

## Base de dados genômica de *Bacillus thuringiensis* subsp. *kurstaki* - estirpe S76





*Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Agrobiologia  
Ministério da Agricultura, Pecuária e Abastecimento*

## **Documentos 279**

### **Base de dados genômica de *Bacillus thuringiensis* subsp. *kurstaki* - estirpe S76**

*Leona Henrique Varial de Melo  
Patrícia de Medeiros Gitahy  
Mauro de Medeiros Oliveira  
Francine Yuriko Otsuka Rocha  
Leonardo Magalhães Cruz  
José Ivo Baldani*

Embrapa Agrobiologia  
Seropédica, RJ  
2011

Exemplares desta publicação podem ser adquiridos na:

**Embrapa Agrobiologia**

BR 465, km 7, CEP 23.851-970, Seropédica, RJ

Caixa Postal 74505

Fone: (21) 3441-1500

Fax: (21) 2682-1230

Home page: [www.cnpab.embrapa.br](http://www.cnpab.embrapa.br)

E-mail: [sac@cnpab.embrapa.br](mailto:sac@cnpab.embrapa.br)

**Comitê de Publicações**

Presidente: Norma Gouvêa Rumjanek

Secretária-Executivo: Carmelita do Espírito Santo

Membros: Bruno José Rodrigues Alves, Carmelita do Espírito Santo, Ednaldo da Silva Araújo, Janaína Ribeiro Costa Rouws, Luc Felicianus Marie Rouws, Luis Cláudio Marques de Oliveira, Luiz Fernando Duarte, Márcia Rodrigues Reed Coelho

Supervisora editorial: Norma Gouvêa Rumjanek

Normalização bibliográfica: Carmelita do Espírito Santo

Tratamento de ilustrações: Maria Christine Saraiva Barbosa

Editoração eletrônica: Marta Maria Gonçalves Bahia

Foto da capa: Leona Henrique Varial de Melo

**1ª edição**

1ª impressão (2011): 50 exemplares

**Todos os direitos reservados.**

A reprodução não-autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

**Dados Internacionais de Catalogação na Publicação (CIP)  
Embrapa Agrobiologia**

---

BASE de dados genômica de *Bacillus thuringiensis* subsp. *kurstaki* - estirpe S76. / Leona Henrique Varial de Melo et al. Seropédica: Embrapa Agrobiologia, 2011. 22 p. (Embrapa Agrobiologia. Documentos 279).

ISSN: 1517-8498

1. Genômica. 2. Bioinformática. 3. Sequenciamento genético. I. Gitahy, Patrícia de Medeiros. II. Oliveira, Mauro de Medeiros. III. Rocha, Francine Yuriko Otsuka. IV. Cruz, Leonardo Magalhães. V. Baldani, José Ivo. VI. Série. VII. Embrapa Agrobiologia.

572.86CDD 23.ed.

# **Autores**

## **Leona Henrique Varial de Melo**

Bolsista de Doutorado em Biotecnologia Vegetal  
UFRJ. E-mail: leonavarial@gmail.com

## **Patrícia de Medeiros Gitahy**

Analista da Embrapa Agrobiologia, BR 465,  
km 7, Seropédica, RJ, CEP 23890-000.  
E-mail: patricia.gitahy@embrapa.br

## **Mauro de Medeiros Oliveira**

Bolsista de Mestrado em Biotecnologia Vegetal  
UFRJ/Embrapa Agrobiologia.  
E-mail: mauromedeiros@agronomo.eng.br

## **Francine Yuriko Otsuka Rocha**

Bolsista de Iniciação Científica - UFRRJ/ Embrapa  
Agrobiologia. E-mail: franotsuka@hotmail.com

## **Leonardo Magalhães Cruz**

Professor da UFPR, Departamento de Bioquímica e  
Biologia Molecular, Caixa Postal 19046, Curitiba,  
PR - CEP 81531-980. E-mail: leonardo@ufpr.br

## **José Ivo Baldani**

Pesquisador da Embrapa Agrobiologia.  
BR 465, km 7, CEP 23890-000 - Seropédica, RJ.  
E-mail: ivo.baldani@embrapa.br



# Apresentação

As atitudes de usar com responsabilidade os recursos naturais (solo, água, ar, flora, fauna, energia), de preservar e conservar a natureza são cada vez mais necessárias para a sociedade moderna acarretando em uma busca constante por sistemas de produção agropecuários apoiados em princípios ecológicos e naturais.

Dentro desse cenário, a Embrapa Agrobiologia construiu o seu atual plano diretor de pesquisa, desenvolvimento e inovação com a seguinte missão “gerar conhecimentos e viabilizar tecnologias e inovação apoiados nos processos agrobiológicos, em benefício de uma agricultura sustentável para a sociedade brasileira”.

A série documentos se constitui em uma linha de publicações que visa disponibilizar informações relevantes das mais diversas etapas dos processos de pesquisa científica e tecnológica. Podem disponibilizar revisões de literatura sobre temas relevantes, relatórios técnicos, um determinado procedimento metodológico, levantamentos de campo, entre outros tipos de conteúdo.

A presente publicação intitulada “Base de dados genômica de *Bacillus thuringiensis* subsp. *kurstaki* - estirpe S76” tem indicação para todos aqueles interessados em conhecer mais sobre o assunto, portanto, boa leitura.





# Sumário

<b>Introdução .....</b>	<b>9</b>
<b>Metodologia aplicada no sequenciamento e na montagem do genoma .....</b>	<b>11</b>
<b>Informações geradas pelo sequenciamento e anotação automática .....</b>	<b>12</b>
<b>Base de dados Genômica de <i>BtkS76</i> .....</b>	<b>14</b>
<b>Perspectivas .....</b>	<b>14</b>
<b>Agradecimentos .....</b>	<b>19</b>
<b>Referências Bibliográficas .....</b>	<b>20</b>
<b>Glossário de termos .....</b>	<b>22</b>



# Base de dados genômica de *Bacillus thuringiensis* subsp. *kurstaki* - estirpe S76

---

*Leona Henrique Varial de Melo*  
*Patrícia de Medeiros Gitahy*  
*Mauro de Medeiros Oliveira*  
*Francine Yuriko Otsuka Rocha*  
*Leonardo Magalhães Cruz*

## Introdução

O sequenciamento de genomas é um campo de estudos relativamente recente, onde o primeiro genoma sequenciado de um organismo não viral foi o da bactéria *Haemophilus influenzae*, em 1995 (FLEISCHMANN et al., 1995), com 1,8 milhões de pares de base (Mpb). Após a publicação deste trabalho, genomas de diversas bactérias foram sequenciados em um curto espaço de tempo (WEISS, 2010).

Na última década, por volta de 2005, com o advento dos sequenciadores de nova geração, e conseqüentemente a rapidez na obtenção de dados, estes estudos tornaram-se muito mais frequentes. Conseqüentemente houve um aumento exponencial na quantidade de dados gerados, dobrando a quantidade de bases depositadas em bancos públicos de seqüências como o *GenBank*, sitiado no Centro Nacional de Informação Biotecnológica ou *National Center for Biotechnology Information* - NCBI (<http://www.ncbi.nlm.nih.gov/>). O NCBI conta atualmente com mais de 350.000 submissões (seqüências) resultando em um total de mais de 100 trilhões de bases depositadas. E de acordo com o ranking mundial realizado em 2009, o Brasil encontrava-se na oitava posição em número de seqüências depositadas (WEISS, 2010).

Existem atualmente 3327 projetos genomas de procariotos no *Genbank*, sendo, 153 do Domínio *Archaea* e 2525 do Domínio *Bacteria* (1017 completos, 961 rascunhos (Drafts) e 547 em andamento) (<http://www.ncbi.nlm.nih.gov/genomes/static/gpstat.html> - Acesso em 07-06-2013). Destes, 198 pertencem ao gênero *Bacillus*, e 42 (21%) são de *B. thuringiensis*, aproximadamente 20% de *B. anthracis* (espécie patogênica) e o restante, 59% pertencentes às mais de 70 espécies do gênero *Bacillus*; reflexo da grande importância e interesse no potencial agrobiotecnológico de *B. thuringiensis*.

O conteúdo genômico de *B. thuringiensis* vem sendo estudado por diversos grupos de pesquisa desde a década de 1970, com interesse principal nos genes *cry*. Diversas proteínas *Cry* com funções distintas já foram descritas e os seus respectivos genes clonados (CRICKMORE, 2006; GONZÁLES et al., 1981; KRONSTAD et al., 1981).

Em trabalho realizado pela Embrapa Agrobiologia, foi selecionada uma estirpe *BtkS76*, capaz de provocar a mortalidade da população de larvas tratadas de *Diatraea saccharalis* (agente etiológico da broca da cana-de-açúcar) com eficiência 10 vezes superior à estirpe comercial HD1, pertencente à mesma subespécie (GITAHY et al., 2007). O estudo demonstrou a presença dos genes *cry1Aa*, *cry1Ab*, *cry2Aa1* e *cry2Ab2*, sugerindo que as toxinas codificadas por estes genes, ou a combinação delas, podem ser as responsáveis pela alta mortalidade provocada por esta estirpe às larvas do inseto. As facilidades de sequenciamento em larga escala a custo baixo e rápida execução permitiram gerar um rascunho do genoma da estirpe *BtkS76* que abriu novas perspectivas para o avanço de conhecimento dos genes *cry* e de centenas de outros genes com função desconhecida.

Este documento descreve a metodologia usada no sequenciamento e montagem das sequências desse genoma, assim como apresenta as principais categorias de genes identificados através do sistema de anotação automática, denominado RAST. Em adição indica o endereço na página da Embrapa Agrobiologia que possibilitará o acesso à base de dados do genoma da estirpe *BtkS76*.

## Metodologia aplicada no sequenciamento e na montagem do genoma

O sequenciamento do DNA total (cromossômico e plasmidial) da estirpe *BtkS76* foi realizado no sequenciador tipo Illumina (SOLEXA) em parceria com a empresa Fasteris, Suíça ([www. Fasteris.com](http://www.Fasteris.com)). As informações de dois genomas completos e seis plasmídeos de estirpes de *B. thuringiensis* (Tab. 1), previamente depositados no banco de genes (*Genbank*) sitiado no NCBI, foram usadas como referência para a montagem do genoma da estirpe *BtkS76* (Tab. 1). Para tal, após o sequenciamento ter sido realizado, as sequências geradas de *BtkS76* foram alinhadas às sequências de referência, a fim de aumentar a qualidade, e facilitar o processo de montagem dos *contigs*. A validação da qualidade das sequências obtidas foi feita através do software MAQ (Mapping and Assembly with Qualities, version 0.7.1, 2008) conforme dados fornecidos pela empresa Fasteris SA.

**Tabela 1.** Sequências genômicas de origem cromossomal e plasmidial de estirpes de *B. thuringiensis* depositadas no *GenBank* e utilizadas como referência no processo de montagem do genoma de estirpe *BtkS76*.

Número de acesso no Banco de genes - NCBI	Microrganismo espécie/estirpe	Material genômico sequenciado
NC_011796.1	<i>B. thuringiensis</i>	Plasmidial pBMB7635
NC_002108.1	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> strain YBT-1520	Plasmidial pBMB2062
NC_001272.2	<i>B. thuringiensis</i> strain YBT-1520	Plasmidial pBMB9741
DQ363750.1	<i>B. thuringiensis</i>	Plasmidial pBMB67
DQ025752.1	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> strain	Plasmidial pAW63
AF050161.1	<i>B. thuringiensis</i> subsp. <i>kurstaki</i>	Plasmidial pBMB2062
AE017355.1	<i>B. thuringiensis</i> subsp. <i>konkukian</i> strain 97-27	cromossomal
CP000485.1	<i>B. thuringiensis</i> strain Al Hakam	cromossomal

As sequências obtidas foram montadas em *contigs* (trechos de sequência maiores) utilizando o programa VELVET (realizado pela empresa Fasteris SA), que é um software para montagem de genomas, concebido para utilizar sequências curtas (pequenos *reads*), como as tecnologias geradas pelo sequenciamento Sollexa ou SolliD. A ferramenta computacional RAST (Rapid Annotation using Subsystem Technology), foi utilizada com o objetivo de: 1) localizar as ORFs (*Open Reading Frames*, ou Regiões de Leitura Aberta, que são os trechos na sequência do DNA que codificam proteínas e RNAs), fazendo uso de métodos matemáticos e estatísticos, e 2) anotar os genes, atribuindo suas funções (AZIZ et al. 2008).

### **Informações geradas pelo sequenciamento e anotação automática**

O sequenciamento do genoma da estirpe *BtkS76*, tomando como base os genomas de referência listados na Tab. 1, mostrou uma cobertura de cerca de 80% do genoma quando os seis plasmídeos foram usados e mais de 92% quando as sequências completas dos genomas de *B. thuringiensis* subsp. *konkukian* estirpe 97-27 (AE017355) e *B. thuringiensis* estirpe *Al Hakam* (CP000485) foram utilizadas (Tab. 2).

O emprego do programa MAQ (Mapping and Assembly with Qualities) permitiu a obtenção de cerca de 5000 *contigs* e de 1660 *scaffolds*, o que reflete num tamanho do genoma da estirpe S76 entre 4,5 e 4,7 milhões de pares de base (Mpb). O tamanho médio de cada *contig* foi estimado em 899 pb enquanto que esse valor foi de 2845 pb para os *scaffolds*. O *contig* de maior tamanho foi calculado possuir 39364 pb enquanto que para o *scaffold* foi de 58282 pb (Tab. 3).

Após a etapa de montagem, foi realizada a validação da mesma, mostrando que cerca de 20% dos *reads* não participaram da montagem final do genoma (dados não mostrados). Este valor é alto para procedimentos de montagem de genomas. O mesmo pode ser atribuído

**Tabela 2.** Percentual total de *reads* na estirpe *BtkS76* mapeados com base nas sequências completas dos genomas e plasmídeos de outros estirpes de *B. thuringiensis* usados como referência.

Número de acesso no <i>Genbank</i> de genomas e plasmídeos usados como referência	Quantidade de <i>reads</i> mapeadas	Percentual de <i>reads</i> (%)	Quantidade de <i>reads</i> alinhados e mapeados	Percentual de <i>reads</i> alinhados e mapeados (%)
AE017355	5.211.812	41,5	4.839.016	92,8
CPO00485	5.182.208	41,2	4.816.459	92,9
6 plasmídeos*	1.164.718	9,3	936.184	80,4

\*Conforme Tab. 1.

**Tabela 3.** Resultados da montagem das sequências da estirpe *BtkS76* provenientes do sequenciamento na plataforma Illumina.

Características analisadas	<i>Contigs</i>	<i>Scaffolds</i>
Comprimento total (pb)	4.513.336	4.737.882
Bases indeterminadas	0	202.891
Nº de <i>contigs</i> com 50 pb	2028	8760
Tamanho médio	899	2845
Maior tamanho	39.364	58.292

às diferenças de cobertura de sequências dos plasmídeos presentes na estirpe *BtkS76* em relação aos usados como referência, sugerindo a presença de plasmídeos diferentes na estirpe *BtkS76*, quando comparados com os genomas utilizados como referência.

A anotação automática realizada pelo programa RAST (Rapid Annotation using Subsystem Technology), um programa que faz uma anotação (atribuição de função biológica - *in silico*), às sequências, mostrou a ocorrência de praticamente todas as categorias de genes encontrados em outras bactérias já sequenciadas e depositadas nos bancos de dados (Fig. 1). As categorias com número de genes superior a 100 foram: metabolismo de carboidratos, metabolismo de aminoácidos e derivados, metabolismo de proteínas, metabolismo

de cofatores e vitaminas, grupos prostéticos e pigmentos, além da categoria "Outros" (Miscellaneous). Outras categorias essencialmente envolvidas em atividades metabólicas em *B. thuringiensis*, mas em menor número de genes foram: resposta a estresse, plasmídeos, virulência e defesa, dormência e esporulação etc.

Somente através da anotação manual será possível confirmar de modo mais preciso todos os genes e as respectivas funções no genoma da estirpe *BtkS76*. Para tanto, faz-se necessário estabelecer uma base de dados que seja "amigável" e que permita a cura individual dos genes anotados pelo sistema automático RAST. Além disso, a mesma deve permitir a exploração do potencial biotecnológico presente no genoma, em especial daqueles 30 a 40% de genes identificados na categoria de função desconhecida. A Fig. 2 apresenta um exemplo de genes encontrados na estirpe *BtkS76* e que estão envolvidos com as etapas de esporulação da bactéria.

## Base de dados Genômica de *BtkS76*

A base de dados contendo as informações genômicas (sequências) da estirpe *BtkS76* encontra-se hospedada nos servidores da Embrapa Agrobiologia. O acesso ao genoma é realizado através do programa GAAT (Genome Assembly and Annotation Tool), gentilmente cedido pela UFPR, no endereço <http://bioinfo-ext.cnpab.embrapa.br/BH>, conforme ilustrado nas Fig. 3 e 4. Para o acesso e a prospecção do conteúdo genômico da estirpe *BtkS76*, o potencial pesquisador parceiro, de outra Unidade da Embrapa e mesmo de Instituição de Ensino e Pesquisa, receberá a senha designada pelo coordenador do projeto, após a aprovação de proposta de pesquisa em parceria com a Unidade.

## Perspectivas

Formulações comerciais à base de *B. thuringiensis* compostas por uma mistura de células, esporos e cristais, que são formados por monômeros de proteínas *Cry* em forma inativa, são usadas há mais



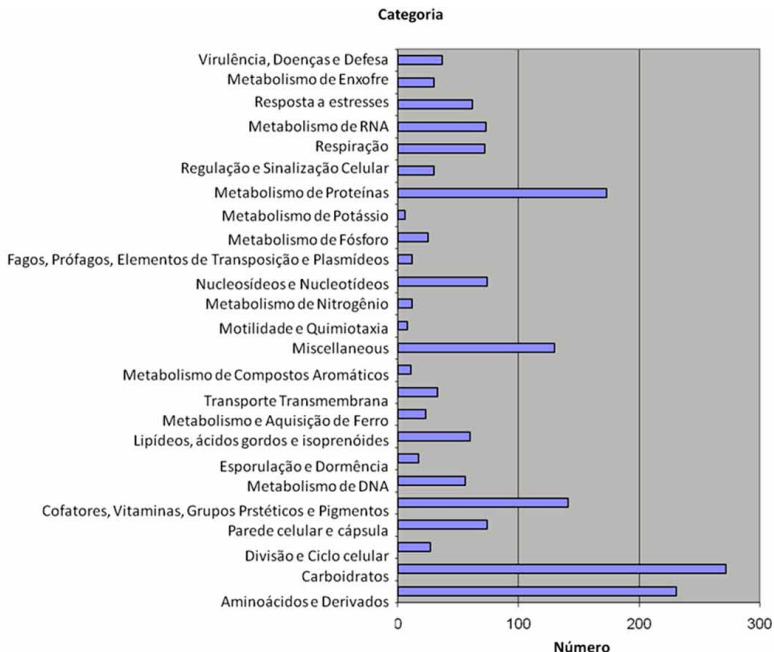
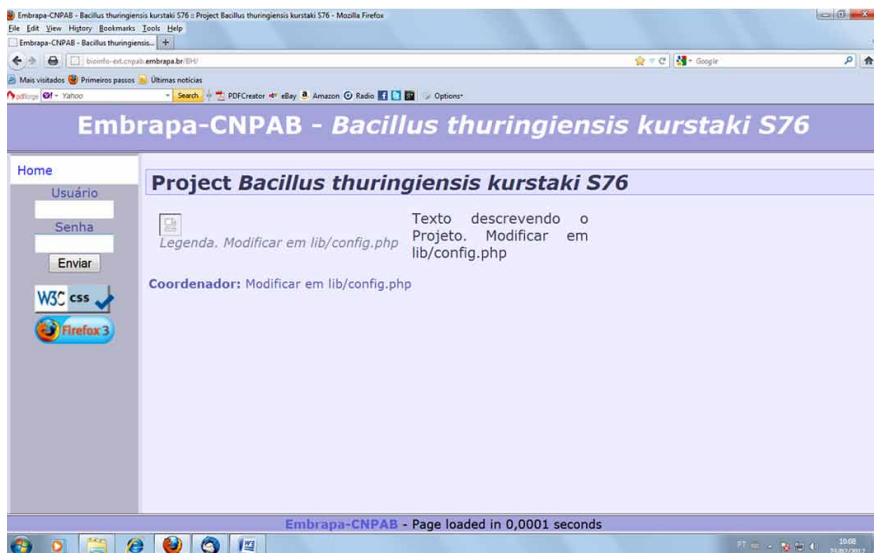


Fig. 1. Categorias de genes identificados no genoma da estirpe *BtkS76* através da anotação utilizando o sistema de anotação RAST.

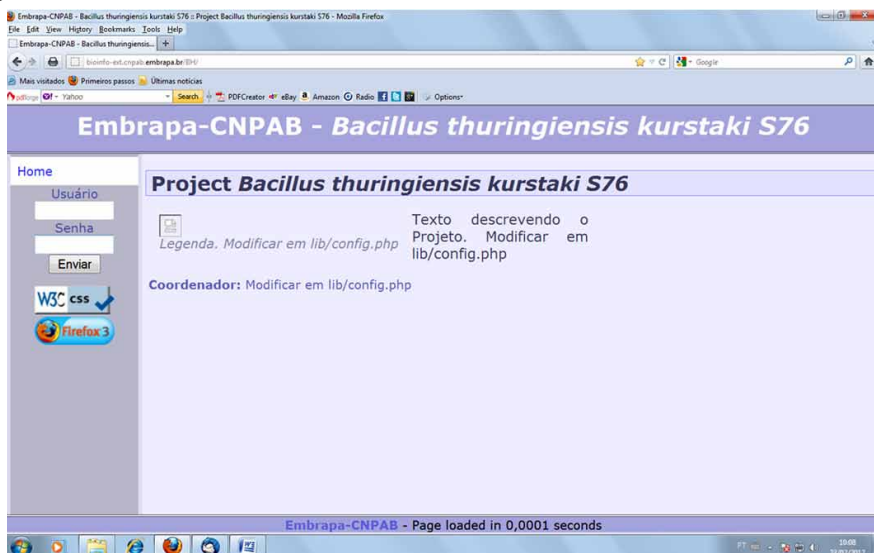
Gene Name	Start	End	Strand	Description
BT9727_5152	1	588	c	conserved hypothetical protein
rpsR	672	863	c	ribosomal protein S18 (30S ribosomal protein S18)
ssh	951	1460	c	single-stranded DNA-binding protein
rpsF	1490	1777	c	ribosomal protein S6 (30S ribosomal protein S6)
yehF	1972	3069	c	conserved GTP-binding protein
BT9727_5157	3188	3382	c	conserved hypothetical protein
BT9727_5158	3406	4284	c	mechanosensitive ion channel family protein
BT9727_5159	4544	5146	c	conserved hypothetical protein
spo0J	5173	5949	c	stage 0 sporulation protein J
soj	6026	6784	c	sporulation initiation inhibitor
spo0J	6978	7847	c	stage 0 sporulation protein J
gidB	7956	8642	c	glucose-inhibited division protein B
<b>gidA</b>	<b>8696</b>	<b>10579</b>	<b>c</b>	<b>glucose-inhibited division protein A</b>
trmE	10629	12002	c	tRNA modification GTPase
jag	12232	12846	c	jag protein (SpoIIIJ-associated protein)
spoIIIJ	12846	13562	c	stage III sporulation protein J
rnpA	13669	14010	c	ribonuclease P
BT9727_2504	14811	15023	c	conserved hypothetical protein, possible chromosomal repl.
dnaN	16310	16957	c	DNA polymerase III, beta subunit.

Fig. 2. Painel evidenciando a presença de diversos genes identificados no genoma da estirpe *BtkS76* após a exportação dos dados para o programa ARTEMIS.

A



B



**Fig. 3.** Página do website da Embrapa Agrobiologia com o endereço de acesso ao programa GAAT (**Fig. 3A**) e com a interface da base de dados do genoma da estirpe *BtkS76* com a lista dos *contigs* (**Fig. 3B**).

A

Embrapa-CNPAB - *Bacillus thuringiensis kurstaki S76*

Home  
Laboratories  
Assembly  
Annotation  
Tools  
Ivo - Admin - [log-off]  
W3C css  
Firefox 3

Statistics  
Contig #34

Size	1565 bp	Annotated ORF's		Additional Information
A:	472	C:	270	Valid: 0
G:	274	T:	548	Pending: 0
GC%	34,76%	Hypothetical:	0	Conserved Hyp.: 0
				tRNAs: 0
				rRNAs: 0
				Colors Code

EMBRAPA-CNPAB  
0034.0002  
0034.0001

Embrapa-CNPAB - Page loaded in 0,0651 seconds

B

Embrapa-CNPAB - *Bacillus thuringiensis kurstaki S76*

ORF annotation page [BthirikS76-BP01]

ORF: 0034.0002 Search

EMBRAPA-CNPAB  
0034.0002  
0034.0001

ORF # 0034.0002

Glimmer 3	RBS	Additional Information
Frame <b>F2</b>	RBS Position <b>162</b>	Base Composition A:29.40% T:35.09%
Gene Start <b>59</b>	RBS Pattern <b>TTTCT</b>	C:17.44% G:18.07%
Gene End <b>1.480</b>	Start Codon <b>ATG (ATG)</b>	Optional Start Codons <b>1</b>
ORF Size <b>1422nt</b>	New Gene Start <b>59</b>	<a href="#">Edit ORF sequence</a>
<b>474aa</b>		

Annotation Fields for ORF 0034.0002  
Status: Unverified  Frameshift  Update  
Gene name: 0034.0002

Embrapa-CNPAB - Page loaded in 0.0219 seconds

Fig. 4. Interface da base de dados do genoma da estirpe *BtkS76* com as informações referentes as ORFs presentes no *contig#34* (Fig. 4A) e o detalhamento de uma das ORFs (Fig. 4B).

de 50 anos na agricultura (BORÉM e GIÚDICE, 2008). No Brasil, o mercado para esses produtos é muito promissor, devido à presença de cerca de 500 pragas com importância econômica, que atacam diversas culturas de interesse agrotecnológico todos os anos.

Estima-se a existência de cerca de 70 produtos já desenvolvidos pelas indústrias, a partir de *B. thuringiensis*, em diversas regiões do mundo, e alguns desenvolvidos por micro indústrias brasileiras (MORAES et al., 2000), além de alguns biopesticidas importados registrados, à base de *B. thuringiensis*, tais como Dipel®, Vectobac® e Bactivec®. Este último é importado pelo governo brasileiro e distribuído em diversos estados para o controle das larvas do *Aedes aegypti* (Díptera) (CAPALBO et al., 2008).

Além dos produtos, existem hoje patentes relacionadas ao uso de genes *cry* e *cyt* para o controle de insetos lepidópteros, coleópteros e dípteros. Os Estados Unidos são líderes mundiais em patentes relacionadas à *B. thuringiensis* e seus genes *cry* e *cyt*, tendo atualmente 7369 depósitos relacionados no United States Patent and Trademark Office (USPTO). No banco de patentes europeu (European Patent Office - EPO) há 2243 patentes, enquanto no Brasil, temos apenas 76 patentes depositadas junto ao Instituto Nacional da Propriedade Industrial (INPI).

Em conclusão, podemos inferir que o sequenciamento genômico da estirpe *BtkS76*, pioneiro para a espécie *B. thuringiensis* no Brasil, abre novas perspectivas para a prospecção de genes de interesse agrobiotecnológico e a consequente exploração comercial com a possibilidade de depósito de patentes. Se considerarmos que existem cerca de 4000 genes no genoma da estirpe *BtkS76* e que 1% possuem potencial biotecnológico, teríamos então 40 genes para exploração comercial. Obviamente, haverá necessidade de esforço contínuo da pesquisa junto ao banco de dados do *BtkS76* para que o potencial biotecnológico seja efetivamente explorado e possa fornecer à sociedade brasileira informações que agregam valor ao agronegócio nacional.

## **Agradecimentos**

Os autores agradecem o apoio financeiro da Embrapa (projeto MP3.03.08.06.050.00), da FAPERJ (processo E-26/101.547/2010) e CNPq (bolsa de produtividade do último autor).

## Referências Bibliográficas

BAZIZ, R. K.; BARTELS, D.; BEST, A. A.; DEJONGH, M.; DISZ, T.; EDWARDS, R. A.; FORMSMA, K.; GERDES, S.; GLASS, E. M.; KUBAL, M.; MEYER, F.; OLSEN, G. J.; OLSON, R.; OSTERMAN, A.; OVERBEEK, R. A.; MCNEIL, L. K.; PAARMANN, D.; PACZIAN, M. T.; PARRELLO, B.; PUSCH, G. D.; GORDON, D.; REICH, C.; STEVENS, R.; VASSIEVA, O.; VONSTEIN, V.; WILKE, A.; ZAGNITKO, O. The RAST Server: rapid annotations using subsystems technology. **BMC Genomics**, v. 1, p. 75, 2008.

BORÉM, A.; del GIÚDICE, M. P. (Ed.). **Biotecnologia e meio ambiente**. Viçosa, 2007. 510 p. Edição do autor.

CAPALBO, D. M. F.; MORAES, I. O.; ARANTES, O. M. N.; REGIS, L. N.; FERNANDEZ-LARREA VEGA, O.; BENINTENDE, G. B.; GUIMARÃES, S. E.; ARRUDA, R. O. M.; MORAES, R. O. Produção de bactérias entomopatogênicas na América Latina. In: ALVES, S. B.; LOPES, R. B. (Ed.). **Controle microbiano de pragas na América Latina**. Piracicaba; FEALQ, 2008. p.239-256.

CRICKMORE, N. Beyond the spore: past and future developments of *Bacillus thuringiensis* as a biopesticide. **Journal of Applied Microbiology**, v. 101, p. 616-619, 2006.

FLEISCHMANN, R. D.; ADAMS, M. D.; WHITE, O.; CLAYTON, R. A.; KIRKNESS, E. F.; KERLAVAGE, A. R.; BULT, C. J.; TOMB, J. F.; DOUGHERTY, B. A.; MERRICK, J. M. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. **Science**, v. 269, n. 5223, p. 496- 512. 1995.

GITAHY, P. M.; SOUZA, M. T de.; MONNERAT, R. G.; ARRIGONI, E. B.; BALDANI, J. I. A Brazilian *Bacillus thuringiensis* strain highly active to sugarcane borer *Diatraea saccharalis* (Lepidoptera: Crambidae). **Brazilian Journal of Microbiology**, v. 38, n. 3, 2007.

GONZÁLES JUNIOR, J. M.; DULMAGE, H. T.; CARLTON, B. C. Correlation between specific plasmids and delta-endotoxin production in *Bacillus thuringiensis*. **Plasmid**, v. 5, p. 351-365, 1981.

MAQ. Mapping and Assembly with Qualities: version 0.6.3 2008. Disponível em: <http://maq.sourceforge.net/index.shtml>] Acesso em: 12 set 2011.

KRONSTAD, J. W.; WHITELEY, H. R. Diversity of locations of *Bacillus thuringiensis* protein genes. **Journal of Bacteriology**, v. 154, p. 419-428, 1981.

PENTEADO, M. I. O. Patentes em biotecnologia no Brasil. **Com Ciência**, 10 maio 2002. Disponível em: < <http://www.comciencia.br/reportagens/transgenicos/trans15.htm> >. Acesso em 12. Jul. 2011.

WEISS, V. A. **Estratégias de finalização da montagem do genoma da bactéria diazotrófica endofítica *Herbaspirillum seropedicae* SmR1**. 2010. Dissertação. (Mestrado em Bioquímica) - Universidade Federal do Paraná, Curitiba, 2010.

ZERBINO, D.; BIRNEY, E. Velvet: algorithms for se novo short read assembly using de bruijn graphs. **Genome Research**, v. 18, p. 821-829, 2008.

## Glossário de termos

ORF - “Open Reading Frame” ou fase aberta de leitura do DNA, que corresponde a sequência com possível região codificadora, identificada por um códon iniciador e um terminador.

*read* - Sequência de DNA obtida a partir de uma reação de sequenciamento.

*contig* - Segmento contíguo de sequências de DNA, resultante da montagem de *reads* obtidos como parte de um projeto de sequenciamento de genoma.

*scaffold* - conjunto de *contigs* organizados em sequência, representam grandes porções de um genoma.

anotação - Atribuição de características funcionais, estruturais primárias através da busca por homologia das sequências de DNA obtidas por sequenciamento ou secundárias e terciárias por predição de sequência de aminoácidos codificados. Descrição de funções e identificação de genes e seus produtos.





**Embrapa**

---

**Agrobiologia**

Ministério da  
**Agricultura, Pecuária  
e Abastecimento**

GOVERNO FEDERAL  
**BRASIL**  
PAÍS RICO É PAÍS SEM POBREZA