

Desenvolvimento de Marcador Molecular para *Bacillus subtilis* Utilizando Proteomics

José Edson Fontes Figueiredo¹
Vinícios Tadeu da Silva Coelho²
Henrique Cestari De Paoli²
Wellington Bressan³
Antônio Alvaro Corcetti Purcino³
Tales Lima Rocha⁴
Maura Vianna Prates⁴

Introdução

Marcadores moleculares estão sendo amplamente utilizados na agricultura. Entre as várias aplicações práticas, destacam-se a identificação de germoplasma, a eliminação de réplicas em bancos de germoplasma, fingerprinting de linhagens e raças, identificação de genes associados com a performance da planta (genes de resistência ou tolerância a doenças e estresse, proteínas de reserva, caracteres quantitativos), entre outros. Em todos esses casos, a definição da estratégia que deverá ser adotada precisa considerar os custos para a obtenção dos marcadores. Ênfase crescente tem sido dada para identificar marcadores menos onerosos.

Marcadores bioquímicos constituem um tipo especial de marcador molecular que se caracteriza pela análise do produto da

expressão gênica. Destacam-se dois tipos principais: isoenzimas e proteínas. Esses marcadores também são definidos como marcadores fenotípicos, pois podem resultar da interação genótipo/ambiente. Embora de aplicação limitada, o uso desses marcadores em algumas áreas ou em casos especiais fornece informações seguras e úteis para a identificação de indivíduos.

Recentemente, as técnicas de SDS-PAGE, RAPD-PCR, ARDRA e seqüenciamento de genes ribossomais foram empregadas no Laboratório de Bioquímica Molecular de Microrganismos da Embrapa Milho e Sorgo, para identificar bactérias endofíticas isoladas do milho. Neste estudo, foi verificada, em todos os isolados, a presença de um polipeptídeo de aproximadamente 42 kDa e cuja expressão era bastante elevada (Figura 1). O objetivo deste trabalho consistiu em aprofundar os

¹Biólogo, Doutor, Embrapa Milho e Sorgo. Caixa Postal 151, CEP 35701-970 Sete Lagoas, MG, Brasil

²Estudante do Curso de Ciências Biológicas do Centro Universitário Metodista Izabela Hendrix, Caixa Postal 46, CEP 30160-012 Belo Horizonte, MG, Brasil

³Eng. Agr., Doutor, Embrapa Milho e Sorgo. Caixa Postal 151, CEP 35701-970 Sete Lagoas, MG

⁴Biólogo, Doutor, Embrapa - Recursos Genéticos e Biotecnologia/DF, Parque Estação Biológica Final Av. W/5 Norte, CEP 70770-900 Brasília-DF, Brasil

estudos sobre o polipeptídeo de 42 kDa (marcador bioquímico) do isolado CMS 23, visando o desenvolvimento de marcadores moleculares de DNA para a identificação de bactérias endofíticas do milho.

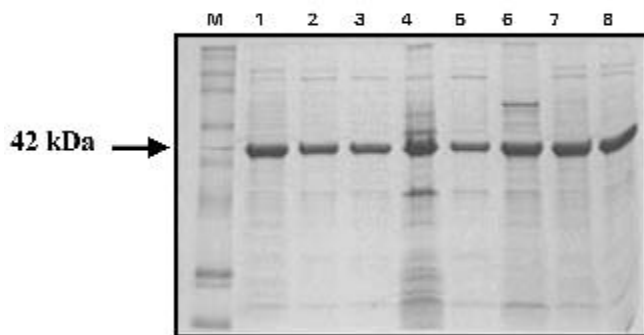


Figura 1. SDS-PAGE de extrato protéico total de bactérias endofíticas isoladas do milho. M = marcador molecular.

Extrato protéico total bacteriano foi inicialmente fracionado em sulfato de amônia, em concentrações que variaram entre 10% e 70%. Amostras dos produtos precipitados foram dialisados contra tampão PBS, concentradas e resolvidas em géis SDS-PAGE 12%, para identificação da fração que continha o polipeptídeo de interesse. Em seguida, foi realizado um gel preparativo e a banda correspondente ao polipeptídeo de 42 kDa foi cortada do gel, eletroeluída e concentrada com microcon. O polipeptídeo purificado foi aplicado novamente em gel SDS-PAGE, para verificar a pureza e a integridade da preparação (Figura 2). O seqüenciamento da região amino-terminal do polipeptídeo purificado foi realizado em seqüenciador automático de proteínas, gerando uma seqüência contendo 25 resíduos de aminoácidos (GLDMASKNAQDGLISLIQTSEGALNET). A análise de seqüências do peptídeo com proteínas depositadas no GenBank revelou alta similaridade com flagelina de *Bacillus subtilis*. Esse dado confirmou a análise prévia de seqüências de rDNA do isolado CMS 23, realizada no Laboratório de Bioquímica Molecular do NBIO, que indicou alta similaridade com seqüências de rDNA

de *B. subtilis* depositadas no GenBank. O alinhamento da seqüência de aminoácidos do peptídeo seqüenciado com 54 proteínas flagelina depositadas no GenBank revelou a presença de resíduos altamente conservados em todas as espécies analisadas, bem como a conservação desse peptídeo entre diferentes estirpes de *B. subtilis* (Figura 3).

A análise das seqüências de aminoácidos também revelou a presença de aproximadamente 50 resíduos de aminoácidos na extremidade amino-terminal da proteína flagelina, em todas as espécies estudadas, mas ausentes em CMS 23. A possibilidade de ocorrência de proteólise inespecífica durante o processo de purificação foi descartada, pois produziria ambigüidade durante o seqüenciamento do polipeptídeo e, no presente estudo, a seqüência gerada teve 100% de confiabilidade. Do mesmo modo, a ocorrência de clivagem específica na região amino-terminal do polipeptídeo de 42 kDa por atividade de protease durante a purificação parece pouco provável. É possível que o polipeptídeo seja processado pela célula bacteriana e os resíduos amino-terminal (~ 50 aa) sejam perdidos durante o processo de maturação da proteína.

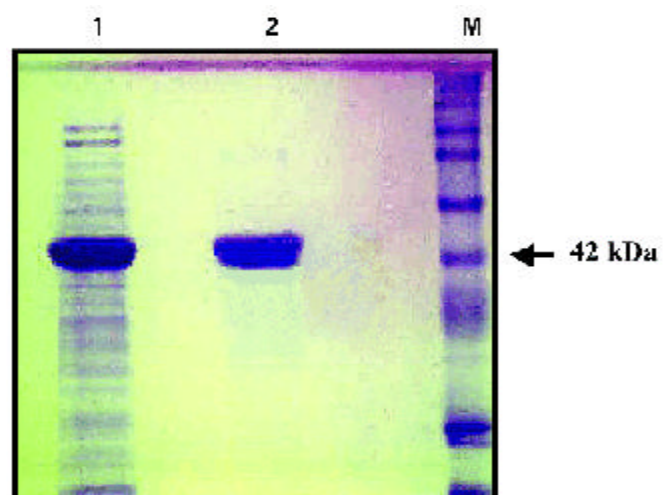


Figura 2. SDS-PAGE de extrato total (1) e da proteína de ~ 42 kDa purificada (2) do isolado CMS23. M = marcador de peso molecular .

GLDMASKNAQDGLISLIQTSEALNET	26	[CMS 23]
.....	76	[<i>Bacillus subtilis</i>] (*****)
..F.....A...T..	76	[<i>Bacillus subtilis</i>]
..E...S.....A...T..	76	[<i>Bacillus subtilis</i>] (**)
..E...S.....A...TD.	76	[<i>Bacillus subtilis</i>]
..Q..R.S.....A...D..	82	[<i>Bacillus halodurans</i>]
..K..ER..M.....M.....MQ..	82	[<i>Bacillus halodurans</i>]
..Q.....A...T..	83	[<i>Clostridium thermocellum</i>]
..E..T.....A.....	83	[<i>Oceanobacillus iheyensis</i>]
..NQ..R.....A.....	83	[<i>Clostridium haemolyticum</i>]
..NQ..R.....A.....	83	[<i>Clostridium tetani</i>]
..NQ..R.....A.....	83	[<i>Clostridium novyi</i>] (***)
..Q..R.....A...S..	83	[<i>Clostridium chauvoei</i>]
..NQ..TR.....A.....	83	[<i>Thermoanaerobacter tengcon</i>]
..Q..QR...S.....A.....	99	[<i>Desulfitobacterium hafnien</i>]
..N..I...A.....A...T..	82	[<i>Thermotoga maritima</i>]
..FQ..R.S.....T.....	83	[<i>Clostridium septicum</i>]
..NQ..S...S.....A.....	83	[<i>Clostridium tyrobutyricum</i>]
.....V...N...IA..A...M...	85	[<i>Vibrio cholerae</i>]
..NK..S.....VA.....	83	[<i>Roseburia cecicola</i>]
..Q..IR..N...M...A...D..	85	[<i>Microbulbifer degradans</i>]
..NQ..R..N...V.VT...Q..	142	[<i>Magnetococcus</i> sp]
..QQ..TR...N...F...T..Y....	83	[<i>Brachyspira</i>
<i>hyodysenteriae</i>]		
..V.VR..N...IA..A...M...	85	[<i>Vibrio parahaemolyticus</i>]
..V.VR..N...IA..A...M...	85	[<i>Vibrio cholerae</i>]
..I.VR..N...IM..A...M...	85	[<i>Vibrio vulnificus</i>]
..V.VR..N...IA..A...M...	85	[<i>Vibrio vulnificus</i>]
..NQ..R...N...F..V...Y.Q..	77	[<i>Treponema maltophilum</i>]
..V.MR..N...IA..A...M..S	85	[<i>Vibrio choicrac</i>]

Figura 3. Alinhamento parcial de seqüência de aminoácidos da proteína flagelina. Os pontos representam aminoácidos idênticos àqueles assinalados no peptídeo seqüenciado (CMS 23) e os resíduos conservados em todas as seqüências foram destacados (sombreado). Os números indicam a posição da seqüência do peptídeo nas proteínas e cada asterisco corresponde ao número de isoformas da espécie representada, cuja seqüência é idêntica ao peptídeo mostrado na figura.

Visando elucidar esta questão, a proteína flagelina do isolado CMS 23 foi novamente purificada por HPLC (Departamento de Química, UFMG) e será novamente seqüenciada.

A análise de genes codificadores para a proteína flagelina revelou a existência de seqüências de DNA específicas para a espécie *B. subtilis* (Figura 4), bem como diferenças entre várias estirpes de *B. subtilis*. Essas informações foram utilizadas para desenhar iniciadores específicos para amplificação de DNA de *B. subtilis* e, juntamente com rDNA e RFLP, poderão ser utilizados para genotipagem de isolados potencialmente úteis em processos agrícolas e biotecnológicos. Iniciadores específicos para o gene da proteína flagelina de *B. subtilis* serão utilizados para amplificação de DNA extraído de plantas de milho e, juntamente com os anticorpos, possibilitarão monitorar a eficácia de inóculos e o processo de colonização da planta hospedeira pela bactéria.

Conclusão

A análise do perfil eletroforético de proteínas totais de bactérias endofíticas isoladas do milho possibilitou a identificação de polipeptídios altamente acumulados. A purificação do polipeptídio com expressão aumentada possibilitou a produção de anticorpos policlonais em camundongos, para a identificação de *B. subtilis* em milho e também a determinação da seqüência de aminoácidos amino-terminal da proteína. Análises de seqüências de aminoácidos e de nucleotídeos codificadores da proteína flagelina possibilitaram a identificação de regiões conservadas em *B. subtilis*, mas variável em outras espécies. Essa informação foi empregada para desenhar iniciadores específicos que serão utilizados para diagnosticar e monitorar plantas de milho inoculadas com estirpes de *B. subtilis* de interesse agrícola. No presente estudo, a análise reversa por proteomics possibilitou a identificação e o desenvolvimento de marcadores moleculares para diagnóstico de *B. subtilis*.


```

Cth*  AAAA----TCAAGGAGCA-ATCAATTA-TGAGAATTAACAACAACATT-ATGGCGTACAA
Cse*  .....G.AAT.....T.....TC....T..GA...CAA..AA.G..-
Cse** .....G.AAT.....T.....TC....T..G....CAA..A..G..-
Cha  C.C.AACC.....C.ACCAGGCACAC. ....T.....TC.T.AT..AA...CAA.CA.T
Cth** .....G-.....G...-.....-.....-.....-.....-.....-
Bs1  .C...--T...-..GA.GT.A.T.CAA.A.....C...C....T..CGCG.CTC.TA.C
Bs4  .....CGCG.CTC.TA.C
Bs22 .....CGCG.CTC.TA.C
Bs5  .....CGC..CTC.TA.C
Bs7  .....CG..-CTC.GA..
-----10+-----20+-----30+-----40+-----50+-----60-

Cth*  CGCCACAGACAATTAAAAGGTAATAAGCGTCAGCGTCAGGAAAAATCCTTGGAGAAATTG
Cse*  ---CATAG.....C.....
Cse** ---CATAC..A.TA.GCC.T-C...CA.AGT...AA.....GTCAA.-...A....A
Cha  GCT.ATAG..A.TA.GGC.A-C...CA.AGT...AA.....TCAAT-...A....A
Cth** .....AGA.GAATTGGTA.....A.AGT.AAA.T.....GGCTAT-...A..G..A
Bs1  ACTAGC.GCGTC.GCTG..T.C.GG.T.AAACT.TG.T.C....ACAT-...A....A
Bs4  ACTAGC.GCGTC.GCTG..T.C.GG.T.AAACT.TG.T.C....ACAT-...A....A
Bs22  ACTAGC.GCGTC.GCTG..T.C.GG.T.AAACT.TG.T.C....ACAT-...A....A
Bs5  ACTAGC.GCGTC.GCTG..T.C.GG.T.AAACT.TG.T.C....ACAT-...A....A
Bs7  AAATGCGTCGTC.GCTG..T.C.GG.T.AAACT.TG.T.C....ACAT-...A....A
-----70+-----80+-----90+-----100+-----110+-----120-

Cth*  TCAIATGACAGGCTACAGAAACAACAGAGCTGGTGACGATGCAGCAGGTCTGTCCATATC
Cse*  .....A.....T.....C.....A...
Cse** AGC.....T.TA....A.....A..T.....AT.A..A..C..
Cha  AGC.....T.TA....A.....A..T.....ATTA..A..C..
Cth** ..T.....T..TCTT.....A..T.....A..T.....TTG.A..T..
Bs1  ..T...C.....TCTTCGC.....C.C.....T..C..T...G.....TG.G..C..
Bs4  ..T...C...TCTTCGC.....C.C.....T..C..T.....GTG.G..C..
Bs22  ..T...C...TCTTCGC.....C.C.....T..C..T.....GTG.G..C..
Bs5  ..T...C...TCTTCGC.....C.C.....T..C..T.....GTG.G..C..
Bs7  ..T...C...TCTTCGC.....C.C.....T..C..T.....GTG.G..C..
-----130+-----140+-----150+-----160+-----170+-----180-

Cth*  CGAAAAAGAGCGCAGATCAGAGCGTCTGCAAATGGCGTCAAAGAATGCCAGGACGGTAT
Cse*  .....T.....
Cse** A.....GA..A.....T-AG..CAA..ATC..GA.C.CA.A..-T.....
Cha  A.....GA..A.....T-AG..CAA..ATC..GA.T.CA.A..-T.....
Cth** A.....G..A.TA.....G-..TA.TCAA..TTC..GA.TGCC.A..-T..A..
Bs1  T.....TT.T..A.T.C.C....-AG.C.....T..A.CGCT.A..-...A..
Bs4  T.....TT.T..A.T.C.C....-AG.C.....T..A.CGCT.A..-...A..
Bs22  T.....TT.T..A.T.C.C....-AG.C.....T..A.CGCT.A..-...A..
Bs5  T.....TT.T..A.T.C.C....-AG.C.....T..A.CGCT.A..-...G..
Bs7  T.....TT.T..A.T.C.C....-AG.C.....T..A.CGCT.A..-...A..
-----190+-----200+-----210+-----220+-----230+-----240-

```

Figura 4. Alinhamento parcial de seqüências codantes para a proteína flagelina. Apenas a região do DNA empregada para desenhar os iniciadores foi representada. Os pontos e tracejados representam, respectivamente, nucleotídeos idênticos à primeira linha e "gaps" para possibilitar o melhor alinhamento das seqüências. A seqüência codificadora do peptídeo seqüenciado (25 resíduos de aminoácidos) encontra-se sombreada. **Cth*** = *Clostridium thermocellum* (acesso GenBank ZAABG0 2000025); **Cse*** = *Clostridium septicum* (acesso GenBank AB058933); **Cse**** = *Clostridium septicum* (acesso GenBank AB058935); **Cha** = *Clostridium haemolyticum* (acesso GenBank AB058939); **Cth**** = *Clostridium thermocellum* (acesso GenBank AB039910); **Bs1** = *Bacillus subtilis* isolado

Figura 4. Continuação

```

Cth* TTCACTCATCCAAACTGCAGAAGGAGCCTTAACTGAAACTCACACGCAATCCTTCCAAAGAA
Cse* .....T..C...C...AAA...
Cse** A...T.A..T.....AA.T.....T.....A.....
Cha A...T.A..T.....AA.T.....T.....A.....
Cth** A...T..T..T.....T...T...TGTG.....
Bs1 C..T..T.....AT..T..G..T..A..G...-.....A..C..T.....AGC..T..
Bs4 C..T..T.....AT..T..G..T..A..G...-.....A..C..T.....AGC..T..
Bs22 C..T..T.....AT..T..G..T..A..G...-.....A..C..T.....AGC..T..
Bs5 C..T..T.....AT..T..G..T..A..G...-.....A..C..T.....AGC..T..
Bs7 C..T..T.....AT..T..G..T..A..G...-.....A..C..T.....AGC..T..
-----250-----260+-----270+-----280-----290-----300+

Cth* TGAGAGAGCTTGCAGTACAGGCACCAAACGATACAAACCGTACAA-ACCGATAGAAGCGC
Cse* .....A.....C..C.....T..
Cse** ..AT..AT.....A...T...T.....GC..AA.....-GA
Cha .....AT..AT.....A...T...T.....GC..AA.....-GA
Cth** ..AT..AT..T.....A...T...T.....AA..AT.....-AA
Bs1 ....C.....TAC...A.....GATTCT...CC..T..T..T
Bs4 ....C.....TAC...A.....GATTCT...CC..T..T..T
Bs22 ....C.....TAC...G..A.....GATTCT...CC..T..TTT
Bs5 ....C.....TAC...A.....GATTCT...CC..T..TTT
Bs7 ....C.....TAC...A..G..A.....GATTCT...CC..T..T..T
-----310-----320+-----330+-----340-----350-----360+

Cth* AAT---TCAGCTCGAAATAAAATGAGTTGGCTGAGGAAGTAACCAGAATAGCCGAAAAGAC
Cse* G..C---.....T.G.....A.....CT...A..T.....GT...CT...G
Cse** .....A...AAG-...T..G..C..A...AA..C..A...A..GAT.....T..T..C..C..A..
Cha .....A...AAG-...T..G..C..A...AA..C..A...A..GAT.....T..T..C..C..A..
Cth** CAAAT-...AAAG.....GC.AC.TAT...A...G..GAT.....GTA.CACT..
Bs1 GAGCT-...AAAA..G..GG.CC.A..A..GTCT...G...A.....CT..TACTG.C..
Bs4 GAGCT-...AAAA..G..GG.CC.A..A..GTCT...C...A.....CT..TACTG.C..
Bs22 GAGCT-...AAAA..G..GG.CC.A..A..TCT...G...A.....CT..TACTG.C..
Bs5 GAGCT-...AAAA..G..GG.CC.A..A..GTCT...G...A.....CT..TACTG.C..
Bs7 GAGCT-...AAAA..G..GG.CC.A..AT..ATCT...G...A.....CT..TACTG.C..
-----370-----380+-----390+-----400-----410-----420+

Cth* CCAGTTCAACAC-CAGAACTCAGTGGCCGGTGGATGGAAAGATGATGGTGGCGGTTACAA
Cse* .....T.....G.....
Cse** A..A.....TAAA.T...TG.A--..AAA---...AA.SS...GA.CCA..TAAAGCT
Cha A..A.....AAA.T...TG.A--..AAA---...AA.SS...GA.CCA..TAAAGCT
Cth** AG..A..T...TAC...T...G...--..AA..ATC...G...CC..GA..GTAAA..AAA..T
Bs1 TG.....G.....C-...AA...AACTGC.C...ACC...C..T..C..AAA..CG
Bs4 AG.....G.....C-...AA...AACTGC.C...ACC...C..T..C..AAA..CG
Bs22 TG.....G.....C-...AA...AACTGC.C...ACC...C..T..C..AAA..CG
Bs5 TG.....C.....C-...A...AACTGC.C...ACC...C..T..C..AAA..CG
Bs7 TG.....G.....C-...A...AACTGC.C...ACC...C..T..C..AAA..CG
-----430-----440+-----450+-----460-----470-----480+

```

Comunicado Técnico, 67

Exemplares desta edição podem ser adquiridos na:
Embrapa Milho e Sorgo
 Caixa Postal 151 CEP 35701-970 Sete Lagoas, MG
 Fone: 0xx31 3779 1000
 Fax: 0xx31 3779 1088
 E-mail: sac@cnpmis.embrapa.br

Comitê de Publicações

Presidente: Ivan Cruz
Secretário-Executivo: Frederico Ozanan Machado Durães
Membros: Antônio Carlos de Oliveira, Arnaldo Ferreira da Silva, Carlos Roberto Casela, Fernando Tavares Fernandes e Paulo Afonso Viana

Ministério da Agricultura, Pecuária e Abastecimento

Expediente

Supervisor editorial: José Heitor Vasconcellos
Revisão de texto: Dilermando Lúcio de Oliveira
Editoração eletrônica: Tânia Mara Assunção Barbosa