UNIVERSIDADE FEDERAL DO PARANÁ

ANA PAULA ANDREAZZA

ANÁLISE METAGENÔMICA DE COMPOSTO LIGNOCELULÓSICO E CARACTERIZAÇÃO FISIOLÓGICA E MOLECULAR DE UMA NOVA ESTIRPE DE *Bacillus* sp.

CURITIBA

2016

ANA PAULA ANDREAZZA

ANÁLISE METAGENÔMICA DE COMPOSTO LIGNOCELULÓSICO E CARACTERIZAÇÃO FISIOLÓGICA E MOLECULAR DE UMA NOVA ESTIRPE DE *Bacillus* sp.

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências-Bioquímica, Setor de Ciências Biológicas da Universidade Federal do Paraná como requisito parcial para a obtenção do grau de Doutor em Ciências-Bioquímica.

Orientador: Profa. Dra. Leda Satie Chubatsu
Co-orientador: Dr. Fábio de Oliveira Pedrosa

CURITIBA
2016

Andreazza, Ana Paula
A557 Análise metagenômica de composto lignocelulósico e caracterização fisiológica e molecular de uma nova estirpe de *bacillus* sp / Ana Paula Andreazza. - Curitiba, 2016.
118 f.: il., tabs, grafs.

Orientador: Leda Satie Chubatsu
Co-orientador: Fábio de Oliveira Pedrosa
Tese (Doutorado) – Universidade Federal do Paraná, Setor de Ciências Biológicas, Curso de Pós-Graduação em Ciências-Bioquímica.

Inclui Bibliografia.

1. Compostagem. 2. *Bacillus*. 3. Genoma. 4. Genômica. 5. Quitinase. I. Chubatsu, Leda Satie. II. Pedrosa, Fábio de Oliveira. III.Título. IV. Universidade Federal do Paraná.

CDD 579.362

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS (BIOQUÍMICA) da Universidade Federal do Paraná foram convocados para realizar a arguição da Tese de Doutorado de ANA PAULA ANDREAZZA, intitulada: "**Análise metagenômica de composto lignocelulósico e caracterização fisiológica e molecular de uma nova estirpe de *Bacillus* sp.",** após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua aprovação.

CURITIBA, 15 de Dezembro de 2016.

LEDA SATIE CHUBATSU
Presidente da Banca Examinadora (UFPR)

DIEVAL GUIZELINI
Avaliador Externo (UFPR)

EMANUEL MALTEMPI DE SOUZA
Avaliador Interno (UFPR)

MARCELO MÜLLER DOS SANTOS
Avaliador Interno (UFPR)

FREDERICO JOSÉ GUEIROS FILHO
Avaliador Externo (USP)

Dedico esse trabalho à minha família
por todo o apoio ao longo desses anos,
por todos os valores que me foram ensinados
e sem a qual nada eu teria conseguido.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por estar sempre comigo e não me deixar desistir mesmo nos momentos mais difíceis ao longo dessa caminhada.

À minha orientadora, Profª. Dra. Leda Satie Chubatsu, por ter me dado a oportunidade de desenvolver este projeto, por seus ensinamentos e, principalmente, pelo exemplo de professora e pesquisadora.

Ao meu co-orientador Dr. Fábio de Oliveira Pedrosa pelo exemplo e dedicação à ciência.

Aos professores Dr. Emanuel Maltempi de Souza e Dr. Marcelo Müller dos Santos pelas contribuições neste trabalho através de sugestões e explicações.

Ao Prof. Dr. Leonardo Magalhães Cruz pelas sugestões e todo o suporte da bioinformática.

A Profa. Dra. Lucélia Donatti pelos experimentos de microscopia eletrônica.

Ao Prof. Dr. Luís Amilton Foerster pelo auxílio nos ensaios envolvendo insetos.

Aos professores Dra. Cyntia M.T. Fadel-Picheth e Dr. Luciano Huergo pelas contribuições no trabalho e sempre foram prestativos e atenciosos.

Aos demais professores do programa de Pós-Graduação em Ciências-Bioquímica, por toda a atenção.

Ao Dr. Helisson Faoro pelos ensinamentos e auxílio nas análises de metagenômica.

Ao Dr. Marco A. Oliveira pelo auxílio e orientações nos ensaios de clonagem dos genes e purificação de proteínas.

Ao Dr. Rodrigo Luis A. Cardoso por todo o trabalho de bioinformática, pelos ensinamentos, pela dedicação e pelas enormes contribuições neste trabalho.

À Dra Michelle Tadra-Sfeir e ao Dr. Eduardo Balsanelli pelo sequenciamento genômico e metagenômico.

À Dra. Sarah Sacks Thimoteo pelo auxílio nos ensaios enzimáticos.

Aos técnicos do grupo de fixação de nitrogênio Roseli Prado, Valter de Baura e Lucinéia Martins.

A Agropecuária Boi Picuá LTDA pelo fornecimento das amostras.

À minha família: meu pai José Fernando, minha mãe Marila, meus irmãos Fernanda e Marcus Felipe e ao sempre companheiro Breno.

Aos amigos Adrian, Alex, Rafaela, André, Caroline, Eloisa, Edileusa, Fernanda, Marcelo, Esther, Paloma, Vivian e Rocio pelos momentos de descontração que fazem a rotina do laboratório ser mais leve e feliz.

À coordenação do curso de pós-graduação em Bioquímica.

Ao CNPq/INCT, CAPES e Fundação Araucária pelo suporte financeiro.

"O papel dos infinitamente pequenos na natureza é infinitamente grande".

(*Louis Pasteur,* 1822-1895)

RESUMO

A compostagem é um processo de decomposição da matéria orgânica caracterizado pela sucessão de comunidades microbianas complexas produtoras de várias enzimas que atuam sinérgica e especificamente ao longo do processo. Nesta tese, a análise metagenômica de um composto rico em material lignocelulósico revelou que gama *Proteobacteria*, representadas pelo gênero *Cellvibrio* foi o filo predominante na amostra seguido pelo filo *Bacteroidetes* representada pelo gênero *Ohtaekwangia*. A presença de organismos com capacidade de degradar compostos aromáticos, degradar celulose, interagir com plantas e fixar nitrogênio (*Paludobacter, Azoarcus, Azospira)* mostrou que a compostagem estudada é uma fonte promissora para a prospecção de enzimas envolvidas na degradação da biomassa vegetal. Utilizando uma cultura de enriquecimento de uma amostra do composto, foi isolada uma bactéria, denominada ABP14, que apresentou atividade celulolítica em meio CMC 0,5%. A análise da sequência parcial do gene 16S rRNA foi identificada como *Bacillus sp.* pertencente ao grupo *Bacillus cereus sensu lato* (s.l.). O isolado ABP14 apresentou atividade inseticida contra *Anticarsia gemmtalis* (lagarta do soja), mas sem a presença do cristal parasporal característico de *Bacillus thuringiensis*. O sequenciamento total do genoma do *Bacillus sp.* ABP14, utilizando as plataformas Illumina MiSeq e Ion Proton System, revelou a presença de dois *replicons*, um cromossomo de 5.141.367pb com conteúdo G+C de 35,4% e um plasmídeo circular (denominado pABP) com 11.199pb e 30,4% G+C. O cromossomo contém 5.238 genes codificantes de proteínas, 13 rRNA operons e 92 tRNA. A análise taxonômica baseada na hibridização DNA-DNA *in silico*, utilizando o programa GGDC 2.1, mostrou que *Bacillus sp.* ABP14 é geneticamente próximo de *B. thuringiensis* serovar *finitimis* YBT020 (*Bt finitimus*) com um GGDH de 81,3%, indicando pertencerem a mesma espécie. A análise genômica comparativa identificou genes específicos do *Bacillus sp*. ABP14 dentro do grupo *B. cereus s.l.* que codificam para enzimas da biossíntese da parede celular, proteínas de membrana e sistema de restrição. Difere exclusivamente do *Bt finitimus* pela presença de uma celulase, pelos genes envolvidos com competência celular, por um conjunto de genes que codificam proteínas que representam uma vantagem para a sobrevivência no ambiente intestinal do hospedeiro e pela ausência dos genes que codificam para as proteínas Cry. *Bacillus sp.* ABP14 e *Bt finitimus* compartilham mecanismos de proteção contra as defesas do hospedeiro, enzimas proteolíticas e adesinas, além de um conjunto de fatores de virulência comum em *B. cereus stricto sensu*. Dois genes identificados no genoma de *Bacillus sp.* ABP14, denominados *chiABP-74* e

chiABP-39, que codificam para quitinases foram clonados em pET29a (+), expressos em *E. coli* BL21(DE3), as proteínas parcialmente purificadas por cromatografia de troca iônica e caracterizadas bioquimicamente. O gene *chiABP-74* codifica para uma proteína de 674 aminoácidos, incluindo um peptídeo sinal de 33 aminoácidos, com massa molecular de 74 kDa e pI de 5,67. O gene *chiABP-39* codifica para uma proteína de 360 aminoácidos, incluindo um peptídeo sinal de 27 aminoácidos, com massa molecular de 39 kDa e pI de 6,21. A busca de similaridade realizada com a ferramenta Blast mostrou que as quitinase ChiABP-74 e ChiABP-39 apresentam uma alta identidade (99%) com as quitinases de *B. cereus e B. thuringiensis*. A análise da sequência de aminoácidos revelou que as quitinases pertencem a família 18 das glicosil-hidrolases, a quitinase ChiABP-39 apresenta um único domínio catalítico (GH18) e a ChiABP-74 apresenta três domínios: o domínio catalítico GH18, um domínio fibronectina tipo-III (Fn3) e um módulo de ligação a carboidrato da família 2 (CBM2). As enzimas apresentaram atividade máxima em pH 5,0 a 50°C, a quitinase ChiABP-39 apresenta maior especificidade por quitina coloidal e sua atividade diminui na presença de SDS 0,5%, mas não sofre alteração na presença de íons divalentes. A quitinase ChiABP-74 é capaz de hidrolisar quitina coloidal e β-quitina, e na presença de Mg^{+2} e Zn^{+2} na concentração de 10mM, bem como de SDS 0,5% apresenta redução da atividade enzimática.

Palavras-chave: Compostagem, *Bacillus*, Genoma, Comparação Genômica, Quitinase.

ABSTRACT

Composting is an organic matter decomposition process characterized by a series of complex microbial communities that produce different enzymes acting in a synergistic and specific manner. In this work, a metagenomic analysis of a lignocellulosic rich compost revealed that *Proteobacteria* from the gamma class, represented by the *Cellvibrio* genus was the predominant phylum in the sample followed by the *Bacteroidetes* phylum, represented by the *Ohtaekwangia* genus. Among those, organisms able to degrade aromatic compounds, cellulose, plant-interact bacteria and nitrogen-fixers such as *Paludobacter, Azoarcus* and *Azospira* showed that this compost is a promising source for prospecting enzymes involved in the degradation of plant biomass. A culture enrichment in medium containing carboxymethyl-cellulose (CMC medium) of a compost sample allowed the isolation of a bacterium, named ABP14, which showed cellulolytic activity in CMC medium. Partial sequence of the 16S rRNA gene identified ABP14 as *Bacillus sp.* belonging to the *B. cereus sensu lato* (s.l.) group. Phenotypic analyses revealed that the isolated ABP14 showed insecticidal activity against *Anticarsia gemmtalis* (soybean caterpillar) but without the presence of parasporal crystal that is characteristic of *Bacillus thuringiensis*. The genome of *Bacillus sp.* ABP14 was sequenced using the Illumina MiSeq and Ion Proton System platforms revealed the presence of two *replicons*, a chromosome of 5,141,367 bp, with 35.4% of G + C content and a circular plasmid (denominated pABP) with 11,199 bp and 30.4% G + C. The chromosome contains 5,238 protein-encoding genes, 13 rRNA operons and 92 tRNA. Taxonomic analysis based on *in silico* DNA-DNA hybridization using the GGDC 2.1 server program showed that *Bacillus sp.* ABP14 is genetically close to *B. thuringiensis* serovar *finitimis* YBT020 (*Bt finitimus*) with a GGDH of 81.3%, indicating that they belong to the same species. The comparative genomic analysis identified specific genes of *Bacillus sp*. ABP14 within the *B. cereus s.l.* group encoding cell wall biosynthesis enzymes, membrane proteins and restriction system. ABP14 differs exclusively from *Bt finitimus* by the presence of genes coding for cellulase, genes involved with cell competence, and a number of genes encoding proteins that represent an advantage for survival in the gut environment of the host and the absence of the genes encoding for the Cry proteins. *Bacillus sp.* ABP14 and *Bt finitimus* share mechanisms of protection against host defenses, proteolytic enzymes and adhesin, in addition to a set of virulence factors that are common in *B. cereus stricto sensu*. Two genes identified in the genome of *Bacillus sp.* ABP14, called *chiABP-74* and *chiABP-39,* encoding chitinase were cloned into pET29a (+), expressed in *E. coli* BL21(DE3), partially purified by ion exchange chromatography and biochemically characterized. The gene *chiABP-*

74 encodes a protein of 674 amino acids including a signal peptide of 33 amino acids with a molecular mass of 74 kDa and pI of 5.67. Gene *chiABP-39* encodes a protein of 360 amino acids including a signal peptide of 27 amino acids with a molecular mass of 39 kDa and pI of 6.21. The similarity search performed with the Blast tool showed that the chitinase ChiABP-74 and ChiABP-39 have a high identity (99%) with chitinase *B. cereus and B. thuringiensis*. The analysis of the amino acid sequence revealed that the chitinases belong to family 18 of the glycosyl hydrolases, chitinase ChiABP-39 has a single catalytic domain (GH18) and ChiABP-74 has three domains: the catalytic domain GH18, a fibronectin-type III (Fn3) domain and a carbohydrate-binding module of family 2 (CBM2). The enzymes showed maximum activity in pH 5.0 at 50°C, chitinase ChiABP-39 has specificity for colloidal chitin and its activity decreases in the presence of 0.5% SDS, but it does not change in the presence of divalent ions. Chitinase ChiABP-74 is able to hydrolyse β-colloidal chitin and chitin in the presence of 10mM Mg^{+2} and Zn^{+2}, as well as of 0.5% SDS, it shows reduced enzymatic activity.

Keywords: Composting, *Bacillus*, Genome, Genomic comparison, Chitinase.

LISTA DE FIGURAS

LISTA DE TABELAS

CAPÍTULO II: Genome analysis of entomopathogenic *Bacillus* sp. ABP14 isolates from lignoellulosic compost

CAPÍTULO III: Cloning, expression and characterization of two chitinase genes *chiABP-39* and *ChiABP-74* from entomopathogenic *Bacillus* sp. ABP14 isolated from compost.

LISTA DE ABREVIATURAS

Ba - *Bacillus antharcis*

Bc - *Bacillus cereus*

BLAST - *Basic Local Alignment Search Tool*

BRIG - BLAST *Ring Image Generator*

Bt – Bacillus thuringiensis

CDS - *coding sequence* (sequência que codifica para uma proteína)

CMC – Carboximethylcellulose

COG - *Cluster of Orthologous Groups*

Cry - crystal protein

CytK - cytotoxin K

DDH - DNA-DNA *hibridization*

EC - *Enzyme Commission*

GGDC - *Genome-to-Genome Distance Calculator*

GGDH - *Genome-to-Genome Distance Hybridization*

GIT - gastrointestinal tract

Hbl - hemolytic enterotoxin

HlyI - type I haemolysin

HlyII - type II haemolysin

InhA - immune inhibitor A

KAAS - KEGG *Automatic annotation server*

Mb - Mega base

MEGA - *Molecular Evolutionary Genetics Analysis*

NCBI - *National Center for Biotechnology Information*

NCBI - *National Center for Biotechnology Information*

Nhe - non-hemolytic enterotoxin

ORF - *open reading frame* (fase aberta de leitura, possível gene)

PC-PLC - phosphatidylcholine-specific phospholipase C

PlcR - pleiotropic transcriptional regulator

s.l. - sensu lato

SPH - sphingomyelinase

SUMÁRIO

1 INTRODUÇÃO

As bactérias são as unidades celulares mais numerosas e antigas da Terra e a diversidade filogenética e funcional que observamos atualmente nas células bacterianas, é resultado de aproximadamente 4 bilhões de anos de alterações evolutivas (MADIGAN *et al.*, 2010). Sendo assim, os genomas bacterianos contém inúmeras rotas para transformações bioquímicas (National Research Council Committee on Metagenomics, 2007). A identificação, caracterização e preservação destes recursos microbianos são fatores de fundamental importância para o avanço biotecnológico e para o desenvolvimento econômico sustentável (OLIVEIRA *et al*., 2006).

Grandes esforços têm sido feito no sentido de desenvolver métodos de triagem e de cultivo de linhagens com propriedades desejadas (AGEITEC). No entanto, grande parte dos microrganismos presentes em uma amostra ambiental permanece desconhecida devido a sua inabilidade em se multiplicar nos meios de cultivo tradicionais (SCHMEISSER *et al*., 2007). Para superar os limites impostos pelos meios de cultivo, uma nova abordagem de trabalho, envolvendo metodologias de bioinformática e biologia molecular, denominada *metagenômica* vem sendo aplicada. A metagenômica permite o acesso direto ao DNA de uma comunidade bacteriana (metagenoma) e técnicas de sequenciamento de DNA de segunda geração permitem conhecer a biodiversidade presente neste ambiente. A metagenômica também permite a prospecção *in silico* dos microrganismos a partir dos dados genômicos depositados em bases de dados, sem a necessidade de isolamento e cultivo (STREIT & SCHMITS, 2004).

Explorar a diversidade microbiana conduzirá a descoberta de novos genes, enzimas e compostos importantes para o desenvolvimento de novas tecnologias com aplicações práticas na medicina, agricultura, energia alternativa, remediação ambiental e na indústria e seus processos (CGEE, 2010).

Dentro deste contexto, este trabalho tem como objetivo identificar os microrganismos predominantes em uma compostagem de resíduos agrícolas, isolar novos microrganismos com potencial aplicação biotecnológica e buscar por genes que codifiquem enzimas hidrolíticas de interesse econômico.

Os resultados deste estudo estão descritos na forma de artigos científicos e organizados em três capítulos. O capítulo I refere-se a abordagem metagenômica utilizada para acessar a composição taxonômica do composto utilizado. O capítulo II compreende o isolamento de uma bactéria celulolítica a partir do cultivo seletivo de uma amostra do composto, a identificação da

sua atividade entomopatogênica e uma análise genômica compartiva com outras estirpes relacionadas. O capítulo III apresenta as etapas de amplificação e clonagem de dois genes que codificam para quitinases, a expressão heteróloga, purificação e caracterização bioquímica das enzimas.

2 Revisão da Literatura

2.1 Compostagem

Na agricultura, a compostagem é um processo de importância econômica, pois permite a transformação de resíduos agrícolas em fertilizante orgânico. Quando aplicados no solo, os fertilizantes orgânicos atuam como condicionadores e melhoradores das propriedades físicas, físico-químicas e biológicas do solo (VINHAL-FREITAS *et al*., 2010).

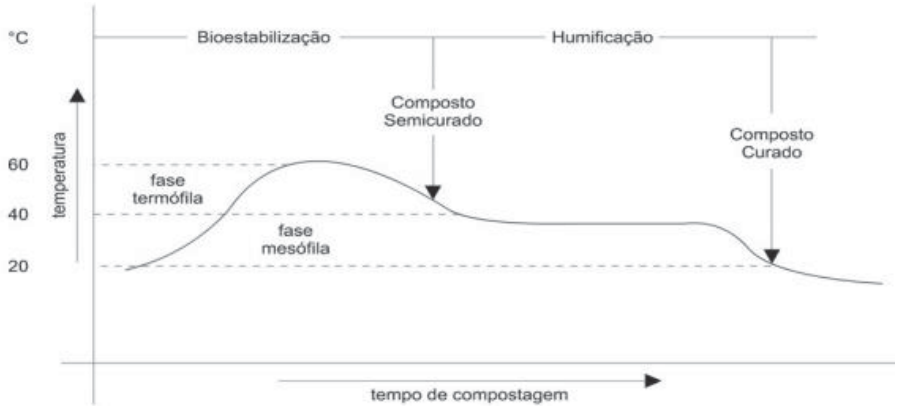
A compostagem é definida como um processo de oxidação biológica no qual resíduos orgânicos sólidos de diferentes origens são transformados pela ação dos microrganismos em dióxido de carbono, água e um material estável, rico em húmus e nutrientes minerais denominado composto (MONTEIRO *et al.* 2001).

A matéria orgânica é decomposta em um processo complexo que envolve a ação cooperativa de comunidades microbianas produtoras de enzimas. A comunidade sofre variações na sua composição de acordo com as condições reacionais (temperatura, pH,umidade, disposição de nutrientes) sendo modificadas em consequência da atividade metabólica destas comunidades (WONGWILAIWALIN et al, 2013). As bactérias são os microrganismos predominantes no processo devido a sua maior diversidade metabólica representada pela variedade de enzimas produzidas (MARTINS *et al.*, 2013).

O processo pode ser dividido em duas fases principais, a bioestabilização e a humificação, de acordo com o perfil térmico do composto. A decomposição inicial é realizada por microrganismos mesofílicos que degradam rapidamente os compostos solúveis, facilmente degradáveis, e o calor gerado pelo seu metabolismo fazem com que a temperatura se eleve. Quando a temperatura excede 40°C, os microrganismos mesófilos se tornam menos competitivos e são substituídos pelos microrganismos termofílicos. As altas temperaturas aceleram a quebra das proteínas, gorduras e carboidratos complexos como celulose e hemicelulose. A medida que o suprimento de nutrientes vai se esgotando, a temperatura do composto diminui gradualmente e os microrganismos termofílicos são substituídos pelos

mesofílicos com habilidade de degradar o material remanescente recalcitrante, como lignina, quitina e celulose (http://cwmi.css.cornell.edu/chapter1).

Figura1. Fases da compostagem e o referente grau térmico.



Fonte: D´ALMEIDA & VILHENA,

Os microrganismos presentes em ecossistemas complexos como o composto são um rico reservatório de genes para aplicação em biotecnologia, principalmente na biocatálise, na qual enzimas são utilizadas como uma alternativa verde à síntese orgânica tradicional (COSTA et al., 2015). Várias enzimas já foram isoladas e caracterizadas a partir de bibliotecas metagenômicas de composto, como amilase (ARIKAN, 2008), celulase (PANG *et al*., 2009), β-glucosidase (UCHIYAMA *et al.*, 2013) e xilanase (VERMA *et al.*, 2013) indicando que a compostagem é um ambiente promissor para a prospecção de enzimas hidrolíticas e que o perfil enzimático de cada ambiente de compostagem depende do material a ser degradado e da planta do processo (PARTANEN *et al.*, 2010).

2.2 Gênero *Bacillus*

Bacillus spp. são conhecidos pela capacidade de produzir e secretar enzimas que degradam uma grande variedade de substratos, permitindo a sua sobrevivência em vários ambientes (SLEPECK & HEMPHILL, 2006). Esta capacidade de secretar enzimas fez dos *Bacillus* os principais produtores de enzimas comerciais homólogas (WESTERS *et al.,* 2004).

O gênero *Bacillus* foi descrito por Ferdinand Cohn (1828-1898) em 1872 e tem como uma de suas principais características a formação de endósporos sob condições ambientais adversas (MADIGAN *et al*., 2010). Suas espécies compreendem células Gram positivas em forma de cilindro ou bastão, com extremidades arredondadas ou retas, que podem estar isoladas, em pares ou filamentos de tamanhos variados (de VOS *et al.*, 2009). Nas espécies móveis, a motilidade é devida aos flagelos peritríquios. Em geral, são quimio-organotróficos, aeróbios ou anaeróbios facultativos, mesofílicos e neutrofílico. Espécies termofílicas e psicrofílicas são capazes de crescer em temperaturas tão elevadas quanto 75°C ou tão baixas como 3°C e algumas espécies suportam extremos de acidez e alcalinidade, com pH variando de 2 a 10. Os esporos são únicos e apresentam a forma oval ou redonda, sendo muito resistente a condições adversas (PUBLIC HEALTH ENGLAND, 2015). A esporulação não é reprimida pela exposição ao ar, diferenciando membros deste gênero dos *Clostridium* spp. (TURNBULL & KRAMER, 1991).

Atualmente (outubro de 2016), no banco de dados LPSN - *List of Prokaryotic names with Standing in Nomenclature* estão descritas 318 espécies e 7 subespécies de *Bacillus* spp. (http://www.bacterio.net). Entre as poucas espécies patogênicas destacam-se o *B. anthracis* causador do antraz em humanos e animais e *B. cereus* responsável por intoxicações alimentares em humanos (SLEPECKY & HEMPHILL, 2006).

O conteúdo G+C do DNA das espécies dentro do gênero pode varias de 32 a 69%. Esta heterogeneidade dos *Bacillus* spp. reflete-se na grande variedade de nichos ecológicos que muitas espécies ocupam e na dificuldade de classificação pelos taxonomistas devido a sua diversidade fenotípica extrema (DROBNIEWSKI, 1993).

As espécies de *Bacillus* foram divididas em três grupos com base na morfologia do esporo e do esporângio. O grupo I é constituído por bacilos Gram positivos, com esporos central ou terminal, esférico ou elipsoide que não deforma o esporângio. O grupo II compreende bacilos Gram variável com esporos elipsoides e esporângio dilatado e o grupo III é Gram variável, esporângio dilatado com esporo terminal ou subterminal (de VOS *et al.*, 2009).

Com o desenvolvimento de métodos moleculares para a identificação das estirpes, especialmente a análise da sequência do gene 16S rRNA/DNA, algumas espécies do gênero *Bacillus* foram reclassificadas em novos gêneros (WISOTZJSEY *et al.*, 1992) e novas espécies foram definidas para alocar novos isolados (GUINEBRETIÈRE *et al*. 2013; JIMÉNEZ *et al*., 2013). Porém algumas espécies são estreitamente relacionadas, não sendo possível fazer a distinção entre as linhagens baseando-se somente em análises da sequência do gene 16S

rRNA/DNA. Assim foram estabelecidos no gênero *Bacillus* dois grupos, o grupo do *B. cereus* e o grupo do *B. subtilis*, composto por espécies altamente relacionadas taxonomicamente. Para o grupo do *B. cereus*, a diferenciação entre as espécies é realizada com base no seu habitat, patogenicidade para mamíferos e insetos, e suas características morfológicas e fisiológicas (FRITZE, 2004). Para o grupo do *B. subtilis*, as espécies podem ser diferenciadas pela análise da composição de ácidos graxos da parede celular, digestão por enzimas de restrição e hibridização DNA-DNA, mas são muito difíceis de serem diferenciadas pelas características fenotípicas (NAKAMURA *et al.*, 1999).

2.2.1 Grupo do *Bacillus cereus*

O grupo do *Bacillus cereus* (*Bacillus cereus sensu lato*) é formado por oito espécies altamente relacionadas taxonomicamente e compreende *B. cereus sensu stricto*, *B. thuringiensis*, *B. anthracis*, *B. mycoides*, *B. pseudomycoides, B. weihenstephanensis, B. cytotoxicus* e *B. toyonensis* (OREM, 2014).

As características clássicas que diferenciam este grupo de todas as outras bactérias aeróbicas formadoras de esporos são sua incapacidade de formar ácido a partir do manitol e sua produção de lecitinase. São geralmente catalase positiva, oxidase variável e metabolizam carboidratos por fermentação. *B. anthracis* é invariavelmente sensível a penicilina enquanto outras espécies são resistentes (PUBLIC HEALTH ENGLAND, 2015).

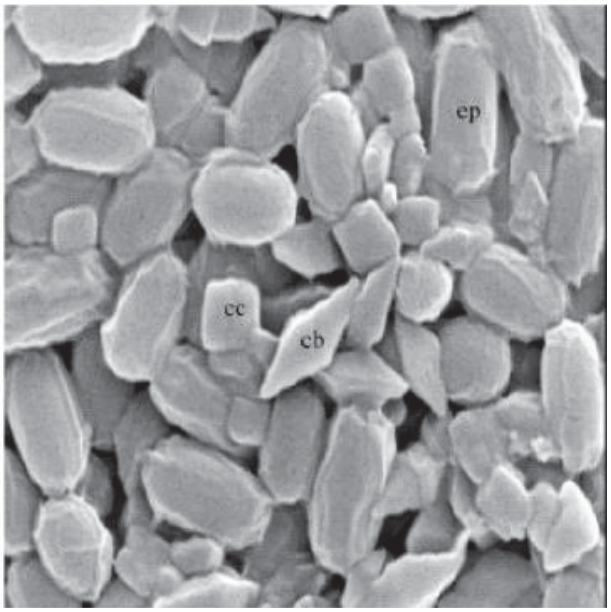
B. anthracis é o agente causal do antraz, uma zoonose que pode ser letal para humanos. Sua virulência está associada a presença da cápsula polipeptídica (ácido poli-D-glutâmico) codificada pelo plasmídeo pXO2 e pela toxina letal (LeTx) codificada pelo plasmídeo pXO1. A exotoxina LeTx é formada por três proteínas termolábeis, o antígeno protetor (PA), o fator edema (EF) e fator letal (LF), responsáveis pelos sintomas clínicos da doença (SHERER *et al.*, 2007).

B. cereus sensu stricto (s.s.) é um patógeno humano oportunista mais comumente associado com intoxicação alimentar que pode se manifestar de duas formas distintas, a Síndrome diarreica causada pela ação de enterotoxinas ou a Síndrome emética pela presença da toxina cereulida. Além da intoxicação alimentar, *B. cerus* também foi associado a casos de infecção local em feridas proveniente de trauma, queimadura ou pós-cirúrgica, em infeções oculares e periodontite (DROBNIEWSKI, 1993). De modo contrário, algumas linhagens de *B. cereus* são utilizadas como probióticos e, por serem consideradas seguras para o emprego médico, são consideradas classe 1 (HONG *et al*., 2005). *B. toyonensis* foi descrito

primariamente como uma subespécie do *B. cereus* s.s. e é utilizado como probiótico na alimentação animal (JIMÉNEZ *et al*., 2013).

B. thuringiensis é uma bactéria de grande interesse agrícola pela sua habilidade inseticida natural contra importantes pragas agrícola. Diferencia-se das demais espécies do grupo pela presença dos cristais parasporais constituídos de proteínas Cry ou Cyt (Figura 2) (SCHNEPF *et al.*, 1998).

Figura 2. Microscopia Eletrônica de Varredura da mistura esporo-cristais da estirpe de *Bacillus thuringiensis* S234.



cd: cristal bipiramidal, **cc**: cristal cuboide, **ep**: esporo, Aumento de 20.000X).
Fonte: PRAÇA *et al*., 2004.

B. mycoides e *B. pseudomycoides* apresentam como principal característica crescimento colonial do tipo rizoide e ausência de motilidade (NAKAMURA & JACKSON, 1995; NAKAMURA 1998). *B. cytotoxicus* é termotolerante e ocasionalmente associado a intoxicação alimentar por apresentar citotoxina K (GUINEBRETIÈRE *et al.*, 2012). *B. weihenstephanensis* é uma espécie psicrotolerante caracterizada pela presença do gene *cspA* (*cold-shock protein*) e pode ser um contaminante de produtos lácteos e alimentos refrigerados (SOUFIANE & CÔTÉ, 2013).

2.2.2 Grupo do *B. subtilis*

O grupo do *B. subtilis* compreende oito espécies estreitamente relacionadas, *B. subtilis* subsp. *subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. atrophaeus*, *B. mojavensis*, *B. vallismortis*, *B. subtilis* subsp. *spizizenii* e *B. sonorensis* (WANG *et al.*, 2007). As espécies desse grupo não são patogênicas e apresentam grande potencial biotecnológico (STRESHINSKAYA *et al.*, 2011). As principais aplicações na indústria são a produção de enzimas, antibióticos e vitaminas (JEYARAM *et al.*, 2011).

2.3 Enzimas quitinolíticas

A quitina é uma fonte de carboidrato natural renovável abundante encontrada em invertebrados marinhos, fungos, algas e insetos (PATIL *et al.*, 2000). Formada por resíduos de N-acetilglucosamina unidos por ligações β-1,4, apresenta-se como longas cadeias não ramificadas dispostas de forma paralela (β-quitina), antiparalelas (α-quitina) ou uma mistura das duas (λ-quitina) (KUMIRSKA *et al*., 2011). Em muitos organismos a quitina apresenta-se associada a outros componentes estruturais como glucanas na parede celular de fungos, proteínas na membrana peritrófica dos insetos e podem apresentar diversos graus de mineralização devido a interação entre compostos fenólicos e lipídicos (GOODAY, 1990).

A hidrólise enzimática da quitina é realizada pelas quitinases (E.C. 3.2.1.14). Estas enzimas são codificadas por uma grande variedade de organismos e exibem diferentes especificidades pelo substrato (AAM *et al*., 2010). Em bactérias, as quitinases exercem funções de nutrição e parasitismo, enquanto que nos fungos e invertebrados estão envolvidas com a morfogênese. Em plantas e vertebrados, as quitinases estão relacionadas com os mecanismos de defesa e nos vírus, como Baculovírus, são consideradas como fator de virulência (KRAMER & MUTHUKRISHNAN, 1997; THOMAS *et al.*, 1998).

As quitinases bacterianas estão agrupadas nas famílias GH18 e GH19. A família GH18 apresenta um domínio catalítico na forma de barril $(\alpha/\beta)_8$, formado por oito folhas beta e oito fitas alfa e a reação catalítica se caracteriza pelo mecanismo de retenção. As quitinases GH19 foram descritas como sendo similares ao modo de ação da lisozima e quitosanase (CHANDRASEKARAN *et al.*, 2012). As quitinases GH18 geralmente apresentam uma estrutura multi-modular, com diferentes rearranjos dos domínios funcionais. Esses domínios funcionais mostram-se envolvidos na ligação com o substrato e exercem importante função na

hidrólise de substratos insolúveis. Eles também podem influenciar no modo de hidrólise, bem como o tamanho dos produtos que estas enzimas produzem (KUDAM *et al*., 2011).

As bactérias quitinolíticas geralmente produzem múltiplas quitinases, provenientes de diferentes genes, que agem de modo cooperativo. As endoquitinases atuam de modo randômico, clivando as ligações glicosídicas internas ao longo do polímero. As exoquitinases atuam sequencialmente a partir das extremidades redutoras e não-redutora do polímero. Os produtos finais da catálise são multímeros de N-acetil-glucosamina solúveis e de baixo peso molecular, sendo quitotetrose, quitotriose e quitobiose os oligossacarídeos predominantes (KRAMER & MUTHUKRISHNAN, 1997).

Os produtos de degradação da quitina, oligossacarídeos e N-acetilglucosamina são de grande interesse para a indústria farmacêutica e de alimentos (SONGSIRIRITTHIGUL *et al.*, 2009) e as quitinases despertam especial interesse na agricultura como agentes para o controle biológico de insetos e fungos patogênicos de plantas (CHANDRASEKARAN *et al.*, 2012).

O sucesso da utilização das quitinases em diferentes aspectos, depende do desenvolvimento de enzimas altamente ativas em processos industriais economicamente viáveis. Assim, o isolamento de microrganismos quitinolóticos, a caracterização bioquímica e molecular das enzimas quitinolíticas, suas relações filogenéticas e o seu melhoramento por engenharia genética, irão torná-las mais eficientes em uma variedade de processos em um futuro próximo.

3 Justificativa

Com o aumento da conscientização mundial em relação à necessidade de preservação dos recursos naturais e utilização de fontes de energia renováveis, a biocatálise é uma área da biotecnologia que vem se destacando, na qual enzimas são utilizadas como uma alternativa ambientalmente correta aos processos industriais tradicionais.

As enzimas podem ser utilizadas na indústria, isoladas ou sendo produzidas *in situ* por microrganismos. As vantagens da biodegradação enzimática e/ou biológica é a utilização de condições mais brandas de operação, elevado controle estéreo-seletivo e reduzida geração de subprodutos. Porém, há pontos que necessitam ser aprimorados como o baixo desempenho e microrganismos pouco adaptados às necessidades industriais, a falta de enzimas para aplicações industriais específicas, a perda de atividade catalítica ao longo do tempo e elevado tempo de reação.

As enzimas mais utilizadas na indústria são hidrolases, enzimas que catalisam a quebra de polímeros utilizando água. Entre as hidrolases encontram-se as celulases e quitinases, capazes de degradar os dois principais biopolímeros estruturais renováveis do planeta, celulose e quitina.

Assim, a busca por novos microrganismos secretores de celulases e quitinases na diversidade de ambientes complexos tornou-se um ponto importante para o desenvolvimento da biotecnologia e em especial na área de biocatálise.

4 Objetivos

4.1 Objetivo geral

- Isolar novas bactérias com atividade hidrolítica sobre CMC, a partir de uma amostra complexa enriquecida com material lignocelulósico e buscar, no genoma dos isolados, genes que codifiquem enzimas hidrolíticas para clonagem, expressão heteróloga e caracterização bioquímica.

4.2 Objetivos específicos

- Identificar os microrganismos predominates na amostra do composto a partir da análise da sequência do gene 16S rRNA;
- Isolar as bactérias com atividade celulolítica do composto por cultivo seletivo;
- Sequenciar o genoma das bactérias isoladas;
- Buscar no genoma das bactérias isoladas genes que codifiquem para enzimas de interesse;
- Amplificar e clonar os genes de interesse;
- Expressar e purificar as enzimas selecionadas;
- Caracterizar as enzimas bioquimicamente;

5 CAPÍTULO I

Metagenomic analysis of the bacterial population of an agricultural waste composting site.

Ana Paula Andreazza[1], Helisson Faoro[2], Michelle Tadra-Sfeir[1], Emanuel M. de Souza, Fábio de Oliveira Pedrosa[1], Leda Satie Chubatsu[1*].

[1]Department of Biochemistry and Molecular Biology, Federal University of Paraná,
Cx. P. (PO Box) 19046; Curitiba 81531-980, Parana, Brazil.
[2] Carlos Chagas Institute, Fiocruz-PR, Curitiba, Paraná, Brazil.
*Author for correspondence: Phone: +55 (41) 3361 1659; Fax: +55 (41) 3266 2042; E-mail: Chubatsu@ufpr.br

Abstract

Composting is a process that is characterized by a series of complex microbial communities that produce different enzymes that act in a synergistic and specific manner in the decomposition of the organic matter. In order to access this biotechnological potential, the metagenomic analysis of a compost rich in lignocellulosic material was conducted in this study with the purpose of determining its prokaryotic communities and gene content involved in the depolymerization of cellulose, hemicellulose and lignin. Proteobacteria from the gamma class, represented by the *Cellvibrio* genus, was the predominant phylum in the sample, followed by phylum *Bacteroidetes*, represented by the *Ohtaekwangia* genus. Bacteria that are capable of degrading aromatic compounds, interacting with the plant and fixing nitrogen are among the dominant communities. *Methanobacterium* and *Methanocella* were the genera that were identified as belonging to the group of methanogenic archaeas. An analysis of the Operational Taxonomic Units (OTUs) revealed that it was not possible to classify 16% of the sequences under the *Bacteria* kingdom, indicating the presence of non-culturable bacteria or bacteria that cannot be grouped together with sequences of known bacterial phyla.

Keywords: Metagenome, Compost, Lignocellulose

1.Introduction

The industrial conversion of lignocellulose into fuel and chemicals is a vital component of the set of renewable energies that can be used to reduce greenhouse gas emissions and the use of fossil fuels (RESCH *et al*., 2013).

Lignocellulose is found in the cell walls of plants and it is made up of three main polymers. Cellulose, which is the main structural polymer (35 to 50% of the biomass), made up of glucose monomers linked by β-1,4 bonds. Lignin (15 to 30% of the biomass), an amorphous polymer that is aromatic in nature (phenylpropanoid) and which gives strength and rigidity to the assembly of cellulose fibers, and hemicellulose (25 to 30% of the dry weight), a heterogeneous polymer made up of pentoses and hexoses, which acts as a connection agent between the cellulose and the lignin (CASTRO, 2009; GROSS, 2012). These three components form a structure that is highly resistant to structural and chemical decomposition, known as recalcitrant.

For production of cellulosic ethanol, it is necessary to deconstruct the plant cell wall, by means of thermochemical methods and/or enzymatic hydrolysis, in order to access the fermentable sugars from cellulose. This process, known as pre-treatment, produces a large number of byproducts (acetate, furfural and lignin), which significantly affect the growth and metabolism of the fermenting microorganisms (MACHADO, 2008).

Currently, the greatest difficulties for the use of biomass in biofuel production are the high cost of the process, the inefficiency of hydrolytic enzymes and microorganisms slight adapted to the lignocellulosic material (LEE *et al*., 2012). Thus, various studies have been developed with the aim of finding enzymes that are more efficient and better adapted for industrial requirements (such as thermal stability, tolerance to organic solvents, pH and salinity) and also isolating microorganisms resistant to the characteristics of the feedstock.

In nature, the cellulosic biomass is decomposed by complex and efficient processes, usually by microbial communities that produce cellulolytic enzymes that act synergistically in the deconstruction of the plant biomass (van der LELIE *et al.*, 2012). To access the genetic potential of these microorganisms, the metagenomic approach makes it possible to predict the biochemical potential of a particular community or prospecting for new biocatalysts of environmental samples, thereby overcoming the limitations of the methodologies based on culture (LORENS & ECK, 2005).

The use of metagenomics for the exploration of environments enriched with lignocellulolytic microorganisms, such as forest soil (KANOKRATANA *et al.*, 2010), the termites digestive system (WARNECKE *et al.*, 2007), rumen of cattle (HESS *et al*, 2011; DAI *et al*, 2012) and composting systems (ALLGAIER *et al*, 2010; DOUGHERTY *et al*. 2012; MARTINS *et al*, 2013;.) revealed a wide range of untapped microorganisms and their metabolic potential for the degradation of plant biomass.

Composting is a process that simulates the degradation of organic matter in nature. The organic matter is decomposed through a complex process involving a cooperative action of microbial communities that produce enzymes (amylases, proteases, cellulases, pectinases, lipases, etc.) successively throughout the process as the reaction conditions (temperature, pH, moisture, nutrient deposition) are modified as a result of the metabolic activity of these communities (WONGWILAIWALIN *et al*., 2013). The bacteria are the predominant microorganisms in the process because of their higher metabolic diversity represented by the variety of enzymes (MARTINS *et al.*, 2013).

Several enzymes have been isolated and characterized through metagenomic libraries of compost, such as amylase (ARIKAN, 2008), cellulase (PANG *et al.*, 2009), β-glucosidase (Uchiyama et al., 2013) and xylanase (VERMA *et al.*, 2013) indicating that composting is a promising environment for the exploration of hydrolytic enzymes and that the enzymatic profile of each composting environment depends on the material to be degraded and on the process plant (PARTANEN *et al.*, 2010).

In this study, a metagenomic approach was carried out to access the taxonomic composition and genetic diversity of an agricultural compost rich in lignocellulosic material, with the aim of finding new enzymes involved in the depolymerization of cellulose, hemicellulose and lignin and improving the conversion of plant biomass into biofuel.

2.Materials and Methods

2.1. Sample collection and DNA extraction

The samples were collected during the maturation phase of a compost pile made of agricultural waste (corn silage, sawdust and cattle manure) in a rural property in the municipality of Tupãsi in the state of Paraná, Brazil.

The sampling method used was intended to obtain a representative sample of the entire pile. Five sub-samples weighing approximately 200 g were collected from the internal

parts (about 20cm from the surface) and external parts of the pile (right and left sides, approximately 5 cm from the surface). These samples were mixed, packaged in plastic bottles and transferred to the lab for the analyses.

The DNA of the sample was extracted using the *Powermax Soil DNA Isolation Kit* (MO BIO Laboratories) according to the manufacturer's instructions.

2.2. Preparation of the metagenomic libraries of the 16S rDNA gene.

The sequence of the 16S rDNA gene was amplified by polymerase chain reaction (PCR) of the total DNA of the compost sample. Primers 515F (5′- GTGCCAGCMGCCGCGGTAA - 3′) and 805R (3′- GGACTACHVGGGTWTCTA AT- 5′) (CAPORASO *et al*., 2011) were used to allow the amplification of a DNA fragment of approximately 290 base pairs that correspond to the V4 region of the 16S rDNA gene of *Archaea* and *Bacteria.*

A reaction system was used for 20μL PCR, consisting of 2U of the Taq DNA polymerase enzyme (Invitrogen), 2μL 10X enzyme buffer, 10pmol of each primer, 200μmol/l of each deoxynucleotide triphosphate (dATP, dCTP, dTTP, and dGTP), 1.5mmol/L of $MgCl_2$ and 30ng of DNA extracted from the compost sample. The thermal cycles used were an initial denaturation of 94°C for 3 minutes followed by 25 cycles of denaturation at 94°C for 45 seconds, annealing at 50°C for 60 seconds and extension at 72°C for 60 seconds, with a final cycle of 72°C for 10 minutes. The PCR reaction was performed in a thermocycler Eppendorf Master Cycler Gradient 5331.

2.3. Sequencing of 16S rDNA and bioinformatics analyses

The PCR products were purified with *UltraClean® PCR Clean-Up Kit* (MO BIO Laboratories) and quantified using *Qubit fluorometer* (Invitrogen, Grand Island, NY). The samples were sequenced using the Illumina MiSeq platform.

The sequences were processed using the QIIME (*Quantitative Insights Into Microbial Ecology*) program (CAPORASO *et al,* 2010). Similar sequences were assembled into operational taxonomic units (OTUs) and the limit of identity between the sequences was 97%. The taxonomy was noted for each OTU using the SILVA database with a minimum of 80% identity.

3.Results and Discussion

3.1. General characteristic of the 16S rRNA library

A total of 17,486 high quality sequences were generated. Most sequences are from prokaryotes of the *Bacteria* kingdom (99.8%) and very few represent microorganisms from the *Archaea* kingdom (0.02%). Some sequences were not classified (0.18%) for bearing no similarity to any known sequence of the 16S rRNA gene in the *Bacteria* or *Archaea* kingdoms. In the sample of the compost, 313 OTUs were found and, of these, 207 were identified to the genus level.

3.2. Composition of the microbial community

The major phyla represented in the compost sample were: *Proteobacteria* (52.57%), *Bacteroidetes* (20.54%), *Firmicutes* (3.72%), *Verrucomicrobia* (3.08%), *Acidobacteria* (1.09%) and *Actinobacteria* (0,5%). 16.29% of the 16S rRNA sequences were not identified within the *Bacteria* kingdom (Figure 1). Other phyla, such as *Armatimonadetes, Chlamydiae, Chloroflexi, Chlorobi, Deferribacteres, Elusimicrobia, Lentisphaerae, Nitrospira, Planctomycetes, Spirochaetes, Tenericutes* and the provisional phyla OD1, BRC1 and TM7 were underrepresented in the sample (0.01 to 0.37% of the sequences).

Proteobacteria was the predominant microbial community in the sample with 104 genera identified. This prevalence contrasts with previous studies where *Proteobacteria* were the smallest communities that were part of a compost, fed with municipal waste and residues from zoo animals (PARTANEN *et al*., 2010; MARTINS *et al*., 2013) but it is in line with the results found by Gannes *et al*. (2013) in the maturation phase of lignocellulosic composts. Proteobacteria were also part of prokaryotic communities that were predominant in the anaerobic decomposition of wood in bioreactors (van der LELIE *et al*., 2012).

Proteobacteria were distributed between the *Gamaproteobacteria* (30.70%), *Alfaproteobacteria* (24.34%), *Betaproteobacteria* (19.64%) and *Deltaproteobacteria* (11.75%) classes, with a small predominance of the gamma class (Figure 2). In the *Gamaproteobacteria* class, 30 genera were identified and *Cellvibrio* was the most prominent genus. Other genera that were detected were *Pseudoxanthomonas, Stenotrophomonas, Steroidobacter and Acinetobacter* (Figure 3). These bacteria are found in the soil or in association with plants that show great metabolic versatility. *Cellvibrio* is a cellulolytic bacterium isolated from the soil (MERGAERT *et al.*, 2003). *Pseudoxanthomonas* is characterized by the ability to reduce nitrate

(WEON *et al.*, 2006). *Stenotrophomonas* has a great biotechnological potential for the promotion of plant growth, plant pathogens control, bioremediation and phytoremediation (RYAN *et al.*, 2009). *Steroidobacter* has the ability to degrade steroid hormones (WANG *et al.*, 2013). *Acinetobacter* has phytostimulation activity based on hormone production, phosphate solubilization and production of siderophores (ROKHBAKHSH-ZAMIN *et al*., 2011).

Among the *Alphaproteobacteria*, of the 40 genera identified, the most prominent ones were *Rhizomicrobium, Sphingomonas, Sphingopyxis Devosia, Phenylobacterium, Caulobacter, Magnetospirillum* and *Azospirillum* (Figure 3). Bacteria such as *Rhizomicrobium, Sphingomonas, Devosia* and *Azospirillum* were identified in association with plant roots (UEKI *et al*., 2010; RIVAS *et al.*, 2002). *Sphingopyxis, Phenylobacterium and Magnetospirillum* have the ability to degrade herbicides, polyethylene glycol, and aromatic compounds, such as toluene, phenol and benzoate respectively (YOSHIYUKI *et al.*, 2015). The study carried out by Lelie *et al*. (2012) *Magnetospirillum* was the predominant organism in the anaerobic decomposition of wood and would be responsible for the degradation of toxic aromatic compounds resulting from the depolymerization of lignin, preventing microbial communities from collapsing due to the accumulation of toxic substances.

The most abundant *Betaproteobacteria* were nitrogen-fixing bacteria, *Azospira* and *Azoarcus* (Figure 3). Among the *Deltaproteobacteria*, the major genera were *Geobacter* and *Byssovorax* (Figure 3). *Geobacter* is a very important anaerobic organism for environmental bioremediation because of its ability to couple the oxidation of organic compounds to the reduction of insoluble Fe(III) and Mn(IV) oxides, to oxidize aromatic hydrocarbons and to precipitate uranium and related compounds by reduction (LOVLEY *et al*., 2011). *Byssovorax was* an organism isolated from the soil and characterized as a degrader of crystalline cellulose (REICHENBACH *et al.*, 2006). *Azoarcus* spp. CIB and *Geobacter mettalireducens* BamVW are capable of degrading aromatic compounds anaerobically (MARTIN-MOULD *et al*., 2015; JUAREZ *et al.*, 2010).

Members of the *Bacteroidetes* phylum are associated with the degradation of complex polymers including cellulose and chitin (MANZ *et al.*, 1996). The predominant genera were *Ohtaekwangia, Paludibacter, Terrimonas* and *Leadbetterella.* The bacteria *Ohtaekwangia* (YOON *et al.*, 2011), *Flavobacterium* and *Leadbetterella* (COUSIN *et al.*, 2007; ABT *et al*., 2011) are derived from the aquatic environment. *Paludibacter* was identified in the gut of

termites (MAKONDE *et al.*, 2015) and some species are diazotrophic (INOUE *et al.*, 2015). *Terrimonas* was isolated from soil, sediments and water .

The *Verrucomicrobia* phylum was a bacterial phylum hardly detected when culture-dependent techniques were used. However, they are often found in metagenomic libraries of the 16S rRNA gene of a wide variety of ecosystems, including water, land and intestinal environments (FREITAS *et al.*, 2012). In the sample, this phylum is represented by genera *Opitutus* and *Subdivision 3*. *Opitutus* are organisms that are adapted to grow in polysaccharides derived from plants (PASSEL *et al.*, 2011) and little is known about the biology of the *Subdivision3* genus.

The *Firmicutes* phylum was mainly represented by the *Clostridia* class, reflecting the thermophilic and semi-static conditions of the composting environment that was studied. Of the 21 genera identified in this phylum, the most prominent ones were *Clostridium III, Acetanaerobacterium* and *Desulfosporosinus.* The *Clostridium* genus is extremely heterogeneous and it was divided by Collins et al (1994) into five groups. Group III comprises efficient cellulose degraders, such as *Clostridium cellulolyticum* and *C. thermocellum. Acetanaerobacterium* was isolated from the residual water in a paper mill (CHEN & DONG, 2004). *Desulfosporosinus* are sulfate-reducing organisms that can be applied in bioremediation. They are capable of degrading toluene, reducing Fe(III) and growing in an acid environment (PESTER *et al.*, 2012).

The *Acidobacteria* phylum was represented by the GP1, GP3 and GP6 genera. These genera were found in soil samples and in the rhizosphere of grasses, potatoes and garlic (ROCHA *et al.*, 2013).

The quantity of sequences found of the *Actinobacteria* phylum was very small (0.5%). This result contrasts with previous studies describing *Actinobacteria* as the most abundant organism in the compost, particularly in the maturation stage (REN *et al.*, 2016). The main representatives of this phylum were *Cellulomonas* and *Conexibacter. Cellulomonas* were already isolated from a compost and their main characteristic is their ability to degrade cellulose (KANG *et al.*, 2007).

The *Archaea* kingdom accounted for a small fraction of the total prokaryotic community of the ABP2 sample, recovering only *Euryarchaeota* sequences. The identified genera were *Methanobacterium* and *Methanocella. Methanobacterium* is an Archaea that is predominant in compost (THUMMES *et al*., 2007). *Methanocella* was isolated from rice fields and it is the only described methanogenic *Archaea* that is able to withstand oxygen (LU & LU, 2012).

4.Conclusion

The metagenomic analysis is an important tool to investigate the microbial ecology of several environments and a useful tool for accessing the genetic diversity for biotechnology applications.

The results indicate that the composition of the prokaryotic communities found in the ABP2 sample is different from the microbial communities already identified in previous studies involving composting environments.

The semi-static characteristic of the cell allowed the development of anaerobic and aerobic microorganisms involved in the degradation of plant biomass and interaction with the plant. The presence of organisms capable of degrading aromatic compounds (*Azoarcus*, *Geobacter*, *Sphingopyxis, Desulfosporosinus, Magnetospirillum*), degrading cellulose (*Clostridium, Cellulomonas, Byssovorax, Paludobacter*), interacting with plants (*Stenotrophomonas*, *Acinetobacter*, *Rhizomicrobium, Sphingomonas, Devosia, Azospirillum, GP1, GP3, GP6*), fixing nitrogen (*Azoarcus, Azospira, Azospirillum)* to differentiate themselves (*Caulobacter*) makes the sample a promising source for the prospecting for enzymes involved in the degradation of biomass and the role of microorganisms associated with plants in the deconstruction of lignocellulose.

References

Abt B, Teshima H, Lucas S, Lapidus A, Rio TGD, Nolan M, Tice H, Cheng JF, Pitluck S, Liolios K, Pagani I, Ivanova I, Mavromatis K, Pati A, Tapia R, Han C, Goodwin L, Chen A, Palaniappan K, Land M, Hauser L, Chang YJ, Jeffries CD, Rohde M, Göker M, Tindall BJ, Detter JC, Woyke T, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Klenk HP, Kyrpides NC, 2011. Complete genome sequence of Leadbetterella byssophila type strain (4M15^T). Stand Genomic Sci. 4, 2-12.

Allgaier M, Reddy A, Park JI, Ivanova N, D'haeseleer P, et al., 2010. Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. PloS ONE 5, e8812.

Amore A, Pepe O, Ventorino V, Birolo L, Giangrande C, Faraco V, 2012. Cloning and recombinant expression of a cellulase from the cellulolytic strain Streptomyces sp. G12 isolated from compost. Microb. Cell Fac. 11,164-176.

Arikan B, 2008. High thermostable, thermophilic, alcaline, SDS and chelator resistant amylase from thermophilic *Bacillus* sp. isolate A3-15. Bioresour. Technol. 99, 3071 – 3076.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, et al., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 7, 335-336.

Castro HF, 2009. Processos Químicos Industriais II: Papel e Celulose. Universidade de São Paulo, Escola de Engenharia de Lorena – EEL, 30 p.

Chen S, Dong X, 2004. Acetanaerobacterium elongatum gen. nov., sp. nov., from paper mill waste water**.** Int. J. Syst. Evol. Microbiol. 54, 2257–2262.

Collins MD, Lawson PA, Willems A, Cordoba JJ, Fernandez-Garayzabal J, Garcia P, CaiI J, Hippe H, Farrow JAE, 1994. The phylogeny of the genus Clostridium: proposal of five new genera and eleven new species combinations. I J Syst Bacteriol. 44, 812-826.

Cousin S, Päuker O, Stackebrandt E, 2007. *Flavobacterium aquidurense* sp. nov. and *Flavobacterium hercynium* sp. nov., from a hard-water creek. International Journal of Systematic and Evolutionary Microbiology. 57, 243-249.

Dai X, Zhu YX, Luo YF, Song L, Liu D, et al., 2012. Metagenomic insights into the fibrolytic microbiome in yak rumen. PloS ONE. 7: e40430

De Angelis KM, Gladden JM, Allgaier M, D'haeseleer P, Fortney JL, et al., 2010. Strategies for enhancing the effectiveness of metagenomic-based enzyme discovery in lignocellulolytic microbial communities. Bioenergy Research 3: 146–158.

Dougherty MJ, D'haeseleer P, Hazen TC, Simmons BA, Adams PD, Hadi MZ, 2012. Glycoside Hydrolases from a targeted Compost Metagenome, activity-screening and functional characterization. BMC Biotechnology. 12:38.

Freitas S, Hatosy S, Fuhrman JA, Huse SM, Welch DBM, Sogin ML, Martiny AC, 2012. Global distribution and diversity of marine Verrucomicrobia. ISME J. 6, 1499–1505.

Gannes V, Euudoxie G, Hickey WJ, 2013. Prokaryotic successions and diversity in composts as revealed by 454-pyrosequencing. Bioresour Technol. 133, 573-580.

Gross AS, 2012. The solvente mediated thermodynamics of celulose deconstruction. Dissertation of Chemical Engineering - University of California, Berkeley.

Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, et al., 2011. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science. 331: 463–467.

Hugenholtz P, Goebel BM, Pace NR, 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. J. Bacteriol. 180, 4765–4774.

Inoue JI, Oshima K, Suda W, Sakamoto M, Iino T, Noda S, Hongoh Y, Hattori M, Ohkuma M, 2015. Distribution and Evolution of Nitrogen Fixation Genes in the Phylum *Bacteroidetes*. Microbes Environ. 30, 44-50.

Juarez JF, Zamarro MT, Barragan MJL, Blazquez B, Boll M, et al., 2010. Identification of the Geobacter metallireducens BamVW Two-Component System, Involved in Transcriptional Regulation of Aromatic Degradation. Appl Env Microbiol 76, 383–385.

Kang MS, Im WT, Jung HM, Kim MK, Goodfellow M, Kim KK, Yang HC, An DS, Lee ST, 2007. *Cellulomonas composti* sp. nov., a cellulolytic bacterium isolated from cattle farm compost. Int J Syst Evol Microbiol 57, 1256–1260.

Kanokratana P, Uengwetwanit T, Rattanachomsri U, Bunterngsook B, Nimchua T, Tangphatsornruang S, Plengvidhaya V, Champreda V, Eurwilaichitr L, 2010. Metagenomic analysis on phylogeneticsand metabolic potential of microbial community in primary tropical peat swamp forest. Microb Ecol. 61, 518–528.

Lee DS, Lee KH, Jung S, Jo EJ, Han KH, Bae HJ, 2012. Synergistic effects of 2A-mediated polyproteins on the production of lignocellulose degradation enzymes in tobacco plants. J. Exp. Bot. 63, 4797–4810.

Lorenz P, Eck J, 2005. Metagenomics and industrial applications. Nature. 3, 510–516.

Lovley DR, Ueki T, Zhang T, Malvankar NS, Shrestha PM, Flanagan KA, Aklujkar M, Butler JE, Giloteaux L, Rotaru AE, Holmes DE, Franks AE, Orellana R, Risso C, Nevin KP, 2011. Geobacter: The Microbe Electric's Physiology, Ecology, and Practical Applications. Adv. Microb. Physiol. 59, 1-100.

Lü Z, Lu Y, 2012. Complete Genome Sequence of a Thermophilic Methanogen, *Methanocella conradii* HZ254, Isolated from Chinese Rice Field Soil. J. Bacteriol.194, 2398–2399.

Machado CMM, 2008. Etanol de 2ª. Geração – desafios para instrumentação e automação. Embrapa Agroenergia. Disponível em: https://www.agencia.cnptia.embrapa.br/Repositorio/Etanol+lignocelul%25F3sico1_000g7ia08fk02wx5ok0wtedt39uhi9tg.pdf

Makonde HM, Mwirichia R, Osiemo Z, Boga HI, Klenk HP, 2015. 454 Pyrosequencing-based assessment of bacterial diversity and community structure in termite guts, mounds and surrounding soils. SpringerPlus. 4, 471-482.

Manz W, Amann R, Ludwig W, Vancanneyt M, Schleifer K, 1996. Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum Cytophaga-Flavobacter-Bacteroides in the natural environment. Microbiology (SGM) 42, 1097–1106.

Martín-Moldes Z, Zamarro MT, Del Cerro C, Valencia A, Gómez MJ, Arcas A, Udaondo Z, Garcia JL, Nogales J, Carmona M, Díaz E, 2015. Whole genome analysis of *Azoarcus* sp. strain CIB provides genetic insights to its different lifestyles and predicts novel metabolic features. Syst Appl Microbiol. 38, 462-471.

Martins LF, Antunes LP, Pascon RC, Oliveira JCF, Digiampietri LA, et al.,2013. Metagenomic Analysis of a Tropical Composting Operation at the São Paulo Zoo Park Reveals Diversity of Biomass Degradation Functions and Organisms. PloS ONE 8, e61928

Mergaert J, Lednicka D, Goris J, Cnockaert MC, De Vos P, Swings J, 2003. Taxonomic study of Cellvibrio strains and description of Cellvibrio ostraviensis sp. nov., Cellvibrio fibrivorans sp. nov. and Cellvibrio gandavensis sp. nov. Int. J. Syst. Evol. Microbiol. 53, 465–471.

Pang H, Zhang P, Duan CJ, Mo XC, Tang JL, Feng JX, 2009. Identification of celulase genes from the metagenomes of compost solis and functional characterization o fone novel endoglucanase. Curr. Microbiol. 58, 404-408.

Partanen P, Hultman J, Paulin L, Auvinen P, Romantschuk M, 2010. Bacterial diversity at different stages of the composting process. BMC Microbiology. 10,94

Passel MWJ, Kant R, Palva A, Copeland A, Lucas S, Lapidus A, Rio TG, Pitluck S, Goltsman E, Clum A, Sun H, 2011. Genome Sequence of the Verrucomicrobium Opitutus terrae PB90-1, an Abundant Inhabitant of Rice Paddy Soil Ecosystems. J. Bacteriol. 193, 2367–2368.

Pester M, Brambilla E, Alazard D, Rattei T, Weinmaier T, Han J, Lucas S, Lapidus A, Cheng JF,Goodwin L, Pitluck S, Peters L, Ovchinnikova G, Teshima H, Detter JC, Han CS, 2012. Complete Genome Sequences of *Desulfosporosinus orientis* DSM765T, *Desulfosporosinus youngiae* DSM17734T, *Desulfosporosinus meridiei* DSM13257T, and *Desulfosporosinus acidiphilus* DSM22704T**.** J. Bacteriol. 194, 6300–6301.

Reichenbach H, Lang E, Schumann P, Spro C, 2006. *Byssovorax cruenta* gen. nov., sp. nov., nom. rev., a cellulose-degrading myxobacterium: rediscovery of '*Myxococcus cruentus*' Thaxter 1897. Int. J. Syst. Evol. Microbiol. 56, 2357–2363.

Ren G, Xu X, Qu J, Zhu L, Wang T, 2016. Evaluation of microbial population dynamics in the co-composting of cow manure and rice straw using high throughput sequencing analysis. World J Microbiol Biotechnol. 32, 101.

Resch MG, Donohoe BS, Baker JO, Decker SR, Bayer EA, Beckham GT, Himmel ME, 2013. Fungal cellulases and complexed cellulosomal enzymes exhibit synergistic mechanisms in cellulose deconstruction. Energy Environ. Sci. 6,1858-1867.

Rivas R, Velazquez E, Willems A, Vizcaino N, Subba-Rao NS, Mateos PF, Gillis M, Dazzo FB, Martinez-Molina E, 2002. A new species of Devosia that forms a unique nitrogen-fixing root-nodule symbiosis with the aquatic legume Neptunia natans (L.f.) Druce. Appl. Environ. Microbiol. 68, 5217–5222.

Rocha UN, Plugge CM, George I, van Elsas JD, van Overbeek LS, 2013. The Rhizosphere Selects for Particular Groups of *Acidobacteria* and *Verrucomicrobia.* PLoS ONE 8: e82443.

Rokhbakhsh-Zamin F, Sachdev D, Kazemi-Pour N, Engineer A, Pardesi KR, Zinjarde S, Dhakephalkar PK, Chopade BA, 2011. Characterization of Plant-Growth-Promoting Traits of Acinetobacter Species Isolated from Rhizosphere of Pennisetum glaucum. J. Microbiol. Biotechnol. 21, 556–566.

Ryan RP, Monchy S, Cardinale M, Taghavi S, Crossman L, Avison MB, Berg G, Lelie DV, Dow JM, 2009. The versatility and adaptation of bacteria from the genus Stenotrophomonas. Nat. Rev. Microbiol. 9, 514-525.

Sambrook J, Fritsch EF, Maniatis T, 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Thummes K, Kämpfer P, Jäckel U, 2007. Temporal change of composition and potential activity of the thermophilic archaeal community during the composting of organic material. Syst Appl Microbiol. 30, 418–429.

Uchiyama T, Miyazaki K, Yaoi K, 2013. Characterization of a novel β-glucosidase from compost microbial metagenome with Strong transglycosylation activity. J Biol Chem. 288, 18325 -18334.

Ueki A, Kodama Y, Kaku N, Shiromura T, Satoh A, Watanabe K, Ueki K, 2010. Rhizomicrobium palustre gen. Nov., sp. nov., a facultatively anaerobic, fermentative stalked bacterium in the class alphaproteobacteria isolated from rice plant roots. J. Gen. Appl. Microbiol. 56, 193-203.

van der Lelie D, Taghavi S, McCorkle SM, Li LL, Malfatti SA, Monteleone D, Donohoe BS, Ding S, Adney WS, Himmel ME, Tringe SG, 2012. The metagenome of an anaerobic microbial community decomposing poplar wood chips. PLoS ONE 7: e36740.

Verma D, Kawarabayasi Y, Miyazaki K, Satyanarayana T, 2013. Cloning, Expression and Characteristics of a Novel Alkalistable and Thermostable Xylanase encoding Gene (Mxyl) Retrieved from Compost-Soil Metagenome. PLOS ONE. 8, e52459

Wang PH, Leu YL, Ismail W, Tang SL, Tsai CY, Chen HJ, Kao AT, Chiang YR, 2013. Anaerobic and aerobic cleavage of the steroid core ring structure by *Steroidobacter denitrificans*. J. Lipid Res. 54, 1493-1504.

Warnecke F, Luginbuhl P, Ivanova N, Ghassemian M, Richardson TH, et al., 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature. 450: 560–565.

Weon HY, Kim BY, Kim JS, Lee SY, Cho YE, Go SJ, Hong SB, Im WT, Kwon SW, 2006. *Pseudoxanthomonas suwonensis* sp. nov., isolated from cotton waste composts. Int J Syst Evol Microbiol. 56, 659–662.

Wongwilaiwalin S, Laothanachareon T, Mhuantong W, Tangphatsornruang S, Eurwilaichitr L, Igarashi Y, Champreda V, 2013. Comparative metagenomic analysis of microcosm structures and lignocellulolytic enzyme systems of symbiotic biomass-degrading consortia. Appl Microbiol Biotechnol. 97:8941–8954.

Yoon JH, Kang SJ, Lee SY, Lee JS, Park S, 2011. Ohtaekwangia koreensis gen. nov., sp. nov. and Ohtaekwangia kribbensis sp. nov., isolated from marine sand, deep-branching members of the phylum Bacteroidetes**.** Int. J. Syst. Evol. Microbiol. 61, 1066–1072.

Yoon JH, Lee ST, Park YH, 1998. Inter- and intraspecific phylogenetic analysis of the genus Nocardioides and related taxa based on 16S rDNA sequences. Int. J. Syst. Bacteriol. 48, 187–194.

Yoshiyuki O, Nagata Y, Numata M,Tsuchikane K, Hosoyama A, Yamazoe A, Tsuda M, Fujita N, Kawai F, 2015. Complete genome sequence of *Sphingopyxis macrogoltabida t*ype strain NBRC 15033, originally isolated as a polyethylene glycol degrader. 3, e01401-15

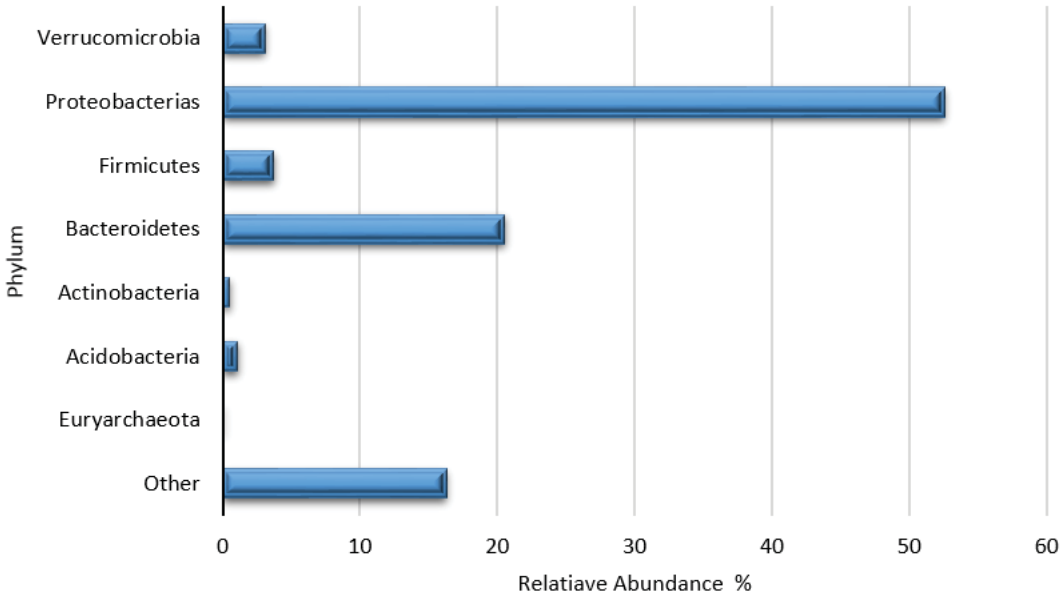


Figure 1. The major bacterial phylum found in the compost sample.

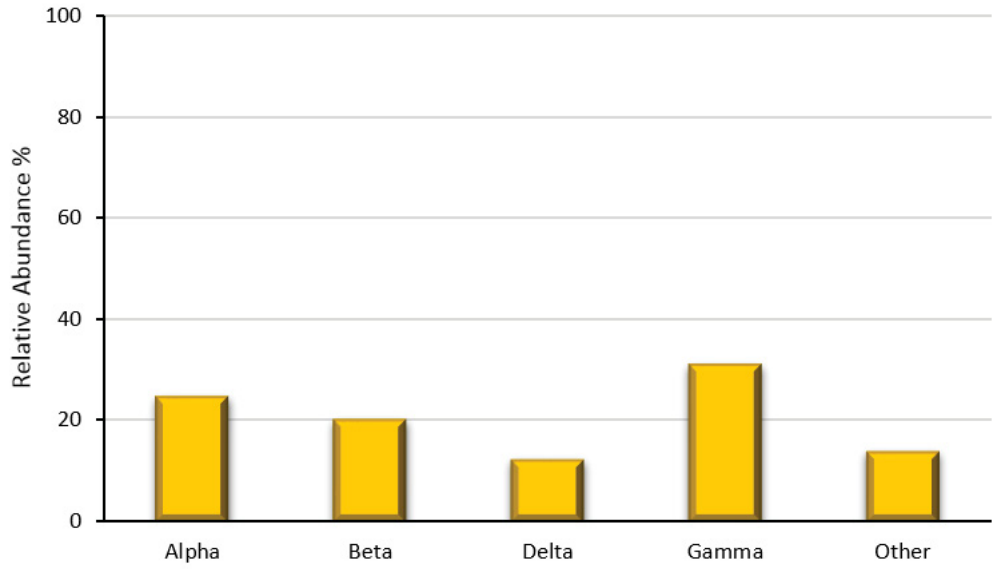


Figure 2. Distribution of different classes of *Proteobacteria* in the compost sample.

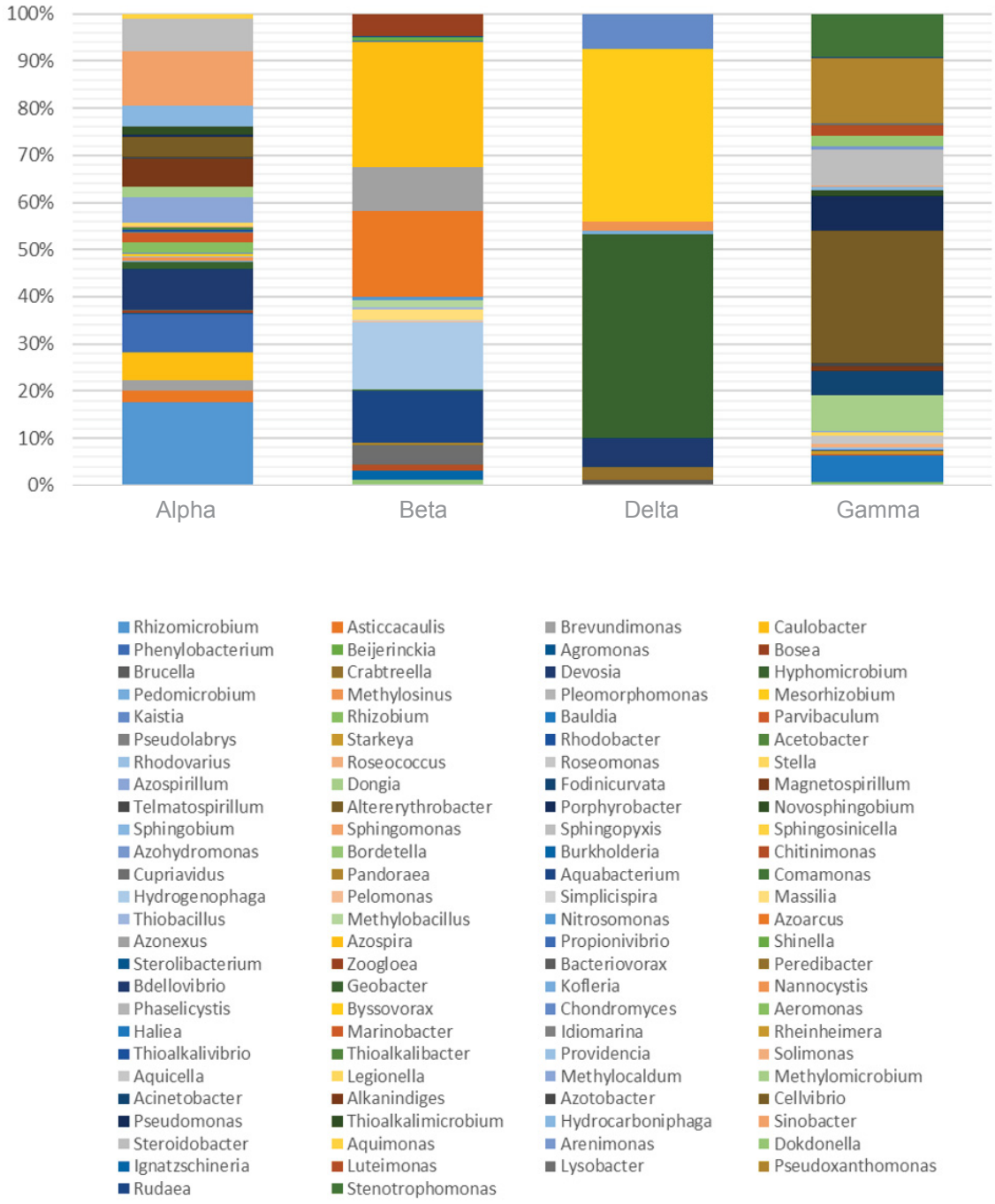


Figure 3. Genus identified in the compost sample belonging to *Proteobacteria* phylum.

6 CAPÍTULO II

Genome analysis of entomopathogenic *Bacillus* sp. ABP14 isolates from lignocellulosic compost

Ana Paula Andreazza[1], Rodrigo L. A. Cardoso[1], Jessica Cocco[2], Dieval Guizelini[3], Helisson Faoro[4], Michelle Tadra-Sfeir[1], Eduardo Balsanelli[1], Leonardo Magalhães Cruz[1], Cyntia M.T. Fadel- Picheth[5], Lucélia Donatti[6], Emanuel M. de Souza[1], Luís Amilton Foerster[2], Fabio de Oliveira Pedrosa[1], Leda Satie Chubatsu[1*].

.

[1]Department of Biochemistry and Molecular Biology, Federal University of Paraná, Cx. P. 19046; Curitiba 81531-980, Paraná, Brazil.

[2]Department of Zoology – Federal University of Paraná, Curitiba, Paraná, Brazil.

[3]Laboratory of Bioinformatics, Professional and Technological Education Sector – Federal University of Paraná, Curitiba, Paraná, Brazil.

[4] Carlos Chagas Institute, Fiocruz-PR, Curitiba, Paraná, Brazil.

[5]Department of Medical Pathology – Federal University of Paraná, Curitiba, Paraná, Brazil.

[6]Department of Cellular Biology - Federal University of Paraná, Curitiba, Paraná, Brazil.

*Author for correspondence: Tel: +55 (41) 3361 1659; Fax: +55 (41) 3266 2042; e-mail: Chubatsu@ufpr.br

Abstract

We report the complete genome of *Bacillus* sp. ABP14 isolated from lignocellulosic compost by its ability in hydrolyzing carboxy-methyl cellulose and was found to produce an insecticidal activity against *Anticarsia gemmtalis* (Lepidoptera). The taxonomic analysis using digital DNA-DNA hybridization (dDDH) revealed that the isolate ABP14 is genetically close to *B. thuringiensis* serovar *finitimus* YBT020, and the phenotypic characterization showed that the ABP14 is a Cry-strain. The comparison of the predicted ABP14 proteome against completed genomes of *B. cereus senso lato* strains using Blast search "all-against-all" identified unique genes related to virulence, metabolic competence, structural components and survival in the insect gut. *Bacillus* sp. ABP14 and *B. thutingiensis finitimus* YBT020 share a set of virulence factors common in *Bacillus cereus stricto sensu*.

Keywords: *Bacillus cereus sensu lato*, insecticidal activity, biological control, complete genome, compost.

1.Introduction

Bacillus spp. comprises Gram-positive, rod-shaped, endospore forming bacteria. Ubiquitous in nature, *Bacillus* species demonstrate great catabolic and biosynthetic versatility and the GC content of 32 to 69% reveal the heterogeneity of the genus. Members of the genus *Bacillus* are able to secrete a wide variety of enzymes and biologically active compounds which are known to be important for environment, medicine, agriculture and industry (Slepecky and Hemphill, 2006; Kunst et al, 1997).

In agriculture, the bacterium *Bacillus thuringiensis* (*Bt*) is the most successful insect pathogen used for insect control (Cawoy et al., 2011). The insects-pest are responsible for destroying one fifth of the world's total crop production annually (FAO http://www.fao.org/3/a-av013e.pdf). *Bt* is active against larval stages of lepidopteran, dipteran and coleopteran insects by the action of insecticidal pore forming proteins known as Cry and Cyt toxins produced during sporulation (Bravo et al., 2011).

The Cry proteins, commonly encoding by c*ry* genes usually located on plasmids, have been divided in four phylogenetically nonrelated protein families: three domains Cry toxins (3D), mosquitocidal Cry toxins (Mtx), the binary-like (Bin) and the Cyt family of toxins (Bravo et al., 2005). In addition to Cry proteins, several other insecticidal factors in *Bt* have been observed, such as heat-stable beta-exotoxin (Liu et al., 2014) and vegetative insecticidal proteins (VIP) (Chakroun et al., 2016).

Despite all insecticide capacity, the *Bt* shows low environmental persistence after topical application limiting their use as a biological control agent (Kupferschmied et al., 2013). To overcome this problem, genetically modified (GM) crops that express variants of the Cry toxins have been developed and commercially used (Bravo et al., 2011). This *Bt*-technology has proved to be a valuable strategy for agricultural pest management with substantial economic and environmental benefits (Pardo-Lópes et al., 2013). However, the discovery that insects can adapt to *Bt*-technology showed the necessity to expand the number of bacterial biocontrol agents and novel proteins with promising insecticidal properties (McGaughey and Whalon, 1992).

Bacillus sp. ABP14 was isolated from lignocellulosic compost and selected by its ability to hydrolyse carboxymethylcellulose. Preliminary screening, ABP14 strain grew on media rich in chitin and protein, showed insecticidal activity against *Lepidoptera* and was not able to produce crystal proteins. In order to investigate the metabolic capacity and the insecticidal activity of *Bacillus* sp. ABP14 the whole-genome sequencing was performed.

2.Material and Methods

2.1. Sample collection

Sample was collected from a pile of agricultural waste (cattle manure, corn silage and sawdust) in over a year semi-static composting process, on a rural estate at municipality of Tupãssi, Paraná, Brazil.

The composite sampling method was used to obtain a representative material of the entire pile. The material was collected from five locations around the pile (5 cm from the surface) and three within the pile (the core of the pile and regions above and below the core). Samples were placed in a clean plastic bucket and mixed thoroughly to ensure homogeneity. The sample was split into four parts and two quarters was taken for analysis. The material was placed in clean plastic bags and transported to the laboratory.

2.2. Screening for cellulose-degrading bacteria

10g of compost sample was added to 190 mL 0.5% CMC Medium (K_2HPO_4 1.6g L^{-1}; KH_2PO_4 0.2g L^{-1}; $(NH_4)_2SO_4$ 1g L^{-1}; $MgSO_4.7H_2O$ 0.2g L^{-1}; $FeSO_4.7H_2O$ 0.01g L^{-1}; NaCl 0.1g L^{-1}; $CaCl_2.2H_2O$ 0.02g L^{-1}; yeast extract 1g L^{-1}; carboxymethyl cellulose (CMC) 5g L^{-1}; pH 7.2) in 500 mL Erlenmeyer flasks and incubated for 4 days in a shaker incubator at 37°C at 120 rpm. Every 5 days, a 2 mL aliquot of the culture was inoculated on fresh medium. The enrichment process was repeated five times. After that dilutions were carried out and inoculated on 0.5% CMC solid medium (CMC medium with 12g L^{-1} bacteriological agar) and incubated at 37°C for 3 days. Bacterial colonies were assayed for their ability to degrade CMC by incubation with 0.1% Congo Red solution for 30 minutes followed by washing with 5M NaCl (Teather & Wood, 1982). Bacterial colonies showing a translucid halo were chosen as positive CMC-degrading strains. Single colonies were picked and checked for purity by repetitive streaking on 0.5% CMC solid medium. Isolates were kept at 4°C on solid medium until analyses and stored in skim milk at -80°C (Cody et al., 2008).

2.3. 16S rRNA partial sequencing

The selected strain was grown in liquid Nutrient medium at 37°C for 16 hours and DNA was purified using phenol-chloroform extraction of cells lysed with lysozyme and SDS (Sambrook et al., 1989). The primers for the *Bacteria* domain, 27F (5´AGAGTTTGATCCTGGCTCAG) and 1492R (5´ACGGCTACCTTGTTACGACTT) (Yoon *et al*, 1998) was used for 16S rRNA gene amplification. PCR mixture (20 µL) contained 0.6 U of *Dream*Taq DNA polymerase (Thermo Fisher Scientific), 2 µL of 10X *Dream*Taq DNA polymerase buffer, 4 pmol of each primer, 200 µM of each deoxynucleoside triphosphate (dNTP), approximately 30 ng of genomic DNA. The thermocycler program was 1 cycle at 95°C for 3 min, followed by 25 sequential cycles of 95°C for 30 seconds, 62°C for 30 seconds, and 72°C for 1 min and a final step at 72°C for 5 min. Amplified fragment was purified using the Ultra Clean PCR Clean-UP kit (MO BIO Laboratories) and sequenced using the BigDye® Terminator Cycle Sequencing kit (Life Technologies) and Applied Biosystems® 3500xL (Thermo Fisher Scientific) DNA sequences were aligned by CAP3 program (Huang & Madan, 1999) and compared to the GenBank database using Blast tool (http://blast.ncbi.nlm.nih.gov/Blast) and Ribosomal Database Project-RDP (https://rdp.cme.msu.edu/) to determine their closest phylogenetic relatives.

2.4. Phenotypic characterization

Phenotypic characterization was performed with colonies grown on Nutrient agar plates (NA) at 37°C overnight. The colony morphology analysis was performed by observing shape, size, margin, elevation, texture, pigmentation and optical property. The cell wall properties were distinguished by Gram staining using optic microscope Olympus BX51. For temperature grown assessment, one fresh colony was inoculated onto NA plates, which were then incubated at 4, 10, 20, 30, 40, 45 and 50°C. The plates were monitored daily up to 14 days until some colonies had formed. Test for motility, rhizoid growth, hemolytic activity and production of crystal toxin was carried out based on Bacteriological Analytical Manual of FDA (http://www.fda.gov).

For scanning electron microscopy (SEM) *Bacillus* sp. ABP14 cultures were harvested and cells fixed with Karnovsky's fixative (Karnovsky,1965). After fixation, the material was dehydrated with increasing alcohol and acetone series, and the critical point obtained at Bal-Tec CPD-030 with carbon dioxide. Then, the gold metallization was done in a Balzers SCD –

030. The material was observed with JEOL-JSM 6360 LV scanning eletron microscope of Eletron Microscopy Center - Federal University of Paraná.

All assays were repeated in duplicate to confirm the final result.

2.5. Insect Bioassay

The free ingestion method was used to determine the toxicity of selected strain culture toward laboratory insect *Anticarsia gemmtalis* (Hübner, 1818) (Lepidoptera: Noctuidae) known as soybean caterpillar. The larvae used in the experiment were obtained from cultures of at the Zoology Department of Federal University of Paraná-Brazil.

Cells were cultivated in GYS medium (Rogof & Yousten, 1969 in a rotary shaker (120 rpm) at 37°C during 24 or 72 hours. Cell culture (500 μL) was loaded onto antibiotic-free insect artificial diet (20g) (Parra, 2001) to feed a third-instar *Anticarsia gemmtalis* larva. Each larva was placed in a container with 20g artificial diet and held at 25°C and 12 hours light:dark photoperiod for about 20 days until pupation. The bioassay comprised three experiments (each containing a single third-instar *A. gemmtalis* larva per container, totaling 6 larvae). The control group was fed with artificial diet added GYS medium. Groups indicated as 24H or 72H corresponded to 24 hours or 72 hours grown bacterial culture, respectively, added to the insect artificial diet. Viability of larval and pupal stages were assessed. The assay was repeated twice to confirm the final result.

2.6. Genome sequencing, assembly and annotation

The whole-genome sequence of selected bacteria was carried out by constructing a genomic DNA library for paired-end sequencing using the *Nextera® DNA Library Preparation Kit* (Illumina) and the Illumina MiSeq platform An additional DNA sequencing was performed using a genomic DNA library constructed with *Ion Express™ Plus Library Kit* (Thermo Fisher Scientific) and the Ion Torrent Technology platform (Thermo Fisher Scientific).

Raw reads quality was checked by FastQC (Andrews, 2010). DNA sequence was *de novo* assembled using CLC Genomics Workbench 6.5.1 (CLC Bio) and GS *de novo* Assembler 2.8 (454 Life Sciences). Large contigs were arranged by using the *Bacillus thuringiensis* serovar *finitimus* YBT020 as the reference sequence with *CLC Genomics Workbench* 6.5.1 (CLC Bio). Gaps were filled with contigs from Ion Torrent assembly and the program *Gfinisher* (Guizelini et al., 2016).

The functional annotation and metabolic reconstruction was carried out by RAST (Rapid Annotation using Subsystem Technology) (Aziz et al., 2008) and KAAS (KEGG Automatic Annotation Server) (Moriya et al., 2007). Assigned functions were checked with pBLAST (Altschul et al., 1997) and InterProScan (Zdobnov & Apweiler, 2001). The rRNA genes was predict by RNAmmer 1.2 (Lagesen et al., 2007).

2.7. Taxonomic classification

Taxonomic analysis was carried out with the isolated strain *Bacillus* sp. ABP14 and closely related species *B. cereus, B. thuringiensis* and *B. anthracis* (Table S1). Genome-to-Genome Distance Hybridization (GGDH) was obtained using the digital DNA-DNA hybridization (dDDH) by the web tool GGDC 2.1 (http://ggdc.dsmz.de) (Meier-Kolthof et al., 2013).

2.8. Comparative genomics

The proteomic set of the *Bacillus* sp. ABP14 and other 27 *Bacillus* strains, retrieved from GenBank, was compared by sequence similarity using an all-against-all Blast search. The homology between proteins was established with 90% amino acid identity over at least 90% of the length of the longest protein (Lukjancenko & Wanessar, 2010). The proteins present in the selected bacteria without homology with other strains analyzed were separated for study. The strains used for genome comparison are described in Table S1 and *B. anthracis str. Ames* was used as type-strain of *B. anthracis* because they are a very isogenic group.

2.9. Nucleotide sequence accession numbers

This *Bacillus* sp. ABP14 genome project has been deposited at GenBank under the accession number CP017016 (chromosome) and CP017017 (pABP, 11Kb).

3.Results and Discussion

3.1. Isolation and identification of cellulolytic bacterium

The ability of microorganisms to degrade polysaccharides such as cellulose is a characteristic of considerable interest in terms of microbial ecology and industrial microbiology (Teather &Wood, 1982). In the present study, microbial colonies were isolated on 0.5% CMC plates after enrichment culture from a lignocellulosic compost of agricultural waste. Those

which showed hydrolysis halos after 0.1% Congo Red staining were selected and identified by molecular methods and characterized from a phenotypic point of view by analysis of cell morphology, Gram reaction, growth temperature, motility and hemolytic activity on Sheep Blood Agar (SBA) plates. Out of 35 colonies, two isolates showed higher CMCase activity and one of them, described in this work showed entomopathogenic activity, a unexpected skills.

The results of phenotypic analysis (Table 1) grouped the isolate, called ABP14, in the genus *Bacillus* and the molecular identification by partial 16S rRNA gene sequence analysis, using universal primers (27F/1492R) identified the isolate ABP14 as *Bacillus sp*, with sequence similarity (99%) for bacterial strains belonging to the *B. cereus sensu lato* group.

The *Bacillus cereus sensu lato* group comprising *B. cereus sensu stricto*, *B. mycoides*, *B. pseudomycoides*, *B. thuringiensis*, *B. weihenstephanensis*, *B. anthracis* and *B. cytotoxicus* (Dréan et al., 2016). Despite the multiple species names, which are often attributed to distinct phenotypic trait conferred by mobile genetic elements, all these organisms share a highly conserved genome and insufficient divergence in 16S rDNA prevented the resolution of strain and species relationships (Maughan et al., 2011).

To differentiating the isolate ABP14 from members of the *B. cereus* group, some phenotypic tests based on Bacteriological Analytical Manual of FDA (U.S. Food & Drug Administration) were performed. The strain ABP14 exhibited hemolytic activity on SBA (Sheep Blood Agar) plates and were able to grow at temperatures of 10°C and 45°C, indicating that it is not *B. anthracis* (non-hemolytic), *B. cytotoxicus* (minimum growth temperature of 20°C) or *B. mycoides* and *B. weihenstephanensis* (incapable of grow at 40°C). The colony morphology was them used to differentiate the isolate ABP14 from *B. pseudomycoides* that show rhizoid colony morphology. The *B. thuringiensis* can be phenotypically distinguished from other *B. cereus sensu lato* strains by production of crystal protein during sporulation phase. These crystal proteins are pore forming protein that kill the insect larval hosts when they are ingested (Ruiu et al., 2013). During sporulation phase of *Bacillus* sp. ABP14, no crystal proteins were observedusing scanning electron microscopy of endospore (Figure 1) but the culture of isolate ABP14 show oral insecticidal activity against *Lepidoptera* (Figure 2).

The toxic effects of ABP14 culture was observed after 15 days of started assay and the larvae showed paralysis, followed by hemocytic melanization and death. This mortality kinetics were not similar to *B. thuringiensis* strains that the larvae began to die 16 hours after infection and mortality reached a plateau 2 days later (Salamitou et al., 2000). The cultures of 24 and 72 hours showed the same effects, with a mortality of 3 and 4 larvae from 5 larvae tested

in each group, respectively. The larvae in control treatment showed normal development. The GYS Medium is an enrichment method for *B. thuringiensis* isolation which allowed propagation of spores (Patel et al., 2013). The 24 hour ABP14 culture consisted of vegetative cells/spores mixture and with 72 hours showed only spores.

3.2. Genome sequencing, assembly and annotation

A culture of *Bacillus* sp. strain ABP14 was grown aerobically in liquid Nutrient medium at 37°C. DNA was purified using phenol-chloroform extraction of cells lysed with lysozyme and SDS.

The genome was sequenced by Illumina MiSeq system (2,968,082 paired-end, 228.92 bp average read length, 131.86 fold coverage) and the reads were assembled with CLC Genomics Workbench 6.5.1 (CLC Bio) and GS *de novo* Assembler 2.8 (454 Life Sciences). The resulting contigs were oriented based on reference genome, *Bacillus thuringiensis* serovar *finitimus* strain YBT020, using *CLC Genomics Workbench* 6.5.1 (CLC Bio). Closing the gaps between scaffolds was achieved by *Gfinisher* (Guizelini et al., 2016) and reads from Ion Proton system (9,882,626 reads, 123.66 bp average read length, 237.18 fold coverage) assembled with *CLC Genomics Workbench* 6.5.1 (CLC Bio).

The functional annotation was carried out by RAST (Rapid Annotation using Subsystem Technology) (Aziz et al., 2008), KEGG (Moriya et al., 2007) and COG (Galperin et al., 2004). Identification of RNA genes were carried out by using RNAmmer 1.2 (Lagesen et al., 2007).

3.3. Genome features

The 5.14-Mb genome of *Bacillus sp.* ABP14 contains two replicons: a circular chromosome with 5,141,367-bp and 35.4% G+C content and a circular plasmid (named pABP14) with 11,199-bp and 30.4% G+C content (Table 2). The chromosome showed 5,238 protein-coding genes, 13 rRNA operons and 92 tRNA that coding for all 20 aminoacyl-tRNA synthetases. The majority of the chromosomal protein coding genes were assigned a putative function (70,79%) while the remaining ones were annotated as hypothetical proteins. The plasmid showed 10 potential ORFs, encoding 5 proteins with assigned functions and 5 with no known function (Table 2). Similarity search performed with Basic Local Alignment Search Tool (BLAST) using the nucleotide sequence of pABP14 as a query revealed 81% identity with pO3 of *B. cereus* FRI-35 (NC_018491.1).

The distribution of genes into COGs functional categories is presented in Table 3. This classification indicated that *Bacillus* sp. ABP14 genome contain many genes involved with amino acid transport and metabolism, follow by carbohydrate transport and metabolism, cell wall/membrane/envelope biogenesis and signal transduction mechanisms. The *Bacillus* sp. ABP14 genes suggest a complex cell-envelope structure and intricate networks of signaling to integrate internal metabolism and environmental conditions.

3.4. Assessment of genome similarity using dDDH

Investigating species affiliation of *Bacillus sp.* ABP14 within *B. cereus sensu lato* group, the complete genome of *Bacillus sp.* ABP14 was subjecting to digital DNA-DNA hybridization (dDDH) against available complete genomes of the *B. cereus sensu lato* strains (Table S1**)**. The values of genomic distances were calculated using Genome-to-Genome Distance Calculator (GGDC) 2.1 server with the recommended settings (Meier-Kolthoff et al., 2013). The results of formula II were used to construct a genomic distance matrices and a heat map was used for ordering the strains in the symmetric matrix and clustered them of species-level at a dDDH threshold of 70% (Figure 3).

Analisys showed that the isolate ABP14 is closely related to *B. thuringiensis* serovar *finitimus* YBT020 (*Bt finitimus* YBT020) with 81.30% dDDH value followed by *B. cereus* AH187, *B. cereus* NC7401 (with 70.60% dDDH) and *B. cereus* Q1 (70.40% dDDH). The others species of *B. cereus*, *B. anthracis* and *B. thuringiensis* showed dDDH values among 40% and 60% suggesting that *Bacillus sp.* ABP14 diverged from the other *Bacillus spp.* studied (Figure 3).

The results show that *Bacillus* sp. ABP14 is the same species group of *B. thuringiensis* and according to Meier-Kolthoff (2014), a value of 79-80% dDDH is the most promising threshold for delineating subspecies. The dDDH values showed the presence of one specie within the group, *Bt finitimus* YBT020, *B. cereus* AH187, *B. cereus* NC7401 and *B. cereus* Q1, suggesting that *Bt finitimus* YBT020 could be reclassified. In the light of genomic data, *B. thuringiensis* can be consider a variety of *B. cereus* that can show a wide range of variations in phenotypes and pathological properties (Arnesen et al., 2008). *Bt finitimus* YBT020 is an insecticidal strain (Zhu et al., 2011), *B. cereus* NC7401 and AH187 (also known as *B. cereus* F4810/72) are cereulide producer strains (Takeno et al., 2012; Rasko et al., 2007) and *B. cereus* Q1 is a non-pathogenic strain with potential for industrial use as surfactant producer (Xiong et

al, 2009). The distinguishing features between the strains are normally encoded by genes located on plasmids, which are well-recognized as high mobile genetic elements.

3.5. *Bacillus sp*. ABP14 and *B. cereus sensu lato* genome comparison

The *Bacillus* sp. ABP14 genome served as a reference for comparison between genomes of *B. cereus sensu lato* strains through BLAST Atlas (Figure 4) in which its possible identify the presence or absence of homologous genes in relation to the order of genes in the *Bacillus* sp. ABP14 genome.

Comparison of the predicted ABP14 proteome against other completed genomes available in GenBank of 12 *B. cereus* strains, 13 *B. thuringiensis* strains and *B. anthracis str.* Ames as type-strain of *B. anthracis* (Table S1), allow us to identify a set of ABP14 strain-specific genes.

Bacillus sp. ABP14 presented 320 proteins not found among the studied *B. cerus sensu lato* strains, of which 149 proteins had functional assignment. Among ABP14 strain-specific genes, those encoding for enzymes involved with the biosynthesis of teichuronic acid, disaccharide linkage unit, major and minor wall teichoic acid (WTA) and teichoic acid translocation system were found (Figure 5). The presence of two types of WTA, major and minor, and the replacement of WTA with teichuronic acid, a phosphorus-free polysaccharide containing uronic acid residues were only described to *B. subtilis* (Mauël et al, 1991; Freymond et al, 2006). The genes that encode to cell wall-associated proteins such as the proteins *phoQ* and *yvcQ*, that are sensing mechanism of cations and antimicrobial peptides respectively (Vésciovi et al., 1997; Mascher et al., 2003), the surface layer protein (Fagan & Fairweather, 2014) and adhesins anchored via LPXTG motif (Davies et al., 2009). The presence of a gene cluster that encoding to DNA restriction-modification type I system (Figure 5), and to the ribonucleases G and E, and the DNA sulfur modification protein DndB. The genes involved in iron acquisition such as enterochelin esterase, a cytoplasmic esterase that release iron from siderophore enterochelin, the most powerful and selective siderophore produced and secreted specifically in response to iron deficiency (Raymond et al., 2003). The genes that encode a bacterial toxin Bacillolysin, a thermolysin-like metalloprotease which play a role during pathogenesis and promote development within the infected host, and they are used to suppress or avoid its innate immune system (Miyoshi & Sinoda, 2000).

3.6. *Bacillus sp*. ABP14 and *Bt finitimus* YBT020 genome comparison

The chromosomes of the *Bacillus* sp. ABP14 and *Bt finitimus* YBT020 have a high degree of collinearity (Figure 6). When the genomes were compared, *Bacillus sp.* ABP14 has 889 proteins not found in *Bt finitimus* YBT020, of which 409 have functional assignment, and *Bt finitimus* YBT020 carry 1408 proteins not found in ABP14, of which 750 have functional assignment. These difference in protein number is related with the genome size between the strains. *Bacillus* sp. ABP14 has 5.14-Mb genome and the *B. thuringiensis* strain YBT-020 contains a circular chromosome (5,355,490 bp) containing 5,477 open reading frames (ORFs) and two circular plasmids, pBMB26 (187,880 bp) and pBMB28 (139,013 bp), carrying 200 and 149 predicted ORFs, respectively (Zhu et al., 2011).

Among strain-specific genes that distinguish *Bacillus* sp. ABP14 from *Bt finitimus* YBT020 encode a cellulase, an enzyme that allow the use of cellulose as a carbon source. The aliphatic amidase AmiE, that catalyzing short-chains amides to produce ammonia that may be used by the bacterium as a source of nitrogen for amino acid synthesis, to neutralize gastric acidity and can cause histological damage (Skouloubris et al., 2001). The arginine deiminase system (ADS) clustered in an operon *arcABDC* enconding arginine deiminase, ornithine carbamoyltransferase, arginine/ornithine antiporter and carbamate kinase respectively. The ADS pathway enables the bacteria to grow anaerobically on arginine and contributes to pH homeostasis protecting the cell from lethal acidification (Liu et al., 2008). The gene that encoding the lactoylglutatione lyase, an important enzyme in methylglyoxal (MG) detoxification and is involved with the survival of the microorganism in MG rich environment like gut milieu (Chakraborty et al., 2015). The competence pseudopili that are encoded by a set of genes in the *comG* operon (*comGABCDEFG)* and the protein that directly control the *com* regulon the alternative sigma factor SigX (Figure 5). The uptake of extracellular DNA from the environment via genetic competence, is a powerful process able to expand and modify the gene inventory of bacteria (Johnston et al., 2014). The copper-containing redox protein namely Azurin that enters into cells by a receptor-mediated endocytic process leading to the induction of apoptosis. This protein was described in *Pseudomonas aeruginosa* and showed preferential entry into cancer compared with normal cells (Zaborina *et al*., 2000). The LexA repressor that regulates the cellular SOS response to DNA damage (Giese et al., 2008). The antibiotic resistance streptothricin acetyltransferase (STAT) and fosmidomycin resistance protein in addition to taurine ABC transport.

3.7. Virulence genes

Bt finitimus YBT020 is a insecticidal strain and this *activity* is related with the presence of crystal protein genes *(cry26Aa* and *cry28Aa) on plasmids (*pBMB26 and pBMB28) (Zhu et al., 2011). In addition to Cry proteins, several others insecticidal factors were found in Bt such as VIP toxin (*Vegetative Insecticidal Protein*), Mtx-like toxin, Thuringesina (Thu) and Zwittermicina A (ZwA), increasing its insecticidal potential (Chakroun et al., 2016; Donovan et al., 2006; Liu et al., 2014; Luo et al, 2011). *Bacillus sp.* ABP14 showed insecticidal ability but no Cry /Cyt (delta-endotoxins) proteins or other insecticidal factors were found in your genomic content. The Thuringesina activity was assayed by feeding *Mythimna sequax* (Franclemont, 1951) *(Lepidoptera: Noctuidae)* with Kikuio grass (*Pennisetum Clandestinum*) inoculated with autoclaveted ABP14 culture and no insecticidal activity was observed (data not showed).

The genome comparison showed that both strains, *Bacillus sp*. ABP14 and *Bt finitimus* YBT020, share a set of virulence factors common in *B. cereus stricto sensu*, such as the non-hemolytic enterotoxin genes (*nheABC*), a perfringolysin-O (HlyI), two channel-forming type III haemolysin (HlyIII), phosphatidylcholine-specific phospholipase C (PC-PLC), sphingomyelinase (SPH) and three microbial collagenases. *Bt finitimus* YBT020 also have genes that encoding haemolysin II (HlyII) and hemolytic enterotoxin (*hblCDA* operon) but not the cytotoxin K (CytK). *Bacillus sp.* ABP14 have a gene encoding cytotoxin K but not the *hbl* gene cluster. The PlcR is a pleiotropic transcriptional regulator of virulence factor gene expression in pathogenic *Bacillus* spp (Agaisse et al., 1999). The genome of *Bt finitimus* YBT020 has two copies of *plcR* gene and *Bacillus sp.* ABP14 has only one.

Both genomes have shown mechanisms of protection against the host defense system such as the immune inhibitor A protein (InhA), which selectively cleaves insect antibacterial peptides (Ivanova et al., 2003) and proteolytic enzymes such as bacillolysin, aureolysin, camelysin and collagenase, that promote the pathogen development within the infected host (Miyoshi & Sinoda, 2000). Genes encoding chitinases were found in both genomes, which are important invasion factors through peritrophic membrane in the insect midgut (Hamid et al, 2013).

3.8. Entomopathogenic activity of *Bacillus* sp. ABP14

Bacillus sp. ABP14 Cry- can be classified as non obligate insect pathogen because it can occur in nature free of host and is culturable on artificial media (Bulla et al., 1975). The

time spent to kill the larvae was higher than *B. thuringiensis* strains. Problably due the lack of the crystal protein. Generally, the Cry toxins are sufficient to kill infected insect larvae because they form transmembrane pores and causing cell lysis. The bacteria in the gut can then invade the haemocoel and cause septicaemia (Salamitou et al., 2000). The genome of *Bacillus* sp. ABP14 show genes encoding to chitinases and several proteases which can damage the peritrophic membrane of larval midgut, allowing the invasion to the haemocoel. These system can be a slower invasion system than Cry toxins, leading to slow death of the insect.

The presence of genes encoding immune inhibitor A protein (InhA), which selectively cleaves insect antibacterial peptides (Fedhila et al., 2002), the aureolysin and bacillolysin that activate the serine proteinase cascade resulting in prophenoloxidase activation that lead to melanization (Altincicek et al., 2007) and phospholipase C, sphingomyelinase and haemolysins with cytolytic acticity againsy haemocyte, are directly involved in establishment and propagation of bacteria within the host. Several strain-specific genes in *Bacillus* sp. ABP14 suggest that it has adapted to live and survive within the gastrointestinal tract (GIT) of insect such as, aliphatic amidase AmiE, arginine deiminase system, lactoylglutatione lyase and cellulase.

4. Conclusion

In this work we report the complete genome of *Bacillus sp.* ABP14 that was isolated from lignocellulosic compost and considered as insect pathogen. The genome-based taxonomy revealed it is closely related to *B. thuringiensis* serovar *finitimus* YBT020 and phenotypic characterization showed the absence of crystal proteins during sporulation process.

Genome comparison reveled differences in terms of virulence, metabolic competence, structural components, restriction modification system and antibiotic resistence supporting the idea that *Bacillus* sp. ABP14 present relevant metabolic features not found in other *B. cereus s.l.* strains.

References

Agaisse H, Gominet M, Økstad OA, KolstØ AB, Lereclus D. 1999. PlcR is a pleiotropic regulator of extracellular virulence factor gene expression in *Bacillus thuringiensis*. Mol. Microbiol. 32: 1043-1053.

Altincicek B, Linder M, Linder D, Preissner KT, Vilcinskas A. 2007. Microbial metalloproteinases mediate sensing of invading pathogens and activate innate immune responses in the Lepidopteran model host *Galleria mellonella.* Infect Immun. 75: 175–183.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. 25: 3389–3402.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available on line at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Arnesen PLS, Fagerlund A, Granum PE. 2008. Fromsoil to gut: Bacillus cereus and its food poisoning toxins. FEMS Microbiol Rev 32: 579–606

Aziz RK, Bartels D, Best AA, De Jongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75.

Bravo A, Likitvivatanavong S, Gill SS, Soberón M. 2011. *Bacillus thuringiensis*: A story of a successful bioinsecticide. Insect Biochem Mol Biol. 41:423–431. doi:10.1016/j.ibmb.2011.02.006.

Bulla LA, Rhodes RA, Julian GS. 1975. Bacteria as insect pathogens. Annu. Rev. Microbiol. 29:163-190.

Cawoy H, Bettiol W, Fickers P, Ongena M. 2011. Bacillus-based biological control of plant diseases, in: Stoytcheva, M. (Ed.), Pesticides in the modern world - Pesticides use and management, ISBN: 978-953-307-459-7, InTech, Available from: http://www.intechopen.com/books/pesticides-in-themodern-world-pesticides-use-and-management/bacillus-based-biological-control-of-plant-diseases.

Chakraborty S, Gogoi M, Chakravortty D. 2015. Lactoylglutathione lyase, a critical enzyme in methylglyoxal detoxification, contributes to survival of *Salmonella* in the nutrient rich environment. Virulence 6:50-65.

Chakroun M, Banyuls N, Bel Y, Escriche B, Ferré J. 2016. Bacterial vegetative Insecticidal Proteins (Vip) from entomopathogenic bacteria. Microbiol Mol Biol Rev. 2016; 80:329-350.

Cody WL, Wilson JW, Hendrixson DR, McIver KS, Hagman KE. Ott CM, Nickerson CA, Schurr MJ. 2008. Skim milk enhances the preservation of thawed -80°C bacterial stocks. J Microbiol Methods. 75:135–138.

Davies JR, Svensäter G, Herzberg MC. 2009. Identification of novel LPXTG-linked surface proteins from *Streptococcus gordonii*. Microbiology 155:1977–1988.

Dréan P, McAuley CM, Moore SC, Fegan N, Fox MF. Characterization of the spore-forming *Bacillus cereus sensu lato* group and *Clostridium perfringens* bacteria isolated from the Australian dairy farm environment. BMC Microbiology. 2015; 15:38. DOI 10.1186/s12866-015-0377-9.

Fagan RP, Fairweather NF. 2014. Biogenesis and functions of bacterial S-layers. Nat Rev Microbiol. 12:211-222.

Fedhila S, Nel P, Lereclus D. 2002. The InhA2 metalloprotease of *Bacillus thuringiensis s*train 407 is required for pathogenicity in insects infected via the oral route. J. Bacteriol. 184: 3296–3304.

Freymond PP, Lazarevic V, Soldo B, Karamata D. 2006. Poly(glucosyl-N-acetylgalactosamine 1-phosphate), a wall teichoic acid of *Bacillus subtilis* 168: its biosynthetic pathway and mode of attachment to peptidoglycan. Microbiology 152: 1709–1718.

Giese KC, Michalowski CB, Little JW. 2008. RecA-Dependent Cleavage of LexA Dimers. J. Mol. Biol. 377:148–161.

Guizelini D, Raittz RT, Cruz LM., Souza EM, Steffens MBR, Pedrosa FO. 2016. GFinisher: a new strategy to refine and finish bacterial genome assemblies. Sci. Rep. 6:34963. DOI: 0.1038/srep34963

Hamid R, Khan MA, Ahmad M, Ahmad MM, Abdin MZ, Musarrat J, Javed S. 2013. Chitinases: an update. J. Pharm Bioallied Sci. 5:21-29.

Huang X, Madan A. 1999. CAP3: A DNA Sequence Assembly Program. Genome Res. 9:868-877.

Ivanova N, Sorokin A, Anderson I, Galleron N, Candelon B, Kapatral V, Bhattacharyya A, Reznik G, Mikhailova N, Lapidus A, Chu L, Mazur M, Goltsman E, Larsen N, D'Souza M, Walunas T, Grechkin Y, Pusch G, Haselkorn R, Fonstein M, Ehrlich SD, Overbeek R, Kyrpides N. 2003 Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis. Nature. 423:87-91.

Johnston C, Martin B, Fichant G, Polard P, Claverys JP. 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat Rev Microbiol. 12:181-196.

Karnovsky MJ. 1965. A formaldehyde-glutaraldehyde fixative of high osmolality for use in electronmicroscopy. J. Cell Biol. 27: A137.

Kupferschmied P, Maurhofer M, Keel C. 2013. Promise for plant pest control: root-associated pseudomonads with insecticidal activities. Frontiers in Plant Science. 4:287. doi: 10.3389/fpls.2013.00287

Kurst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessières P, Bolotin A, Borchert S, Borriss R, Boursier L, Brans A, Braun M, Brignell SC, Bron S, Brouillet S, Bruschi CV, Caldwell B, Capuano V, Carter NM, Choi SK, Codani JJ, Connerton IF, Cummings NJ, Daniel RA, Denizot F, Devine KM, Düsterhöft A, Ehrlich SD, Emmerson PT, Entian KD, Errington J, Fabret C, Ferrari E, Foulger D, Fritz C, Fujita M, Fujita Y, Fuma S, Galizzi A, Galleron N, Ghim SY, Glaser P, Goffeau A, Golightly EJ, Grandi G, Guiseppi G, Guy BJ, Haga K, Haiech J, Harwood CR, Hénaut A, Hilbert H, Holsappel S, Hosono S, Hullo MF, Itaya M, Jones L, Joris B, Karamata D, Kasahara Y, Klaerr-Blanchard M, Klein C, Kobayashi Y, Koetter P, Koningstein G, Krogh S, Kumano M, Kurita K, Lapidus A, Lardinois S, Lauber J, Lazarevic V, Lee SM, Levine A, Liu H, Masuda S, Mauël C, Médigue C, Medina N, Mellado RP, Mizuno M, Moestl D, Nakai S, Noback M, Noone D, O'Reilly M, Ogawa K, Ogiwara A, Oudega B, Park SH, Parro V, Pohl TM, Portetelle D, Porwollik S, Prescott AM, Presecan E, Pujic P, Purnelle B, Rapoport G, Rey M, Reynolds S, Rieger M, Rivolta C, Rocha E, Roche B, Rose M, Sadaie Y, Sato T, Scanlan E, Schleich S, Schroeter R, Scoffone F, Sekiguchi J, Sekowska A, Seror SJ, Serror P, Shin BS, Soldo B, Sorokin A, Tacconi E, Takagi T, Takahashi H, Takemaru K, Takeuchi M, Tamakoshi A, Tanaka T, Terpstra P, Tognoni A, Tosato V, Uchiyama S, Vandenbol M, Vannier F, Vassarotti A, Viari A, Wambutt R, Wedler E, Wedler H, Weitzenegger T, Winters P, Wipat A, Yamamoto H, Yamane K, Yasumoto K, Yata K, Yoshida K, Yoshikawa HF, Zumstein E, Yoshikawa H, Danchin A. 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. Nature. 390:249-256.

Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35: 3100-3108.

Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35: 3100-3108.

Liu X, Ruan L, Peng D, Li L, Sun M, Yu Z. 2014. Thuringiensin: A Thermostable Secondary Metabolite from *Bacillus thuringiensis* with Insecticidal Activity against a Wide Range of Insects. Toxins. 6:2229-2238.

Liu Y, Dong Y, Chen YM, Burne RA. 2008. Environmental and growth phase regulation of the *Streptococcus gordonii* arginine deiminase genes. Appl. Environ. Microbiol. 74: 5023–5030.

Luft JH. 1961. Improvements in epoxy resin embedding methods. J. Biophys. Biochem.Cytol. 9: 409–414. doi:10.1083/jcb.9.2.409

Lukjancenko O, Wassenaar TM, Ussery DW. 2010. Comparison of 61 sequenced *Escherichia coli* genomes. Microb Ecol 60:708–720.

Mascher T, Margulis NG, Wang T, Ye RW, Helmann JD. 2003. Cell wall stress responses in *Bacillus subtilis*: the regulatory network of the bacitracin stimulon. Mol Microbiol 50: 1591–1604.

Mauël C, Young M, Karamat D. 1991. Genes concerned with synthesis of poly(glycero1 phosphate), the essential teichoic acid in *Bacillus subtilis* strain 168, are organized in two divergent transcription units. J. Gen. Microbiol. 137:929-941.

Maughan H, Auwera GV. 2011. *Bacillus* taxonomy in the genomic era finds phenotypes to be essential though often misleading. Infect Genet Evol 11: 789–797.

McGaughey WH, Whalon ME. 1992. Managing Insect Resistance to *Bacillus thuringiensis* toxins. Science. 258:1451-1455.

Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, Rohde C, Rohde M, Fartmann B, Goodwin LA, Chertkov O, Reddy TBK, Pati A, Ivanova NN, Markowitz V, Kyrpides NC, Woyke T, Göker M, Klenk HP. 2014. Complete genome sequence of DSM 30083T, the type strain (U5/41T) of Escherichia coli, and a proposal for delineating subspecies in microbial taxonomy. Standards in Genomic Sciences; 9:2.

Meier-Kolthoff JP, Auch AF, Klenk HP, Goker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC bioinformatics. 14:60. Epub 2013/02026. doi: 10.1186/1471-2105-14-60.

Miyoshi S, Shinoda S. 2000. Microbial metalloproteases and pathogenesis. Microbes Infect. 2:91−98.

Moriya Y, Itoh M, Okuda S, Yoshizawa A, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35:W182-W185.

Pardo-López L, Soberón M, Bravo A. 2013. *Bacillus thuringiensis* insecticidal three-domain Cry toxins: mode of action, insect resistance and consequences for crop protection. FEMS Microbiol. Rev. 37: 3–22.

Parra JRP. 2001. Técnicas de criação de insetos para programas de controle biológico, sixth ed. Piracicaba, ESALQ/FEALQ, 134p.

Patel KD, Bhanshali FC, Chaudhary AV, Ingle, SS. 2013. A new enrichment method for isolation of *Bacillus thuringiensis* from diverse sample types. Appl Biochem Biotechnol 170:58–66.

Rasko DA, Rosovitz MJ, Økstad OA, Fouts DE, Jiang L, Cer RZ, Kolstø AB, Gill SR, Ravel J. 2007. Complete sequence analysis of novel plasmids from emetic and periodontal *Bacillus cereus* isolates reveals a common evolutionary history among the *B. cereus*-group plasmids, including *Bacillus anthracis* pXO1. J. Bacteriol. 189:52–64.

Raymond KN, Dertz EA, Kim SS. 2003. Enterobactin: An archetype for microbial iron transport. PNAS 100: 3584–3588. DOI:10.1073/pnas.0630018100

Reynolds ES. 1963. The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. J. Cell Biol. 17: 208–212. doi:10.1083/jcb. 17.1.208

Rogoff M, Yuosten AA. 1969. *Bacillus thuringiensis*: Microbiological considerations. Ann. Rev. Microbiol. 23:257-386.

Ruiu L, Satta A, Floris I. 2013. Emerging entomopathogenic bacteria for insect pest management. B. Insectol. 66: 181-186.

Salamitou S, Ramisse F, Brehélin M, Bourguet D, Gilois N, Gominet M, Hernandez E, Lereclus D. 2000. The *plcR* regulon is involved in the opportunistic properties of *Bacillus thuringiensis* and *Bacillus cereus* in mice and insects. *Microbiology* 146: 2825–2832.

Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2ed. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press., 1989.

Skouloubris S, Labigne A, Reuse H. 2001. The AmiE aliphatic amidase and AmiF formamidase of *Helocobacter pylori*: natural evolution of two enzymes paralogues. Mol. Microbiol. 40:596-509.

Slepecky RA, Hemphill HE. 1991. The genus *Bacillus*-nonmedical the procaryotes, in: Dworkin, M., Falkow, S., Rosemberg, E., Schleifer, K.-H, Stackebrandt, E. (Eds), The Procaryotes, Springer Verlag, New York, pp. 1663–1696.

Takeno A, Okamoto A, Tori K, Oshima K, Hirakawa H, Toh H, Agata N, Yamada K, Ogasawara N, Hayashi T, Shimizu T, Kuhara S, Hattori M, Ohta M. 2012. Complete Genome Sequence of *Bacillus cereus* NC7401, Which Produces High Levels of the Emetic Toxin Cereulide. J. Bacteriol. 194: 4767–4768. doi:10.1128/JB.01015-12.

Teather RM, Wood PJ. 1982. Use of Congo Red-Polysaccharide Interactions in Enumeration and Characterization of Cellulolytic Bacteria from the Bovine Rumen. Appl. Environ. Microbiol. 43:777-780.

Véscovi EG, Ayala YM, Cera ED, Groisman EA. 1997. Characterization of the bacterial sensor protein PhoQ. J. Biol. Chem. 272: 1440-1443.

Watson ML. 1958. Stain in gof tissue sections for electron microscopy with heavy metals. J. Biophys. Biochem. Cytol. 4:475–478. doi:10.1083/jcb.4.4.475

Xiong Z, Jiang Y, Qi D, Lu H, Yang F, Yang J, Chen L, Sun L, Xu X, Xue Y, Zhu Y, Jin Q. 2009. Complete genome sequence of the extremophilic *Bacillus cereus s*train Q1 with industrial applications. J. Bacteriol. 109:1120–1121.

Yoon JH, Lee ST, Park YH. 1998. Inter- and intraspecific phylogenetic analysis of the genus Nocardioides and related taxa based on 16S rDNA sequences. Int. J. Syst. Bacteriol. 48:187–194.

Zaborina O, Dhiman N, Chen ML, Kostal J, Holder IA, Chakrabarty AM. 2000. Secreted products of a nonmucoid *Pseudomonas aeruginosa* strain induce two modes of macrophage killing: external-ATP-ependent, P2Z-receptor-mediated necrosis and ATP-independent, caspase-mediated apoptosis. *Microbiology* **146:** 2521–2530.

Zdobnov EM, Apweiler R. 2001. InterProScan – an integration plataforma for the signature recognition methods in InterPro. Bioinformatics 17: 847-848.

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. 2011. Complete Genome Sequence of *Bacillus thuringiensis* serovar *finitimus* strain YBT-020. J. Bacteriol. 193: 2379–2380.

Table 1. Phenotype of the isolated bacteria ABP14 from enrichment culture of lignocellulosic compost and screened to cellulolytic activity by Congo Red method on CMC 0,5% plates.

Phenothype	**Bacterial isolate ABP14**
Colony morphology*	White cream regular with filamentous margins
Gram reaction	Gram positive rod
Endospore formation	Subterminal
Motility	Motile
Pigment production	Negative
Hemolytic activity	Positive
Oxidase activity	Positive
Catalase activity	Positive
Growth temperature	10 - 45°C

* The bacterial colony formation was analysed on Nutrient Agar

Table 2. Sequence features of replicons from *Bacillus sp.* ABP14

	Replicons	
Features	*Chromosome*	*Plasmid (pABP14)*
Size, (bp)	5,141,367	11,199
GC content (%)	35.4	30.4
Total number of genes	5,369	10
Coding sequences	5,238	10
Average ORFs size (bp)	823.8	769.2
Protein coding regions (%)	83.9	68.7
Genes with functional assignment	3708	5
Function unknown	1530	5
rRNA operon	13	0
tRNA	92	0

Table 3. Number of genes associated with the general COG functional categories

Code	Description	Value	%
	Metabolism	**2236**	
E	Amino acid transport and metabolism	472	7,99
C	Energy production and conversion	277	4,69
H	Coenzyme transport and metabolism	268	4,53
I	Lipid transport and metabolism	193	3,27
G	Carbohydrate transport and metabolism	357	6,04
F	Nucleotide transport and metabolism	155	2,62
P	Inorganic ion transport and metabolism	309	5,23
Q	Secondary metabolites biosynthesis, transport	205	3,47
	Cellular processes and signaling	**1379**	
O	Post-translational modification, protein turnover, chaperones	245	4,14
T	Signal transduction mechanisms	337	5,70
N	Cell motility	79	1,34
M	Cell wall/membrane/envelope biogenesis	350	5,92
U	Intracellular trafficking, secretion, and vesicular transport	70	1,18
V	Defense mechanisms	185	3,13
D	Cell cycle control, cell division, chromosome	94	1,59
Z	Cytoskeleton	3	0,05
W	Extracellular structures	16	0,27
	Information storage and processing	**1006**	
J	Translation, ribosomal structure and biogenes	356	6,02
L	Function unknown	189	3,20
K	Transcription	457	7,73
A	RNA processing and modification	2	0,03
B	Chromatin structure and dynamics	2	0,03
	Poorly characterized	**1007**	
R	General function prediction only	555	9,39
S	Function unknown	452	7,65
X	**Unclassified**	283	4,79

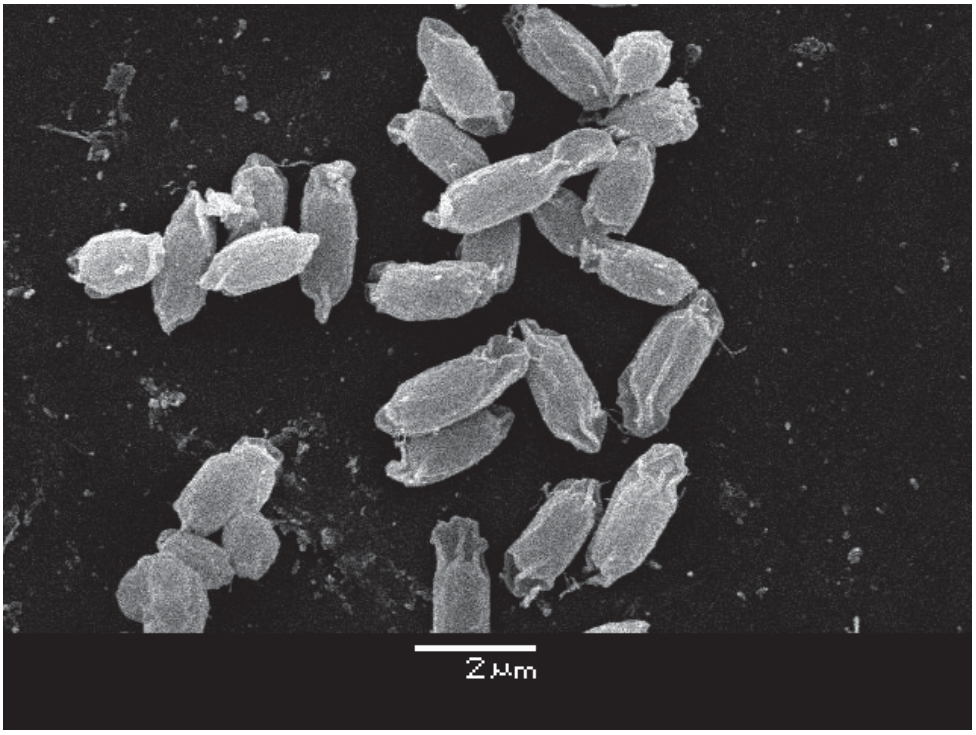


Figure 1. Scanning electron microscopy (SEM) of Bacillus sp. ABP14 endospores. The ABP14 endospores was prepared according Bacteriological Analytical Manual of FDA, and the cells were fixed with Karnovsky's fixative (Karnovsky,1965), dehydrated with increasing alcohol and acetone series, the critical point was obtained at Bal-Tec CPD-030 with carbon dioxide. The gold metallization was done in a Balzers SCD – 030. The cells were observed with JEOL-JSM 6360 LV scanning eletron microscope with 15KV tension and x8,000 increase.

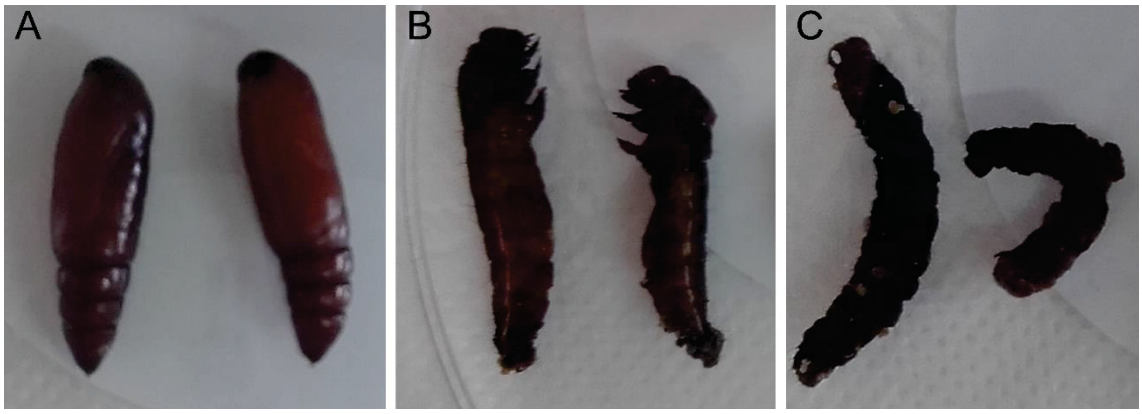


Figure 2. Effect of the isolate ABP14 on *Anticarsia gemmatalis*. The insects, third-instars, were feeding with artificial diet added (500µL) of 24h and 72h ABP14 cultures on GYS medium. A) Effect of 24h culture; B) Effect of 72h culture; C) Control that was added only GYS medium.

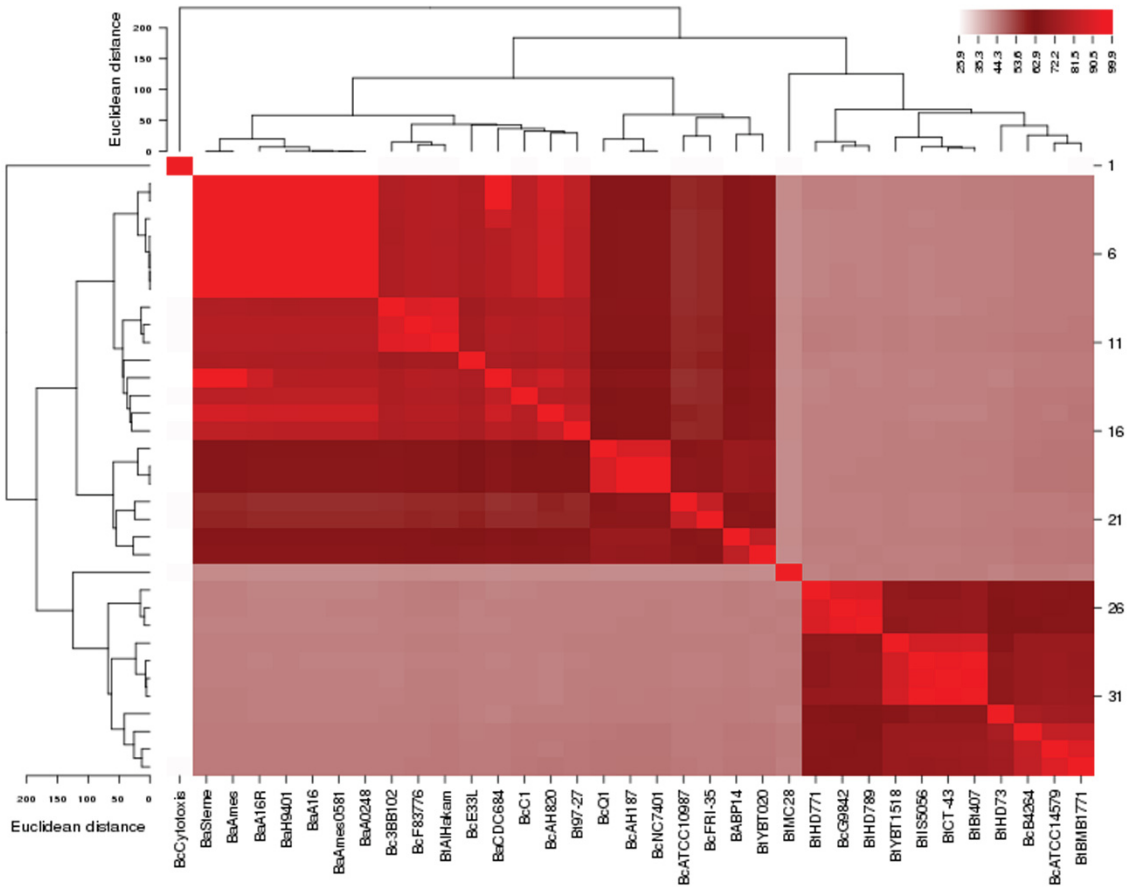


Figure 3. Heatmap of dDDH values using GGDC method. Pairwise comparison of *Bacillus sp.* ABP14 genome and 34 sequenced genomes of *B. cereu sebsu lato*. The phylogenomic tree was used for ordering strains in the symmetric matrix and clustered them of species-level at a dDDH threshold of 70%. Strains names: *Bacillus cytotoxicus* NVH 391-98 (BcCytotox), *B. anthracis* str.Sterne (BaSterne), *B. anthracis* str. Ames (BaAmes), *B. anthracis* str. A16R (BaA16R), *B. anthracis* str. H9401 (BaH9401), *B.anthracis* str. A16 (BaA16), *B. anthracis* str. Ames Ancestor (BaAmes0581), *B. anthracis* str. AO248 (BaAO248), *B. cereus* 03BB102 (Bc03BB102), *B. cereus* F837/76 (BcF837/76), *B. thuringiensis* str. Al Hakam (BtAlHakam), *B. cereus* E33L (BcE33L), *B. anthracis* str. CDC684 (BaCDC684), *B. cereus* biovar *anthracis* str. CI (BcCI), *B. cereus* AH820 (BcAH820), *B. thuringiensis* serovar *konkukian* str.97-27 (Bt97-27), *B. cereus* Q1 (BcQ1), *B. cereus* AH187 (BcAH187), *B.cereus* NC7401 (BcNC7401), *B. cereus* ATCC10987 (BcATCC10987), *B. cereus* FRI-35 (BcFRI-35), *Bacillus* sp. ABP14 (BABP14), *B. thuringiensis* serovar *finitimus* YBT020 (BtYBT020), *B. thuringiensis* MC28 (BtMC28), *B. thuringiensis* HD-771 (BtHD771), *B. cereus* G9842 (BcG9842), *B. thuringiensis* HD-789 (BtHD789), *B. thuringiensis* YBT-1518 (BtYBT1518), *B. thuringiensis* serovar *thuringiensis* str. IS5056 (BtIS5056), *B. thuringiensis* serovar *chinensis* CT-43 (BtCT-43), *B. thuringiensis* Bt407 (BtBt407), *B. thuringiensis* serovar *kurstaki* str. HD73 (BtHD73), *B. cereus* B4264 (BcB4264), *B. cereus* ATCC14579 (BcATCC14579), *B. thuringiensis* BMB171 (BtBMB171). The accession numbers are listed in Table S1.

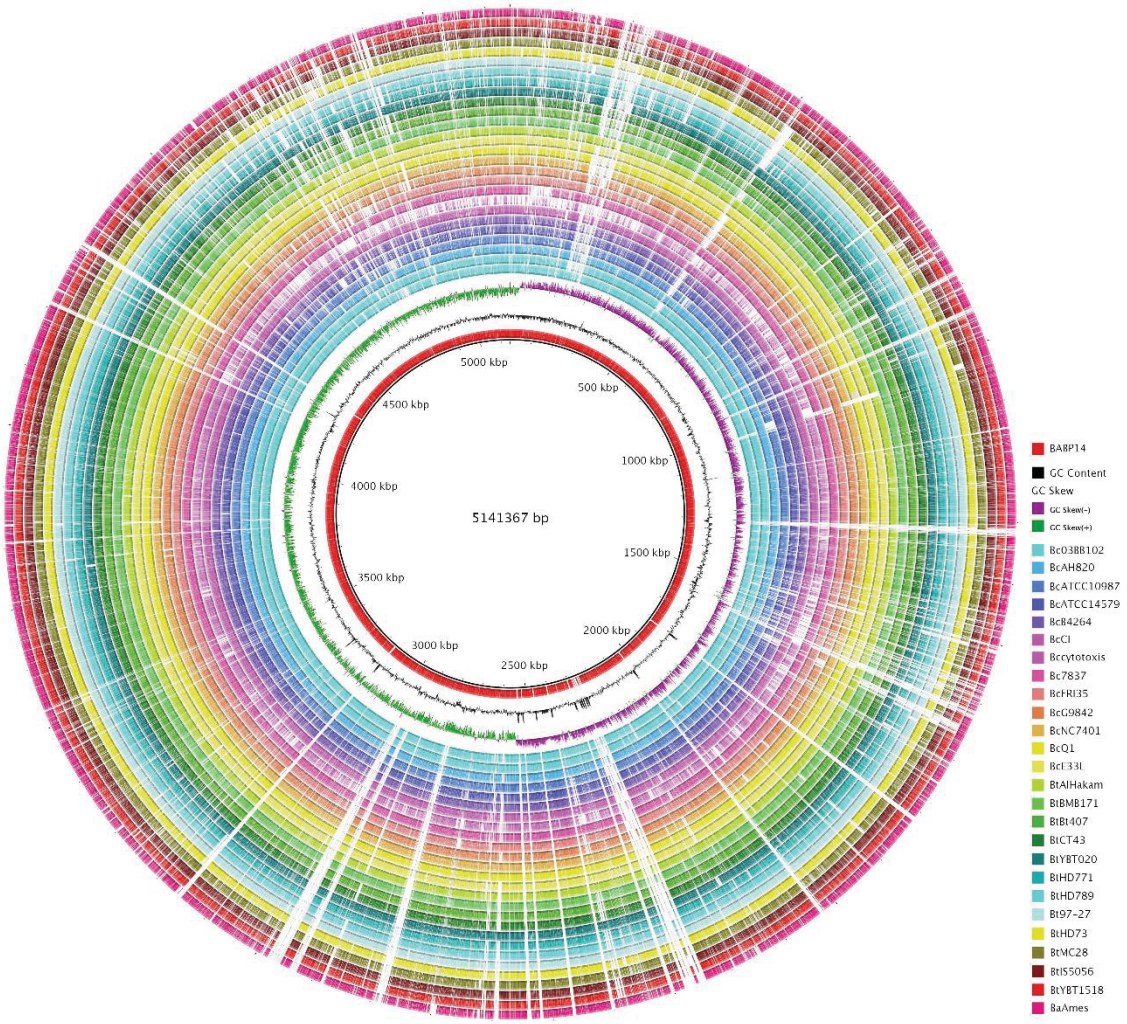


Figure 4. BLAST ATLAS of *Bacillus* sp. ABP14 and *B. cereus sensu lato* genomes. Circles (from inside to outside) 1 represent *Bacillus* sp. ABP14 genome as reference strain, 2 and 3 represent GC content and GC skew, circles 4-29 (colored as show in figure) represent the proteomic sets of other genomes in homology with the reference. Strains names and accession numbers are listed in Table S1.

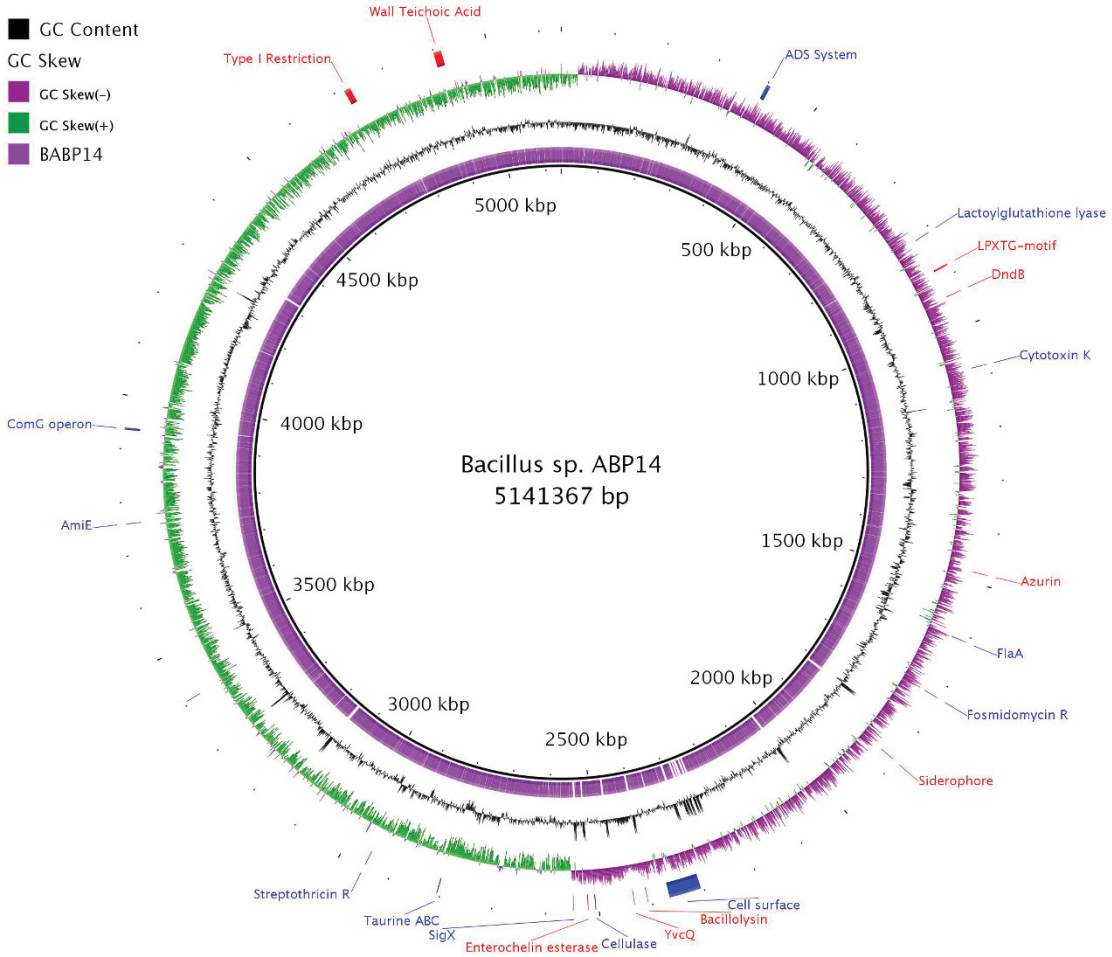


Figure 5. Localization of genes enconding unique proteins studied onto *Bacillus sp.* ABP14 genome. The proteomic set of *Bacillus* sp. ABP14 and *Bacillus cereus s.l.* strain was compared by sequence similarity using an *all-against-all* Blast search considering 90% amino acid identity and 90% cover the largest protein. The unique proteins identified among *Bacillus* sp. ABP14 and *Bacillus cereus s.l.* strain comparison was represented in red. The unique proteins identified between *Bacillus sp.* ABP14 and *B. thuringiensis* serovar *finitimus* YBT020 comparison was represented in blue.

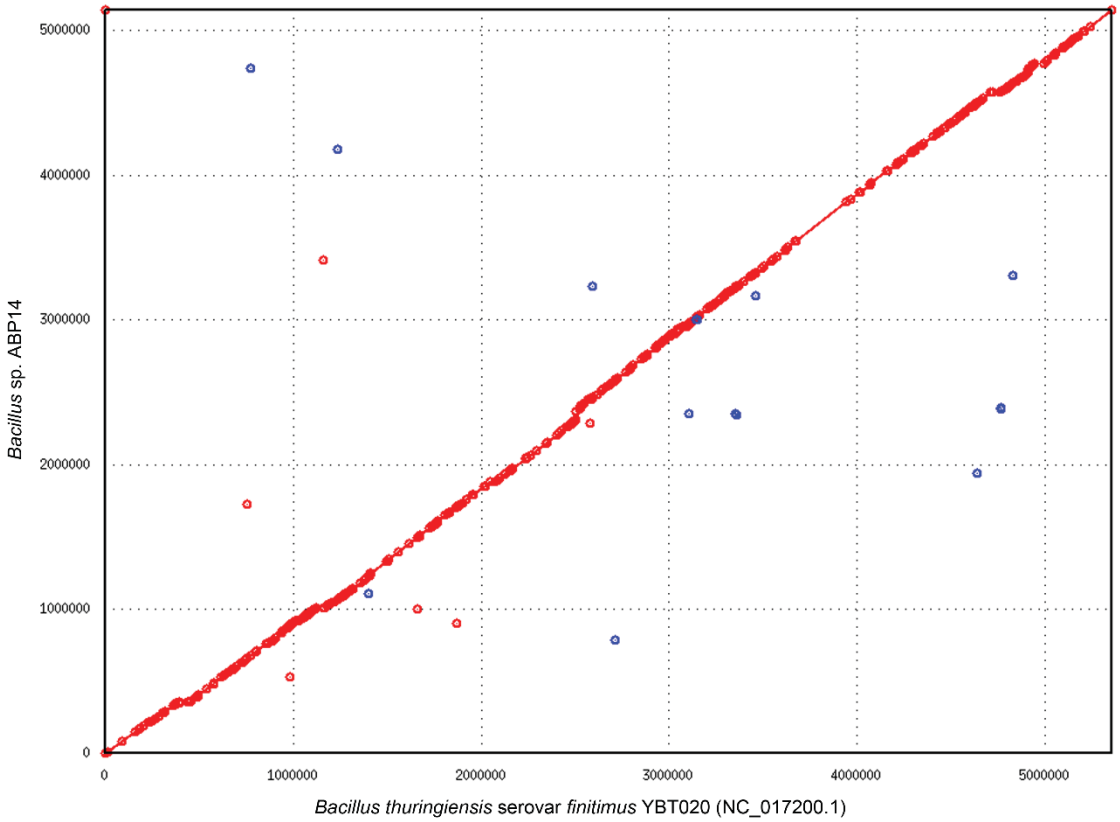


Figure 6. Nucleotide alignment between *Bacillus sp.* ABP14 and *B. thuringiensis* serovar *finitimus* YBT020. The *B. thuringiensis* serovar *finitimus* YBT020 was used as reference.

Table S1. Genome sequence data of *Bacillus cereus sensu lato* strains used to dDDH and comparison.

Microorganisms	Strain	Ref. Seq. Chrom.	Chrom. size (Mb)	Plasmids: size (Mb)	Genome size (Mb)
***Bacillus anthracis* str. Ames**	BaAmes	NC_003997.3	5.23	-	5.23
***Bacillus anthracis* str.Sterne**	BaSterne	NC_005945.1	5.23	-	5.23
***Bacillus anthracis* str. Ames Ancestor**	BaAmes0581	NC_007530.2	5.23	pXO2 (NC_007323.3: 0.09483), pXO1 (NC_007322.2: 0.181677)	5.50
***Bacillus anthracis* str. CDC684**	BaCDC684	NC_012581.1	5.23	pXO2 (NC_012577.1: 0.094875), pXO1(NC_012579.1: 0.181773)	5.50
***Bacillus anthracis* str. AO248**	BaAO248	NC_012659.1	5.23	pXO2(NC_012655.1:0.09483), pXO1 (NC_012656.1: 0,181677)	5.50
***Bacillus anthracis* str. H9401**	BaH9401	NC_017729.1	5.22	BAP1 (NC_017726.1:0,1817), BAP2 (NC_017727.1: 0.094824)	5.40
***Bacillus anthracis* str. A16R**	BaA16R	NZ_CP001974.1	5.23	pXO1 (NZ_CP001975.1: 0,181763)	5.40
***Bacillus anthracis* str. A16**	BaA16	NZ_CP001970.1	5.22	pXO1 (NZ_CP001971.1: 0.181764), pXO2 (NZ_CP001972.1: 0.094839)	5.50
***Bacillus cereus* ATCC14579**	BcATCC14579	NC_004722.1	5.41	pBClin15 (NC_004721.2: 0.015274)	5.43
***Bacillus cereus* ATCC10987**	BcATCC10987	NC_003909.8	5.22	pBc10987 (NC_005707.1: 0.208369)	5.43
***Bacillus cereus* E33L**	BcE33L	NC_006274.1	5.3	pE33L466 (NC_007103.1: 0.46637), pE33L5 (NC_007104.1: 0.005108), pE33L54 (NC_007105.1: 0.053501), pE33L8 (NC_007106.1: 0.008191), pE33L9 (NC_007107.1: 0,00915)	5.84
***Bacillus cereus* Q1**	BcQ1	NC_011969.1	5.21	pBc239 (NC_011973.1:0.239246), pBc53 (NC_011971.1: 0.052766)	5.50
***Bacillus cereus* B4264**	BcB4264	NC_011725.1	5.42	-	5.42
***Bacillus cereus* AH187**	BcAH187	NC_011658.1	5.27	pAH187_12 (NC_011654.1: 0.012481), pAH187_270 (NC_011655.1: 0.270082), p AH187_45 (NC_011656.1: 0.045173), pAH187_3 (NC_011657.1: 0.003091)	5.60
***Bacillus cereus* G9842**	BcG9842	NC_011772.1	5.39	pG9842_209 (NC_011774.1: 0.209488), pG9842_140 (NC_011775.1: 0.140001)	5.73

***Bacillus cereus* AH820**	BcAH820	NC_011773.1	5.3	pAH820_3 (NC_011776.1: 0.003091), pAH820_272 (NC_011771.1: 0.272145), pAH820_10 (NC_011771.1: 0.010915)	5.59
***Bacillus cereus* 03BB102**	Bc03BB102	NC_012472.1	5.27	P03BB102_179 (NC_012473.1: 0.17968)	5.45
***Bacillus cereus* biovar *anthracis* str. CI**	BcCI	NC_014335.1	5.2	pCI-XO1 (NC_014331.1: 0.181907), pCI-XO2 (NC_014332.1: 0.094469), pBAslCI14 (NC_014333.1: 0.014219)	5.49
***Bacillus cereus* F837/76**	BcF837/76	NC_016779.1	5.22	pF837_55 (NC_016794.1: 0.055304), pF837_10 (NC_016780.1: 0.010288)	5.23
***Bacillus cereus* NC7401**	BcNC7401	NC_016771.1	5.22	pNCcld (NC_016792.1: 0.270082), pNC1 (NC_016772.1: 0.047972), pNC2 (NC_016773.1: 0.005436), pNC3 (NC_016793.1: 0.003869), pNC4 (NC_016774.1: 0.003091)	5.55
***Bacillus cereus* FRI-35**	BcFRI-35	NC_018491.1	5.08	p01 (NC_018492.1: 0.218786), p02 (NC_018493.1: 0.040993), p03 (NC_018499.1: 0.036273), p04 (NC_018494.1: 0.003091)	5.40
***Bacillus cytotoxicus* NVH 391-98**	BcCytotox	NC_009674.1	4.09	pBC9801 (NC_009673.1: 007135)	4.10
***Bacillus thuringiensis* serovar *konkukian* str.97-27**	Bt97-27	NC_005957.1	5.24	pBT9727 (NC_006578.1: 0.077112)	5.31
***Bacillus thuringiensis* YBT-1518**	BtYBT1518	NC_022873.1	6.0	pBMB00228 (NC_020124.1: 0.017706), pBMB0229 (NC_022874.1:0.045206), pBMB0230 (NC_022875.1: 0.049195), pBMB0231 (NC_022876.1: 0.146276), pBMB0232 (NC_022877.1:0.171593), pBMB0233 (NC_022882.1: 0.240661)	6.67
***Bacillus thuringiensis* str. Al Hakam**	BtAlHakam	NC_008600.1	5.26	pALH1 (NC_008598.1: 0.055939)	5.31
***Bacillus thuringiensis* BMB171**	BtBMB171	NC_014171.1	5.33	pBMB171 (NC_014172.1: 0.312963)	5.64
***Bacillus thuringiensis* serovar *finitimus* YBT020**	BtYBT020	NC_017200.1	5.36	pBMB26 (NC_017201.1: 0.18788), pBMB28 (NC_017199.1: 0.139013)	5.68
***Bacillus thuringiensis* serovar *chinensis* CT-43**	BtCT-43	NC_017208.1	5.49	pCT127 (NC_017202.1: 0.127885), pCT14 (NC_017209.1: 0.01486), pCT281 (NC_017203.1: 0.281231), pCT51 (NC_017204.1: 0.051488), pCT6880 (NC_017210.1: 0.00688), pCT72 (NC_017205.1: 0.072074), pCT8252 (NC_017211.1:0.008252),	6.15

				pCT83 (NC_017206.1: 0.08359), pCT8513 (NC_017207.1: 0.008513), pCT9547 (NC_017212.1: 0.009547)	
***Bacillus thuringiensis* HD-771**	BtHD771	NC_018500.1	5.89	p01 (NC_018486.1: 0,17103), p02 (NC_018501.1: 0.168999), p03 (NC_018487.1: 0.069876), p04 (NC_018488.1: 0.06547), p05 (NC_018502.1: 0.045262), p06 (NC_018489.1: 0.014056), p07 (NC_018503.1: 0.00907), p08 (NC_018490.1: 0.008574)	6.33
***Bacillus thuringiensis* MC28**	BtMC28	NC_018693.1	5.41	pMC8 (NC_018648.1: 0.007826), pMC54 (NC_018694.1: 0.054484), pMC95 (NC_018685.1: 0.095433), pMC183 (NC_018686.1: 0.18321), MC319 (NC_018688.1: 0.31971), pMC429 (NC_018689.1: 0.429674)	6.69
***Bacillus thuringiensis* Bt407**	BtBt407	NC_018877.1	5.5	BTB_502p (NC_018878.1: 0.501911), BTB_78p (NC_018879.1: 0.077895), BTB_15p (NC_018880.1: 0.015189), BTB_8p (NC_018881.1: 0.00824), BTB_7p (NC_018882.1: 0.007635), BTB_6p (NC_018883.1: 0.00688), BTB_5p (NC_018884.1: 0.005518), BTB_2p (NC_018885.1: 0.002062), BTB_9p (NC_018886.1: 0.008513)	6.13
***Bacillus thuringiensis* serovar *kurstaki* str. HD73**	BtHD73	NC_020238.1)	5.65	pHT73 (NC_020249.1: 0.077351), pHT77 (NC_020239.1: 0.07649), pAW63 (NC_020240.1: 0.071777), pHT11 (NC_020250.1: 0.011769), pHT8_1 (NC_020241.1: 0.008513), pHT8_2 (NC_020242.1: 0.008241), pHT7 (NC_020243.1: 0.007635)	5.91
***Bacillus thuringiensis* HD-789**	BtHD789	NC_018508.1	5.5	p01 (NC_018516.1: 0.349599), p02 (NC_018509.1: 0.235425), p03 (NC_018510.1: 0.224872), p04 (NC_018517.1: 0.014935),	6.33

				p05 (NC_018511.1: 0.007697), p06 (NC_018512.1: 0.006824)	
***Bacillus thuringiensis* serovar *thuringiensis* str. IS5056**	BtIS5056	NC_020376.1	5.49	pIS56-6 (NC_020390.1: 0.00688), pIS56-8 (NC_020377.1: 0.008251), pIS56-9 (NC_020378.1: 0.009671), pIS56-11 (NC_020391.1: 0,011331), pIS56-15 (NC_020379.1: 0.015185), pIS56-16 (NC_020380.1: 0.016206), pIS56-39 (NC_020381.1: 0.039749), pIS56-63 (NC_020392.1: 0.063864), pIS56-68 (NC_020382.1: 0.068616), pIS56-85 (NC_020383.1: 0.085134), pIS56-107 (NC_020393.1: 0,107431), pIS56-233 (NC_020394.1: 0,23373), pIS56-285 (NC_020385.1: 0.328151)	6.77

Table S2. Plasmid (pABP) annotation of *Bacillus sp.* ABP14 .

Feature_id	Location	Strand	Predicted Function
pABP.1	1311_283	-	Recombinase
pABP.2	2844_2260	-	Hypothetical protein
pABP.3	3611_3045	-	Transcriptional regulator, XRE family
pABP.4	4005_4577	+	DNA integration/recombination/inversion protein
pABP.5	5953_4646	-	Hypothetical protein
pABP.6	8180_5946	-	Hypothetical protein
pABP.7	8353_8553	+	Hypothetical protein
pABP.8	8702_9187	+	CAAX amino terminal protease family protein
pABP.9	9629_9775	+	Hypothetical protein
pABP.10	10481_9921	-	Putative mercury resistance protein

7 CAPÍTULO III

Cloning, expression and characterization of two chitinase genes *chiABP-39* and c*hiABP-74* from entomopathogenic *Bacillus* sp. ABP14 isolated from compost.

Ana Paula Andreazza[1], Fabio de Oliveira Pedrosa[1], Emanuel M. Souza[1], Leda Satie Chubatsu[1].

[1] *Department of Biochemistry and Molecular Biology, Center for Biological Fixation of Nitrogen, UFPR (Federal University of Paraná), Curitiba, Brazil.*

§ Corresponding author
Chubatsu@ufpr.br
Phone: + 55 (41) 3361 1659
Fax: +55 (41) 3266 2042

Abstract

Two genes encoding potential chitinases, c*hiABP-39* and c*hiABP-74*, were identified in the genome of the entomopathogenic bacterium *Bacillus* sp. ABP14. Molecular analyses showed that gene c*hiABP-39* encodes a protein of 360 amino acids, including a signal peptide of 27 amino acids, with a molecular mass of 39 kDa and pI of 6.2. Gene c*hiABP-74* encodes a 674 amino acids protein including a signal peptide of 33 amino acids with a molecular mass of 74 kDa and pI of 5.7. Primary structure analysis indicated that the ChiABP-39 and ChiABP-74 enzyme can be classified as glycosyl hydrolase from family 18 and have a high identity with the chitinases of *B. cereus* and *B. thuringiensis*. The heterologous expression of the enzymes was performed in *E.coli* BL21 (DE3) and the recombinant chitinases were purified by ion exchange chromatography. The biochemical characterization of the ChiABP-39 and ChiABP-74 showed that the enzymes have maximum activity at pH 5.0 to 50°C. The enzymes were active on colloidal chitin and chitin flakes, but only ChiABP-74 was able to hydrolyze β-chitin. The presence of Mg^{+2} and Zn^{+2} in the 10mM concentration reduced the ChiABP-74 activity to 67.6% and 53.7% of the maximum activity, respectively, and the detergent sodium dodecyl sulfate (SDS) at 0.5% reduced the enzyme activity to 45.6% of its maximum activity. The ChiABP-39 produce chitobiose as product of colloidal chitin hydrolysis.

Keyword: chitinase, *Bacillus*, chitinolytic bacteria, compost, colloidal chitin, β-chitin.

1.Introduction

Chitinases (EC 3.2.1.14) are enzymes that degrade chitin, a polymer formed by units of N-acetyl-β-D-glucosamine (GlcNAc) linked by β1→4 bonds. Chitin is found in terrestrial and aquatic environment and it plays structural and protective roles in fungi, insects, shellfish, protozoan cysts and nematode eggs (GOOGAY, 1990). In insects, this polymer is present in several tissues such as the parietal layer, exo- and endo-cuticle, respiratory tract, gastrointestinal tract and peritrophic membrane (MERZENDORFER *et al.*, 2003). During its life cycle, the insects are heavily dependent on the construction and remodeling of these chitin structures, which makes them an excellent target for the development of new technologies for insect control. This strategy is reinforced by the lack of chitin in vertebrates and plants, making it a safer alternative (KRAMER & MUTHUKRISHNAN, 1997).

One of the main sources of chitinase are microorganisms, particularly bacteria, which can use chitin as a source of carbon and nitrogen (GOODAY, 1990). The enzymatic degradation of chitin occurs in several stages. One endochitinase (EC 3.2.1.14) breaks the glycoside bonds in a random way along the chain and reduces the polymers to oligomers, which are subsequently degraded to monomers by the exoquitinases (EC 3.2.1.52) (JUNG & PARK, 2014). The products of the degradation of chitin, oligosaccharides and N-acetylglucosamine, are of great interest to the pharmaceutical and food industry (SONGSIRIRITTHIGUL *et al.*, 2009).

A microorganism capable of degrading chitin may have one or more chitinases. It is believed that the presence of multiple chitinases enables a more efficient degradation of the polymer due to the synergistic interaction between the enzymes or the affinity of the enzymes for different substrates (SVITIL *et al*., 1997). An example of this biochemical complexity is the microorganism *Serratia marcescens*, which presents a complex and efficient chitinolytic machinery composed of three chitinases (ChiA, ChiB and ChiC) and a monoxogenase polysaccharide (CBP21) (VAAJE-KOLSTAD *et al*, 2013).

Bacillus spp. are one of the main decomposers of chitin, and they usually have two or more chitinolytic enzymes like *B. circulans* WL-12, which has three chitinases (ChiA, ChiB and ChiC) (ALAM *et al*, 1996) and *B. pumilus* SG2, *B. cereus* and *B.*

thuringiensis subsp. *colmeri* that have two chitinases (HERAVI *et al*, 2014; WANG & LIANG, 2012; HUANG & CHEN, 2005; MABUCHI & ARAKI, 2001; LIU *et al*., 2010). Some *Bacillus* chitinases stand out as probable biological control agents of pests like ChiA chitinase of *B. thuringiensis* colmeri (Liu et al., 2010) and ChiCW of *Bacillus cereus* 28-9 (HUANG *et al*., 2005), which have fungicidal activity, and the ChiA74 chitinase of *B. thuringiensis* serovar *kenyae* which showed a synergistic effect with the Cry insecticidal proteins (BARBOSA-CORONA *et al.*, 2003; BARBOZA-CORONA *et al.*, 2008). Although several chitinases have been identified in *Bacillus spp.*, few have been well characterized biochemically and only one has industrial application.

Bacillus sp. ABP14 is an entomopathogenic bacterium isolated from compost, which has the chitinase enzyme as the main factor in the host invasion. In the genome of *Bacillus* sp. ABP14, it was possible to find two genes encoding chitinases (*chiABP-39 and chiABP-74*), two chitin binding proteins (*cbpABP1* and *cbpABP2*) and a cellulase (*celABP1*). This complex biochemical machinery makes this organism a good study tool to better understand the process of degradation of chitin by the *Bacillus*. In this study, the chitinases genes, *chiABP-39* and *chiABP-74,* were cloned, expressed in Escherichia coli, partially purified and biochemically characterized. The chitinases were termed ChiABP-39 and ChiABP-74, respectively, in accordance with the theoretical molecular mass of the proteins.

2.Material and Methods

2.1. Microorganisms, culture conditions and vectors

Bacillus sp. ABP14 was isolated from a compost pile through the culture in a CMC 0.5% medium (K_2HPO_4 1.6g/L; KH_2PO_4 0.2g/L; $(NH_4)_2SO_4$ 1g/L; $MgSO_4$. $7H_2O$ 0.2g/L; $FeSO_4$. $7H_2O$ 0.01g/L; NaCl 0.1g/L; $CaCl_2.2H_2O$ 0.02g/L; yeast extract 1g/L; carboxymethylcellulose 5g/L; pH 7.2). After isolation, the strain was maintained on plates containing Nutrient agar at 4°C. *Escherichia coli* TOP10 and the vector pET29a (+) (Novagen, Madison, WI, USA) was used in the step that involved the cloning of genes c*hiABP-39* and c*hiABP-74*. *E. coli* BL21 (DE3) was used for expression of the recombinant protein. Strains of *E. coli* were maintained in Luria-Bertani (LB) medium

with appropriate antibiotics. The culture media used was prepared as described in the literature (AUSUBEL *et al*., 2003).

2.2. Reagents and enzymes

For genomic DNA extraction, we used lysozyme, proteinase K, equilibrated phenol (Invitrogen), chloroform, isoamyl alcohol, 99% ethanol (Sigma-Aldrich, St. Louis, MO, USA). For DNA amplification by PCR, we used *Pfu* DNA polymerase and deoxyribonucleotide triphosphates (Invitrogen Life Technologies, USA). The restriction enzymes, T4 DNA ligase, shrimp alkaline phosphatase (SAP) and molecular weight protein markers were purchased from Fermentas (Waltham, Massachusetts, USA). Q Sepharose Fast Flow (GE Healthcare, USA) columns of 1 mL were used for the chromatography test. For testing of the enzyme activity, reagents (potassium ferricyanide, sodium carbonate), the standard (N-acetyl-glucosamine) and the substrate (chitin, Avicel, carboxymethylcellulose, chitosan) were obtained from Sigma (St. Louis, MO, USA). The α- and β-chitin were from Mahtani Chitosan (Veraval, India) and the colloidal chitin was prepared in the laboratory as described by Hsu and Lockwood (1975).

2.3. Cloning of genes c*hiABP-39* and *chiABP-74* of *Bacillus* sp. ABP14

Genes *chiABP-39* and *chiABP-74*, encoding chitinases, were amplified from the genomic DNA of *Bacillus* sp. ABP14 by PCR using specific primers that introduce cloning restriction sites (Table 1)

The PCR reaction was performed in a volume of 20µL containing 2U of *Pfu* DNA polymerase, 2µL of 10X buffer, 10pmol of each primer, 200µmol/l each of deoxynucleotide triphosphate (dATP, dCTP, dTTP, and dGTP), 1.5 mmol/L $MgCl_2$ and 30ng of DNA of *Bacillus* sp. ABP14. The PCR reaction was performed in a thermocycler Eppendorf Master Cycler Gradient 5331 using the following conditions: initial denaturation temperature of 98°C for 1min, followed by 25 cycles of denaturation at 95°C for 15s, annealing at 55°C for 15s and extension at 72°C for 45s, with a final extension of 72°C for 7min. The PCR products were purified with the kit *Ultra Clean PCR Clean-up kit* (MO BIO Laboratories), digested with restriction enzymes *NcoI* and *BamHI* then cloned in *NcoI* and *BamHI* sites of vector pET29a (+), producing the recombinant expression vectors pET29a-*chiABP-39* and pET29a-*chiABP-74*. These recombinant

vectors were transformed into *E. coli* TOP10 and the transformant clones were selected by growth in LB medium containing kanamycin (100µg/mL). The presence of recombinant plasmids was confirmed by double digestion (*NcoI* and *BamHI*) and DNA sequencing.

2.4. Expression and purification of recombinant chitinases

The cells of *E. coli* BL21 (DE3) transformed with the recombinant plasmid (pET29a-*chiABP-36* or pET29a-*chiABP-70*) grew in an orbital shaker (200rpm) at 37°C in LB medium until $O.D_{600}$ of 0.6 and were induced by addition of IPTG up to the final concentration of 0.5mM. The induced culture was incubated for 18h at 16°C and the cells were collected by centrifugation at 9,000×g for 5min at 4°C. The cell pellet was resuspended in 20mL of lysis buffer (50mM Tris-HCl pH 8.0, 100mM NaCl, 10% (v/v) glycerol) and lysed by ultra-sonication in ice bath (30 cycles with pulses of 15 seconds, and interval of 15 seconds, 35% amplitude) to the SONICATOR® XL 2020 (Heat Systems-Ultrasonics Inc., New Highway, Farmingdale, NY, USA) yielding the crude extract (CE).

The crude extract was centrifuged (15,000×g for 20 min) at 4°C to separate insoluble and soluble components. The supernatant, containing the soluble extract (ES), was transferred to a anion exchange Q-Sepharose Fast Flow (GE Healthcare, USA) equilibrated with buffer C (50mM sodium phosphate pH 7.5, 100mM NaCl, 10% (v/v) glycerol). The enzyme was collected in the non-adsorbed fraction, the column and the contaminating proteins were eluted with buffer D (50mM sodium phosphate pH7.5, 1M NaCl, 10% (v/v) glycerol). The elution of the protein was monitored at 280nm and by SDS-PAGE of the fractions. The unadsorbed fraction was concentrated 10X using Amicon® Ultra-15 (Merck KGaA, Darmstadt, Germany) and stored at 4°C.

2.5. Determination of the protein concentration and electrophoresis

The protein concentration was determined by the Bradford method (Bradford, 1976) using BSA (bovine serum albumin) as standard. Electrophoresis of the protein samples was done by SDS-PAGE 12% (LAEMMLI, 1970) and the gel stained with Coomassie brilliant blue R-250 and destained in a solution of methanol/acetic acid/ water (5/1/4).

2.6. Chitinase activity assay

The enzyme activity was determined by the formation of reducing ends as described by Imoto and Yagishita (1971). The colloidal chitin was used as a substrate and N-acetylglucosamine as standard. The standard assay (200µL), consisting of 5µg of enzyme and 100 µL of colloidal chitin 2% in 50mM of sodium citrate buffer (pH 5.0) was incubated at 50°C for 30min. After this period, 1.5 mL of the reagent solution (0.5g/L of potassium ferricyanide in 0.5 M sodium carbonate) was added and heated at 80°C for 15min. The mixture was centrifuged at 15.000xg for 1 minute to remove excess colloidal chitin and the absorbance at 420nm was measured immediately. All experiments were performed in triplicate, calibration curves for N-acetylglucosamine were done for each reaction condition and the effect of non-enzymatic hydrolysis of the substrate was subtracted. One unit of enzyme activity was defined as 1 µmol of N-acetylglucosamine produced per minute.

2.7. Effect of temperature and pH on the enzyme activity

The maximum temperature of conversion of the substrate for a 30min reaction time was determined by incubating 5µg of the enzyme in 100µL of colloidal chitin 2% in 50 mM of sodium citrate (pH 5.0) in a 200µL reaction volume at various temperatures between 20°C and 90°C. To determine the optimum pH, 5µg of enzyme was added to 100µL of 2% colloidal chitin in 50 mM of buffer with different pHs (3.0 to 12.0). The buffers used were sodium citrate (pH 3.0 to 6.0), sodium phosphate (pH 6.0 to 8.0), glycine-NaOH (pH 8.0 to 10.0), NaH_2PO_4-NaOH (pH 10.0-12.0).

2.8. Enzyme thermostability

The enzyme thermostability was determined by an analysis of the residual activity of chitinase after incubation of the enzyme at the temperatures of 50°C, 60°C, 70°C and 80°C during different time intervals (30, 60, 90, 120 and 720 min). 10µL of the enzyme was incubated for 30 minutes in 200µL tubes in a thermocycler (Eppendorf, Hamburg, Germany) operating in gradient mode. After the incubation period, the tubes were cooled on ice and the residual activity was determined with 5µg of enzyme and 100µL of colloidal chitin in the reaction mixture.

2.9. Effect of surfactants and ions on the enzyme activity

On the activity of the chitinase, we tested the following surfactants: Tween 20, Tween 40, Tween 80, Triton X-100, sodium dodecyl sulfate (SDS) and N-lauroyl sarcosine (NLS). The effect of these substances on enzyme activity was determined according to the standard assay, except that the surfactants were previously added to the reaction mixture at concentrations of 0.1% and 0.5%, and incubated at 4°C for 60min. The metal ions KCl, $CaCl_2$, $MgSO_4$, $ZnSO_4$ were added in the final concentrations of 1mM, 5 mM and 10 mM; and NaCl in concentrations of 0.5M, 1.0M and 1.5M.

2.10. Determination of specificity by the substrate

The specificity of the chitinase by the substrate was tested on colloidal chitin 2%, chitin from shrimp shells (flakes), α-chitin, β-chitin, avicel (microcrystalline cellulose) and carboxymethylcellulose 2%. The reaction mixture was made up of 5µg of enzyme, 5mg of substrate in 50mM of citrate buffer (pH 5.0) at 50°C for 30min. The chitinase activity was determined by the Schales method by measuring the reducing ends formed.

2.11. MALDI-TOF analysis of the reaction products

The products of the hydrolysis of the colloidal chitin by the ChiABP-39and ChiABP-74 chitinases were identified by mass spectrometry. A 200µL reaction system consisting of 100µL of 2% colloidal chitin and 5µg of enzyme in 20mM ammonium acetate buffer (pH 5.0) was incubated at 40°C for 24 hours. The system was centrifuged at 15.000rpm for 2min. 2µL of the supernatant was mixed with 2µL of the matrix (2.5-dihydroxybenzoic acid (DHB) 10mg.ml^{-1} in acetonitrile 50% (v/v)) and 2µL of this mixture was deposited in the analysis plate and dried in the air. Mass spectra MALDI-TOF (*Matrix-assisted laser desorption/ionization* MALDI *time-of-flight* TOF *mass spectra* MS) were obtained on a MALDI-TOF/TOF II Autoflex spectrometer (Bruker Daltonics, Bremen, Germany) in reflector positive ion mode with an acceleration voltage of 2 kV, 20ns delay time and the mass acquisition *boundaries between 0 and 3000 m/z*. The spectra are the sum of 200 shots. Peaks were identified by comparing the masses obtained with the standards N-acetylglucosamine, chitotriose, chitopentose and chitohexose the masses (Sigma-Aldrich, St. Louis, Germany).

2.12. Statistical analysis

All experiments were performed in technical and biological triplicate. The data presented are the mean of data obtained with the standard deviation.

2.13. The nucleotide sequence accession number

The nucleotide sequence data described in this study have been deposited in the GenBank database under accession number AOY14035.1 (ChiABP-74) and AOY17019.1 (ChiABP-39).

3. Results and Discussion

3.1. Analysis of the nucleotide sequences of the *chiABP-39* and *chiABP-74* chitinases

The use of chitinases for the enzymatic digestion of chitin can be presented as a new strategy for biological control of insects. Chitinases cause injuries in the peritrophic membrane of the insect allowing the microorganisms to invade the hemocoel (BAHAR *et al.*, 2012) and they also have a synergistic effect with the delta-endotoxins of *B. thuringiensis* (REGEV *et al.*, 1996). However, few chitinases of entomopathogenic bacteria have been biochemically characterized.

Genome analysis of *Bacillus* sp. ABP14, which was isolated from a lignocellulosic compost, identified two chitinase coding genes, named c*hiABP-39* and *chiABP-7*. These genes were amplified and cloned for expression.

Gene c*hiABP-39* has 1082 base pairs (bp) and encodes a protein of 360 amino acids with a calculated molecular mass of 39,401.8 Da and pI of 6.2. A ribosome binding site (Shine-Dalgarno) (5′ AGGAGT 3′) is located 6bp upstream of the initiator codon (ATG) and a probable promoter region of the -35/-10 type (-35 AAGAAT and -10 TTAAAT) located at 47pb upstream of the ATG. Downstream of the stop codon (TAA), a potential transcription termination region was found. It is made up of a palindromic sequence of 18bp, which corresponds to a hairpin structure (Figure S1).

Gene c*hiABP-74* has 2025 base pairs (bp) and encodes a protein of 674 amino acids with a calculated molecular mass of 74,312 Da and pI of 5.7. We found one ribosome binding site (Shine-Dalgarno) (5′ AGGAGA 3′) at 8bp upstream of the initiator codon (ATG) and a probable promoter region (-35 TTGATT and -10 TTAAAT) located

at 30pb upstream of the ATG. Downstream of the stop codon (TAA), a likely transcription termination region was found. It is made up of a palindromic sequence of 16bp, which corresponds to a hairpin structure (Figure S2).

Analyses of the c*hiABP-39* and *chiABP-74* genes indicated a probable promoter region similar to promoters recognized by factor σ^A of the polymerase RNA of *B. subtilis* (-35 TTGAca and -10 TAtAAT) (HELMANN *et al.*, 1995). The promoter regions of *chiABP-74* gene is homologous to promoter regions of genes *chi255, chiA74* and *chiA71* that encode for chitinases of *B. thuringiensis* subsp. *kurstaki* and *B. thuringiensis* serovar *kenyae* and *B. thuringiensis* serovar *pakistani* respectively(DRISS *et al*., 2005; BARBOSA-CORONA *et al*., 2003; THAMTHIANKUL *et al.*, 2001) and the promoters found in the chitinases of five isolated Mexicans of *B. thuringiensis* (ROSAS-GARCIA *et al.*, 2013). These results suggest that the genes *chiABP-39* and *chiABP-74* and the genes encoding chitinase in *B. thuringiensis*, have a conserved regulatory system that is similar to vegetative promoters of *B. subtilis.*

3.2. Analysis of the amino acid sequences of the ChiABP-39 and ChiABP-74 chitinases.

The analysis of the amino acid sequence of the chitinase ChiABP-39 revealed that the protein has a single domain, similarly to the catalytic domain of family 18 of the glycosyl hydrolases. The sequence shows a potential signal peptide of 27 amino acids with a cleavage site located between the alanine-27 and alanine-28 residues. The similarity search performed with the tool BLAST (Basic Local Alignment Search Tools) showed that ChiABP-39 shows a high identity (99%) with various chitinases of *B. thuringiensis* and *B. cereus,* 74.1% with ChiD *B. circullans*, and 57.4% with ChiA of *Flavobacterium johnsoae.*

The amino acid sequence of ChiABP-74 showed that the protein has three domains: the GH18 catalytic domain, a fibronectin type-III domain (Fn3) and a carbohydrate binding module of family 2 (CBM2). The sequence shows a potential signal peptide of 33 amino acids with a cleavage site located between the alanine-32 and aspartate-33 residues. The similarity search performed on the GenBank and UniProt databases showed that ChiABP-74 shows a high identity (99%) of chitinase ChiB of *B.*

thuringiensis and *B. cereus,* 48.8% with ChiD *B. circullans*, and 50.9% with chitinase A of *B. pumilus*.

Primary structure of ChiABP-39 and ChiABP-74 revealed that the two enzymes belong to family GH18 but have different tertiary structures, suggesting that there are different functions in the degradation of chitin. The ChiABP-39 chitinase showed a single GH18 domain while the ChiABP-74 chitinase showed a modular structure consisting of three domains, the catalytic domain GH18 in the N-terminal portion, followed by the fibronectin type-III (Fn-III) domain and a carbohydrate binding module of family 2 (CBM2) in the C-terminal portion. The presence of these functional domains presented in this order is characteristic of chitinases with greater molecular mass of *B. thuringiensis*.

3.3. Investigation of ChiABP-39 and ChiABP-74 similarity with related chitinases

To knowing the similarity of cloned chitinases from *Bacillus* sp. ABP14 in relation to characterized *Bacillus* spp. chitinases from CAZy database (*Carbohydrate-Active enZyme Database*), a phylogenetic tree was made with MEGA software (http://www.megasoftware.net). The phylogenetic tree showed that ChiABP-39 is homologous to *B. thuringiensis* serovar *colmeri* chitinase (ABM05818.1) and ChiABP-74 is homologous to *B . cereus* E33L chitinase (AAU19880.1).

The alignment revealed that the sequence of ChiABP-39 is different at some sites from those of five other chitinases which share the highest similarity shown by alignment (Figure S3). The sequence similarity of ChiABP-39 was 96% with chitinase from *B. thuringiensis* subsp. *colmeri* (ABM05818.1), 95% with chitinases from *B. cereus* CH (BAB16890.1), *B. cereus* NCTU2 (ACY39278.1), *B. cereus* ATCC14579 (AAP10651.1) and 94% with chitinase from *B. cereus* strain 28-9 (AAP47142.1).

The alignment of ChiABP-74 complete sequence is different at several amino acids from those eight other chitinases which share the highest similarity shown by alignment (Figure S4 and S5). The sequence similarity of ChiABP-74 was 99% with *B. cereus* E33L (AAU19880.1) and 94% with chitinases from *B. cereus* CH (BAB16891.1), *B. cereus* 28-9 (AAM48520.2), *B. thuringiensis* serovar *kenyae* (AAL17867.1), *B. thuringiensis* subsp. *colmeri* (ABF57674.1), *B. thuringiensis* subsp. *kurstaki* (AAO34713.1), *B. thuringiensis* 97243-1(ACA34287.1) and *B. thuringiensis* BT218/010 (ABO63962.1) .

3.4. Cloning, expression and purification of chitinases

The coding sequences for the chitinases c*hiABP-39* and *chiABP-74* were cloned without the 5' end encoding the signal peptide, corresponding to amino acids 27 and 33 respectively, into the pET29a . *E. coli* BL21 (DE3) carrying the recombinant plasmids pET29a-c*hiABP-39* and *pET29a-chiABP-74* were induced with 0.5mM IPTG for 18 h at 16°C and the extracts obtained after cell lysis were analyzed by SDS-PAGE. The SDS-PAGE gel showed that the chitinases were expressed in a soluble form showing a migration rate of approximately 45 kDa and 74 kDa, respectively (Figure.1A and Figure.2A). The calculated molecular masses with ProtParam tool (http://web.expasy.org/protparam/) for the ChiABP-39 and ChiABP-74 chitinases, without the signal peptide and with additional amino acids derived from the expression vector, were 42 kDa and 76 kDa respectively corresponding to the values observed in the electrophoresis gel.

The purification of the chitinase was performed using an anion exchange chromatography Q-Sepharose due the cloned chitinases were not in-frame and as consequence, terminal His-tag was not added. Both proteins were present on the flow-through, and most of the contaminants were removed. (Figure.1B and Figure.2B). The proteins partially purified were then concentrated 10X by ultrafiltration using the Amicon® system and stored at 4°C.

3.5. Effect of pH and temperature on the enzyme activity

When the chitinase activity assays were carried out at different pH values, the enzymes showed a maximum activity at pH 5.0. The maximum activity to ChiABP-39 and ChiABP74 was 44.5 and 46.3 U/min/mg of protein, respectively. For ChiABP-39, the activity remained above 95% in the range between pH 4.0 to 6.0 in 50mM of sodium citrate buffer. The enzyme maintained 70% of its maximum activity between pH 6.0 to 8.0 in sodium phosphate buffer 50mM, 50% of its maximum activity between pH 9.0 and 10.0 in glycine-NaOH buffer, and showed less than 30% of its maximum activity between pH 10.0 and 11.0 in NaH_2PO_4-NaOH buffer (Figure 3A).

For ChiABP-74 the activity remained above 90% in the range between pH 4.0 and 6.0 in 50mM of sodium citrate buffer. The enzyme had an activity of 50% between pH 6.0 and 8.0 in sodium phosphate buffer 50mM, and the activity decreased from 50% to

30% of its maximum activity between pH 8.0 and 10.0 in glycine-NaOH buffer and between pH 10.0 and 12.0 in NaH_2PO_4-NaOH (Figure 3B).

The chitinases, in 50mM of sodium citrate buffer (pH 5.0), showed maximum activity at 50°C. The chitinase ChiABP-39 was able to maintain more than 70% of its activity in the range of temperatures between 30°C and 60°C (Figure 6A). For ChiABP-74 the activity decreased drastically at temperature above 60ºC. These results indicate thatthe ChiABP-39 and ChiABP-74 chitinases of *Bacillus* sp. ABP14 can be characterized as mesophilic and acidic chitinases.

Of already been reported chitinases, only four revealed similarity with ChiABP-39 as ChiA of *B. cereus* CH, ChiCH of *B. cereus* 28-9, ChiNCTU2 of *B. cereus* NCTU2 and ChiA of *B. thuringiensis* subsp. *colmeri*. All these chitinases have a molecular mass of 36 kDa and although they have few different amino acids between their sequences, they differ in their biochemical characteristics. ChiABP-39 is more active in an acid medium (pH 5.0) and temperature of 50°C. ChiNCTU2 has maximum activity at pH 6.0 and 37°C temperature (HSIEH *et al.*, 2010). The ChiCH of *B. cereus* 28-9 was characterized at 37°C and pH 5.0 (HUANG & CHEN, 2005). The ChiA of *B. cereus* CH, and ChiA of *B. thuringiensis* subsp. *colmeri* have maximum activity at pH 5.0 and temperatures of 50°C and 60°C respectively (MABUCHI & ARAKI, 2001; LIU *et al*., 2010). Among the chitinases similar to Chi ABP-74, just ChiA74 of *B. thuringiensis* serovar *kenyae* was characterized biochemically and differs from ChiABP-74 for having a bimodal activity at pH 4.7 and pH 6.3, and for the optimum temperature of 57°C at pH6.

3.6. Thermal stability of chitinases

The residual activity of ChiABP-39 and ChiABP-74 was determined after incubating the enzyme at different temperatures ranging from 50°C to 80°C for different time intervals in 50mM of sodium citrate buffer (pH 5.0). The enzymes exhibited high thermal stability at 50°C, retaining 70% of the activity after 24 hours of incubation (Figure 4). The control reaction was done with no heated enzyme and considered as 100% activity.

3.7. Determination of the specificity of the chitinases by the substrate

To evaluate the specificity of the chitinases by the substrate, the hydrolytic activity was compared using different substrates such as chitin from shrimp shells (flakes), α-chitin, β-chitin, 2% colloidal chitin, avicel and CMC 2%. The enzyme activity was performed according to standard assay. The results showed that the enzymes showed maximal activity on the colloidal chitin (100%), but they differ for other substrates. ChiABP-39 showed 33.52% activity for chitin flakes and 5.42% for β-chitin. ChiABP-74 has 48.86% of activity for β-chitin and 38.20% for chitin flakes. The ChiABP-39 and ChiABP-74 have no activity on α-chitin, avicel and CMC 2% (Table 2).

Additional modules, CBM and Fn-III, are related with the binding of the enzyme to the substrate and with the correct structural conformation of the chitinase, thus contributing to the processive action of the enzyme. Thus, it was observed that the ChiABP-74 chitinase showed a higher activity on β-chitin (48.9%) than ChiABP-39 (5.4%), but none of the enzymes showed activity on α-chitin (the most crystalline form of chitin) when tested alone.

3.8. Effects of chemical agents on the activity of the chitinases

The effect of surfactants and ions on the enzyme activity was investigated after incubation of the enzymes with different agents in 50 mM of sodium citrate buffer (pH 5.0) for 1 hour at 4°C. The results show that the detergents Tween 20, Tween 80, Triton X-100 and N-lauryl sarcosine, at concentrations of 0.1% and 0.5%, did not affect the activity of the ChiABP-39 chitinase. The sodium dodecyl sulfate reduced the activity of the ChiABP-39 to 87% and 45% when used at concentrations of 0.1% and 0.5% respectively (Table 3). Neither the presence of salts, such as K^{+}, Ca^{+2}, Mg^{+2} and Zn^{+2}, at concentrations of 1mM, 5mM and 10mM, nor the presence of NaCl at concentrations of 0.5M, 1.0M and 1.5M, interfere significantly with the activity of the ChiABP-39 chitinase and (Table 4).

For the ChiABP-74 chitinase, the detergents Tween 20, Tween 80 and Triton X-100 at concentrations of 0.1% and 0.5%, did not have a significant impact on the activity. The sodium dodecyl sulfate reduced the enzyme activity to 66.6% and 45.6% when used at concentrations of 0.1% and 0.5% respectively. Similarly, the presence of N-lauryl sarcosine reduced the activity of the ChiABP-74 to 82% and 49.2% of the maximum

activity at concentrations of 0.1% and 0.5% (Table 3). Neither the presence of salts, such as K^{+}, Ca^{+2}, Mg^{+2} and Zn^{+2}, at concentrations of 1 mM and 5 mM, nor the presence of NaCl at concentrations of 0.5M, 1.0M and 1.5M, interfere significantly with the activity of the ChiABP-74 chitinase (Table 4). The presence of Mg^{+2} and Zn^{+2}, at a concentration of 10mM reduced the activity of ChiABP-74 to 67.6% and 53.7% of the maximum activity, respectively.

3.9. Analysis of the MALDI-TOF reaction products

The products of the hydrolysis of the colloidal chitin by the ChiABP-39 chitinase after 24 hours of reaction at 40°C were analyzed by MALDI-TOF-MS. The results show that ChiABP-39 produces predominantly dimers (*m/z* 447.190) of N-acetylglucosamine (Figure 8). The reaction products of ChiABP-74 on colloidal chitin were not conclusive, in these work, when analyzed by MALDI-TOF-MS.

Despite studied chitinases of 36kDa showed very similar amino acid sequence, the enzymatic activity products are different**.** The enzymes ChiNCTU2 of B. cereus NCTU and ChiCH of *B. cereus* 28-9 are considered exochitinases, while the ChiA of *B. cereus* CH is considered an endochitinase. The hydrolysis products of the ChiA of *B. thuringiensis colmeri* were not studied. The ChiABP-39 chitinase of *Bacillus* sp. ABP14 was able to release the N-acetylglucosamine dimers from the colloidal chitin, suggesting the exochitinase activity.

4.Conclusion

In the genome of *Bacillus* sp. ABP14, genes c*hiABP-39* and *chiABP-74* code for two different and functional chitinases, called ChiABP-39 and ChiABP-74 respectively. Chitinases have been characterized as mesophilic and acidic with high thermal stability at 50°C, indicating an advantage for applications in biotechnology. The amino acid sequence of the enzyme has high similarity (99%) with chitinases of *B. cereus* and *B. thuringiensis* but ChiABP-39 and ChiABP-74 differ in their biochemical characteristics compared with the characterized chitinases of those species.

These interesting results should lead to more exhaustive studies not only for a more detailed characterization of these enzymes but also the joint activity of them on the crystalline substrate.

References

AUSUBEL, F.M.; BRENT, R.; KINGSTON, R..; MOORE, D.D.; SEIDMAN, J.G.; SMITH, J.A.; STRUHL, K. **Current Protocols in Molecular Biology**. New York: John Wiley & Sons, 2003.

BAHAR, A.A.; SEZEN, K.; DEMIRBAĞ, Z.; NALÇACIOĞLU, R. The relationship between insecticidal effects and chitinase activities of Coleopteran-originated entomopathogens and their chitinolytic profile. **Ann Microbiol.** v.62, p. 647–653, 2012.

BARBOZA-CORONA, J.E.; NIETO-MAZZOCCO, E.; VELA´ZQUEZ-ROBLEDO, R.; SALCEDO-HERNANDEZ, R.; BAUTISTA, M.; JIMÉNEZ, B.; IBARRA, J.B. Cloning, Sequencing, and Expression of the Chitinase Gene *chiA74* from *Bacillus thuringiensis.* **Appl. Environ. Microbiol.** v.69, p.1023–1029, 2003.

BARBOZA-CORONA, J.E.; REYES-RIOS, D.M.; SALCEDO-HERNÁNDEZ, R.; BIDESHI, D.K. Molecular and biochemical characterization of na endochitinase (ChiA-HD73) from Bacillus thuringiensis subsp. kurstaki HD-73. **Mol. Biotech**. v.39, p.29-37, 2008.

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

DRISS, F.; KALLASSY-AWAD, M.; ZOUARI, N.; JAOUA, S. Molecular characterization of a novel chitinase from *Bacillus thuringiensis* subsp. *kurstaki*. **J. Appl. Microb.** v.99, p.945–953, 2005.

GOODAY, G.W. Physiology of microbial degradation of chitin and chitosan. **Biodegradation.** v.1, p.177–190, 1990.

GOODAY, G.W. **The Ecology of Chitin Degradation**. K. C. Marshall (ed.), Advances in Microbial Ecology, Plenum Press, New York, 1990.

HELMANN, J.D. Compilation and analysis of Bacillus subtilis σ^A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic Acids Res v.23, p.2351-2360, 1995.

HERAVI, K.M.; SHALI, A.; NAGHIBZADEH, N.; AHMADIAN, G. Characterization of cis-acting elements residing in the chitinase promoter of Bacillus pumilus SG2. **World J Microbiol Biotechnol.** v.30, p.1491–1499, 2014.

HSIEH, Y.C.; WU, Y.J.; CHIANG, T.Y.; KUO, C.Y.; SHRESTHA, K.L.; CHAO, C.F.; HUANG, Y.C.; CHUANKHAYAN, P.; WU, W.; LI, Y.K.; CHEN, C.J. Crystal Structures of *Bacillus cereus* NCTU2 Chitinase Complexes with Chitooligomers Reveal Novel Substrate Binding for Catalysis. **J. Biol. Chem**. v.285, p.31603–31615, 2010.

HUANG, C.J.; CHEN, C.Y. High-level expression and characterization of two chitinases, ChiCH and ChiCW, of *Bacillus cereus* 28-9 in Escherichia coli. **Biochem. Biophys. Res. Commun.** v.327, p.8–17, 2005.

HSU, S.C.; LOCKWOOD, J.L. Powdered chitin agar as a selective medium for enumeration of actinomycetes in water and soil. **Appl Microbiol**. v.29, p.422–6, 1975.

IMOTO, T.; YAGISHITA, K. A simple activity measurement of lysozyme. **Agric Biol Chem**. v.35, p.1154–1156, 1971.

JUNG, W.J.; PARK, R.D. Bioproduction of Chitooligosaccharides: Present and Perspectives. **Mar. Drugs** v.12, p.5328-5356, 2014.

KRAMER, K.J.; MUTHUKRISHNAN, S. Insect Chitinases: Molecular Biology and Potential Use as Biopesticides. **Insect Biochem. Molec. Biol.** v.27, p.887–900, 1997.

LAEMMLI, U.K. Cleavage of structural proteins during the assembling of the head of the bacteriophage T7. **Nature**, v. 277, p. 680-685, 1970.

LIU, D.; CAI, J.; XIE, C.; LIU, C.; CHEN, Y. Purification and partial characterization of a 36-kDa chitinase from *Bacillus thuringiensis* subsp. *colmeri*, and its biocontrol potential. **Enzyme Microb. Technol**. v.46, p.252–256, 2010.

MABUCHI, N.; ARAKI, Y. Cloning and sequencing of two genes encoding chitinase A and B from *Bacillus cereus* CH. **Can. J. Microbiol**. v.47, p. 895–902, 2001.

MERZENDORFER, H.; ZIMOCH, L. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. **J Exp Biol** 206: 4393–4412, 2003.

REGEV, A.; KELLER, M.; STRIZHOV, N.; SNEH, B.; PRUDOVSKY, E.; CHET, I.; GINZBERG, I.; KONCZ-KALMAN, Z.; KONCZ, C.; SCHELL, J.; ZILBERSTEIN, A. Synergistic Activity of a *Bacillus thuringiensis* d-Endotoxin and a Bacterial Endochitinase against *Spodoptera littoralis* Larvae. **Appl Environ Microbiol**. v.62, p. 3581–3586, 1996.

ROSAS-GARCÍA, N.M.; FORTUNA-GONZÁLEZ, J.M.; BARBOZA-CORONA, E. Characterization of the chitinase gene in *Bacillus thuringiensis* Mexican isolates. **Folia Microbiol.** v.58, p.483–490, 2013.

SONGSIRIRITTHIGUL, C.; PESATCHA, P.; EIJSINK, V.G.H.; YAMABHAI, M. Directed evolution of a *Bacillus* chitinase. **Biotechnol. J.** v.4, p.501–509, 2009.

SVITIL, A.L.; CHADHAIN, S.M.N.; MOORE, J.A.; KIRCHMAN, D.L. Chitin degradation proteins produced by the marine bacterium *Vibrio harveyi* growing on diferente forms of chitin. **Appl. Environ.Microbiol**. v.63, p.408–413, 1997.

THAMTHIANKUL, S.; SUAN-NGAY, S.; TANTIMAVANICH, S.; PANBANGRED, W. Chitinase from *Bacillus thuringiensis* subsp. *pakistani.* **Appl Microbiol Biotechnol**. v.56, p.395–401, 2001.

VAAJE-KOLSTAD, G.; HORN, S.J.; SORLIE, M.; EIJSINK, V.G.H. The chitinolytic machinery of *Serratia marcescens* – a model system for enzymatic degradation of recalcitrant polysaccharides. **FEBS Journal**. v.280, p.3028–3049, 2013.

WANG, S.L.; LIU, C.P.; LIANG, T.W. Fermented and enzymatic production of chitin/chitosan oligosaccharides by extracellular chitinases from *Bacillus cereus* TKU027. **Carbohydrate Polymers**. v.90, p.305– 1313, 2012.

Table 1. Pair of primers used in the amplification and cloning of the genes c*hiABP-39* and *chiABP-74*.

Gene	**Primer**	**Sitio Restrição**
*ChiABP-39*F	5′- CAAGCACCCATGGCAAACAATTT- 3′	NcoI
*ChiABP-39*R	5′- TGTTTTGGATCCTTTTTGCAAGGG - 3′	*BamHI*
*ChiABP-74*F	5′- GCATCCATGGTATCACAAAAGCAA - 3′	*NcoI*
*ChiABP-74*R	5′- TTA GGATCCGTTTTCGCTAATGGA - 3′	*BamHI*

Table 2. Specificity of chitinase ChiABP-39 and ChiABP-74 by the substrate.

	Relative activity (%)	
Substrate	ChiABP-39	ChiABP-74
Colloidal chitin	100 ± 0.005	100±0.003
α-Chitin	0,0	0,0
β-Chitin	5.42 ± 0.005	48.86±0.009
Chitin from shrimp shells (flakes)	33.52 ± 0.007	38.20±0.012
Avicel®	0.26 ± 0.002	0.0
Carboxymethylcellulose (CMC)	0.0	0.0

Table 3. Effect of different surfactants on the activity of chitinase ChiABP-39 and ChiABP-74.

Surfactant	***Relative activity*** (%)			
	ChiABP-39		ChiABP-74	
	Concentration		Concentration	
	0.1%	**0.5%**	**0.1%**	**0.5%**
Control	100 ± 0.002	100 ± 0.002	100±0.002	100±0.002
Tween 20	101.01 ± 0.001	101.23 ± 0.012	99.73±0.001	115.45±0.016
Tween 80	101.21 ± 0.001	102.33 ± 0.002	100±0.001	114.60±0.040
Triton X-100	100.8 ± 0.003	102.19 ± 0.002	99.6±0.002	106.62±0.005
Sodium Dodecyl Sulfate	87.85 ± 0.064	45.81 ± 0.008	66.66±0.019	45.67±0.009
N-Lauroylsarcosine	100.53 ± 0.001	100.53 ± 0.001	82.0±0.002	49.23±0.022

Table 4. Effect of different ions on the activity of chitinase ChiABP-39.

Íons	***Relative activity*** (%)		
	Concentration		
	1mM	**5mM**	**10mM**
Control	100 ± 0.001	100 ± 0.001	100 ± 0.001
KCl	97.8 ± 0.005	102.63± 0.003	99.26± 0.006
CaCl	97.98 ± 0.002	97.36± 0.043	98.9± 0.003
$MgSO_4$	98.16 ± 0.011	101.5± 0.012	97.25± 0.007
$ZnSO_4$	98.16 ± 0.011	101.5± 0.012	97.25± 0.007
	0.5M	**1.0M**	**1.5M**
NaCl	100± 0.004	101.15± 0.004	99.42± 0.004

Table 5. Effect of different ions on the activity of chitinase ChiABP-74.

Íons	***Relative activity*** (%)		
	Concentration		
	1mM	**5mM**	**10mM**
Control	100±0.012	100±0.012	100±0.012
KCl	102.69±0.041	100.36±0.001	105.18±0.013
CaCl	104.14±0.041	100.18±0.001	80.91±0.110
$MgSO_4$	104.77±0.032	100.73±0.002	67.63±0.005
$ZnSO_4$	110.37±0.017	93.76±0.004	53.73±0.069
	0.5M	**1.0M**	**1.5M**
NaCl	104.3±0.003	100.58±0.001	99.41±0.009

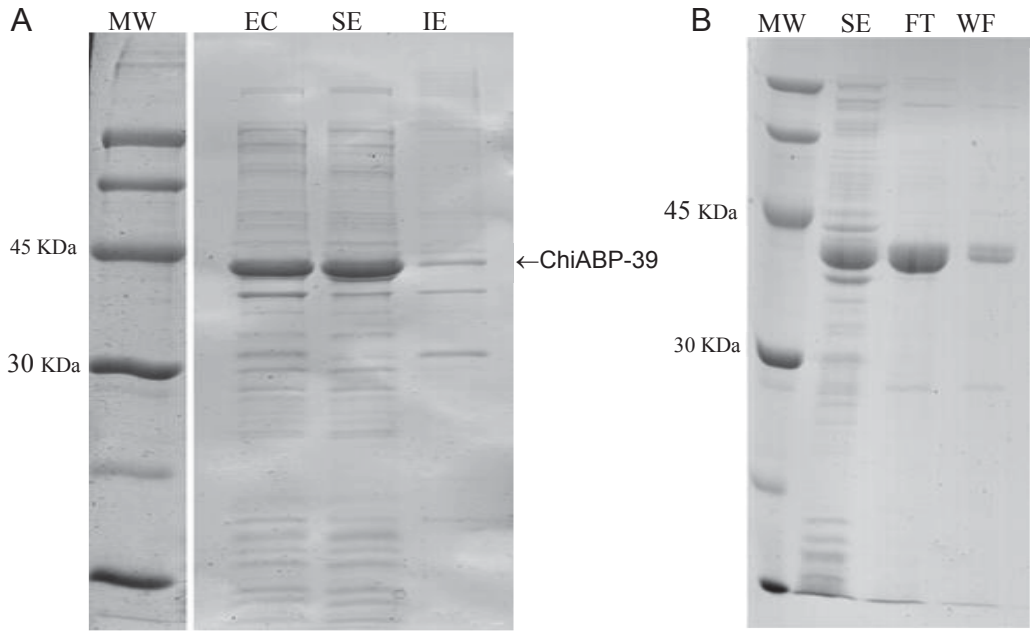


Figure 1. Expression and purification of the ChiABP-39 chitinase. Electrophoresis SDS-PAGE in 12% polyacrylamide gel. **A**) Fractions obtained from the expression test and solubility of the ChiABP-39 chitinase. *E. coli* BL21 (DE3) transformed with the plasmid pET29a-*chiABP-39* was induced with 0.5mM IPTG at 16°C for 18h in LB medium. After harvested, cells were lysed by sonication and the crude extract (EC) obtained was centrifuged (15,000 x g / 20 min / 4°C) yielding the supernatant (SE) and the insoluble fraction (IE) of the crude extract. 5µg of each fraction was loaded onto each line. Arrow indicates the expressed ChiABP-39 chitinase. **B**) Purification of chitinase ChiABP-39 with the Q-Sepharose Fast Flow column. SDS-PAGE electrophoresis on 12% polyacrylamide gels. The ChiABP-39 recombinant protein appears partially purified in the fraction not bound to the Q-Sepharose column and the contaminating proteins were eluted with the elution buffer gradient. MW = molecular weight marker; SE = soluble extract; FT = flow through fraction; WF = wash fraction. Gels were stained with Coomassie blue.

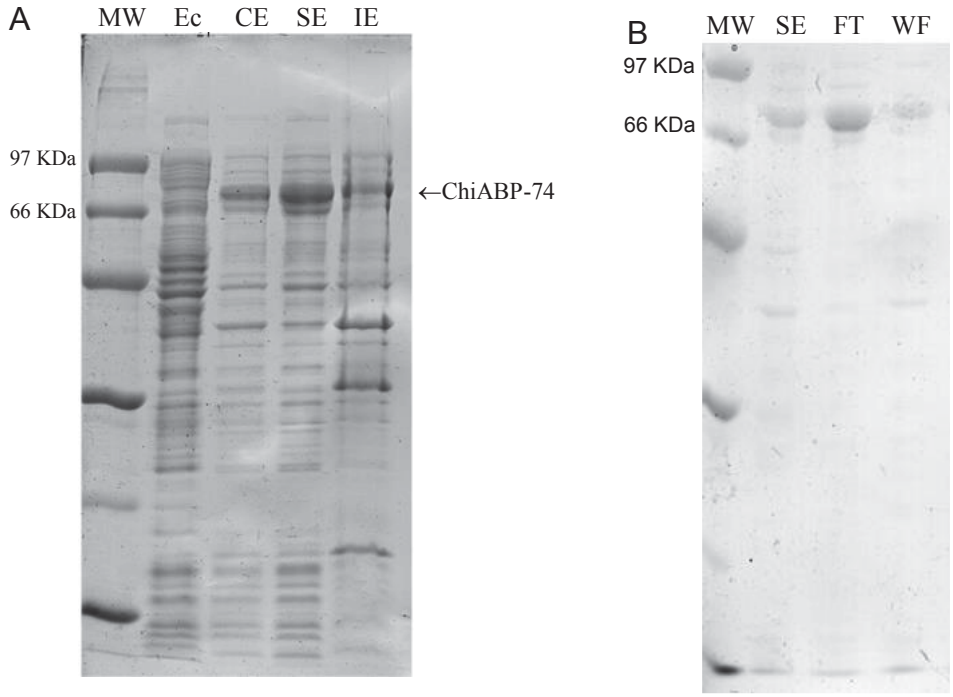


Figure 2. Expression and purification of the ChiABP-74 chitinase. Electrophoresis SDS-PAGE in 12% polyacrylamide gel. **A**) Fractions obtained from the expression test and solubility of the ChiABP-74 chitinase. *E. coli* BL21 (DE3) transformed with the plasmid pET29a-*chiABP-74* was induced with 0.5mM IPTG at 16°C for 18h in LB medium. After harvested, cells were lysed by sonication and the crude extract (EC) obtained was centrifuged (15,000 x g / 20 min / 4°C) yielding the supernatant (SE) and the insoluble fraction (IE) of the crude extract. 5µg of each fraction was loaded onto each line. Arrow indicates the expressed ChiABP-74 chitinase. **B**) Purification of chitinase ChiABP-74 with the Q-Sepharose Fast Flow column. SDS-PAGE electrophoresis on 12% polyacrylamide gels. The ChiABP-74 recombinant protein appears partially purified in the fraction not bound to the Q-Sepharose column and the contaminating proteins were eluted with the elution buffer gradient. MW = molecular weight marker; SE = soluble extract; FT = flow through fraction; WF = wash fraction. Gels were stained with Coomassie blue.

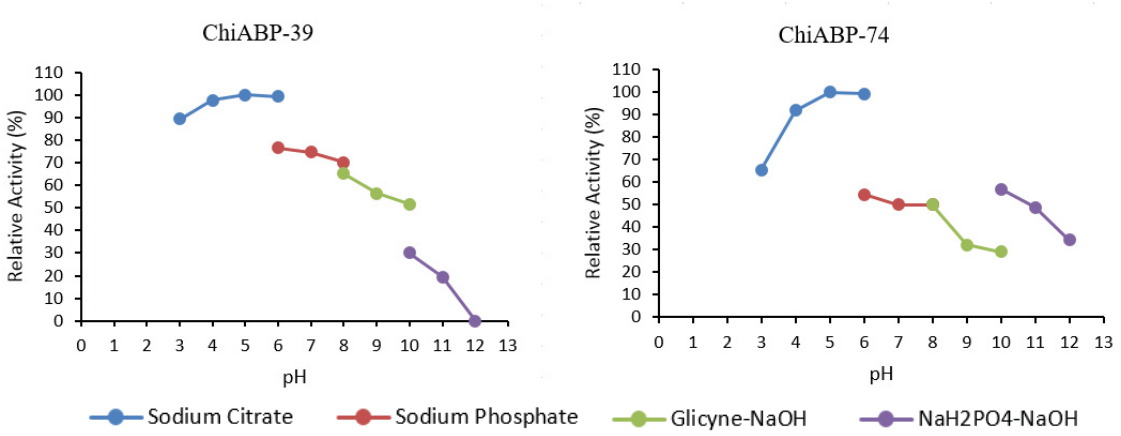


Figure 3. Effect of pH on the activity of the ChiABP-39 and ChiABP-74 chitinases. The activity due to the optimal pH range for activity of chitinases was determined using 100µl of colloidal chitin and 5µg of enzyme in a 200µL reaction system during 30 minutes at 50°C. Different buffers were used at a concentration of 50 mM with a pH ranging from 3.0 to 12.0. The ChiABP-39 and ChiABP-74 chitinases showed maximum activity in sodium citrate buffer at pH 5.0.

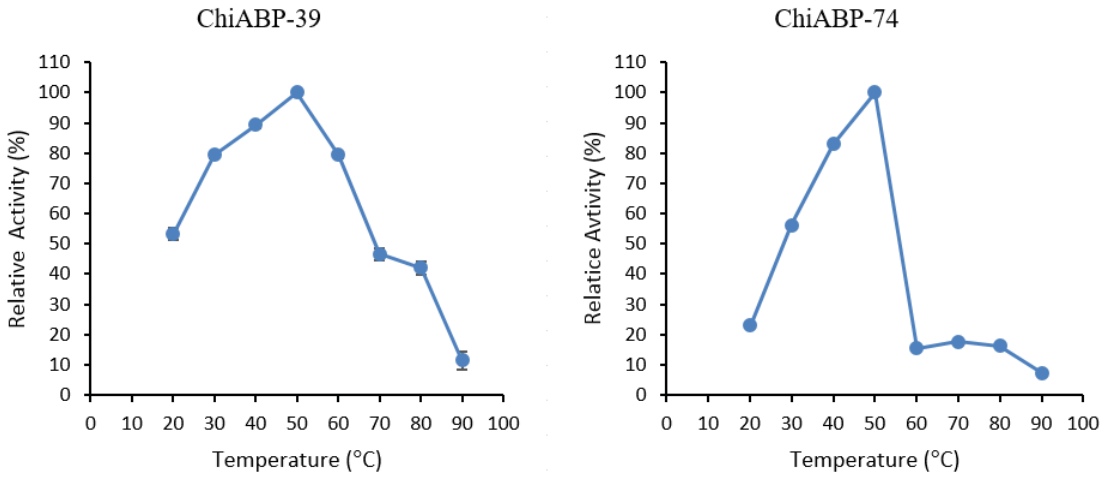


Figure 4. The effect of the temperature on the activity of the ChiABP-39 and ChiABP-74. The chitinases activity was determined using 5µg of enzyme and 100µl of colloidal chitin in 50mM of sodium citrate buffer (pH 5.0) for 30 minutes at different temperatures (20-90°C). The chitinases ChiABP-39 and ChiABP-74 showed maximum activity at 50°C.

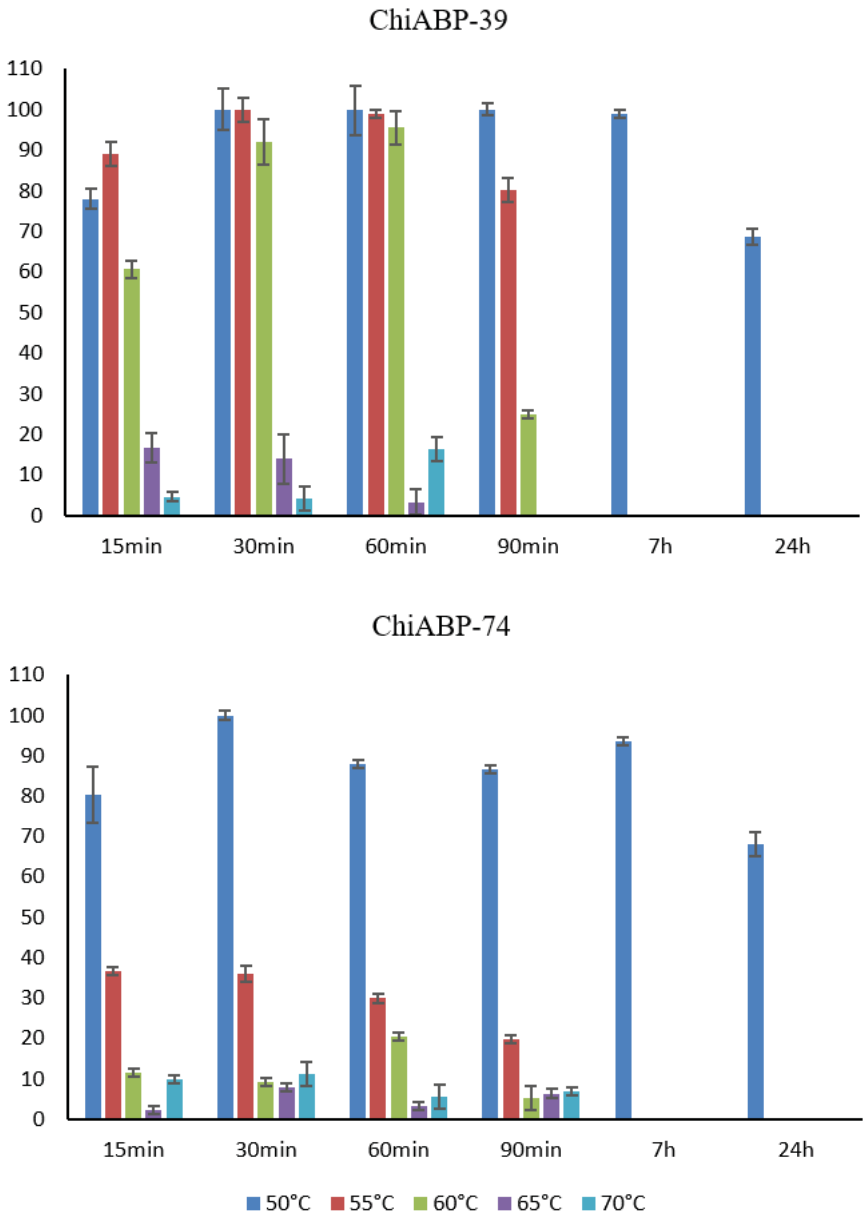


Figure 5. Thermal stability of chitinases ChiABP-39 and ChiABP-74. The enzyme was incubated at different temperatures (50°C to 80°C) for several time intervals (15 min to 24 hours) in 50mM sodium citrate buffer (pH 5.0) and its activity was tested using 5µg of chitinase and 100µl of colloidal chitin in 50mM sodium citrate buffer (pH 5.0) for 30min. The chitinases ChiABP-39 and ChiABP-74 showed a high thermal stability at 50°C.

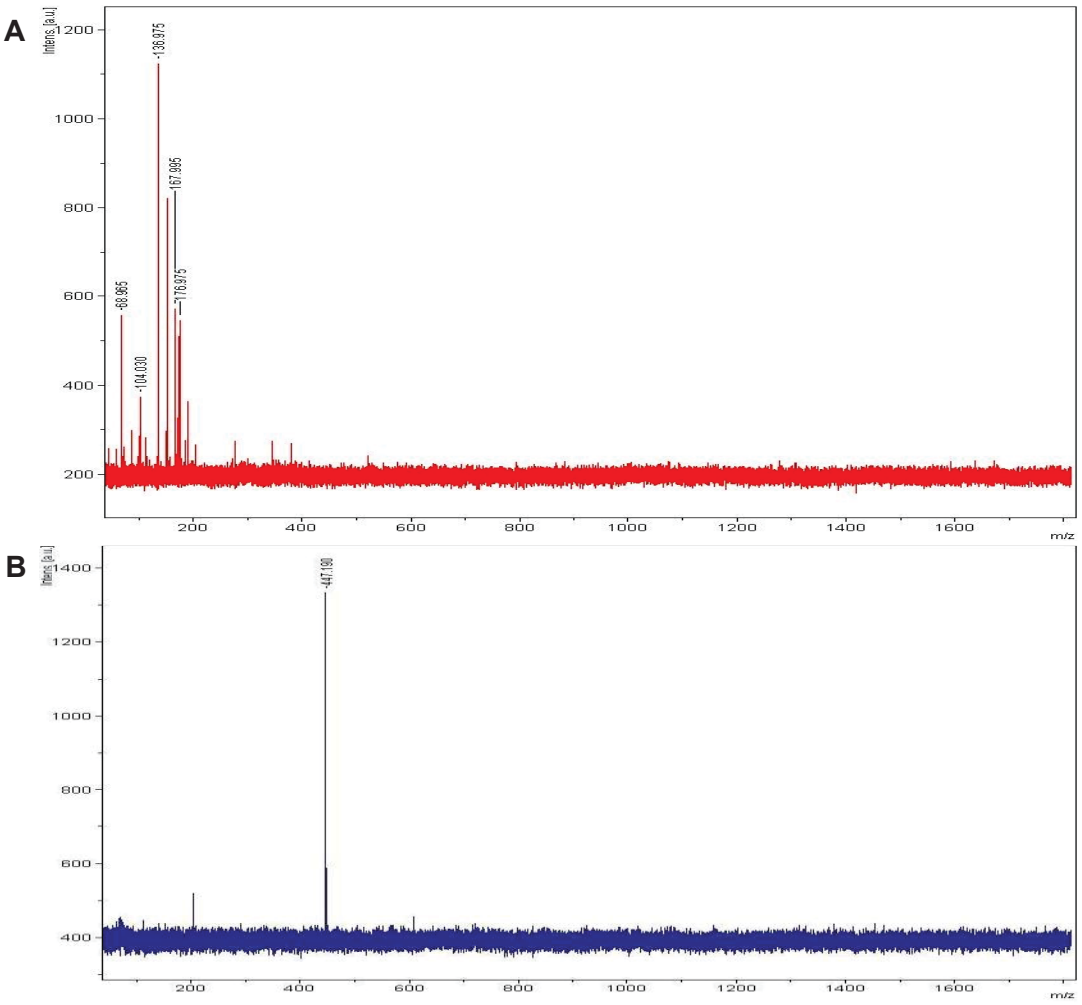


Figure 6. MALDI-TOF mass spectra of hydrolysis products of 2% chitin colloidal after 24 hours of reaction with ChiABP-39 at 40°C. **A**) Control reaction: colloidal chitin without the ChiABP-39 enzyme. **B**) Hydrolysis reaction: 2% colloidal chitin and 5µg of ChiABP-39. The *m/z* 447.190 corresponds to $(GlcNAc)_2[H^+]_4$.

Gene *chiABP-39* of *Bacillus* sp. ABP14

```
1       caaacaaaccagctcctattacgccgccaattgagatcattgtaatgtgacgtattttga
61      ggtctttcttaagatttttatccaactcttatacatccccttccaaacgtaattttgaaa
121     tttaccccgcattaacgggcagtaaaacccccacctcaaaatttagcggaagcaaagaag
181     ttaggcgggaatctactgcccgtaaatgcccgattggtgagggctgataatcagcggggg
241     atgcccccactgattaaagtttcactttatgtaagagtattatgcgaaaaaccttggcct
301     acactttaaaatatatcaaaatattctgaatttaacaaatgtaatacaaactcctacaaa
361     aattattacattaggtttactaattttccatatacaattgttgagaaaattatagtcacc

                -35                     -10
421     ttacaaaagaatgaattttcagatattattttaaataaacaagattatacatctagacaact
                  RBS
481     ttttgtataggagtgttgatATGTTAAACAAGTTCAAATTCATTTGTTGTACGTTAGTAA

541     TTTTTTTACTACTACCACTAGCCCCCTTTCAAGCACAAGCAGCAAACAATTTAGGATCAA
601     AACTACTCGTTGGATACTGGCATAACTTTGATAACGGTACTGGCATTATTAAATTAAGAG
661     ACGTTTCACCAAAATGGGATGTAATCAATGTATCTTTCGGGGAAACTGGTGGTGATCGTT
721     CTACTGTTGAATTTTCTCCTGTGTATGGTACAGATGTAGAATTCAAATCAGATATTTCTT
781     ATTTAAAAAGTAAAGGAAAGAAAGTAGTTCTTTCAATAGGTGGACAAAACGGGGTCGTTT
841     TACTTCCTGACAATGCTGCTAAGCAACGTTTCATTAATTCCATACAATCTCTAATCGATA
901     AATACGGTTTTGATGGAATAGATATTGACCTTGAATCAGGTATTTACTTAAACGGAAATG
961     ACACTAATTTCAAAAACCCAACTACTCCTCAAATCGTAAATCTTATTTCGGCTATTCGAA
1021    CAATCTCAGATCATTATGGTCCCGACTTTCTATTAAGTATGGCTCCTGAAACAGCTTATG
1081    TTCAAGGAGGTTATAGCGCATACGGAAGCATCTGGGGGGCATATTTACCGATTATTTACG
1141    GAGTGAAAGACAAACTGACATACATTCACGTTCAACACTACAACGCTGGCAGCGGGGTTG
1201    GAATGGACGGTAATAACTACAATCAAGGTACTGCAGACTACGAGGTCGCTATGGCAGATA
1261    TGCTCTTACATGGTTTTCCTGTAGGTGGTAATGCAAATAACATGTTCCCAGCTCTTCGTT
1321    CGGATCAAGTAATGATTGGACTTCCTGCAGCACCAGCGGCCGCTCCAAGTGGTGGGTATA
1381    TTTCTCCAACTGAAATGAAAAAAGCTTTAGATTATATCATTAAGGGAATTCCTTTCGGAG
1441    GAAAGTATAAACTTTCAAATGAGAGTGGCTATCCTGCATTCCGCGGTCTAATGTCTTGGT
1501    CTATTAATTGGGATGCAAAAAACAACTTTGAATTCTCTAGTAACTATAGAACATATTTCG

1561    ATGCGATCCCCTTGCAAAAATAAtaaaaaacaacaataggtttatatgacctattgttg

1621    tttttttatttaccatcaattaacatcgctcgttcatcaaaatcaaattaagaatgatgat
```

Figure S1. The 1681bp nucleotide sequence of *Bacillus* sp. ABP14 with probable coding and regulatory regions of *chiABP-39* gene. The start and stop codons of transcription are identified by a vertical arrow. A potential ribosome binding site (RBS) and the -10 e -35 regions of the promoter sequence are in bold underlined. The palindromic structure dowstream the termination codon (TAA) is marked by inverted arrows.

Figure S2. The 2025bp nucleotide sequence of *Bacillus* sp. ABP14 with probable coding and regulatory regions of *chiABP-74* gene. The start and stop codons of transcription are identified by a vertical arrow. A potential ribosome binding site (RBS) and the -10 e -35 regions of the promoter sequence are in bold underlined. The palindromic structure dowstream the termination codon (TAA) is marked by inverted arrows

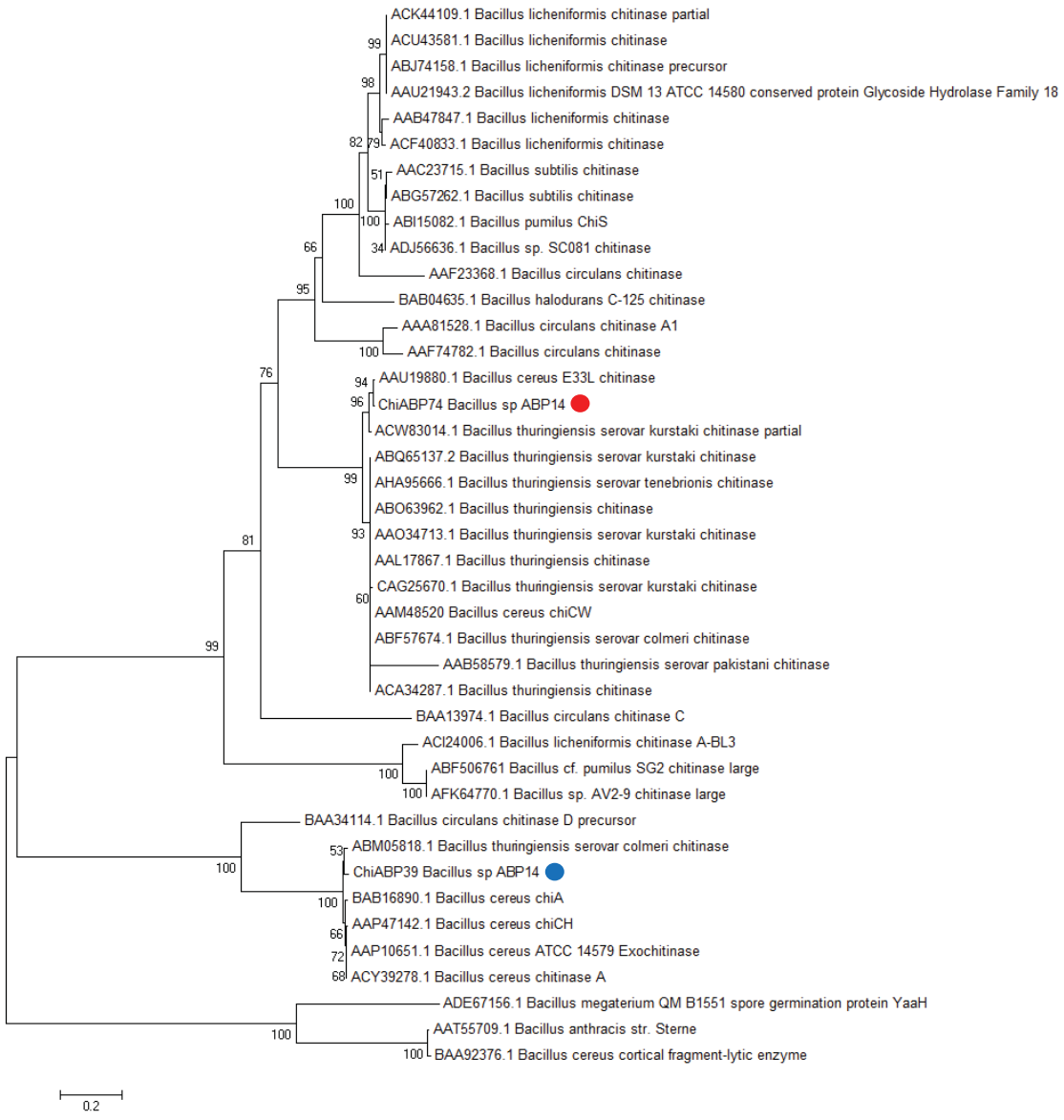


Figure S3. A neighbour-joining tree based on *Bacillus* spp characterized GH18 showing the phylogenetic position of ChiABP-39 and ChiABP-74 of *Bacillus* sp. ABP14. Bootstrap values based on 1000 replications are shown at branch nodes. The sequences were obtained from the CAZy database and the entries are indicated first the strains.

Alignment of GH18 domain of the ChiABP-39

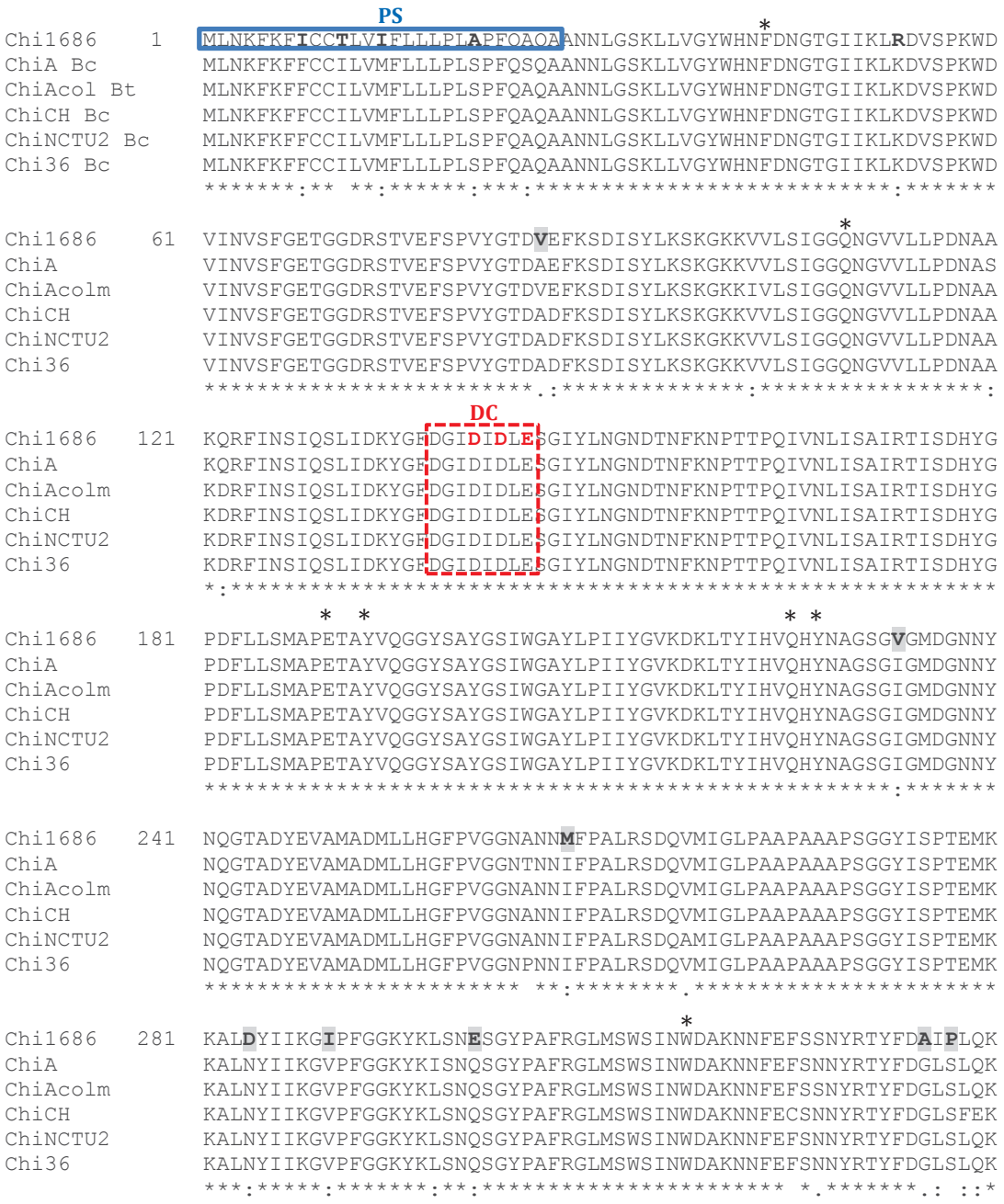


Figure S4. Alignment of the amino acid sequence of the ChiABP-39 with the most similar chitinases from *Bacillus* spp. at CAZy database. The grey amino acids differentiate the ChiABP-39 to the others sequences. The dotted region locates the catalytic site of GH18 and amino acids in red identify aspartate (D) and glutamate (E) in active site of the enzyme. The residues Phe-43, Gln-109, Asp-143, Glu-190, Tyr-193, Gln-225, Tyr-227, Tyr-333 which are involved with catalytic functions are marked with an asterisks (*). The abbreviations indicated the chitinases - ChiABP-39: this work; ChiA *B. cereus* CH (BAB16890.1), ChiAcolm *B. thuringiensis colmeri* 15A3 (ABF57674.1), ChiCH *B. cereus* strain 28-9 (AAP47142.1), ChiNCTU2 *B. cereus* NCTU2 (ACY39278.1), Chi36 *B. cereus* ATCC14579 (AAP10651.1)

Alignment of GH18 domain of the ChiABP-74

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5          15         25         35         45         55
ChiABP-74       QKIVGYFPSW GIYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC
ChiB            QKIVGYFPSW GVYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC
ChiCW           QKIVGYFPSW GVYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC
ChiA74          QKIVGYFPSW GVYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC
ChiBcol         QKIVGYFPSW GVYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC
Chi97           QKIVGYFPSW GVYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC
ChiK            QKIVGYFPSW GVYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC
ChiBt           QKIVGYFPSW GVYGRNYQVA DIDASKLTHL NYAFADICWN GKHGNPSTHP DNPNKQTWNC

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    65         75         85         95        105        115
ChiABP-74       KESGVPLQNK EVPNGTLVLG EPWADVTKSY PGSGTTWEDC DKYARCGNFG ELKRLKAKYP
ChiB            KESGVPLQNK EVPNGTLVLG EPWADVTKSY PGSGTTWEDC DKYARCGNFG ELKRLKAKYP
ChiCW           KESGVPLQNK EVPNGTLVLG EPWADVTKSY PGSGTTWEDC DKYARCGNFG ELKRLKAKYP
ChiA74          KESGVPLQNK EVPNGTLVLG EPWADVTKSY PVSGTTWEDC DKYARCGNFG ELKRLKAKYP
ChiBcol         KESGVPLQNK EVPNGTLVLG EPWADVTKSY PGSGTTWEDC DKYARCGNFG ELKRLKAKYP
Chi97           KESGVPLQNK EVPNGTLVLG EPWADVTKSY PGSGTTWEDC DKYARCGNFG ELKRLKAKYP
ChiK            KESGVPLQNK EVPNGTLVLG EPWADVTKSY PGSGTTWEDC DKYARCGNFG ELKRLKAKYP
ChiBt           KESGVPLQNK EVPNGTLVLG EPWADVTKSY PGSGTTWEDC DKYARCGNFG ELKRLKAKYP

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   125        135        145        155        165        175
ChiABP-74       HLKTIISVGG WTWSNRFSDM AADEKTRKVF ADSTVDFLRE YGFDGVDLDW EYPGVETIPG
ChiB            HLKTIISVGG WTWSNRFSDM AADEKTRKVF AESTVAFLRA YGFDGVDLDW EYPGVETIPG
ChiCW           HLKTIISVGG WTWSNRFSDM AADEKTRKVF AESTVAFLRA YGFDGVDLDW EYPGVETIPG
ChiA74          HLKTIISVGG WTWSNRFSDM AADEKTRKVF AESTVAFLRA YGFDGVDLDW EYPGVETIPG
ChiBcol         HLKTIISVGG WTWSNRFSDM AADEKTRKVF AESTVAFLRA YGFDGVDLDW EYPGVETIPG
Chi97           HLKTIISVGG WTWSNRFSDM AADEKTRKVF AESTVAFLRA YGFDGVDLDW EYPGVETIPG
ChiK            HLKTIISVGG WTWSNRFSDM AADEKTRKVF AESTVAFLRA YGFDGVDLDW EYPGVETIPG
ChiBt           HLKTIISVGG WTWSNRFSDM AADEKTRKVF AESTVAFLRA YGFDGVDLDW EYPGVETIPG

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   185        195        205        215        225        235
ChiABP-74       GSYRPEDKQN FTLLLQDVRN ALTKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW
ChiB            GSYRPEDKQN FTLLLQDVRN ALNKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW
ChiCW           GSYRPEDKQN FTLLLQDVRN ALNKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW
ChiA74          GSYRPEDKQN FTLLLQDVRN ALNKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW
ChiBcol         GSYRPEDKQN FTLLLQDVRN ALNKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW
Chi97           GSYRPEDKQN FTLLLQDVRN ALNKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW
ChiK            GSYRPEDKQN FTLLLQDVRN ALNKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW
ChiBt           GSYRPEDKQN FTLLLQDVRN ALNKAGAEDG KQYLLTIASG ASQRYADHTE LKKISQILDW

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   245        255        265        275        285        295
ChiABP-74       INIMTYDFHG GWEATSNHNA ALYKDPNDPA ADTNFYVDGA IDIYTNEGVP ADKLVLGVPF
ChiB            INIMTYDFHG GWEATSNHNA ALYKDPNDPA ANTNFHVDGA INVYTNEGVP VDKLVLGVPF
ChiCW           INIMTYDFHG GWEATSNHNA ALYKDPNDPA ANTNFYVDGA INVYTNEGVP VDKLVLGVPF
ChiA74          INIMTYDFHG GWEATSNHNA ALYKDPNDPA ANTNFYVDGA INVYTNEGVP VDKLVLGVPF
ChiBcol         INIMTYDFHG GWEATSNHNA ALYKDPNDPA ANTNFYVDGA INVYTNEGVP VDKLVLGVPF
Chi97           INIMTYDFHG GWEATSNHNA ALYKDPNDPA ANTNFYVDGA INVYTNEGVP VDKLVLGVPF
ChiK            INIMTYDFHG GWEATSNHNA ALYKDPNDPA ANTNFYVDGA INVYTNEGVP VDKLVLGVPF
ChiBt           INIMTYDFHG GWEATSNHNA ALYKDPNDPA ANTNFYVDGA INVYTNEGVP VDKLVLGVPF

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
```

```
                305        315        325        335        345        355
ChiABP-74     YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV
ChiB          YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV
ChiCW         YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV
ChiA74        YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV
ChiBcol       YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV
Chi97         YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV
ChiK          YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV
ChiBt         YGRGWKSCGK ENNGQYQPCK PGSDGKLASK GTWDDYSTGD TGVYDYGDLA ANYVNKNGFV

              ....|....| ....|....| ....|....| ....|....| ....|....| ..
                365        375        385        395        405
ChiABP-74     RYWNDVAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LSGAMFWELS GD
ChiB          RYWNDTAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LNGAMFWELS GD
ChiCW         RYWNDTAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LSGAMFWELS GD
ChiA74        RYWNDTAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LSGAMFWELS GD
ChiBcol       RYWNDTAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LSGAMFWELS GD
Chi97         RYWNDTAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LSGAMFWELS GD
ChiK          RYWNDTAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LSGAMFWELS GD
ChiBt         RYWNDTAKVP YLYNATTGTF ISYDDNESMK YKTDYIKTKG LSGAMFWELS GD
```

Figure S5. Alignment of the amino acid sequence of GH18 domain of the ChiABP-74 with the catalytic domain of the most similar chitinases from *Bacillus* spp. at CAZy database. The grey amino acids differentiate the GH18 domains of ChiABP-74 to the others sequences. The dotted region locates the catalytic site of GH18 and amino acids in red identify aspartate (D) and glutamate (E) in active site of the enzyme. The abbreviations indicated the chitinases - ChiABP-74: this work; ChiB: *Bacillus cereus* CH (BAB16891.1), ChiCW: *Bacillus cereus* 28-9 (AAM48520.2), ChiA74: *Bacillus thuringiensis* LBIT-82 serovar *kenyae* (AAL17867.1), ChiBcol: *Bacillus thuringiensis* serovar *colmeri* (ABF576741), Chi97: *Bacillus thuringiensis* 97243-1 (ACA34287.1), ChiK: *Bacillus thuringiensis* serovar *kurstaki* (AAO34713.1), ChiBt: *Bacillus thuringiensis* BT7218/010 (ABO63962.1).

Alignment of Fn3-like domain

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   5         15         25         35         45         55
ChiABP-74       PTNVKNIAVT NKNSNSVQLN WTASTDNVGV TEYEITAGEE KWSTTTNSIA IKNLKPNTEY
ChiCW           PTNVKNVVVT NKNSNSVQLN WTASTDNVGV MEYEITAGEE KWSTTTNSIT IKNLKPNTEY
ChiB            PTNVKNIVVT NKNSNSVQLN WTVSTDNVGV TEYEITAGEE KWSTTTNSIT IKNLKPNTEY
ChiBt           PTNVKNIVVT NKNSNSVQLD WTASTDNVGV TEYEITAGEE KWSTTTNSIT IKNLKPNTEY
ChiA74          PTNVKNIVVT NKNSNSVQLN WTASTDNVGV TEYEITAGEE KWSTTTNSIT IKNLKPNTEY
ChiBcol         PTNVKNIVVT NKNSNSVQLN WTASTDNVGV TEYEITAGEE KWSTTTNSIT IKNLKPNTEY
Chi97           PTNVKNIVVT NKNSNSVQLN WTASTDNVGV TEYEITAGEE KWSTTTNSIT IKNLKPNTEY
ChiK            PTNVKNIVVT NKNSNSVQLN WTASTDNVGV TEYEITAGEE KWSTTTNSIT IKNLKPNTEY

                ....|....| ...
                    65
ChiABP-74       TFSILAKDAA GNK
ChiCW           TFSIIAKDAA GNK
ChiB            KFSVIAKDAA GNK
ChiBt           KFSIIAKDAA GNK
ChiA74          TFSIIAKDAA GNK
ChiBcol         KFSIIAKDAA GNK
Chi97           KFSIIAKDAA GNK
ChiK            KFSIIAKDAA GNK
```

Alignment of CBM2 domain

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   5         15         25         35         45         55
ChiABP-74       SNATFSVTSN WGSGYNFSIV IKNSGTTSIK NWKLEFDYNG NLTQVWDSKI SSKINNHYVI
ChiB            ~TATFSVTSN WGSGYNFSII IKNNGTTPIK NWKLEFDYSG NLTQVWDSKI SSKTNNHYVI
ChiCW           GTATFSVTSN WGSGYNFSII IKNNGTTPIK NWKLEFDYSG NLTQVWDSKI SSKTNNHYVI
ChiA74          GTATFSVTSN WGSGYNFSII IKNNGTTPIK NWKLEFDYSG NLTQVWDSKI SSKTNNHYVI
ChiBcol         GTATFSVTSN WGSGYNFSII IKNNGTNPIK NWKLEFDYSG NLTQVWDSKI SSKTNNHYVI
Chi97           GTATFSVTSN WGSGYNFSII IKNNGTNPIK NWKLEFDYSG NLTQVWDSKI SSKTNNHYVI
ChiK            GTATFSVTSN WGSGYNFSII IKNNGTNPIK NWKLEFDYSG NLTQVWDSKI SSKTNNHYVI
ChiBt           GTATFSVTSN WGSGYNFSII IKNNGTNPIK NWKLEFDYSG NLTQVWDSKI SSKTNNHYVI

                ....|....| ....|....| ....|....| ....
                    65         75         85
Chi3702         TNAGWNREIP PGGSITIGGA GTGTPAELVN PS.
ChiB            TNAGWNGEIP PGGSITIGGA GTGNPAELLN AVIS
ChiCW           TNAGWNGEIP PGGFITIGGA GTGNPAELLN TVIS
ChiA74          TNAGWNGEIP SGGSITIGGA GTGNPAELLN AVIS
ChiBcol         TDAGWNGEIP PGGSITIGGA GTGNSAELLN AVIS
Chi97           TNAGWNGEIP PGGSITIGGA GTGNPAELLN AVIS
ChiK            TNAGWNGEIP PGGSITIGGA GTGNPAELLN AVIS
ChiBt           TNAGWNGEIP PGGSITIGGA GTGNPAELLN AVIS
```

Figure S6 Alignment of the peptide sequences of the Fn3 and CBM2 domains of ChiABP-74 with the most similar chitinases from *Bacillus* spp. at CAZy database. The amino acids in grey differenciate the Fn3 and CBM2 domains of ChiABP-74 to the others sequences. The abbreviations indicated the chitinases - ChiABP-74: this work; ChiB: *Bacillus cereus* CH (BAB16891.1), ChiCW: *Bacillus cereus* 28-9 (AAM48520.2), ChiA74: *Bacillus thuringiensis* LBIT-82 serovar *kenyae* (AAL17867.1), ChiBcol: *Bacillus thuringiensis* serovar *colmeri* (ABF576741), Chi97: *Bacillus thuringiensis* 97243-1 (ACA34287.1), ChiK: *Bacillus thuringiensis* serovar *kurstaki* (AAO34713.1), ChiBt: *Bacillus thuringiensis* BT7218/010 (ABO63962.1).

8 CONCLUSÕES

A análise metagenômica utilizando a sequência partial do gene 16S rRNA mostrou que o composto estudado é uma fonte promissora para a prospecção de microrganismos envolvidos na degradação da biomassa vegetal, biorremediação, interação com planta e fixação de nitrogênio.

O *Bacillus* sp. ABP14 isolado do composto é uma bactéria celulolítica, quitinolítica e entomopatogênica não-obrigatória que apresenta genes únicos que a diferencia de todos os outros organismos do grupo do *Bacillus cereus*.

Os genes *chiABP-39* e *chiABP-74* presentes no genoma do *Bacillus* sp. ABP14 codificam para duas quitinases diferentes e funcionais, denominadas ChiABP-39 e ChiABP-74 respectivamente. As quitinases foram caracterizadas como mesofílicas e acídicas, com grande estabilidade térmica a 50°C, o que indica uma vantagem para aplicações em biotecnologia. A sequência de aminoácidos das enzimas apresenta alta silimaridade com as quitinases de *B. cereus* e *B. thuringiensis,* mas as enzimas ChiABP-39 e ChiABP-74 diferem nas características bioquímicas quando comparadas com as quitinases já caracterizadas dessas espécies.

REFERÊNCIAS BIBLIOGRÁFICAS

AAM, B.B.; HEGGSET, E.B.; NORBERG, A.L.; SORLIE, M.; VARUM, K.M.; EIJSINK, V.G.H. Production of chitooligosaccharides and their potential applications in medicine. **Mar Drugs**. v. 8, p.1482–1517, 2010.

AGEITEC- *Agência Embrapa de Informação Tecnológica*. MELO, I.S. Recursos Genéticos disponível em http://www.agencia.cnptia.embrapa.br/gestor/agricultura_e_meio_ambiente/arvore/CONTAG01_26_299200692526.html. Acesso em: setembro de 2016.

ARIKAN, B. High thermostable, thermophilic, alcaline, SDS and chelator resistant amylase from thermophilic *Bacillus* sp. isolate A3-15. **Bioresour Technol**. 99, 3071 – 3076, 2008.

CGEE (2010). *Química verde no Brasil: 2010-2030*. Brasília, DF: Centro de Gestão e Estudos Estratégicos. (http://www.cgee.org.br)

CHANDRASEKARAN, R.; REVATHI, K.; NISHA, S.; KIRUBAKARAN,S.A.; SATHISH-NARAYANAN, S.; SENTHIL-NATHAN, S. Physiological effect of chitinase purified from Bacillus subtilis against the tobacco cutworm Spodoptera litura Fab. **Pest Biochem Physiol** v.104, p. 65–71, 2012.

COSTA, C.Z.; ALBUQUERQUE,M.C.C.; BRUM, M.C.; CASTRO, A.M. Degradação microbiológica e enzimática de polímeros: uma revisão. **Quim Nova**, v. 38, no. 2, p. 259-267, 2015.

D'ALMEIDA, M. L. O.; VILHENA, A. *Lixo municipal: manual de gerenciamento integrado*. 2. ed. São Paulo: IPT/CEMPRE, p. 370, 2000.

DE VOS, P.; GARRITY, G.; JONES, D.; KRIEG, N.R.; LUDWIG, W.; RAINEY, F.A.; SCHLEIFER, K.-H.; WHITMAN, W.B. (Eds.). ***Bergey's Manual of Systematic Bacteriology***. Volume 3: The Firmicutes. Second Ed., 2009.

DROBNIEWSKI, F.A. *Bacillus cereus* and Related Species. **Clin Microbiol Rev** v. 6, p. 324-338, 1993.

FRITZE D. Taxonomy of the genus *Bacillus* and related genera: the aerobic endospore-forming bacteria. **Phytopathology** v. 94, p.1245-1248, 2004.

GOODAY, G.W. Physiology of microbial degradation of chitin and chitosan. **Biodegradation** v. 1, p. 177–190, 1990.

GUINEBRETIÈRE, M.H.;, AUGER, S.; GALLERON, N.; CONTZEN, M.; SARRAU, B.; BUYSER, M.L.; LAMBERET, G.; FAGERLUND, A.; GRANUM, P.E.; LERECLUS, D.; VOS, P.; NGUYEN-THE, C.; SOROKIN, A. *Bacillus cytotoxicus* sp. nov. is a new thermotolerant species of the *Bacillus cereus* group occasionally associated with food poisoning. **Int J Syst Bacteriol** v. 63, p. 31 - 40, 2012.

HONG, H.A.; DUC, L.H.; CUTTING, S.M. The use of bacterial spores formers as probiotics. **FEMS Microbiol Rev** v. 29, no. 4, p. 813-835, 2005.

JEYARAM, K.; ROMI, W.; SINGH, T.A.; ADEWUMI, G.A.; BASANTI, K.; OGUNTOYINBO, F.A. Distinct differentiation of closely related species of *Bacillus subtilis* group with industrial importance. **J Microbiol Methods** v. 87, no. 2, p.161-164, 2011.

JIMÉNEZ, G.; URDIAIN, M.; CIFUENTES, A.; LÓPEZ-LÓPEZ, A.; BLANCH, A.R.; TAMAMES, J.; KÄMPFER, P.; KØLSTO, A.-B.; RAMÓN, D.; MARTÍNEZ, J.F.; CODOÑER, F.M.; ROSSELLÓ-MÓRA, R. Description of *Bacillus toyonensis* sp. nov., a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the

species of the group by means of ANI calculations**. Syst Appl Microbiol** v. 36, no. 6, p. 383-391, 2013.

KRAMER,K.J.; MUTHUKRISHNAN, S. Insect Chitinases: Molecular biology and potential use as biopesticides. **Insect Biochem Molec Biol** v. 27, no. 11, p. 887–900, 1997.

KUDAN, S.; KUTTIYAWONG, K.; PICHYANGKURA, R. Carboxy-terminus truncations of Bacillus licheniformis SK1 CHI72 with distinct substrate specificity. **BMB Rep** v.44, p. 375-380, 2011.

KUMIRSKA, J.; WEINHOLD, M.X.; THÖMING, J.; STEPNOWSKI, P. Biomedical Activity of Chitin/Chitosan Based Materials— Influence of Physicochemical Properties Apart from Molecular Weight and Degree of N-Acetylation. **Polymers** v.3, p. 1875-1901, 2011.

MADIGAN, T.M.; MARTINKO, J.M.; PARKER, J. ***Microbiologia de Brock***. 12 ed., São Paulo: Prentice Hall, p.1160, 2010.

MARTINS, L.F.; ANTUNES, L.P.; PASCON, R.C.; OLIVEIRA, J.C.F.; DIGIAMPIETRI, L.A., BARBOSA,D.; PEIXOTO, B.M.; VALLIM, M.A.; VIANA-NIERO, C.; OSTROSKI, E.H.; TELLES, G.P.; DIAS, Z.; CRUZ, J.B.; JULIANO, L.; VERJOVSKI-ALMEIDA, S.; SILVA, A.M.; SETUBAL, J.C. Metagenomic Analysis of a Tropical Composting Operation at the São Paulo Zoo Park Reveals Diversity of Biomass Degradation Functions and Organisms. **PloS ONE** 8, e61928, 2013.

MONTEIRO, J. P. *Manual de Gerenciamento Integrado de resíduos sólidos*. Coordenação técnica Victor Zular Zveibil. Rio de Janeiro: IBAM, 200 p., 2001.

NAKAMURA, L.K.; JACKSON, M.A. Clarification of the Taxonomy of *Bacillus mycoides*. **Int J Syst Bacteriol** v. 45, p.46-49, 1995.

NAKAMURA, L.K. *Bacillus pseudomycoides* sp. nov. **Int J Syst Bacteriol** v. 48, p.1031-1035, 1998.

NAKAMURA,L.K.; ROBERTS, M.S.; COHAN, F.M. Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: a proposal for *Bacillus subtilis* subsp. *subtilis* subsp. nov, and *Bacillus subtilis* subsp. *spizizenii* subsp. nov. **Int J Syst Bacteriol** v. 49, p. 1211-1215, 1999.

National Research Council (US) Committee on Metagenomics. The dawning of a new microbial age. In: **The New Science of Metagenomic: Revealing the secrets of our microbial planet.** Washington (DC): National Academic Press, 2007. ISBN-13:978-0-309-10676-4ISBN-10: 0-309-10676-1.

OLIVEIRA, V.M.; SETTE, L.D.; FANTINATTI-GARBOGGINI, F. Preservação e prospecção de recursos microbianos. *MultiCiência: construindo a história dos produtos naturais*, p.19, 2006.

OREM, J.C. *Filogenia e perfil plasmidial de Bactérias aeróbias formadoras de endósporos isoladas de solo.* 2014. Dissertação (Mestrado em Biologia Microbiana) - Universidade de Brasília, Brasília. 2014.

PANG, H.; ZHANG, P.; DUAN, C.J.; MO, X.C.; TANG, J.L.; FENG, J.X. Identification of celulase genes from the metagenomes of compost solis and functional characterization o fone novel endoglucanase. **Curr Microbiol** v. 58, p. 404-408, 2009.

PARTANEN, P.; HULTMAN, J.; PAULIN, L.; AUVINEN, P.; ROMANTSCHUK, M. Bacterial diversity at different stages of the composting process. **BMC Microbiol** v. 10, p. 94, 2010.

PATIL, R.S.; GHORMADE, V.; DESHPANDE, M.V. Chitinolytic enzymes: an exploration. **Enzyme Microb Technol** v. 26, p. 473–483, 2000.

PRAÇA,L.B.; BATISTA, A.C.; MARTINS, E.S.; SIQUEIRA, C.B.; DIAS, D.G.S.; GOMES, A.C.M.M.; FALCÃO, R.; MONNERAT, R.G. Estirpes de *Bacillus thuringiensis* efetivas contra insetos das ordens Lepidoptera, Coleoptera e Diptera. **Pesq Agropec Bras**, v.39, n.1, p.11-16, 2004.

PUBLIC HEALTH ENGLAND. (2015). Identification of *Bacillus* species. UK Standards for Microbiology Investigations. ID 9 Issue 3. https://www.gov.uk/uk-standards-for-microbiology-investigations-smi-quality-and-consistency-in-clinical-laboratories.

SCHMEISSER, C.; STEELE, H.; STREIT, W.R. Metagenomics, biotechnology with non-culturable microbes. **Appl Microbiol Biotechnol** v.75, p.955-962, 2007.

SCHMID, A.; DORDICK, J.S.; HAUER,B.; KIENER, A.; WUBBOLTS, M.; WIITHOLT, B. Industrial biocatalysis today and tomorrow. **Nature** v.409, p.258-268, 2001.

SCHNEPF, E.; CRICKMORE, N.; VAN RIE, J.; LERECLUS, D.; BAUM, J.; FEITELSON, J; ZEIGLER, D.R.; DEAN, D.H. 1998. *Bacillus thuringiensis* and Its Pesticidal Crystal Proteins. **Microbiol Mol Biol R** v. 62, no. 3, p. 775-806, 1998.

SHERER,K.; LI, Y.; CUI, X.; EICHACKER, P.Q. Lethal and edema toxins in the pathogenesis of *Bacillus anthracis* septic shock. **Am J Respir Crit Care Med** v. 175, no. 3, p. 211–221, 2007.

SLEPECKY, R.A.; HEMPHILL, H.E. The Genus Bacillus - Nonmedical. In: DWORKIN, M.; FALKOW, S.; ROSEMBERG, E.; SCHLEIFER, K.H.; STACKEBRANDT, E. **The Prokaryotes,** Ed. Springer, v. 4, pp. 530–562, 2006.

SONGSIRIRITTHIGUL, C.; PESATCHA, P.; EIJSINK, V.G.H.; YAMABHAI, M. Directed evolution of a *Bacillus* chitinase. **Biotech J** v. 4, p. 501-509, 2009.

SOUFIANE, B.; CÔTÉ, J.C. *Bacillus weihenstephanensis* characteristics are present in *Bacillus cereus* and *Bacillus mycoides* strains. **FEMS Microbiol Lett** v. 341, p. 127 - 137, 2013.

STREIT, W.R.; SCHMITZ, R.A. Metagenomics – the key to the uncultured microbes. **Cur Opin Microbiol**. v.7, p.492-498, 2004.

STRESHINSKAYA, G.M.; SHASHKOV, A.S.; POTHEKINA, N.V.; KOZLOVA, Y.I.; TUL'SKAYA, E.M.; SENCHENKOVA, S.N.; KUDRYASHOVA, E.B.; ANAN'INA, L.N. Carbohydrate-Containing Cell Wall Polymers of some Strains of the *Bacillus subtilis* Group. **Microbiology** v. 80, no. 1, p. 21-29, 2011.

THOMAS,C.J.; BROWN, H.L.; HAWES, C.R.; LEE, B.Y.; MIN, M.K.; KING, L.A.; POSSEE, R..D. Localization of a Baculovirus-Induced Chitinase in the Insect Cell Endoplasmic Reticulum. **J Virol** v.72, no. 12, p. 10207–10212, 1998.

TURNBULL, P. C. B.; KRAMER, J.M. Bacillus, p. 296-303. In BALOWS, A.; HAUSLER,W.J.; HERRMANN, K.L.; ISENBERG, H.D.; SHADOMY, H.J. (ed.): ***Manual of clinical microbiology*****,** 5th ed. American Society for Microbiology, Washington, D.C, 1991.

UCHIYAMA, T.; MIYAZAKI, K. Functional metagenomics for enzyme discovery: challenges to efficient screening. **Cur Opinion Biotech** v.20, p.616-622, 2009.

UCHIYAMA, T.; MIYAZAKI, K.; YAOI, K. Characterization of a novel β-glucosidase from compost microbial metagenome with Strong transglycosylation activity. **J Biol Chem** v. 288, p. 18325 -18334, 2013.

VERMA, D.; KAWARABAYASI, Y.; MIYAZAKI, K.; SATYANARAYANA, T. Cloning, expression and characteristics of a novel alkalistable and thermostable xylanase encoding gene (Mxyl) retrieved from compost-soil metagenome. **PloS ONE** 8, e52459, 2013.

VINHAL-FREITAS, I.C.; WANGEN, D.R.B.; FERREIRA, A.S.; CORRÊA, G.F.; WENDLING, B. Microbial and enzymatic activity in soil after organic composting. **R Bras Ci Solo** v. 34, p. 757-764, 2010.

WANG, L.T.; LEE, F.L.; TAI, C.J.; KASAI, H. Comparison of *gyrB* gene sequences, 16S rRNA gene sequences and DNA—DNA hybridization in the *Bacillus subtilis* group. **Int J Syst Evol Microbiol** v.57, p. 1846 - 1850, 2007.

WESTERS,L.; WESTERS,E.; QUAX, W.J. Bacillus subtilis as cell factory for pharmaceutical proteins:a biotechnological approach to optimize the host organism. **Biochim Biophys Acta** v. 1694, p. 299– 310, 2004.

WISOTZKEY, J.D.; JURTSHUK, P.; FOX, G.E.; DEINHARD, G.; PORALLA, K. Comparative sequence analyses on the 16S rRNA (rDNA) of *Bacillus acidocaldarius*, *Bacillus acidoterrestris*, and *Bacillus cycloheptanicus* and proposal for creation of a new genus, *Alicyclobacillus* gen. nov. **Int J Syst Bacterio** v. 42, p. 263-269, 1992.

WONGWILAIWALIN, S.; LAOTHANACHAREON, T.; MHUANTONG, W.; TANGPHATSORNRUANG, S.; EURWILAICHITR, L.; IGARASHI, Y.; CHAMPREDA, V. Comparative metagenomic analysis of microcosm structures and lignocellulolytic enzyme systems of symbiotic biomass-degrading consortia. **Appl Microbiol Biotechnol** v. 97, p. 8941–8954, 2013.