES, P.; BINNS, D.; CHANG, H. Y.; FRASER, M.; LI, W.; MCANULLA, C.; MCWILLIAM, H.; MASLEN, J.; MITCHELL, A.; NUKA, G.; PESSEAT, S.; QUINN, A. F.; SANGRADOR-VEGAS, A.; SCHEREMETJEW, M.; YONG, S. Y.; LOPEZ, R.; HUNTER, S. InterProScan 5: genome-scale protein function classification. **Bioinformatics**, v. 30, p. 1236-1240, 2014.

JOSEPH, S.; DESAI, P.; JI, Y.; CUMMINGS, C. A.; SHIH, R.; DEGORICIJA, L.; RICO, A.; BRZOSKA, P.; HAMBY, S. E.; MASOOD, N.; HARIRI, S.; SONBOL, H.; CHUZHANOVA, N.; MCCLELLAND, M.; FURTADO, M. R.; FORSYTHE, S. J. Comparative analysis of genome sequences covering the seven *Cronobacter* species. **Plos One**, v. 7, p. e49455, 2012.

KIM, W.; HONG, Y. P.; YOO, J. H.; LEE, W. B.; CHOI, C. S.; CHUNG, S. I. Genetic relationships of *Bacillus anthracis* and closely related species based on variable-number tandem repeat analysis and BOX-PCR genomic *fingerprinting*. **FEMS Microbiol Lett**, v. 207, p. 21-27, 2002.

KLOEPPER, J. W.; RYU, C. M.; ZHANG, S. Induced systemic resistance and promotion of plant growth by *Bacillus* spp. **Phytopathology**, v. 94, p. 1259-1266, 2004.

KOBAYASHI, T.; SUZUKI, M.; INOUE, H.; ITAI, R. N.; TAKAHASHI, M.; NAKANISHI, H.; MORI, S.; NISHIZAWA, N. K. Expression of iron-acquisition-related genes in iron-deficient rice is co-ordinately induced by partially conserved iron-deficiency-responsive elements. **J Exp Bot**, v. 56, p. 1305-1316, 2005.

KOONIN, E. V. Computational genomics. **Curr Biol**, v. 11, p. R155-158, 2001.

KOTZE, A. C.; O'GRADY, J.; GOUGH, J. M.; PEARSON, R.; BAGNALL, N. H.; KEMP, D. H.; AKHURST, R. J. Toxicity of *Bacillus thuringiensis* to parasitic and free-living life-stages of nematode parasites of livestock. **Int J Parasitol**, v. 35, p. 1013-1022, 2005.

KUNST, F.; OGASAWARA, N.; MOSZER, I.; ALBERTINI, A. M.; ALLONI, G.; AZEVEDO, V.; et al. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. **Nature**, v. 390, p. 249-256, 1997.

KUZU, S. B.; GUVENMEZ, H. K.; DENIZCI, A. A. Production of a thermostable and alkaline chitinase by *Bacillus thuringiensis* subsp. *kurstaki* strain HBK-51. **Biotechnol Res Int**, v. 2012, p. 135498, 2012.

KYRPIDES, N. C. Fifteen years of microbial genomics: meeting the challenges and fulfilling the dream. **Nat Biotechnol**, v. 27, p. 627-632, 2009.

LAGESEN, K.; HALLIN, P.; RODLAND, E. A.; STAERFELDT, H. H.; ROGNES, T.; USSERY, D. W. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. **Nucleic Acids Res**, v. 35, p. 3100-3108, 2007.

LAING, C.; BUCHANAN, C.; TABOADA, E. N.; ZHANG, Y.; KROPINSKI, A.; VILLEGAS, A.; THOMAS, J. E.; GANNON, V. P. Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions. **BMC Bioinformatics**, v. 11, p. 461, 2010.

LAMB, T. G.; TONKYN, D. W.; KLUEPFEL, D. A. Movement of *Pseudomonas aureofaciens* from the rhizosphere to aerial plant tissue. **Can J Microbiol**, v. 42, p. 1112-1120, 1996.

LASLETT, D.; CANBACK, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. **Nucleic Acids Res**, v. 32, p. 11-16, 2004.

LEE, H.; CHUREY, J. J.; WOROBO, R. W. Biosynthesis and transcriptional analysis of thurincin H, a tandem repeated bacteriocin genetic locus, produced by *Bacillus thuringiensis* SF361. **FEMS Microbiol Lett**, v. 299, p. 205-213, 2009a.

LEE, K. D.; GRAY, E. J.; MABOOD, F.; JUNG, W. J.; CHARLES, T.; CLARK, S. R.; LY, A.; SOULEIMANOV, A.; ZHOU, X.; SMITH, D. L. The class IId bacteriocin thuricin-17 increases plant growth. **Planta**, v. 229, p. 747-755, 2009b.

LEE, S. J.; PARK, S. Y.; LEE, J. J.; YUM, D. Y.; KOO, B. T.; LEE, J. K. Genes encoding the N-acyl homoserine lactone-degrading enzyme are widespread in many subspecies of *Bacillus thuringiensis*. **Appl Environ Microbiol**, v. 68, p. 3919-3924, 2002.

LIBRADO, P.; ROZAS, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, p. 1451-1452, 2009.

LIU, G.; SONG, L.; SHU, C.; WANG, P.; DENG, C.; PENG, Q.; LERECLUS, D.; WANG, X.; HUANG, D.; ZHANG, J.; SONG, F. Complete genome sequence of *Bacillus thuringiensis* subsp. *kurstaki* strain HD73. **Genome Announc**, v. 1, p. e0008013, 2013.

LIU, M.; CAI, Q. X.; LIU, H. Z.; ZHANG, B. H.; YAN, J. P.; YUAN, Z. M. Chitinolytic activities in *Bacillus thuringiensis* and their synergistic effects on larvicidal activity. **J Appl Microbiol**, v. 93, p. 374-379, 2002.

LIU, X.; ZHOU, R.; FU, G.; ZHANG, W.; MIN, Y.; TIAN, Y.; HUANG, D.; WANG, K.; WAN, Z.; YAO, J.; YANG, Z. Draft genome sequence of *Bacillus*

thuringiensis NBIN-866 with high nematocidal activity. **Genome Announc**, v. 2, p.1-2, 2014.

LOPES, R. B. M. **Potencial de inibição de quorum sensing e análise dos perfis de macro-restrição de bactérias endofíticas das folhas do feijoeiro comum (*Phaseolus vulgaris*)**. 55 p. Monografia (Graduação em Ciências Biológicas). Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa. 2013.

LÓPEZ-LÓPEZ, A.; ROGEL, M. A.; ORMEÑO-ORRILLO, E.; MARTÍNEZ-ROMERO, J.; MARTÍNEZ-ROMERO, E. *Phaseolus vulgaris* seed-borne endophytic community with novel bacterial species such as *Rhizobium endophyticum* sp. nov. **Syst Appl Microbiol**, v. 33, p. 322-327, 2010.

LOWE, T. M.; EDDY, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. **Nucleic Acids Res**, v. 25, p. 955-964, 1997.

LUKASHIN, A. V.; BORODOVSKY, M. GeneMark.hmm: new solutions for gene finding. **Nucleic Acids Res**, v. 26, p. 1107-1115, 1998.

LUO, C.; HU, G.; ZHU, H. Genome reannotation of *Escherichia coli* CFT073 with new insights into virulence. **BMC Genomics**, v. 10, p. 552, 2009.

MACLEAN, D.; JONES, J. D.; STUDHOLME, D. J. Application of 'next-generation' sequencing technologies to microbial genetics. **Nat Rev Microbiol**, v. 7, p. 287-296, 2009.

MADUPU, R.; BRINKAC, L. M.; HARROW, J.; WILMING, L. G.; BOHME, U.; LAMESCH, P.; HANNICK, L. I. Meeting report: a workshop on best practices in genome annotation. **Database (Oxford)**, v. 2010, p. baq001, 2010.

MARCHLER-BAUER, A.; LU, S.; ANDERSON, J. B.; CHITSAZ, F.; DERBYSHIRE, M. K.; DEWEESE-SCOTT, C.; FONG, J. H.; et al. CDD: a Conserved Domain Database for the functional annotation of proteins. **Nucleic Acids Res**, v. 39, p. D225-229, 2011.

MARCO, G.; MANUEL, P. Ecological mysteries: Is *Bacillus thuringiensis* a real insect pathogen? **Bt Res.**, v.3, p. 1-2, 2012.

MARCON, J.; TAKETANI, R. G.; DINI-ANDREOTE, F.; MAZZERO, G. I.; SOARES, F. L. J.; MELO, I. S.; AZEVEDO, J. L.; ANDREOTE, F. D. Draft genome sequence of *Bacillus thuringiensis* strain BrMgv02-JM63, a chitinolytic bacterium isolated from oil-contaminated mangrove soil in Brazil. **Genome Announc**, v. 2, p. 2014.

MARDIS, E. R. The impact of next-generation sequencing technology on genetics. **Trends Genet**, v. 24, p. 133-141, 2008.

MARTIN, R. G.; ROSNER, J. L. The AraC transcriptional activators. **Curr Opin Microbiol**, v. 4, p. 132-137, 2001.

MARTÍNEZ, L.; CABALLERO-MELLADO, J.; OROZCO, J.; MARTÍNEZ-ROMERO, E. Diazotrophic bacteria associated with banana (*Musa* spp.). **Plant Soil**, v. 257, p. 35-47, 2003.

MARTÍNEZ-ROMERO, E. Coevolution in *Rhizobium*-legume symbiosis? **DNA Cell Biol.**, v. 28, p. 361-370, 2009.

MARTÍNEZ-ROMERO, E. Diversity of *Rhizobium-Phaseolus vulgaris* symbiosis: overview and perspectives. **Plant Soil**, v. 252, p. 11-23, 2003.

MCINROY, J. A.; KLOEPPER, J. W. Population dynamics of endophytic bacteria in field-grown sweet corn and cotton. **Can J Microbiol**, v. 41, p. 895-901, 1995.

MEADOWS, M. P. *Bacillus thuringiensis* in the environment: ecology and risk assessment. In: ENTWISTLE, P. F.; CORY, J. S.; BAILEY, M. J.; HIGGS, S. (ed). **Bacillus thuringiensis an environment biopesticide: theory and practice**, Chichester: John Wiley & Sons, p. 193-220, 1993.

MÉDIGUE, C.; MOSZER, I. Annotation, comparison and databases for hundreds of bacterial genomes. **Res Microbiol**, v. 158, p. 724-736, 2007.

MEDINI, D.; DONATI, C.; TETTELIN, H.; MASIGNANI, V.; RAPPUOLI, R. The microbial pan-genome. **Curr Opin Genet Dev**, v. 15, p. 589-594, 2005.

MEURER, G.; GERLITZ, M.; WENDT-PIENKOWSKI, E.; VINING, L. C.; ROHR, J.; HUTCHINSON, C. R. Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. **Chem Biol**, v. 4, p. 433-443, 1997.

MISAGHI, I. J.; DONNDELINGER, C. R. Endophytic bacteria in symptom-free cotton plants. **Phytopathology**, v. 80, p. 808-811, 1990.

MISHRA, P. K.; MISHRA, S.; SELVAKUMAR, G.; BISHT, J. K.; KUNDU, S.; GUPTA, H.S. Coinoculation of *Bacillus thuringiensis*-KR1 with *Rhizobium leguminosarum* enhances plant growth and nodulation of pea (*Pisum sativum* L.) and lentil (*Lens culinaris* L.). **World J Microbiol Biotechnol**, v. 25, p. 753-761, 2009.

MISHRA, R.; SINGH, R.; JAISWAL, H.; KUMAR, V.; MAURYA, S. *Rhizobium*-mediated induction of phenolics and plant growth promotion in rice (*Oryza sativa* L.). **Curr Microbiol**, v. 52, p. 383-389, 2006.

MONNERAT, R. G.; SANTOS, R.; BARROS, P.; BATISTA, A.; BERRY, C. **Isolamento e caracterização de estirpes de *Bacillus thuringiensis* endofíticas de algodão**, Brasília: Embrapa Recursos Genéticos e Biotecnologia, 4 p., 2003.

MURAWSKA, E.; FIEDORUK, K.; BIDESHI, D. K.; SWIECICKA, I. Complete genome sequence of *Bacillus thuringiensis* subsp. *thuringiensis* strain IS5056, an isolate highly toxic to *Trichoplusia ni*. **Genome Announc**, v. 1, p. e0010813, 2013.

NAVON, A. *Bacillus thuringiensis* application in agriculture. In: CHARLES, J. F.; DELÉCLUSE, A.; NILESEN-LE ROUX, C. (ed.). **Entomopathogenic bacteria: from laboratory to field application**, Netherlands: Kluwer academic Publishers, p. 355-367, 2000.

OVERBEEK, R.; BEGLEY, T.; BUTLER, R. M.; CHOUDHURI, J. V.; CHUANG, H. Y.; COHOON, M.; et al.. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. **Nucleic Acids Res**, v. 33, p. 5691-5702, 2005.

PAIK, H. D.; BAE, S. S.; PARK, S. H.; PAN, J. G. Identification and partial characterization of tochicin, a bacteriocin produced by *Bacillus thuringiensis* subsp. *tochigiensis*. **J Ind Microbiol Biotechnol**, v. 19, p. 294-298, 1997.

PALMA, L.; MUNOZ, D.; BERRY, C.; MURILLO, J.; CABALLERO, P. Draft genome sequences of two *Bacillus thuringiensis* strains and characterization of a putative 41.9-kDa insecticidal toxin. **Toxins (Basel)**, v. 6, p. 1490-1504, 2014.

PARK, R. Y.; CHOI, M. H.; SUN, H. Y.; SHIN, S. H. Production of catechol-siderophore and utilization of transferrin-bound iron in *Bacillus cereus*. **Biol Pharm Bull**, v. 28, p. 1132-1135, 2005.

PATEL, P. S.; HUANG, S.; FISHER, S.; PIRNIK, D.; AKLONIS, C.; DEAN, L.; MEYERS, E.; FERNANDES, P.; MAYERL, F. Bacillaene, a novel inhibitor of prokaryotic protein synthesis produced by *Bacillus subtilis*: production, taxonomy, isolation, physico-chemical characterization and biological activity. **J Antibiot (Tokyo)**, v. 48, p. 997-1003, 1995.

PAULSEN, I. T.; PRESS, C. M.; RAVEL, J.; KOBAYASHI, D. Y.; MYERS, G. S.; MAVRODI, D. V.; DEBOY, R. T.; SESHADRI, R.; REN, Q.; MADUPU, R.; DODSON, R. J.; DURKIN, A. S.; BRINKAC, L. M.; DAUGHERTY, S. C.; SULLIVAN, S. A.; ROSOVITZ, M. J.; GWINN, M. L.; ZHOU, L.; SCHNEIDER, D. J.; CARTINHO, S. W.; NELSON, W. C.; WEIDMAN, J.; WATKINS, K.; TRAN, K.; KHOURI, H.; PIERSON, E. A.; PIERSON, L. S., 3RD; THOMASHOW, L. S.; LOPER, J. E. Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. **Nat Biotechnol**, v. 23, p. 873-878, 2005.

POLANCZYK, R.; ALVES, S. *Bacillus thuringiensis*: uma breve revisão. **Agrociência**, v. 8, p. 1-10, 2003.

PRAÇA, L. B.; GOMES, A. C. M. M.; CABRAL, G.; MARTINS, E. S.; SUJII, E. R.; MONNERAT, R. G. Endophytic colonization by Brazilian strains of *Bacillus thuringiensis* on cabbage seedlings grown *in vitro*. **Bt Res**, v. 3, p. 11-19, 2012.

PRICE, A. L.; JONES, N. C.; PEVZNER, P. A. De novo identification of repeat families in large genomes. **Bioinformatics**, v. 21, p. i351-358, 2005.

PRIEST, F. G.; BARKER, M.; BAILLIE, L. W.; HOLMES, E. C.; MAIDEN, M. C. Population structure and evolution of the *Bacillus cereus* group. **J Bacteriol**, v. 186, p. 7959-7970, 2004.

QUADT-HALLMANN, A.; KLOEPPER, J. W. Immunological detection and localization of the cotton endophyte *Enterobacter asburiae* JM22 in different plant species. **Can J Microbiol**, v. 42, p. 1144-1154, 1996.

QUIÑONES, B.; DULLA, G.; LINDOW, S. E. Quorum sensing regulates exopolysaccharide production, motility and virulence in *Pseudomonas syringae*. **Mol Plant Microbe Interact**, v 18, p. 682-693, 2005.

RADDADI, N.; CHERIF, A.; OUZARI, H.; MARZORATI, M.; BRUSETTI, L.; BOUDABOUS, A.; DAFFONCHIO, D. *Bacillus thuringiensis* beyond insect biocontrol: plant growth promotion and biosafety of polyvalent strains, **Ann Microbiol**, v. 57, p. 481-494, 2007.

RASKO, D. A.; ROSOVITZ, M. J.; MYERS, G. S.; MONGODIN, E. F.; FRICKE, W. F.; GAJER, P.; CRABTREE, J.; SEBAIHIA, M.; THOMSON, N. R.; CHAUDHURI, R.; HENDERSON, I. R.; SPERANDIO, V.; RAVEL, J. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. **J Bacteriol**, v. 190, p. 6881-6893, 2008.

READ, T. D.; AKMAL, A.; BISHOP-LILLY, K. **Annotation of the *Bacillus thuringiensis* BGSC4Y1 genome.** Available at: http://www.ncbi.nlm.nih.gov/nuccore/NZ_CM000746.1, 2009.

REESE, M. G. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. **Comput Chem**, v. 26, p. 51-56, 2001.

REEVES, G. A.; TALAVERA, D.; THORNTON, J. M. Genome and proteome annotation: organization, interpretation and integration. **J R Soc Interface**, v. 6, p. 129-147, 2009.

RISSMAN, A. I.; MAU, B.; BIEHL, B. S.; DARLING, A. E.; GLASNER, J. D.; PERNA, N. T. Reordering contigs of draft genomes using the Mauve aligner. **Bioinformatics**, v. 25, p. 2071-2073, 2009.

ROH, J. Y.; CHOI, J. Y.; LI, M. S.; JIN, B. R.; JE, Y. H. *Bacillus thuringiensis* as a specific, safe, and effective tool for insect pest control. **J Microbiol Biotechnol**, v. 17, p. 547-559, 2007.

ROJAS-AVELIZAPA, L. I.; CRUZ-CAMARILLO, R.; GUERRERO, M. I.; RODRIGUEZ-VAZQUEZ, R.; IBARRA, J. E. Selection and characterization of a proteo-chitinolytic strain of *Bacillus thuringiensis*, able to grow in waste media. **World J Microbiol Biotechnol**, v. 15, p. 261-268, 1999.

ROULI, L.; M, M. B.; ROBERT, C.; NDIAYE, M.; LA SCOLA, B.; RAOULT, D. Genomic analysis of three African strains of *Bacillus anthracis* demonstrates that they are part of the clonal expansion of an exclusively pathogenic bacterium. **New Microbes New Infect**, v. 2, p. 161-169, 2014.

ROSEN BLUETH, M.; MARTINEZ-ROMERO, E. Bacterial endophytes and their interactions with hosts. **Mol Plant Microbe Interact**, v. 19, p. 827-837, 2006.

RUDE, M. A.; KHOSLA, C. Engineered biosynthesis of polyketides in heterologous hosts. **Chem Eng Sci**, v. 59, p. 4693–4701, 2004.

RUST, A. G.; MONGIN, E.; BIRNEY, E. Genome annotation techniques: new approaches and challenges. **Drug Discov Today**, v. 7, p. S70-76, 2002.

RUTHERFORD, K.; PARKHILL, J.; CROOK, J.; HORSNELL, T.; RICE, P.; RAJANDREAM, M. A.; BARRELL, B. Artemis: sequence visualization and annotation. **Bioinformatics**, v. 16, p. 944-945, 2000.

RYAN, R. P.; GERMAINE, K.; FRANKS, A.; RYAN, D. J.; DOWLING, D. N. Bacterial endophytes: recent developments and applications. **FEMS Microbiol Lett**, v. 278, p. 1-9, 2008.

SADFI, N.; CHERIF, M.; FLISS, I.; BOUDABBOUS, A.; ANTOUN, H. Evaluation of *Bacillus* isolates from salty soils and *Bacillus thuringiensis* strains for the biocontrol of *Fusarium* dry rot of potato tubers. **J Plant Pathol**, v. 83, p. 101-118, 2001.

SALAMA, H. S.; MORRIS, O. N. The use of *Bacillus thuringiensis* in development countries. In: ENTWISTLE, P. F.; CORY, J. S.; BAILEY, M. J.; HIGGS, S. (ed). ***Bacillus thuringiensis, an environmental biopesticide: theory and practice***, Chichester: John Wiley & Sons, p. 237-253, 1993.

SALZBERG, S. L. Genome re-annotation: a wiki solution? **Genome Biol**, v. 8, p. 102, 2007.

SANGER, F.; AIR, G. M.; BARRELL, B. G.; BROWN, N. L.; COULSON, A. R.; FIDDES, C. A.; HUTCHISON, C. A.; SLOCOMBE, P. M.; SMITH, M. Nucleotide sequence of bacteriophage phi X174 DNA. **Nature**, v. 265, p. 687-695, 1977a.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proc Natl Acad Sci USA**, v. 74, p. 5463-5467, 1977b.

SCHNEPF, E.; CRICKMORE, N.; VAN RIE, J.; LERECLUS, D.; BAUM, J.; FEITELSON, J.; ZEIGLER, D. R.; DEAN, D. H. *Bacillus thuringiensis* and its pesticidal crystal proteins. **Microbiol Mol Biol Rev**, v. 62, p. 775-806, 1998.

SCOTT, B. Epichloë endophytes: fungal symbionts of grasses. **Curr Opin Microbiol**, v. 4, p. 393-398, 2001.

SHEPPARD, A. E.; POEHLEIN, A.; ROSENSTIEL, P.; LIESEGANG, H.; SCHULENBURG, H. Complete genome sequence of *Bacillus thuringiensis* strain 407 cry-. **Genome Announc**, v. 1, p. 1-2, 2013.

SMITH, R. A.; COUCHE, G. A. The phylloplane as a source of *Bacillus thuringiensis* variants. **Appl Environ Microbiol**, v. 57, p. 311-315, 1991.

SMITH, R. A.; COUCHE, G. A. The phylloplane as a source of *Bacillus thuringiensis* variants. **Appl Environ Microbiol**, v. 57, p. 311-315, 1991.

SMITS, T. H.; JAENICKE, S.; REZZONICO, F.; KAMBER, T.; GOESMANN, A.; FREY, J. E.; DUFFY, B. Complete genome sequence of the fire blight pathogen *Erwinia pyrifoliae* DSM 12163T and comparative genomic insights into plant pathogenicity. **BMC Genomics**, v. 11, p. 2, 2010.

SNEH, B.; SCHUSTER S.; GROSS, S. Improvement of the insecticidal activity of *Bacillus thuringiensis* var. *entomocidus* on larvae of *Spodoptera fittoralis* (Lepidoptera, Noctuidae) by addition of chitinolytic bacteria, a phagostimulant and a UV-protectant. **Z Angew Entomol**, v. 96, p. 77-83, 1983.

SPRENT, J.; FARIA, S. Mechanisms of infection of plants by nitrogen fixing organisms. **Plant Soil**, v. 110, p. 157-165, 1988.

STAHLY, P.; ANDREWS, R.E.; YOUSTEN, A.A. The genus *Bacillus* – Insect Pathogens. In: BALOWS, A.H.G.; TRUPPER, M.; DWORKIN, W; HERDER, K.H. **The Prokaryotes**, 2nd ed., New York: Springer-Verlag, p. 4770, 1991.

STRAGIER, P.; KUNKEL, B.; KROOS, L.; LOSICK, R. Chromosomal rearrangement generating a composite gene for a developmental transcription factor. **Science**, v. 243, p. 507-512, 1989.

STEIN, L. Genome annotation from sequence to biology. **Nat Rev**, v. 2, p. 493-505, 2001.

STUDHOLME, D. J.; DIXON, R. Domain architectures of sigma54-dependent transcriptional activators. **J Bacteriol**, v. 185, p. 1757-1767, 2003.

SUBRAHMANYAN, P.; REDDY, M. N.; RAO, A. S. Exudation of certain organic compounds from seeds of groundnut. **Seed Sci Technol**, v. 11, p. 267-27, 1983.

SUGINTA, W.; ROBERTSON, P. A.; AUSTIN, B.; FRY, S. C.; FOTHERGILL-GILMORE, L. A. Chitinases from *Vibrio*: activity screening and purification of *chiA* from *Vibrio carchariae*. **J Appl Microbiol**, v. 89, p. 76-84, 2000.

SULLIVAN, T. J.; RODSTROM, J.; VANDOP, J.; LIBRIZZI, J.; GRAHAM, C.; SCHARDL, C. L., BULTMAN, T. L. Symbiont-mediated changes in *Lolium arundinaceum* inducible defenses: evidence from changes in gene expression and leaf composition. **New Phytol**, v. 176, p. 673-679, 2007.

SUN, Y.; ZHAO, Q.; XIA, L.; DING, X.; HU, Q.; FEDERICI, B. A.; PARK, H. W. Identification and characterization of three previously undescribed crystal proteins from *Bacillus thuringiensis* subsp. *jegathesan*. **Appl Environ Microbiol**, v. 79, p. 3364-3370, 2013.

SUZUKI, M. T.; HERNÁNDEZ-RODRÍGUEZ, C. S.; ARAÚJO, W. L.; FERRÉ, J. Characterization of an endophytic *Bacillus thuringiensis* strain isolated from sugar cane. In: **Proceedings of 41st Annual Meeting of the Society for Invertebrate Pathology and 9th International Conference on *Bacillus thuringiensis***, Coventry, United Kingdom, 3-7 August 2008.

SUZUKI, M. T.; LERECLUS, D.; ARANTES, O. M. N. Fate of *Bacillus thuringiensis* strains in different insect larvae. **Can J Microbiol**, v. 50, p. 973-975, 2004.

TAMURA, K.; STECHER, G.; PETERSON, D.; FILIPSKI, A.; KUMAR, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. **Mol Biol Evol**, v. 30, p. 2725-2729, 2013.

TAN, F.; ZHENG, A.; ZHU, J.; WANG, L.; LI, S.; DENG, Q.; WANG, S.; LI, P.; TANG, X. Rapid cloning, identification, and application of one novel crystal protein gene *cry30Fa1* from *Bacillus thuringiensis*. **FEMS Microbiol Lett**, v. 302, p. 46-51, 2010.

TAN, F.; ZHU, J.; TANG, J.; TANG, X.; WANG, S.; ZHENG, A.; LI, P. Cloning and characterization of two novel crystal protein genes, *cry54Aa1* and *cry30Fa1*, from *Bacillus thuringiensis* strain BtMC28. **Curr Microbiol**, v. 58, p. 654-659, 2009.

TANG, Y.; ZOU, J.; ZHANG, L.; LI, Z.; MA, C.; MA, N. Anti-fungi activities of *Bacillus thuringiensis* H3 chitinase and immobilized chitinase particles and their effects to rice seedling defensive enzymes. **J Nanosci Nanotechnol**, v. 12, p. 8081-8086, 2012.

TANUJA, R.; BISHT, S. C.; MISHRA, P. K. Ascending migration of endophytic *Bacillus thuringiensis* and assessment of benefits to different legumes of N.W. Himalayas. **Eur J Soil Biol**, v. 56, p. 56-64, 2012.

TETTELIN, H.; MASIGNANI, V.; CIESLEWICZ, M. J.; DONATI, C.; MEDINI, D.; WARD, N. L.; et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". **Proc Natl Acad Sci USA**, v. 102, p. 13950-13955, 2005.

THOMAS, D. J. I.; ALUN, J.; MORGAN, W.; WHIPPS, J. M.; SAUNDERS, J. R. Plasmid transfer between the *Bacillus thuringiensis* subspecies *kurstaki* and *tenebrionis* in laboratory culture and soil and in Lepidopteran and Coleopteran larvae. **Appl Environ Microbiol**, v. 66, p. 118-124, 2000.

TICKNOR, L. O.; KOLSTO, A. B.; HILL, K. K.; KEIM, P.; LAKER, M. T.; TONKS, M.; JACKSON, P. J. Fluorescent amplified fragment length polymorphism analysis of Norwegian *Bacillus cereus* and *Bacillus thuringiensis* soil isolates. **Appl Environ Microbiol**, v. 67, p. 4863-4873, 2001.

TILLIER, E. R.; COLLINS, R. A. Genome rearrangement by replication-directed translocation. **Nat Genet**, v. 26, p. 195-197, 2000.

TURNBAUGH, P. J.; LEY, R. E.; HAMADY, M.; FRASER-LIGGETT, C. M.; KNIGHT, R.; GORDON, J. I. The human microbiome project. **Nature**, v. 449, p. 804-810, 2007.

VAN FRANKENHUZEN, K. The challenge of *Bacillus thuringiensis*. In: ENTWISTLE, P. F.; CORY, J. S.; BAILEY, M. J.; HIGGS, S. (ed). **Bacillus**

***thuringiensis*, an environment biotpesticide: theory and practice**, Chichester: John Wiley & Sons, p. 1-35. 1993.

VAN LANEN, S. G.; SHEN, B. Microbial genomics for the improvement of natural product discovery. **Curr Opin Microbiol**, v. 9, p. 252-260, 2006.

VENTER, J. C.; ADAMS, M. D.; MYERS, E. W.; LI, P. W.; MURAL, R. J.; SUTTON, G. G.; SMITH, H. O.; et al.. The sequence of the human genome. **Science**, v. 291, p. 1304-1351, 2001.

VILAS-BÔAS, G. F. L. T.; VILAS-BÔAS, L. A.; LERECLUS, D.; ARANTES, O. M. N. *Bacillus thuringiensis* conjugation under environmental conditions. **FEMS Microbiol Ecol**, v. 25, p. 369-374, 1998.

VILAS-BÔAS, L. A.; VILAS-BÔAS, G. F. L. T.; SARIDAKIS, H. O.; LEMOS, M. V. F.; LERECLUS, D.; ARANTES, O. M. N. Survival and conjugation of *Bacillus thuringiensis* in a soil microcosm. **FEMS Microbiol Ecol**, v. 31, p. 225-259, 2000.

VON BODMAN, S. B.; BAUER, W. D.; COPLIN, D. L. Quorum sensing in plant-pathogenic bacteria. **Annu Rev Phytopathol**, v. 41, p. 455-482, 2003.

WANG, A.; PATTEMORE, J.; ASH, G.; WILLIAMS, A.; HANE, J. Draft genome sequence of *Bacillus thuringiensis* strain DAR 81934, which exhibits molluscicidal activity. **Genome Announc**, v. 1, p. 1-2, 2013.

WANG, S. L.; HWANG, J. Microbial reclamation of shellfish wastes for the production of chitinases. **Enzyme Microb Technol**, v. 28, p. 376-382, 2001.

WANG, S. L.; LIN, T. Y.; YEN, Y. H.; LIAO, H. F.; CHEN, Y. J. Bioconversion of shellfish chitin wastes for the production of *Bacillus subtilis* W-118 chitinase. **Carbohydr Res**, v. 341, p. 2507-2515, 2006.

WEI, J. Z.; HALE, K.; CARTA, L.; PLATZER, E.; WONG, C.; FANG, S. C.; AROIAN, R. V. *Bacillus thuringiensis* crystal proteins that target nematodes. **Proc Natl Acad Sci U S A**, v. 100, p. 2760-2765, 2003.

WEISER, J. Impact of *Bacillus thuringiensis* on applied entomology in eastern Europe and in Soviet Union. In: KRIEG, A.; HUGER, A. M. **Mitteilungen aus der biologischen bundesanstalt für land und forstwirtschaft Berlin-Dahlem heft**, Berlin: Paul Parey, p. 37-50, 1986.

WILSON, K. E.; FLOR, J. E.; SCHWARTZ, R. E.; JOSHUA, H.; SMITH, J. L.; PELAK, B. A.; LIESCH, J. M.; HENSENS, O. D. Difficidin and oxydifficidin: novel broad spectrum antibacterial antibiotics produced by *Bacillus subtilis*. II. Isolation and physico-chemical characterization. **J Antibiot (Tokyo)**, v. 40, p. 1682-1691, 1987.

WILSON, M. K.; ABERGEL, R. J.; RAYMOND, K. N.; ARCENEUX, J. E.; BYERS, B. R. Siderophores of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*. **Biochem Biophys Res Commun**, v. 348, p. 320-325, 2006.

WIWAT, C.; THAITHANUN, S.; PANTUWATANA, S.; BHUMIRATANA, A. Toxicity of chitinase-producing *Bacillus thuringiensis* ssp. *kurstaki* HD-1 (G) toward *Plutella xylostella*. **J Invertebr Pathol**, v. 76, p. 270-277, 2000.

YE, W.; ZHU, L.; LIU, Y.; CRICKMORE, N.; PENG, D.; RUAN, L.; SUN, M. Mining new crystal protein genes from *Bacillus thuringiensis* on the basis of mixed plasmid-enriched genome sequencing and a computational pipeline. **Appl Environ Microbiol**, v. 78, p. 4795-4801, 2012.

ZAWADZKA, A. M.; KIM, Y.; MALTSEVA, N.; NICHIPORUK, R.; FAN, Y.; JOACHIMIAK, A.; RAYMOND, K. N. Characterization of a *Bacillus subtilis* transporter for petrobactin, an anthrax stealth siderophore. **Proc Natl Acad Sci USA**, v. 106, p. 21854-21859, 2009.

ZGHAL, R. Z.; JAOUA, S. Evidence of DNA rearrangements in the 128-kilobase pBtoxis plasmid of *Bacillus thuringiensis israelensis*. **Mol Biotechnol**, v. 33, p. 191-198, 2006.

ZHANG, G.; DENG, A.; XU, Q.; LIANG, Y.; CHEN, N.; WEN, T. Complete genome sequence of *Bacillus amyloliquefaciens* TA208, a strain for industrial production of guanosine and ribavirin. **J Bacteriol**, v. 193, p. 3142-3143, 2011.

ZHANG, W.; CHEN, J.; YANG, Y.; TANG, Y.; SHANG, J.; SHEN, B. A practical comparison of de novo genome assembly software tools for next-generation sequencing technologies. **Plos One**, v. 6, p. e17915, 2011.

ZHOU, Y.; CHOI, Y. L.; SUN, M.; YU, Z. Novel roles of *Bacillus thuringiensis* to control plant diseases. **Appl Microbiol Biotechnol**, v. 80, p. 563-572, 2008.

ZHU, Y.; SHANG, H.; ZHU, Q.; JI, F.; WANG, P.; FU, J.; DENG, Y.; XU, C.; YE, W.; ZHENG, J.; ZHU, L.; RUAN, L.; PENG, D.; SUN, M. Complete genome sequence of *Bacillus thuringiensis* serovar *finitimus* strain YBT-020. **J Bacteriol**, v. 193, p. 2379-2380, 2011.

ZIMMERMAN, S. B.; SCHWARTZ, C. D.; MONAGHAN, R. L.; PELAK, B. A.; WEISSBERGER, B.; GILFILLAN, E. C.; MOCHALES, S.; HERNANDEZ, S.; CURRIE, S. A.; TEJERA, E.; ET AL. Difficidin and oxydifficidin: novel broad spectrum antibacterial antibiotics produced by *Bacillus subtilis*. I. Production, taxonomy and antibacterial activity. **J Antibiot (Tokyo)**, v. 40, p. 1677-1681, 1987.

ZWEERINK, M. M.; EDISON, A. Difficidin and oxydifficidin: novel broad spectrum antibacterial antibiotics produced by *Bacillus subtilis*. III. Mode of action of difficidin. **J Antibiot (Tokyo)**, v. 40, p. 1692-1697, 1987.

ZWICK, M. E.; JOSEPH, S. J.; DIDELOT, X.; CHEN, P. E.; BISHOP-LILLY, K. A.; STEWART, A. C.; WILLNER, K.; NOLAN, N.; LENTZ, S.; THOMASON, M. K.; SOZHAMANNAN, S.; MATECZUN, A. J.; DU, L.; READ, T. D. Genomic

characterization of the *Bacillus cereus sensu lato* species: backdrop to the evolution of *Bacillus anthracis*. **Genome Res**, v. 22, p. 1512-1524, 2012.